Biochimica et Biophysica Acta 1850 (2015) 996–1005

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Coupled protein–ligand dynamics in truncated hemoglobin N from atomistic simulations and transition networks☆

Pierre-André Cazade [a,1], Ganna Berezovska [a], Markus Meuwly [a,b,*]

[a] Department of Chemistry, University of Basel, Klingelbergstrasse 80 4056 Basel, Switzerland
[b] Department of Chemistry, Brown University, Providence/RI, USA

ARTICLE INFO

Article history:
Received 5 July 2014
Received in revised form 1 September 2014
Accepted 3 September 2014
Available online 16 September 2014

Keywords:
Truncated hemoglobin
Network
Ligand dynamics

ABSTRACT

Background: The nature of ligand motion in proteins is difficult to characterize directly usingexperiment. Specifically, it is unclear to what degree these motions are coupled.
Methods: All-atom simulations are used to sample ligand motion in truncated Hemoglobin N. A transition network analysis including ligand- and protein-degrees of freedom is used to analyze the microscopic dynamics.
Results: Clustering of two different subsets of MD trajectories highlights the importance of a diverse and exhaustive description to define the macrostates for a ligand-migration network. Monte Carlo simulations on the transition matrices from one particular clustering are able to faithfully capture the atomistic simulations. Contrary to clustering by ligand positions only, including a protein degree of freedom yields considerably improved coarse grained dynamics. Analysis with and without imposing detailed balance agree closely which suggests that the underlying atomistic simulations are converged with respect to sampling transitions between neighboring sites.
Conclusions: Protein and ligand dynamics are not independent from each other and ligand migration through globular proteins is not passive diffusion.
General significance: Transition network analysis is a powerful tool to analyze and characterize the microscopic dynamics in complex systems. This article is part of a Special Issue entitled Recent developments of molecular dynamics.

1. Introduction

The dynamics of small ligands in proteins is relevant from a number of viewpoints. First, the dynamics and reactivity of molecules such as dioxygen (O_2), nitric oxide (NO) or carbon monoxide (CO) is highly relevant from a physiological perspective. Also, ligand migration and binding is known to affect ligand binding affinities at distal sites (allostery) [1]. Finally, the coupling of ligand and protein motion is still largely uncharacterized and may fundamentally influence the relationship between protein dynamics and function.

Recent computational work using accurate electrostatics for the unbound CO-ligand has established that, depending on the overall structure of the protein from which the free energy for migration was determined by umbrella sampling simulations, the barrier height for ligand passage differs [2]. This, together with the finding that the acceptance ratio for ligand insertion from Monte Carlo simulations is very small, suggests that ligand migration is greatly assisted by protein fluctuations. Hence, ligand transport and migration in proteins such as myoglobin is probably not purely diffusive. The relevance of fluctuations among conformational substates for the passage of ligands through the protein matrix has been inferred before based on experimental observations and simulations [3,4].

Truncated hemoglobin (trHbN) of *Mycobacterium tuberculosis* is a topical protein to investigate multi-ligand dynamics in proteins and coupling of protein and ligand motion for several reasons. First, the function of the protein is in detoxification of nitric oxide by converting it to NO_3^- following the general reaction scheme Fe(II)–(NO/O_2) + (O_2/NO) → Fe^+(III) + NO_3^-, where Fe(II)–(NO/O_2) refers to either Fe(II)NO or Fe(II)O_2. Hence, both ligands (NO and O_2) are likely to be present in unbound form at the same time in the protein. Physiologically, trHbN operates at microaerobic conditions [5] and which ligand binds first to the heme is determined by a) the concentration of the ligands and b) their individual affinities to binding to the heme. Starting from Fe(II)NO + O_2 the heme-bound NO is displaced by O_2 (Fe(II)NO + O_2 → Fe(II)O_2), followed by the reaction of NO with Fe(II)O_2 to give nitrate [6]. Given this sequence of events and the typically higher affinity of Fe(II) towards NO than to O_2 [7,8], it is likely that Fe(II)NO + O_2 is one of the early (if not the first) physiologically relevant

☆ This article is part of a Special Issue entitled Recent developments of molecular dynamics.

* Corresponding author at: Department of Chemistry, University of Basel, Klingelbergstrasse 80, 4056 Basel, Switzerland. Tel.: +41 61 267 38 21; fax: +41 61 267 38 55.
E-mail address: m.meuwly@unibas.ch (M. Meuwly).
[1] Current address: Department of Physics and Energy, University of Limerick, Castletroy, Ireland.

http://dx.doi.org/10.1016/j.bbagen.2014.09.008
0304-4165/

states in a sequence of events in NO detoxification. Second, trHbN is known to exhibit structurally well defined sites (Xenon pockets) in which unbound diatomic molecules can be potentially bound. The transition between such pockets is a thermally activated process and coupling to the protein motion was found for similar situations in myoglobin (Mb), another topologically related globular protein [2,4]. Finally, ligand migration in trHbN was characterized from transition networks which is an effective way of investigating available states and their populations in complex systems [9,10].

In the present work we address both issues, the coupling of protein and ligand motion and the dynamics of multiple (two) ligands in the same protein. To gain insight into whether and how protein dynamics is coupled to ligand motion a recently completed study on ligand motion in trHbN was extended to include the protein degrees of freedom [10] and the dynamics of multiple unbound ligands is characterized from explicit atomistic simulations. Both aspects employ state-of-the art simulation and analysis techniques suitable for quantitative investigations of condensed-phase systems.

2. Computational methods

2.1. Atomistic simulations

All standard simulations are performed with NAMD [11]. The CHARMM22 forcefield [12] is used for the protein while the NO and O_2 parametrizations are those from previous work [13,14]. The TIP3P model is used for water [15]. The total system consists of 22,197 atoms. Periodic boundary conditions are used and all bonds involving H-atoms are kept rigid with the SHAKE algorithm [16]. The system is first minimized, heated and equilibrated for 200 ps in the *NPT* ensemble with the Berendsen algorithm [17] at 300 K and 1 atm, resulting in a box of size of 78.25 × 52.55 × 52.11 Å. Subsequent simulations are performed in the *NVT* ensemble using Langevin dynamics. Non-bonding interactions are calculated within a cutoff of 12 Å. The electrostatics is handled employing Particle Mesh Ewald with a 80 × 80 × 54 mesh for the wave vector in reciprocal space. A time step of 1 fs is used and the individual trajectories are 10 ns long. Ten simulations are performed with initially either NO or O_2 in the reactive site and the second ligand in one of the experimentally identified [18] docking sites: Xe1 to Xe5.

Two types of umbrella sampling simulations are performed with CHARMM [19]. On the one hand, for the sake of comparison with our previous study [20], the same data set was used by removing the bond between NO and Fe and changing the heme forcefield from 6-liganded to a 5-liganded model [21]. On the other hand, data from the current simulations were used. A transition of interest was identified in one of the trajectories. The coordinates of the 25 frames which compose this transition are extracted to provide the initial coordinates of the 25 umbrella sampling windows. This transition is gate-controlled by the Phe62 residue and the rotation of the phenyl ring (CA–CB–CG–CD2) is a meaningful progression coordinate [20]. The umbrella sampling is carried out for dihedral angles from almost 180° (gate closed) to 0 (gate closed again). The gate is open at around 60°. The force constant is the same for all the windows (10 kcal/mol/Å) and the step between two consecutive windows is 0.06 rad. Alternatively, the Fe–O_2 distance was used as a reaction coordinate for the umbrella sampling. In this case, the sampling extends from 14.0 Å(Xe1a) to 9.0 Å with a step of 0.1 Å and with the same force constant. Each simulation is 51 ps long with the first ps reserved for equilibration. Such a protocol was found to converge the umbrella sampling simulations which was established by means of bootstrapping in earlier work on the same system [20]. Before the production phase, the H atoms are relaxed in order to ensure compatibility between NAMD and CHARMM. The carbon alpha of the proteins is constrained similarly to the NAMD simulations, except the one of Phe62. The Weighted Histogram Analysis Method is used to obtain the unbiased barrier [22,23].

