Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Molecular Graphics and Modelling 22 (2004) 369–376
Journal of Molecular Graphics and Modelling
www.elsevier.com/locate/JMGM

Enhanced sampling by multiple molecular dynamics trajectories: carbonmonoxy myoglobin 10 μs $A_0 \rightarrow A_{1-3}$ transition from ten 400 picosecond simulations

Anne E. Loccisano [a,b], Orlando Acevedo [a,b], Jason DeChancie [a,b], Brita G. Schulze [c], Jeffrey D. Evanseck [b,*]

[a] *The National Energy and Technology Laboratory, Pittsburgh, PA 15236-0940, USA*
[b] *Center for Computational Sciences and the Department of Chemistry and Biochemistry, Duquesne University, 600 Forbes Avenue, Pittsburgh, PA 15282-1530, USA*
[c] *Munich Biotech AG, Forstenriederstr 10, 82061 Neuried, Germany*

Abstract

The utility of multiple trajectories to extend the time scale of molecular dynamics simulations is reported for the spectroscopic A-states of carbonmonoxy myoglobin (MbCO). Experimentally, the $A_0 \rightarrow A_{1-3}$ transition has been observed to be 10 μs at 300 K, which is beyond the time scale of standard molecular dynamics simulations. To simulate this transition, 10 short (400 ps) and two longer time (1.2 ns) molecular dynamics trajectories, starting from five different crystallographic and solution phase structures with random initial velocities centered in a 37 Å radius sphere of water, have been used to sample the native-fold of MbCO. Analysis of the ensemble of structures gathered over the cumulative 5.6 ns reveals two biomolecular motions involving the side chains of His64 and Arg45 to explain the spectroscopic states of MbCO. The 10 μs $A_0 \rightarrow A_{1-3}$ transition involves the motion of His64, where distance between His64 and CO is found to vary up to 8.8 ± 1.0 Å during the transition of His64 from the ligand (A_{1-3}) to bulk solvent (A_0). The His64 motion occurs within a single trajectory only once, however the multiple trajectories populate the spectroscopic A-states fully. Consequently, multiple independent molecular dynamics simulations have been found to extend biomolecular motion from 5 ns of total simulation to experimental phenomena on the microsecond time scale.

1. Introduction

1.1. Background

Thermodynamic and kinetic description of complex biological molecules, such as proteins or nucleic acids, requires a thorough sampling of configurations. Proper sampling is difficult and an area of intense methodological development. The source of the problem resides in the fact that native folds of proteins have rugged potential energy surfaces consisting of multiple minima corresponding to similar, but slightly different conformations, separated by barriers of differing size and shape [1–10]. Small changes in conformation typically result in unpredictable and erratic changes in energy [11]. Large entropic or enthalpic barriers intrinsic to the protein energy landscape prevent frequent crossing between regions on the time scale accessible by simulation, which result in incomplete sampling and biasing of computed properties. If the time average of a computed quantity depends on the initial conditions, then the system is considered to be nonergodic on the time scale and temperature of the simulation [12,13]. Consequently, equilibrium distributions are difficult to obtain by standard implementations of molecular dynamics and Monte-Carlo techniques. Nonergodic simulations yield incomplete or misleading dynamical information, since representative sampling is not achieved and important minima essential for the description of protein function are missed [14–18].

Advances in molecular dynamics [19,20] and Monte-Carlo [21] simulations have resulted in the enhanced sampling of phase space for complex biological systems [22–28]. Several methods attempt to probe more of the molecule's conformational space, such as the extension of the simulation to longer time periods [17,29–32], guided enhanced sampling [33], multiple short-time trajectories [11,34–38], simulated annealing [39], conformational flooding [40,41], and locally enhanced sampling [42–44]. In this study, we review the simulation work of multiple trajectories previously reported, [34] since recent studies indicate that

* Corresponding author.
E-mail address: evanseck@duq.edu (J.D. Evanseck).

1093-3263/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2003.12.004

this approach improves the sampling of conformational space; [11,23,34–36,40,45–48] however, a more quantitative assessment of the actual time extension has not been presented.