2.2. Network clustering

An important step for a realistic network representation of atomistic trajectories is the choice of a suitable data clustering method. In the following two different methods are employed: the *structure*-based *k*-means clustering and the *kinetic*-based complex network analysis rooted in the MCL algorithm. *k*-means clustering [24,25] was extensively employed in our previous work which focused on the study of ligand coordinates only [10] and is used here as a benchmark for defining possible states of the system. The complex network analysis previously employed [26] allowed a straightforward incorporation of several degrees of freedom [27] that turns out to be useful here for the study of coupling between ligand and protein motion. For the details of the *k*-means clustering methods we refer the reader to the literature [24,25]. The protocol for the complex network analysis with MCL clustering is described below.

Analysis of the ligand migration in trHbN is based on a discretization of conventional order parameter time series by following *local fluctuations* and subsequent generation of a Markov-state-model. This approach was initially applied to single molecule experiments [28–30] and was recently extended for studying conventional order parameter time series [26]. Also this framework was successfully applied to the description of Fip35 folding mechanism [27]. The methodology of building a Markov-state-model contains several steps: (1) defining the microstates; (2) building a transition network, and (3) lumping together kinetically similar states. The methodology was extensively described elsewhere [26], however in what follows we recall the essentials. Software for the analysis can be found at raolab.com

2.2.1. Microstate definition

Each microstate was associated with a time snapshot and was defined according to the local fluctuations of the order parameter in the time window t_w centered around this snapshot. The comparison along the trajectory was done backwards using the leader algorithm. Given a current snapshot, the cumulative RMSD-distribution in the time window $t_w/2$ left and right of the snapshot is determined. This distribution is compared to the set of distributions – suitable for microstate definition – found so far up to this point along the trajectory. Two cumulative distributions are considered to belong to the same microstate if they pass the Kolmogorov–Smirnov test [31]. Here, the maximum difference between the two cumulative distributions, D, should fulfill the condition $D \leq \zeta\sqrt{(2/t_w)}$, where ζ is a confidence level (in the present analysis $\zeta = 0.5$ was used and t_w was variable). If the comparison fails for all distributions collected so far, a new microstate is defined [29].

2.2.2. Configuration-space-network

As a result of the discretization of the trajectory described above a time series of microstates is obtained. From this a configuration-space-network can be built, where the nodes represent microstates and the link between two nodes indicates that they were visited successively along the trajectory. From this, transitions probabilities P_{ij} between states i and j can be evaluated. In the limit of infinite sampling, detailed balance must hold which is here imposed by averaging the number of transitions in both directions.

2.2.3. Network clustering

In the next step we defined the states of the system by lumping together kinetically similar microstates. For this we applied the Markov-Clustering-Algorithm (MCL) [32]. The granularity of the clustering is characterized by the parameter p. Larger values of p increase the number of clusters, each separated by smaller free energy barriers. Hence, the interest in the present work is to decrease the value of p in order to minimize the number of states such that only the most significant of them are retained. This also minimizes noise in the clustering. Using the obtained states a Markov-state-model was built.

2.2.4. First passage time distribution

In order to justify the kinetic similarity between the original trajectory and the Markov-state-model we compared the first-passage-time (fpt) distributions for particular states. The fpt is a distribution of times which characterizes the duration to reach a given target state from any other snapshot along the trajectory. For the MD trajectory the fpt distribution can be calculated in a straightforward way by taking as a target one of the states known from *k*-means clustering [10]. For the Markov-state-model a suitable random walk on the transition matrix was generated. As a target here an MCL state was taken, which mainly contains corresponding *k*-means state. Arrival times depend only on the target state and not on the decomposition of the trajectory.

2.3. Application of local fluctuations analysis to trHbN

Transition network analysis was previously used to characterize the ligand O_2 migration in truncated hemoglobin trHbN [10]. Clustering of the states included only the O_2-positions and neglected the influence of protein fluctuation. However, as also observed previously, protein dynamics and ligand migration are probably coupled which necessitates a combined analysis [2]. Hence, a network analysis using the approach described above was employed for the O_2 coordinates and for the protein motion *separately*. Afterwards the obtained states for each of the "subsystems" were merged to a "combined state" which provided a reduced kinetic model for the complete system of ligand and protein. Specifically, if for a given time snapshot the O_2 ligand was in state *A* and the protein was in state *B*, then a combined state for the protein–ligand system was (*A*, *B*). Such a scheme was already successfully used for the analysis of Fip35 folding mechanism [27].

In the previous study of O_2 migration in truncated hemoglobin trHbN [10], 32 independent MD simulations of 2 ns each were performed. However, 3 out of the original 32 runs were not included in the final analysis, since they were found to be not representative of the dynamics exhibited by the remaining 29 trajectories which were analyzed. The choice of the run duration was justified in Ref. [20,14,9]. The simulations were started with O_2 in one of the 5 pockets identified experimentally [33,18] or from previous atomistic simulations [14], namely from the pockets Xe1, Xe2, Xe3, Xe4, and Xe5. In the latter study we analyzed part of these MD data considering 10 runs with the ligand initially in pocket Xe1, thus having a total of 20 ns trajectory. For testing the role of sampling, an additional set of 18 ns of MD data originating from 9 runs with O_2 initially in pockets Xe1, Xe2 and Xe3 (3 runs initiated in each of the pockets) was analyzed.

For the O_2 "subsystem" analysis the MCL data from the previous study was employed [10]. First, Euclidean space was discretized into cubic cells of side length 1 Å, with each cell representing a microstate of the O_2 molecule. The discretization represented a fine grained description of the molecular process. The analysis resulted in a total of 24,664 microstates and 98,163 links visited by O_2 along the complete 58 ns trajectory containing $5.8 \cdot 10^5$ snapshots. The discrete microstates time series built from the continuous O_2 trajectories were then mapped onto a transition network (see Fig. 1) as it was described above. Afterwards the MCL algorithm was applied with granularity parameter $p = 1.6$ in order to clusterize the network, resulting in a total of 2358 clusters. For the trajectory primarily investigated here (20 ns built of 10 separate runs starting from pocket Xe1 and containing $2 \cdot 10^5$ snapshots) 168 clusters were found.

For the combined protein plus ligand analysis the protein root mean square deviation (RMSD) relative to the X-ray crystal structure (Protein Bank entry 1IDR) was used as an order parameter [33]. While the protein consists of 131 amino acid residues and a prosthetic heme group, the RMSD was not determined by superimposing all residues in order to avoid the influence of strongly fluctuating parts of protein. Based on the results for root mean square fluctuations (RMSF) (see Fig. 2), the root mean square deviation was determined by superimposing residues 25–36 and 53–105 of the protein backbone (see the regions highlighted in gray in Fig. 2). The obtained RMSD time series were then analyzed according to the above described scheme. First the discretization of trajectories was done with $t_w = 501$ frames and 50 bins for the histogram leading to the network of the total of 5764 microstates and 20,124 links between them. Since the 20 ns trajectory studied here was built from 10 separate 2 ns runs, a straightforward application of the methodology without careful extraction in the time intervals around the merging points of separate runs would lead to artifacts. Therefore, in order to avoid "spurious" transitions at the merging points no discretization of the trajectory around the merging points t_m was carried out, i.e. no microstates were defined in the regions $(t_m - t_w/2, t_m + t_w/2)$. Finally, the link between the last microstate of a given MD run and the first microstate of the next one was removed. After these preparations, the MCL algorithm was applied. For most of the study the MCL granularity parameter was $p = 1.12$ and $t_w = 501$ was used, providing a total of 3 clusters with probabilities 0.50, 0.33, and 0.17. However, also other combinations of (t_w, p) parameters were discussed.

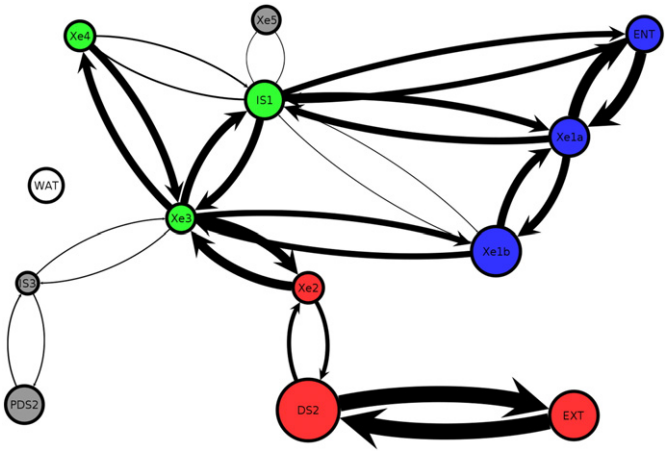

Fig. 1. Migration network for the O_2 ligand in trHbN with the most relevant sites labeled and their connectivities indicated by arrows. The network is built from *k*-means clustering of the 29 trajectories used in the previous study [10]. The size of the pockets and of the arrows is proportional to their weight in the network. Only connectivities with a relative weight larger than 0.3% are shown. Pockets belonging to channel I are colored in blue, to channel II in green, to the reactive site in red. Pockets aside from the main network are displayed in gray and the water side is colored in white.

2.4. Analysis of the multi-ligand MD trajectories

The new set of 10 trajectories, each 10 ns in length, performed to study multi-ligand dynamics were analyzed in a similar fashion as was already done in a previous study [20]. All trajectories are first reoriented on the iron atom of the heme. Then, they are reoriented by aligning the Fe – NB bond with *x*-axis and the (Fe–NB–NA) plane with *xy*-plane.

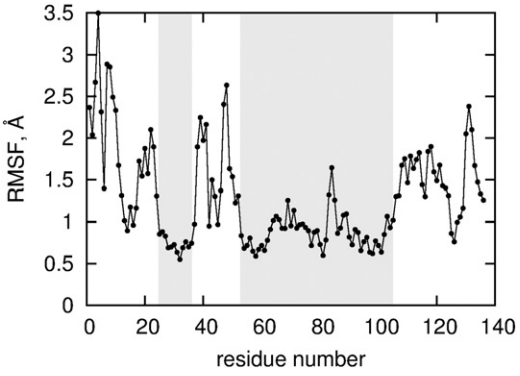

Fig. 2. C_α–RMSF from the MD simulations. The regions shown in gray represent the range of residues taken for the RMSD calculation of protein.

Finally, the Fe – His81 bond is directed towards the positive *z*-axis. After reorienting, the position of the NO- and O_2-ligand is monitored in order to follow their migration through the protein interior. The distance between Fe and the center of mass of the ligands is also analyzed as it provides a straightforward way to identify the transitions.

3. Results

3.1. Local fluctuation analysis and transition network

First, the overall dynamics of the system from the total of 20 ns simulations is analyzed. This provides the necessary information for the subsequent analysis and for building the transition network. Next, in order to have an overview of the ensuing dynamics, structural analysis was carried out using 2-dimensional cloud plots based on the ligand–protein order parameters R and ρ within each of the clusters. Here, R is the distance of the diatomic ligand to the Fe-atom and ρ is the RMSD of the protein backbone relative to X-ray crystal structure [33]. Prior to the analysis all snapshots were reoriented with respect to this reference structure.

Considering protein dynamics, an RMSD is the most common choice for an order parameter. To obtain a first insight into the correlation between the position and motion of the O_2 ligand and the protein motion 20,000 frames from a trajectory of total length 2 ns starting with O_2 from Xe1 were considered. Fig. 3 shows a cloud plot for the O_2–Fe distances and protein-RMSD of the backbone. The data clearly exhibits multiple states both for the O_2 coordinates and the protein-RMSD. Having this qualitative indication we turned to a more quantitative approach to the problem and followed a standard scheme for its analysis as was described in the Computational methods section.

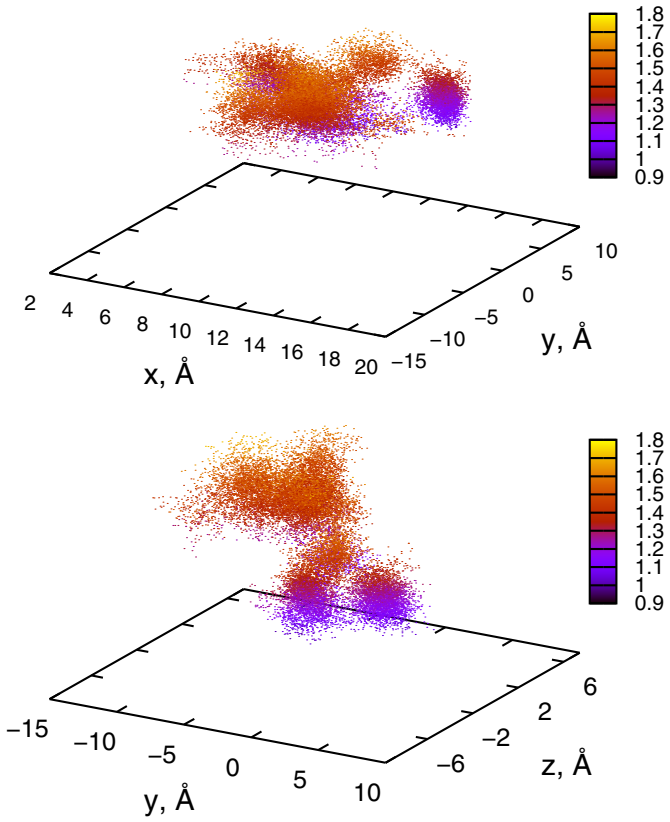

Fig. 3. The point cloud plot for correlation between the O_2–Fe distance coordinates (shown in the plane) and the protein-RMSD ρ (color coded from gray (0.9 Å) to yellow (1.8 Å)). The upper plot corresponds to the (x,y)–projection of the O_2–Fe distance and the lower one is the (x,z)–projection.

3.1.1. State definition

The definition of a "state" is an essential step in any clustering. As the clustering in MCL depends on the parameter p, a suitable value for p needs to be found by trial and error. Furthermore, sensitivity to the parameter t_w, which defines the width over which a cumulative probability distribution is calculated and thus influences the definition of microstates, is also explored. Examples for such analyses are presented in Fig. 4. Here the ρ – time series from one MD run (2 ns) and the corresponding states detected by the MCL algorithm are compared. The data was obtained for a time window $t_w = 501$ and granularity parameters $p = 1.12$ and $p = 1.15$, which results in 3 and 89 clusters (states), respectively. As it was noted in the Computational methods section, larger values of p correspond to sensitivity to smaller barriers which are detected by MCL. Thus, for $p = 1.15$ (in blue) the protein fluctuations more strongly influence the clustering and more states are detected (see also Fig. 6 for detail).

In the next step the protein states obtained were used to build "combined" states by merging them with the ligand states. The data for the ligand-only states was that from the study in Ref. [10]. Fig. 5 shows the cumulative probability distribution of the "combined" states obtained from the analysis of 10 trajectories starting from Xe1. The combined states were first ranked according to their population. The black curve corresponds to the MCL clustering with p = 1.6 applied to the discretized trajectory of the O_2 coordinates only [10]. For the combined states the following parameter pairs (t_w, p) were considered: (1001,1.15) in orange, (501,1.15) in green, (1001,1.12) in red, (501,1.12) in blue. The emerged networks of combined states contain a larger total number of nodes compared to clustering by O_2 – states only. Moreover, due to the increased number of clusters with the cumulative population ≤ 85 %, the plot indicates a more detailed description of the protein–ligand motion. Also the obtained cumulative probability distribution is sensitive to the MCL parameter value p, while the size of the window t_w does not influence the analysis.

In order to complete the discussion about how the granularity parameter p influences the MCL network, Fig. 6 reports the total number of states in the network as a function of p. The analyzed trajectory stretch contained 10 MD trajectories and the local fluctuations analysis was carried out with $t_w = 501$. The red circles both in the inset and in the main plot report the number of clusters as a function of p for the protein only, while the black circles show the case of combined states, including both the ligand coordinate and the protein-RMSD. While the network for O_2 was built with p = 1.6 and provided 168 states, the RMSD trajectory shows a strong susceptibility to small p-values. This is due to the fact that the timescales for the O_2 – motion and that for the protein are vastly different. In the MD simulations the protein structure only changes slightly during O_2 migration, and thus the sampled energy

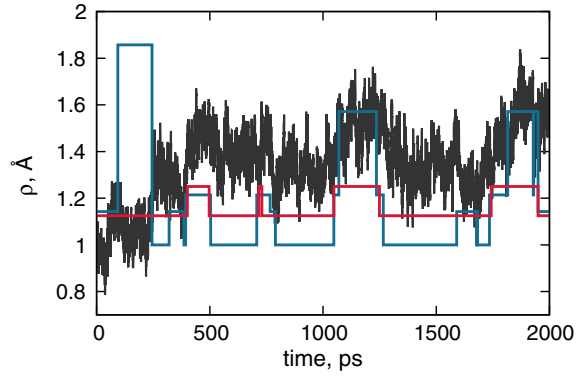

Fig. 4. The dependence of ρ on time calculated for the 2 ns MD run (in black). The step curves depict the corresponding states which have resulted from MCL clusterization for two different values of the granularity parameter p: $p = 1.12$ and $p = 1.15$ are shown in red and blue, respectively.