Due to the abundance of structural, kinetic and spectroscopic information, carbonmonoxy myoglobin (MbCO) has been selected to study the enhanced sampling of multiple trajectories [49,50]. The bound CO ligand in MbCO exhibits four IR absorption bands centered at 1965, 1949, 1942 and 1932 cm^{-1}, which are identified as the A-states ($A_{0–3}$) [51–55]. The A_2 frequency typically appears as a small shoulder of the larger A_1 peak, thus both are collectively referred to as $A_{1,2}$. Kinetic properties of sperm whale MbCO spectroscopic A-states have been investigated by Springer and co-workers using flash photolysis with infrared (IR) monitoring [49,50]. Rate coefficients for the interconversion processes between the A-states have been extracted using the maximum-entropy method, which yielded a non-Arrhenius temperature dependence. The interconversion time was found to be 10 μs at 300 K for the $A_0 \leftrightarrow A_1 + A_3$ transition. The challenge is to discover the structural variations responsible for this transition on the 10 μs time scale.

Given the ubiquitous problem of sampling complex biological systems, it is of interest to investigate the time scale enhancement resulting from multiple trajectories. The advantages of using multiple trajectories to sample phase space more aggressively have been exploited in proteins [8,11,34,35,45,47,56–70], nucleic acids [32,37,48,71–76], ion channels [77,78], and lipid bilayers [79]. Despite the general recognition of sampling shortcomings in standard computational simulations, a quantitative assessment of the time scale extension has not been presented before. MbCO is a unique protein with the spectroscopic A-state transitions reported on the microsecond scale, which will provide the target time scale of the multiple trajectory methodology.

2. Methodology

2.1. Atomic coordinates

As previously reported, [34] five high-resolution structures of sperm-whale MbCO were selected from the Brookhaven Protein Data Bank [80]. The structures utilized were two conformations of an X-ray structure at 1.5 Å resolution where seven residues and the CO ligand were modeled in two conformations (1MBC) [81], a neutron diffraction structure at 1.8 Å resolution (2MB5) [82,83], and the first two NMR structures from a set of 12 conformations obtained at pH 5.7 (1MYF) [84]. Missing hydrogens were generated with the HBUILD algorithm [85]. Of the 12 His residues, 8 were modeled in their protonated form (i.e., charge +1) consistent with the NMR structures (residues 12, 36, 48, 81, 97, 113, 116 and 119) [84]. The other histidines residues, 24, 82, proximal (93) and distal (64), were assumed to have the proton at the N_δ nitrogen (HSD) [84]. The structures were used as starting points for five trajectories, which are called X-ray-A_s, X-ray-B_s, $Neutron_s$, $NMR1_s$ and $NMR2_s$. In addition, the distal His64 was rotated in each of the PDB structures about its Cβ–Cγ bond and five additional trajectories were started from these conformations, called X-ray-$A_{s–r}$, X-ray-$B_{s–r}$, $Neutron_{s–r}$, $NMR1_{s–r}$ and $NMR2_{s–r}$.

2.2. Simulation protocol

Simulations were executed with the CHARMM all-atom parameter set [86] in combination with the EGO_VIII software [87] which employs a combined structure-adapted multipole method and multiple time step algorithm to calculate electrostatic forces efficiently enabling a cut-off to be exempt. The crystal waters (137 in the X-ray structures, 89 in the neutron diffraction structure) were retained and each conformation was enveloped in a 37 Å sphere of equilibrated TIP3P water molecules [88]. No attempt was made to explore system size influences upon computed molecular properties. The system was selected to allow for sufficient hydration of MbCO balanced against expected computer resource usage [89]. The surface region of the water was subjected to an SBOUND potential [85] to confine water molecules within the sphere. There are known computed differences when using stochastic boundary and periodic boundary conditions [90]. One conclusion is that less global motion is permitted with stochastic boundaries. This paper does not address the differences between the two methods, however one should be aware that computed differences could arise. The system was equilibrated in several steps. First, all water molecules were relaxed with a gradient minimizer (400 steps) and then equilibrated for 10 ps at 300 K (protein constrained). Next, the whole system was minimized (400 steps) and subsequently equilibrated for 30 ps at 300 K. During equilibration, the system was coupled to a heat bath with a time constant of 10^{-13} s in order to achieve the desired temperature of 300 K. For the subsequent molecular dynamics simulation (370 ps), a weaker coupling (time constant 10^{-11} s) was employed for infrequent and slight adjustment of the temperature. Structural characterization is based only on the 370 ps time period. Our decision to terminate the simulations after 400 ps stems from our observation that the RMSD of each MbCO trajectory stops increasing after 400 ps. The choice of simulation time is consistent with the work reported by Loncharich and Brooks on MbCO [89]. For visualization in conformational space, structures from the whole 400 ps time were included. Two trajectories (X-ray-B_s and $Neutron_s$) were extended to 1200 ps to compare the sampling of conformational space between the shorter and longer time trajectories. Thus, a total simulation time of 5.6 ns was computed, and analysis of the trajectories was carried out using standard tools of CHARMM [91] and XPLOR. [92]