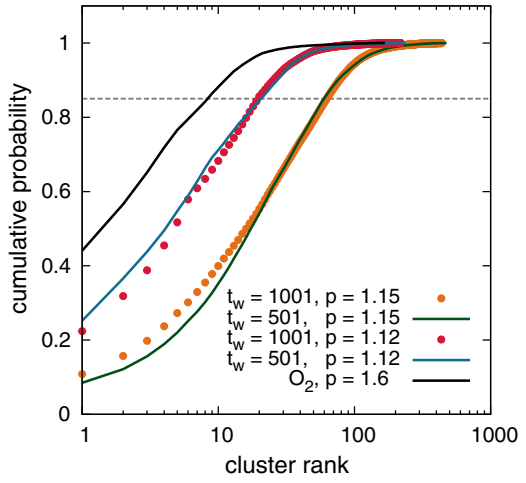

Fig. 5. The cumulative probability distribution of "combined" MCL clusters, ranked according to their population, depending on t_w and p. The following parameter pairs (t_w, p) were considered: (1001,1.15) (orange), (501,1.15) (green), (1001,1.12) (red), (501, 1.12) (blue). The black curve shows the distribution of states obtained from analysis of the ligand dynamics only (p = 1.6) [10]. The gray dashed line defines 85% threshold for the cumulative probability.

landscape contains several isoenergetic barriers for O_2 −migration. Since this study is concerned with finding differences in the network description of ligand motion kinetics by including protein motion, a less detailed decomposition of the protein states is further used with p = 1.12.

3.2. Structural analysis

A numerical analysis of the cluster populations is given in Table 1. The structures of the most populated combined clusters for $t_w = 501$ and p = 1.12 are displayed in Fig. 7. The x-axis is the protein-RMSD and the y-axes is the ligand coordinate. The corresponding weights w of the clusters are reported in each of the panels. Clusters with larger weights ($w = 0.13$, $w = 0.12$ and $w = 0.11$) are found to have broader distributions in RMSD and/or ligand position while clusters with smaller weight are more compact. By careful study of the data in Table 1 one can, however, notice that the most populated state detected by MCL (state 1) is in fact a mixture of three states obtained with k-means clustering: it contains 51% of state Xe1a, 31% of state IS1 and 16% of state ENT. States 2 and 3 are rather pure containing 77.4% of state Xe1b and 90.3% of state Xe2 respectively.

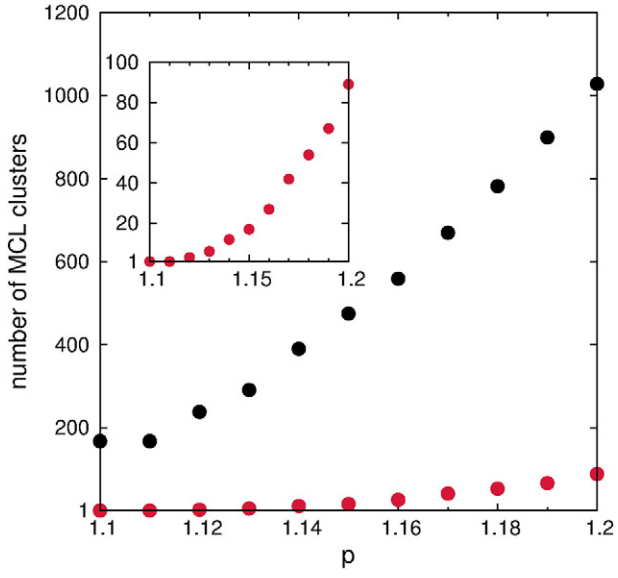

Fig. 6. The number of clusters depending on the MCL parameter p for $t_w = 501$. The red circles are the number of states with p for protein-RMSD only, and the black circles display the case of combined protein–ligand states. The inset shows identical "zoomed in" data for the protein states.

The structural difference between the states can be readily seen in Fig. 8. Here the physical structure of states 1 and 4 from Fig. 7 is reported. The data shown originates from 50 randomly sampled structures of each of the states (1 and 4) from one 2 ns simulation. The strongly fluctuating N- and C-terminal residues were removed and only residues 20 to 110 are retained. The structures were aligned relative to the first frame by superimposing again residues 25–36 and 53–105. Using MOLMOL, the average coordinates and corresponding standard deviation were calculated [34]. The results for the protein are shown in the upper panels. The widths of the tubes are proportional to the magnitude of the standard deviation for the corresponding residue. In the lower panels the ligand positions from the same 50 snapshots are reported. From a structural point of view the states are rather well defined. As was already indicated in Fig. 7, state 1 is larger both in terms of the protein-RMSD and in the ligand positions.

As a last remark about the structure of states obtained with MCL clusterization, we compared the most populated states obtained from different subsets of the MD data. In addition to the MCL clusterization performed on the 10 MD runs starting with O_2 in Xe1, a separate, independent local fluctuation analysis and clusterization for data built from 3 runs starting with O_2 in Xe1, 3 runs in Xe2 and 3 runs in Xe3 with a total simulation time of 18 ns. A new MCL–k-means correspondence table (see Table 2) for the obtained states highlights the importance of extensive sampling of the free energy landscape. As an example, state EXT appears in Table 2 whereas in Table 1 it did not contribute to any of the most populated MCL states. Also, state 1 from the first clustering (Table 1) has a considerably different composition in terms of k-means states than state 1 from the second clustering (Table 2). All this highlights the important role of sampling in view of the obtained networks of states.

3.3. First passage time distribution

The network and the corresponding clustering can be tested for Markovianity. This is done by comparing the original MD dynamics with a suitable random walk on the transition matrices by using ftp distributions. For this, 1000 independent random walks each with 10^6 steps was carried out on the network. As targets to compare with, states 1 and 10 from Table 1 were considered. These two states can be associated with the Xe1a and DS2 cavities obtained from the k-means clustering and known from experiment and previous simulations. The resulting distributions for the original MD trajectory and from the random walks are shown in Fig. 9. The black lines are exemplary random walks on the network (with averages over the 1000 random walks and standard deviations as gray lines and shaded regions, respectively) on the network and the red lines are the fpt from all 29 MD trajectories obtained from the analysis of the ligand coordinate only in our previous study [10]. The right-hand side panels in Fig. 9 report the results of studying the network based on ligand coordinates only [10] whereas the left-hand side panels contain fpts for the combined protein–ligand network. The improvement of the description by including the protein coordinates is remarkable. Even performing a rather non-detailed network analysis of the protein energy landscape yields very favorable agreement between the fpt of the original trajectory and the Markov-state-model. The short- and long-time dynamics are correctly captured by the combined protein–ligand network and the standard deviations are reduced by almost an order of magnitude. Although the ligand-only based network is able to qualitatively reproduce the explicit MD trajectory, including the protein degree of freedom has a non-negligible effect on the quality of the network. It is also of interest to

Table 1
Correspondence between the MCL–RMSD clusters with a population of at least 1% and *k*-means clusters. This clustering is from 10 trajectories starting from Xe1 and constitutes the main clustering employed in the remainder of the present work.