2.3. Collection of structures and analysis

Principal component analysis (PCA) as a form of multivariate analysis was applied to the extraction of functionally relevant data from the large set of molecular dynamics trajectories. One structure every 4 ps was collected for each of the 10 short trajectories (7 from equilibration, 93 from simulation) to produce 1000 structures. In addition, 400 snapshots from the extended X-ray-B_s and Neutron$_s$ trajectories (400–1200 ps), and the five PDB structures were added to yield a total of 1405 conformations. The creation of the ensemble was carried out to visualize the evolution of the molecular dynamics trajectories, compare the extent of conformational space covered by the short and long trajectories, evaluate the effect of the His64 orientation on global and local protein structure, and determine any significant global or local protein motion. No differences were found when structures were collected more frequently (every 1 ps) or averaged over the 4 ps time period. Average structures of each simulation were computed from molecular dynamics transient conformations over 30–400 ps. An overall average structure was computed from the transients of all trajectories. A PDB average from the five PDB structures served as a reference for all structural comparisons. Before analysis, structures were superimposed onto the PDB average to eliminate differences due to rotation or translation.

3. Results and discussion

3.1. Molecular dynamics structural comparison

In order to evaluate how the trajectories evolve and move away from their respective starting experimental structures, each of the five individual PDB structures and 12 trajectory averages were compared to the computed PDB average. The root-mean square deviations (RMSDs) [93] for several atom selections are summarized in Table 1. The RMSDs pertaining to the NMR1 structure and corresponding trajectories deviate strongly from those of the other PDB structures and trajectories, which is mainly due to variations in the D helix, as well as in the EF and FG loops. Consistent with other work, we find that the majority of differences originate from changes in the loop regions [5]. In addition, the table indicates that the orientation of the distal histidine does not influence the evolution of the global or local structure, since the respective heavy atom and His64 RMSDs from the s- and s–r-trajectories do not deviate significantly. The total average structure, formed from all molecular dynamics transients, shows best agreement with the PDB reference (0.9 Å). A possible explanation is that each individual trajectory becomes caught in one substate (even if simulated for 1.2 ns) and averaging over an ensemble of independent trajectories yields better agreement with experiment. Thus, the average structure from the molecular dynamics simulations is interpreted to better represent the MbCO native-fold, as compared to a single long-time trajectory.

Table 1
RMSDs in Å from the reference structure (PDB average)

Trajectory	Heavy atoms	Sphere[c]	His64
X-ray-A_s	1.3	1.2	2.4
X-ray-B_s	1.4	0.9	1.3
Neutron$_s$	1.5	1.3	1.8
NMR1$_s$	2.1	0.9	1.1
NMR2$_s$	1.4	0.8	1.1
Average$_s$[a]	1.5	1.0	1.5
X-ray-A_{s-r}	1.6	0.9	1.4
X-ray-B_{s-r}	1.3	0.9	1.2
Neutron$_{s-r}$	1.4	1.0	1.3
NMR1$_{s-r}$	2.2	1.6	2.6
NMR2$_{s-r}$	1.5	1.0	1.5
Average$_{s-r}$[a]	1.6	1.1	1.8
Average (all trajectories)	1.6	1.0	1.6
Total average structure[b]	0.9	0.8	1.5
Neutron$_s$ (1.2 ns)	1.4	1.3	2.0
X-ray-B_s (1.2 ns)	1.5	0.8	1.1

[a] The trajectory averages of the five respective trajectories are considered.
[b] Computed from *one* structure obtained by averaging over all molecular dynamics structures.
[c] Centered at oxygen atom of CO, $r = 5$ Å.