State	Pop. (%)	ENT	Xe1a	Xe1b	Xe2	DS2	EXT	Xe3	IS1	IS3	Xe4	Xe5	PDS2	WAT	DUM
1	13.07	15.95	51.08	0.94	0.00	0.00	0.00	0.41	31.49	0.01	0.00	0.12	0.00	0.00	0.00
2	12.13	0.01	9.88	77.39	0.07	0.00	0.00	0.00	1.50	0.00	0.00	11.15	0.00	0.00	0.00
3	11.39	0.00	0.00	1.10	90.29	0.18	0.00	0.00	0.00	0.00	0.00	8.44	0.00	0.00	0.00
4	7.19	8.80	51.59	0.05	0.00	0.00	0.00	3.96	35.57	0.00	0.01	0.02	0.00	0.00	0.00
5	5.73	0.00	0.00	0.00	0.00	0.00	0.00	65.88	15.67	0.00	18.43	0.03	0.00	0.00	0.00
6	5.21	99.83	0.10	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.00
7	4.12	16.52	51.55	0.01	0.00	0.00	0.00	4.71	27.14	0.00	0.00	0.06	0.00	0.00	0.00
8	3.86	0.00	9.27	87.45	0.03	0.00	0.00	0.00	0.49	0.00	0.00	2.76	0.00	0.00	0.00
9	3.62	0.00	8.72	85.17	0.25	0.00	0.00	0.00	0.79	0.00	0.00	5.07	0.00	0.00	0.00
10	2.87	0.00	0.00	0.00	6.74	89.78	3.48	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	1.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	95.42	0.00	0.00	0.00	4.58
12	1.92	98.50	0.77	0.00	0.00	0.00	0.00	0.00	0.59	0.13	0.00	0.00	0.00	0.00	0.00
13	1.86	34.25	0.00	0.00	0.00	0.00	0.00	3.05	59.89	2.53	0.28	0.00	0.00	0.00	0.00
14	1.59	99.58	0.03	0.00	0.00	0.00	0.00	0.00	0.29	0.10	0.00	0.00	0.00	0.00	0.00
15	1.57	0.00	0.00	0.00	0.00	0.00	0.00	0.29	0.00	0.00	99.71	0.00	0.00	0.00	0.00
16	1.21	0.00	0.00	6.72	0.00	0.00	0.00	0.89	2.72	0.00	0.00	89.67	0.00	0.00	0.00
17	1.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	89.92	0.00	0.00	0.00	10.08
18	1.15	0.00	0.00	0.00	0.22	97.20	2.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
19	1.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	96.25	0.00	0.00	0.00	3.75
MCL combined		8.7	24.4	19.6	11.4	4.0		5.7	1.9		5.7	1.2			
k-Means		7.8	8.6	10.8	9.7	14.2	12.8	6.3	8.5	2.7	7.2	4.4	4.6	5.5	
MCL ligand-only [10]		5.9	15.3	11.6	8.9	15.6	7.2	9.1	0.8		3.1	1.7	3.4		

consider briefly the effect of imposing detailed balance in constructing the transition matrix. The analysis of the fpt distributions was repeated without explicitly symmetrizing the transition matrix (blue curves in Fig. 9) and it is found that they closely follow the analysis in which detailed balance is explicitly imposed. Hence, one can also conclude that the underlying MD simulations are close to detailed balance which suggests that sampling of the transitions is exhaustive.

4. Multi-ligand dynamics

Truncated Hemoglobin belongs to a family of proteins which require two ligands to carry out their function. As both ligands, NO and O_2 have an appreciably binding affinity towards heme-Fe, both ligand-bound states are possible. However, before binding to the chemically relevant site, both ligands have to diffuse towards the heme-unit. Hence, it is

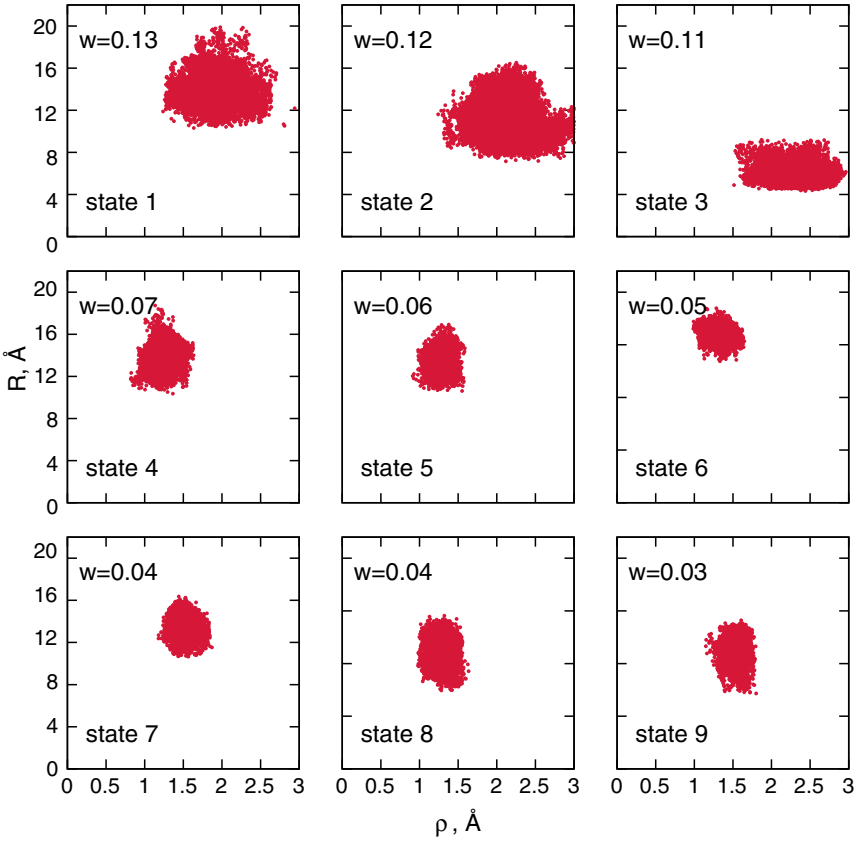

Fig. 7. The structure of the most populated combined states for $t_w = 501$ and $p = 1.12$. The weights are indicated in each panel.

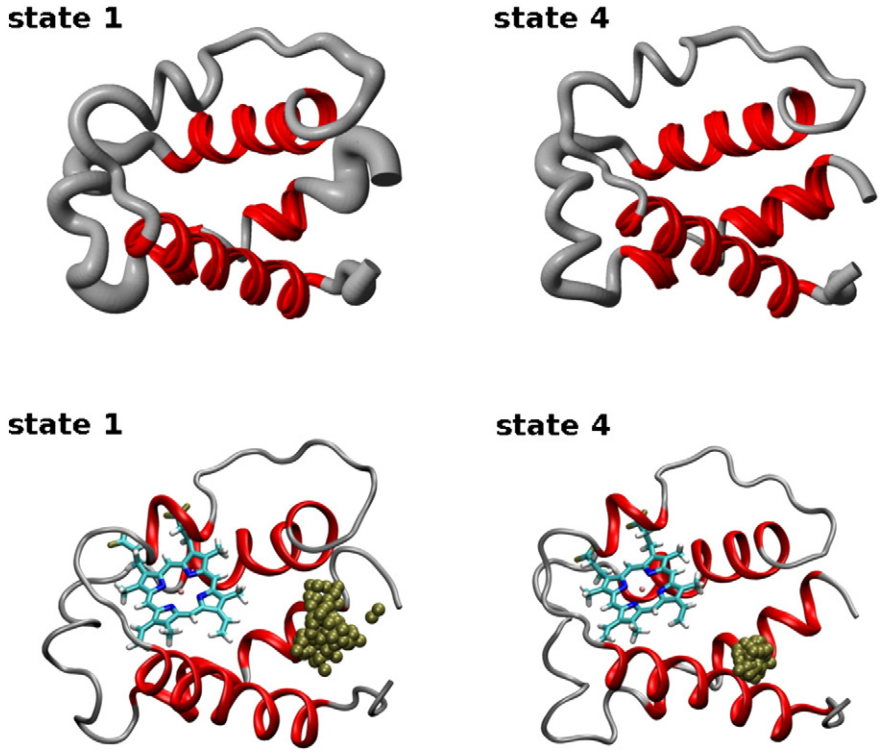

Fig. 8. Degree of fluctuations shown as a superposition of structures corresponding to states 1 (left panels) and 4 (right panels). The upper panels visualize the fluctuations of the protein and the lower panels report ligand positions within the protein.

relevant to characterize the unbound motion of the two diatomic molecules. To this end, a total of 100 ns simulations (10 trajectories of 10 ns each) were carried out from different initial populations. Indeed 5 docking sites are experimentally identified (Xe1 to Xe5) [18], but due to the relevance of Xe1b for ligand migration, it is used in place of Xe5 [20]. Then, for O_2 populating one of these pockets, NO is placed in the reactive site and vice versa. A summary of the initial states considered is given in Table 3.

Analyses are performed as described in the Computational methods section. The most striking observation of the current multi-ligand simulations is how the dynamics of the ligand is hindered when both are free in the protein. This is true in particular for NO in the reactive site (RS) and O_2 in the network. This contrasts with previous results in which one of the ligands was bound to the heme-Fe [14,20] and the dynamics of the second, unbound ligand was rapid and facile throughout the protein network. In previous work, equilibrium simulations of 2 ns length

Table 2
Projection of the MCL–RMSD clusters with a population of at least 1% onto *k*-means clusters. This clustering is from 3 runs starting from pocket Xe1, 3 runs from pocket Xe2 and 3 runs from pocket Xe3.

State	Pop. (%)	ENT	Xe1a	Xe1b	Xe2	DS2	EXT	Xe3	IS1	IS3	Xe4	Xe5	PDS2	WAT	DUM
1	15.37	0.06	3.35	24.86	0.00	1.96	49.94	0.00	0.15	0.00	0.00	0.28	0.00	5.79	13.61
2	11.31	0.02	1.52	15.94	0.75	52.80	9.37	0.06	0.04	0.01	0.00	0.17	0.00	7.73	11.60
3	8.28	2.70	37.67	0.01	0.08	2.57	0.00	7.07	22.56	0.00	11.50	0.04	0.00	0.01	15.80
4	6.72	16.43	36.19	0.54	0.11	2.04	0.00	3.56	39.18	0.39	0.33	0.08	0.31	0.00	0.83
5	4.77	0.00	0.93	14.13	4.66	70.52	1.73	2.94	0.07	0.00	1.34	2.46	0.00	0.01	1.20
6	4.65	0.00	0.00	0.00	0.02	1.04	0.00	42.16	20.83	0.00	29.95	0.40	0.00	0.00	5.58
7	3.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	95.78	0.00	0.00	0.00	4.22
8	3.15	0.02	0.14	2.21	4.72	3.44	0.31	2.37	0.25	2.70	1.94	1.00	1.30	70.71	8.90
9	2.86	0.00	0.00	0.00	0.00	0.00	0.00	0.62	0.00	0.00	99.36	0.00	0.00	0.00	0.02
10	2.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	92.47	0.00	0.00	0.00	7.53
11	2.06	0.00	0.19	1.19	38.56	1.00	0.00	0.47	0.00	0.00	9.19	2.93	0.00	18.49	27.98
12	2.01	99.49	0.28	0.00	0.00	0.03	0.00	0.14	0.03	0.00	0.03	0.00	0.00	0.00	0.00
13	1.86	0.00	0.00	0.00	4.08	94.54	1.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	1.84	0.00	7.75	83.84	0.37	0.00	0.00	0.09	0.43	0.00	2.56	3.55	0.00	0.09	1.30
15	1.74	0.00	0.00	0.00	0.42	0.00	0.00	67.88	0.00	0.00	31.01	0.20	0.00	0.00	0.49
16	1.50	0.00	0.19	0.04	0.04	1.25	0.00	9.85	53.51	0.00	0.30	0.15	0.00	0.00	34.67
17	1.48	0.00	0.00	0.00	2.23	81.71	2.61	12.37	0.00	0.00	1.08	0.00	0.00	0.00	0.00
18	1.25	0.18	0.00	0.00	3.88	37.65	0.00	37.11	18.30	0.00	2.83	0.00	0.00	0.00	0.05
19	1.18	0.00	0.00	0.00	21.94	34.19	0.00	41.94	0.00	0.00	1.94	0.00	0.00	0.00	0.00
20	1.14	0.00	0.00	0.00	7.43	10.12	5.93	11.81	1.30	10.07	12.21	1.94	5.23	2.04	31.90
21	1.02	0.00	1.12	81.63	0.06	0.00	0.00	0.00	0.67	0.00	0.00	16.53	0.00	0.00	0.00
MCL combined		2.0	15.0	1.8	2.0	19.4	15.4	5.8	8.2		8.3			1.1	
k-Means		7.8	8.6	10.8	9.7	14.2	12.8	6.3	8.5	2.7	7.2	4.4	4.6	5.5	
MCL ligand-only [10]		5.9	15.3	11.6	8.9	15.6	7.2	9.1	0.8		3.1	1.7	3.4		

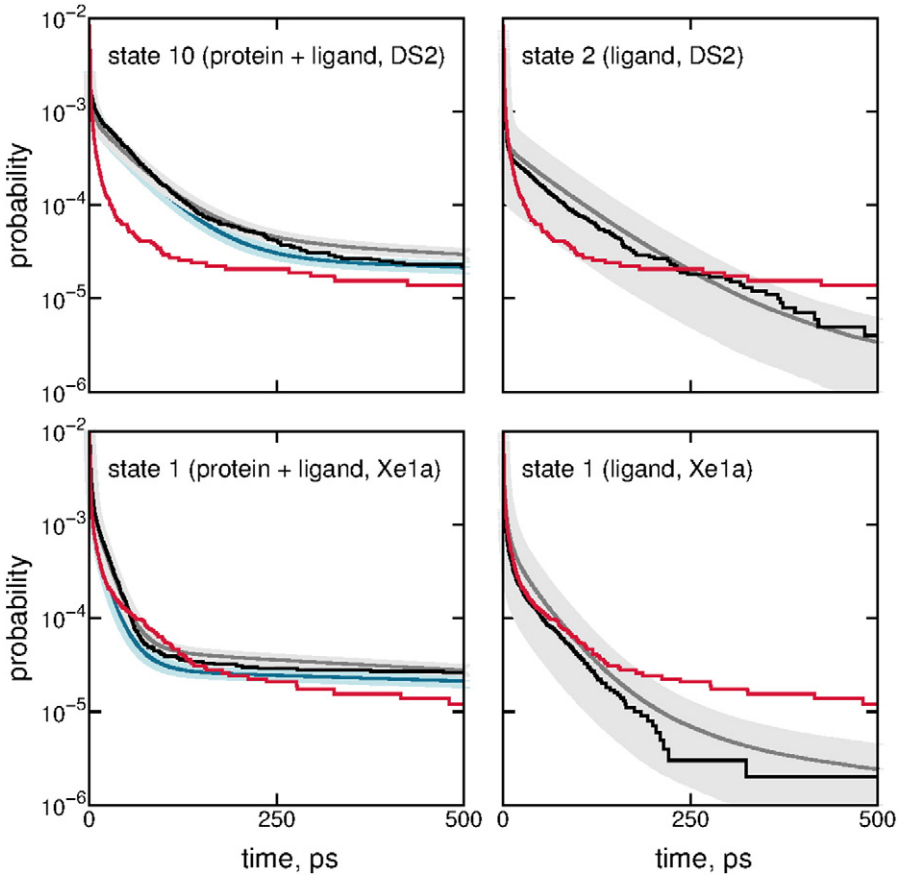

Fig. 9. The first passage time distribution for states DS2 (top) and Xe1a (bottom). The left-hand panels report the results of the present study (network built from protein and ligand coordinates), whereas the right-hand panels show the results by clustering the ligand coordinates only [10]. The labels "state 1" and "state 10" in the left hand column refer to the states found in the clustering as shown in Table 1. The black lines are the results from sampling the transition matrix together with the standard deviation indicated by the gray background. Results from the explicit MD simulations (red curves) and analysis of the transition matrix without imposing detailed balance (blue curves, see text for explanation) are also given.

exhibited a large number of transitions and sampling almost all the docking sites, while in the present simulations, such rapid dynamics and facile migration is not observed, although the simulation time is 5 times larger (10 ns for each trajectory).

Fig. 10 shows the distance of the two ligands to the heme iron atoms. In particular, Fig. 10A and B illustrate the hindrance of the ligand dynamics as either O_2 or NO is in Xe1a. They are both localized in this docking site for nearly 9 ns which has never been observed for a single free ligand. Fig. 10C and D illustrate two different cases of coupled O_2 and NO dynamics, one with NO initially in the RS and O_2 in Xe1b, the other one with O_2 initially in RS and NO in Xe3. In the first case, the coupled dynamics occurs for many ns around Phe62 with O_2 in Xe1a and NO in Xe1b. The second one is also on the ns time scale with O_2 in Xe2 and NO in Xe1b. Fig. 10E shows the fast dynamics NO exhibits after O_2 departure from the protein. From RS, NO samples the entire protein network going first into channel II, then channel I and again channel II before leaving the protein through Xe4. Fig. 10F depicts an interesting event as the O_2 is replaced in RS by NO. Observing this event is important as such exchanges are likely to occur experimentally for the system in its reactive configuration. The different migration steps of each ligand and for each trajectory are summarized in Table 3.