3.2. PCA plots

To visualize global and local motions, PCA was carried out on the His64, ligand and heme atomic coordinates. The projections of the original 1405 structures onto the first two eigenvectors are shown in Fig. 1. The first two PCs display 61.8% of all structural variation. The plot displays the projections of the s-trajectories and the s–r-trajectories. In Fig. 1, three clusters are identified. Visual inspection of representative structures from each substate shows that the position of His64 differs significantly in these groups. In the lower left cluster, the distal histidine is oriented such that its N_ϵ atom points toward the oxygen of the CO ligand.

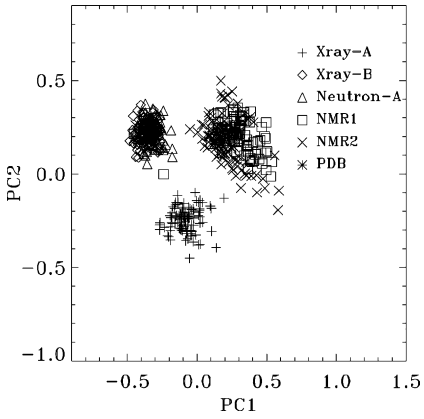

Fig. 1. PCA plots of molecular dynamics transients based on the atomic coordinates of His64.

Table 2
Classification of His64 geometries based on distances (Å) between His64 and CO

Trajectory	$H(N_\delta)$H64–O(CO)[a] MD average	(N_ϵ) H64–O(CO) MD average	His64 position[b]
X-ray-A_s	5.9 ± 0.5	3.9 ± 0.3	In/N_ϵ
	8.0 + 0.7	7.1 ± 0.9	Out (72–400 ps)
X-ray-B_s	6.6 ± 1.0	4.8 ± 1.3	In/N_ϵ
	3.6 ± 0.6	5.1 ± 0.5	In/N_δ (105–668 ps)
	7.6 ± 1.4	6.2 ± 2.2	Out (669–1200 ps)
$Neutron_s$	6.6 ± 0.8	6.2 ± 1.0	Out
$NMR1_s$	6.3 ± 0.7	4.2 ± 0.5	In/N_ϵ
$NMR2_s$	6.5 ± 0.9	4.6 ± 0.8	In/N_ϵ
X-ray-$A_{s–r}$	6.8 ± 0.5	4.3 ± 0.5	In/N_ϵ
X-ray-$B_{s–r}$	3.6 ± 0.4	4.9 ± 0.4	In/N_ϵ
$Neutron_{s–r}$	3.5 ± 0.4	4.8 ± 0.4	In/N_ϵ
$NMR1_{s–r}$	6.6 ± 0.5	5.8 ± 0.5	Out
	4.0 ± 0.5	6.9 ± 0.5	In/N_ϵ (120–400 ps)
$NMR2_{s–r}$	5.9 ± 1.2	8.8 ± 1.0	Out
	4.6 ± 0.6	7.2 ± 0.5	In/N_ϵ (250–400 ps)

[a] Distance measured from the hydrogen to the ligand oxygen.
[b] Classification based upon the average structure over the time interval indicated.

This is called the in/N_ϵ conformation and is found in all five PDB structures. In the upper left group of conformations, the His64 has rotated around its Cβ–Cγ bond and now the hydrogen atom of the N_δ atom points toward O(CO). This geometry is called the in/N_δ conformation. In the right cluster, His64 has undergone a translational movement away from the CO ligand (both N_δ and N_ϵ are more than 5 Å from CO) and is referred to as the out conformation. Both in the s- and s–r-trajectories, all three clusters are sampled which indicates that the His64 rotation in the s–r-trajectories was not an irreversible modification and that all three conformations can convert into each other. A few trajectories move between the clusters, where X-ray-A_s and $NMR2_s$ move between the out and in/N_ϵ clusters, and X-ray-B_s samples all three groups. In addition, the $NMR1_{s–r}$ and $NMR2_{s–r}$ trajectories move between the out and in/N_δ clusters.

3.3. His64 dynamics

The conformational changes of His64 were studied in more detail, since the position and orientation of this residue relative to the CO ligand is thought to be the most decisive factor for the occurrence of the spectroscopic A-states in MbCO [22,23,94–107]. In addition, PCA in Fig. 1 indicates that distinct substates were sampled in the simulations. This high mobility of the distal histidine is also known from crystallography [108] and from previous molecular dynamics simulations [105,109]. The distances of the His64 N_δ and N_ϵ atoms to the ligand's oxygen atom are summarized in Table 2 and shown in Fig. 2. For the *s*-trajectories, the starting geometry (PDB) is in/N_ϵ with a N_ϵ(His64)-O(CO) distance between 2.7 and 3.9 Å. Only two trajectories remained in this conformation for the duration of the

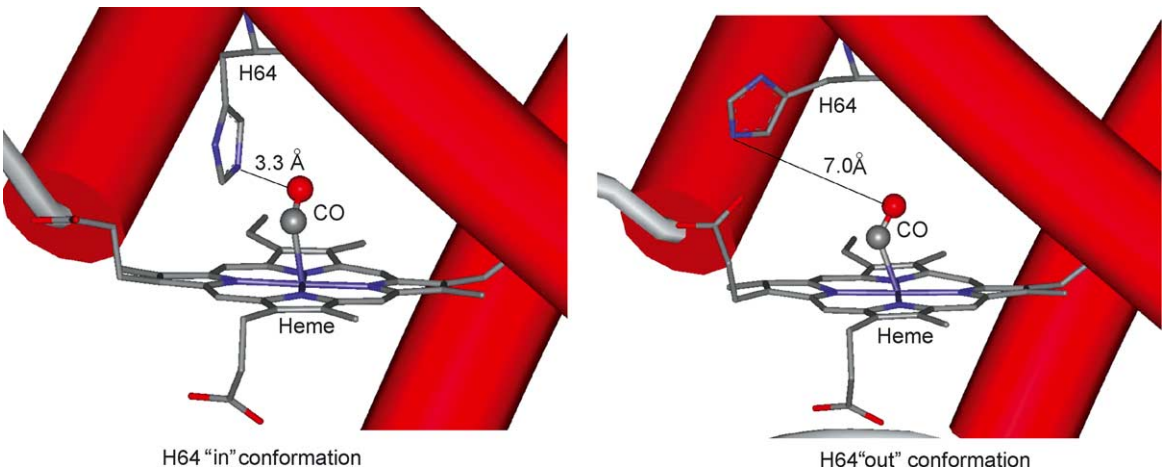

Fig. 2. Schematic of MbCO heme pocket created from the simulations representing the A_0 and $A_{1–3}$ states using WebLab Viewer.

simulations, while three simulations ended up in the out conformation. In one trajectory (X-ray-B_s), the His64 ring rotated by approximately 120° about its Cβ–Cγ bond and formed an in/N_δ geometry for approximately 500 ps. Although a complete ring flip was not achieved, the simulation time was sufficient to bring the protonated N_δ closer to O(CO) than the N_ϵ. We hypothesize that if all the simulations were extended to longer times, such ring flips could occur more often and result in a hydrogen bond between the proton of the N_δ and the ligand. This was the rationale for us to perform the second set of simulations starting from the PDB structures with a 180° ring flip of His64 (s–r-trajectories). In the s–r-trajectories, the starting geometry is in/N_δ with an HN_δ(His64)–O(CO) distance between 3.2 and 4.4 Å (see Table 2). Here the distal histidine remained in the in/N_δ conformation in two trajectories, in two it moved outward and then back to in/N_δ, and only in one trajectory it reverted to in/N_ϵ. Thus, His64 is observed to be very flexible which is in agreement with previous work [105,108,109].

Experimentally, the transition rates between the A_0 and $A_{1–3}$ states have been reported to occur between $1.4 \times 10^6\ s^{-1}$ (aqueous solution, 273 K) and $2.3 \times 10^4\ s^{-1}$ (75% glycerol, 273 K) [110]. Therefore, the transition time between the His64 "open" and "closed" states range from 0.71 to 43 μs at 273 K. His64 was computed to switch between the in and out conformations four times over the 5.6 ns of total simulation time at 300 K (Table 2). Three of the transitions originated from starting structures which could be accelerated due to a strained conformation. Only one true transition (X-ray-B_s from 105 to 668 ps) was computed and its time constant was 563 ps (Table 2). The *ca*. 10^3 discrepancy between the kinetically determined transition times (0.71 μs at 273 K) and the simulated values reported here (563 ps at 300 K) and by other groups [109,111] demands further examination. To bridge an appropriate comparison between theory and experiment, it is imperative to realize that the experimental interconversion times between the "open" and "closed" states of MbCO are known to be dependent upon solvent composition and temperature [112].