A practical way to illustrate how protein and ligand dynamics are potentially coupled is to consider the RMSD per residue as it reports on the local dynamics of the protein and can be correlated with the ligand dynamics. One possibility is to use a so-called heatmap which reports the RMSD per residue as a function of time with a color code.

Table 3
Ligand dynamics in TrHbN. RS stands for Reactive Site. The left two columns give the initial position of each ligand at the beginning of the simulation. The next two columns summarize the ligand-migration path.

	NO	O_2	NO dynamics	O_2 dynamics	remarks
1	RS	Xe1a	RS → Xe2 → Xe3 → IS1 → Xe4	Xe1a ↔ IS1	O_2 mainly in Xe1a
2	RS	Xe2	RS → Xe2 → Xe3 → IS1 → Xe4 → water	Xe2 → Xe3 → IS1 → Xe4 → water	O_2 leaves first
3	RS	Xe3	RS → Xe2 → Xe3 → IS1 → Xe4 → IS1 → Xe1a → ENT	Xe3 → IS1 → Xe4 → water	
4	RS	Xe4	RS → Xe2 → Xe1b → Xe1a	Xe4 → water	
5	RS	Xe1b	RS ↔ Xe1b	Xe1b ↔ Xe1a	Coupled dynamics
6	Xe1a	RS	Xe1a ↔ Xe1b	RS → Xe2	NO mainly in Xe1a
7	Xe2	RS	Xe2 → Xe3 → IS1 → Xe4 → water	RS → Xe2	
8	Xe3	RS	Xe3 → IS1 → Xe1a → Xe1b → Xe1a	RS → Xe2 → DS2	Coupled dynamics
9	Xe4	RS	Xe4 → IS1 → Xe1*a* → Xe1*b* → Xe2 → RS	RS → Xe2 → Xe1b → Xe1a	NO/O_2 exchange in Xe2
10	Xe1b	RS	Xe1b ↔ Xe1a	RS ↔ Xe2	

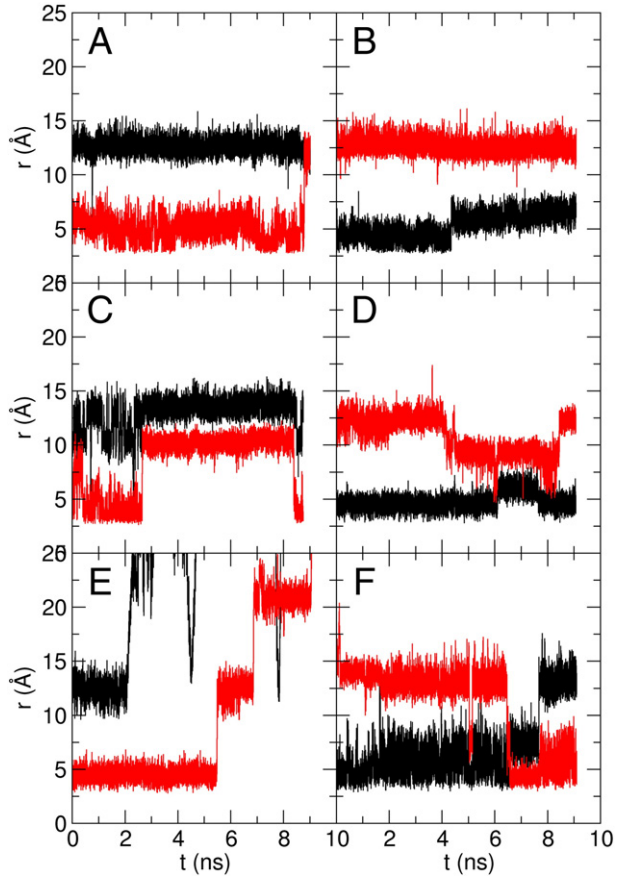

Fig. 10. Distance between the heme iron atom and O_2 (black curve) and NO (red curve) for 6 representative trajectories with different initial conditions. (A) O_2 in Xe1a and NO in RS, (B) O_2 in RS and NO in Xe1a, (C) O_2 in Xe1b and NO in RS, (D) O_2 in RS and NO in Xe3, (E) O_2 in Xe3 and NO in RS, and (F) O_2 in RS and NO in Xe4.

The abscissa corresponds to the time while the ordinate to the residue number. The RMSD of a particular residue is determined relative to itself with its structure at $t = 0$ as reference. Fig. 12 shows heatmaps for trajectories one (O_2 in Xe1a and NO in RS) and three (O_2 in Xe3 and NO in RS). As reported in Table 3 the first trajectory does not exhibit much dynamics of the ligands contrary to trajectory three.

Fig. 11. Representation of TrHbN with sequence of residues is color coded. White: 1 to 10 and 131 to 136 (no particular role), brown: 11 to 20 (channel I entrance), magenta: 21 to 40 (RS, Xe2), black: 41 to 50 (EXT), green: 51 to 70 (channel I), yellow: 71 to 80 (Xe5), blue: 81 to 90 (PDS), orange: 91 to 110 (channel II), red: 111 to 130 (channels I and II).

Fig. 11 shows the protein sequence colored according to the docking sites they are part of. Schematically, residues 20–30 (magenta) are found around the RS, Xe2 and to a lesser extent Xe1b. Residues around 40 (black) are close to EXT, and 50–60 (green) are part of channel I. Those around 70 (yellow) are close to Xe5, residues around 80 (blue) close to the heme and PDS. From 90–100 (orange) the residues are part of channel II and from 110–120 (red) they are shared by channels I and II. According to this, the hindrance of O_2 can be linked to the absence of dynamics of Phe62 and its surrounding residues as migration through channel I is controlled by this residue. Some residues in channel II also exhibit relatively large fluctuations but in this particular trajectory the ligands sample protein locations which do not assist migration. In the second trajectory considered, O_2 is found in Xe3 (part of channel II) which is known to be a hub of the protein network [35, 20]. Migration from this docking site has been found to be easy, in particular towards channel II. For the second trajectory considered, the protein exhibits more intense dynamics spread all over the backbone, in particular along channel I. Channel II also exhibits increased dynamics, in particular at early times high fluctuations are found. This, combined with the facile migration through channel II (exit route for O_2 [20]) is sufficient to explain O_2 escape from the protein. On the other hand, the large fluctuations observed throughout the protein and, in particular, through channel I assist the rapid migration and complete sampling of the protein network performed by NO.

A more quantitative assessment of the ligand migration barriers can be obtained from umbrella sampling simulations, as was done previously for this system [20]. This was repeated here in an exploratory fashion for one situation in which NO is in the RS and O_2 samples the transition between Xe1a and Xe1b. The barrier for migration in this example is similar ($\approx$2 kcal/mol) to that reported previously [20]. However, as was found in one of the explicit MD simulations (trajectory 1), O_2 can localize for extended times in Xe1a which indicates considerably larger barriers for migration. On the other hand this transition is known to be strongly coupled with protein dynamics as it requires Phe62 to be in a certain conformation to take place (gate controlled migration). Hence, despite a favorable barrier, O_2 does not migrate due to insufficient fluctuations exhibited by the Phe62 side chain and the surrounding residues as depicted by the heatmaps in Fig. 12. Preliminary umbrella sampling simulations suggest that under such circumstances the barrier for ligand migration is high and not easily surmounted in unbiased simulations at room temperature, which is what is observed in the explicit simulations. To what extent the decreased protein and side chain flexibility is related to free versus heme-iron-bound NO requires additional simulations.

5. Outlook

The present work provides insight into the coupled dynamics between ligand and protein motion. In particular, it is found that for NO bound to the heme-Fe, the O_2 dynamics couples to protein fluctuations. For a quantitative investigation of this, an approach based on clustering and transition networks was used. The obtained Markov-state-model built on a combined analysis of protein structural changes together with the O_2 dynamics was able to faithfully reproduce the underlying kinetics, see Fig. 9. The states found by such a clustering differ in size for both, the protein and the ligand degrees of freedom. From the data reported in the current work, analysis of ligand migration networks based on transition networks appears to be a meaningful procedure. Future applications of such an approach could be potentially beneficial to investigate modifications in the connectivity of such a network upon protein mutation. Concomitantly, experimental investigations which directly report on small ligand diffusion will be particularly relevant in probing and testing the present proposals and the influence of the amount of sampling required needs to be carefully assessed.

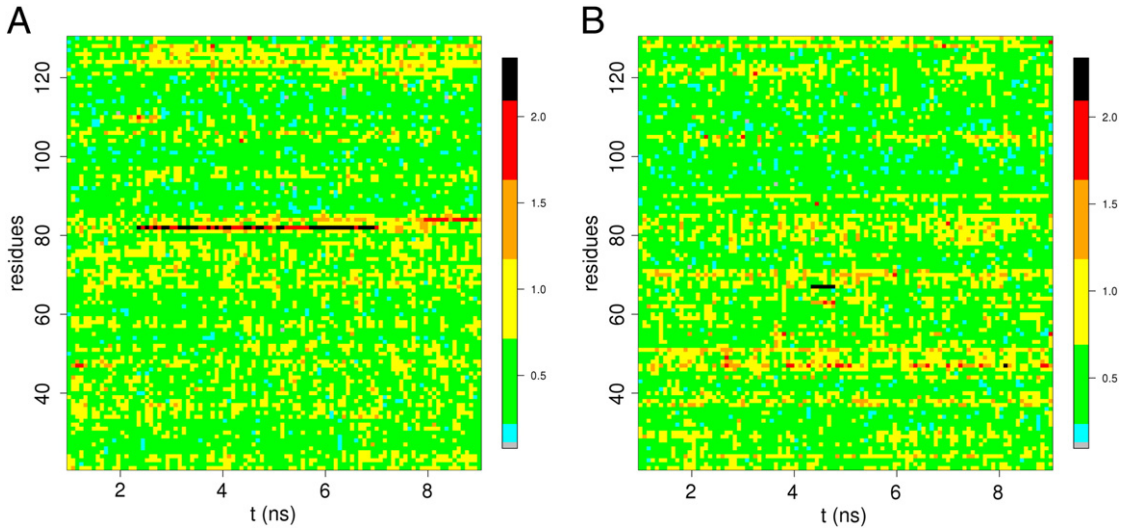

Fig. 12. Heatmaps (RMSD per residues) for two representative trajectories: (A) O_2 in Xe1a and NO in RS, (B) O_2 in Xe3 and NO in RS. Color coding is white, cyan, green, yellow, orange, red, black, from 0 to 2.3 Å respectively. The scale is uniform.

Directly observing ligand dynamics in protein is still a challenging endeavor. Using 2-dimensional vibrational echo spectroscopy, it has recently been possible to quantitatively determine the time constant for substate switching in CO-bound myoglobin [36]. In this case, interconversion between the A_1 and A_3 state is characterized by a time constant of 47 ps. It may also be possible to follow the dynamics of unbound ligands in globular proteins using similar techniques. However, the existence of multiple substates makes this more challenging. Under such circumstances, the combined analysis using experiment and suitable MD simulations will be a meaningful approach as has been recently shown for the structural dynamics of solvated CN- and N-methylacetamide [37,38].

The dynamics involving two ligands (NO and O_2 unbound) is sufficiently different from the dynamics in which one of the ligands is coordinated to the heme-iron to warrant further and more in-depth analysis. Simulations discussed so far in the present work highlight again, that protein and ligand motions are coupled and mutually influence each other. The question what consequences this observation has for the physiological function of the protein deserves and requires additional in-depth analysis based on both, extended MD simulations with suitable force fields and corresponding analysis of the underlying motions.

Acknowledgements

This work was supported by the Swiss National Science Foundation through grant 200021-117810, the NCCR MUST, and the University of Basel.

References

[1] Q. Cui, M. Karplus, Protein Sci. 17 (2008) 1295–1307.
[2] N. Plattner, M. Meuwly, Biophys. J. 102 (2012) 333–341.
[3] P.J. Steinbach, et al., Biochemistry 30 (1991) 3988–4001.
[4] J.Z. Ruscio, D. Kumar, M. Shukla, M.G. Prisant, T.M. Murali, A.V. Onufriev, Proc. Natl. Acad. Sci. 105 (2008) 9204–9209.
[5] H. Ouellet, Y. Ouellet, C. Richard, M. Labarre, B. Wittenberg, J. Wittenberg, M. Guertin, Proc. Natl. Acad. Sci. 99 (2002) 5902.
[6] P. Ascenzi, M. Bolognesi, P. Visca, Biochem. Biophys. Res. Commun. 357 (2007) 809–814.
[7] O. Chen, S. Grob, A. Liechty, D.P. Ridge, J. Am. Chem. Soc. 121 (1999) 11910.
[8] C. Rovira, K. Kunc, J. Hutter, P. Ballone, M. Parrinello, J. Phys. Chem. A 101 (1997) 8914.
[9] S. Mishra, M. Meuwly, Biophys. J. 99 (2010) 3969–3978.
[10] P.A. Cazade, W. Zheng, D. Prada-Gracia, G. Berezovska, F. Rao, C. Clementi, M. Meuwly, J. Chem. Phys. (2014) (under revision).
[11] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhor shid, E. Villa, C. Chipot, R.D. Skeel, L. Kale, K. Schulten, J. Comput. Chem. 26 (2005) 1781–1802.
[12] A.D. MacKerell Jr., et al., J. Phys. Chem. B 102 (1998) 3586–3616.
[13] D. Nutt, M. Meuwly, ChemPhysChem 5 (2004) 1710–1718.
[14] S. Mishra, M. Meuwly, Biophys. J. 96 (2009) 2105–2118.
[15] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. 79 (1983) 926–935.
[16] W. van Gunsteren, H. Berendsen, Mol. Phys. (1977) 1311–1327.
[17] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, J. Chem. Phys. 81 (1984) 3684–3690.
[18] M. Milani, A. Pesce, Y. Ouellet, S. Dewilde, J. Friedman, P. Ascenzi, M. Guertin, M. Bolognesi, J. Biol. Chem. 279 (2004) 21520–21525.
[19] B.R. Brooks, et al., J. Comput. Chem. 30 (2009) 1545–1614.
[20] P.-A. Cazade, M. Meuwly, ChemPhysChem 3 (2012) 4276–4286.
[21] M. Meuwly, O.M. Becker, R. Stote, M. Karplus, Biophys. Chem. 98 (2002) 183–207.
[22] S. Kumar, J.M. Rosenberg, D. Bouzida, R.H. Swendsen, P.A. Kollman, J. Comp. Chem. 13 (1992) 1011–1021.
[23] B. Roux, Comput. Phys. Commun. 91 (1995) 275–282.
[24] H. Steinhaus, Bull. Acad. Polon. Sci 1 (1956) 801–804.
[25] J. Hartigan, Clustering algorithms, John Wiley & Sons, Inc., 1975
[26] G. Berezovska, D. Prada-Gracia, S. Mostarda, F. Rao, J. Chem. Phys. 137 (2012) 194101.
[27] G. Berezovska, D. Prada-Gracia, F. Rao, J. Chem. Phys. 139 (2013) 035102.
[28] A. Baba, T. Komatsuzaki, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 19297.
[29] P. Schuetz, R. Wuttke, B. Schuler, A. Caflisch, Phys. Chem. Chem. Phys. 114 (2010) 15227.
[30] A. Baba, T. Komatsuzaki, Phys. Chem. Chem. Phys. 13 (2011) 1395.
[31] N.V. Smirnov, Rec. Math. (Mat. Sbornik) (NR) 6 (1939) 3.
[32] A.J. Enright, S. Van Dongen, C.A. Ouzounis, Nucleic Acids Res. 30 (2002) 1575.
[33] M. Milani, A. Pesce, Y. Ouellet, P. Ascenzi, M. Guertin, M. Bolognesi, EMBO J. 20 (2001) 3902–3909.
[34] R. Koradi, M. Billeter, K. Wüthrich, J. Mol. Graphics 14 (1996) 51–55.
[35] S. Mishra, M. Meuwly, J. Am. Chem. Soc. 132 (2010) 2968–2982.
[36] H. Ishikawa, K. Kwak, S.K. Chung, M.D. Fayer, Proc. Natl. Acad. Sci. 105 (2008) 8619–8624.
[37] M.W. Lee, J.K. Carr, M. Göllner, P. Hamm, M. Meuwly, J. Chem. Phys. 139 (2013) 054506.
[38] P.-A. Cazade, T. Bereau, M. Meuwly, J. Phys. Chem. B (2014) 8135.