The simulations were carried out at 300 K, as compared to the experiments, which were measured at 273 K. The strong dependency of the interconversion times between the "open" and "closed" states on temperature has been previously shown by a study in 75% glycerol solution [112]. Transition rates of 1.2×10^5, 2.3×10^4, and $8.1 \times 10^3\ s^{-1}$ at 293, 273, and 264 K were measured, respectively [112] Simple logarithmic extrapolation of the data to 300 K gives a transition rate of $2.4 \times 10^5\ s^{-1}$, or a transition time of 4.2 μs in 75% glycerol. Thus, the transition time reduces by a factor of ten ($2.3 \times 10^4\ s^{-1}/2.4 \times 10^5\ s^{-1}$) when considering an increase of temperature from 273 to 300 K in the 75% glycerol solution. By analogy, if one assumes that the same reduction applies in aqueous solution, then the 0.71 μs transition time ($1.4 \times 10^6\ s^{-1}$, aqueous solution, 273 K) adjusted to 300 K is estimated to be 69 ns. As a result, the computed transition of 563 ps is *ca*. 10^2 times faster than the 69 ns value provided by experiment at 300 K. Many factors, such as the ability of the solvent model to represent frictional forces and viscosity effects, incomplete sampling of full conformational change, and the short time scale of the simulations can rationalize the computed and observed time difference. Due to the limited number of observed transitions in this study, and absence of direct experimental measurement in aqueous solution at 300 K, explanation of the transition time difference between theory and experiment is not attempted in this work.

3.4. A_0 substrate

From mutation studies, it was found that when His64 is replaced by a smaller residue, only the A_0 band remains [96,113]. Consequently, the interaction between His64 and CO is minimal in the A_0 substate. X-ray crystallography at low pH values (4.5) confirmed this hypothesis, and showed that His64 actually swings out of the heme pocket [108]. Both our results and previous calculations [105,107] indicate that the neutral His64 is also very mobile and can move out of the heme pocket, as shown in Fig. 2. The outward movement of His64 is thought to be biologically relevant because it might create a ligand entry pathway to the heme [108,114–118]. Thus, it seems plausible that His64 actually swings out before becoming protonated, making its protonation easier and leading to an increased intensity of the A_0 band. From our calculations, the His64(N_δ)–CO(O) distance is estimated to be greater than 4.5 Å for the A_0 state.

The model is in qualitative agreement with the time scales reported, where the $A_0 \rightarrow A_{1–3}$ interconversion times have been estimated to range between 10^{-6} and 10^{-4} s depending upon solvent composition and temperature [112,119], and the $A_{1,2} \rightarrow A_3$ time at 10^{-9} s [119]. Thus, the movement of His64 would take *ca*. 3–5 orders of magnitude longer than the movement of Arg45, which is plausible because His64 is more buried than Arg45 and attached to helix E, which is a more rigid unit than the CD loop where Arg45 is located. Also, the distances to be traveled by His64 are larger. The simulations also indicate that the Arg45 interaction with His64 is fast and frequent, since its energy of interaction is weakened due to Arg45 solvation.

It is well-known that solvent can alter the time evolution of atomic motions and molecular properties [90,120]. Molecular dynamics simulations have been used to show that solvent slows protein motions by random collisions between the solvent and protein, and by introducing friction through a Stoke's-like dissipative force dependent upon the solvent's viscosity, velocity, and size. Consequently, as the residues participating in the dynamical process become more solvent exposed and travel longer distances, the more the solvent composition influences their motion. The observed $A_{1–3} \rightarrow A_0$ transition suggests that solvent plays a significant role in the process, since the rates measured for water and water/glycerol (75/25) mixtures differ by two orders of magnitude [112]. The observed solvation effect for

the $A_{1–3} \rightarrow A_0$ transition is supported qualitatively by our data, where His64 makes contact with solvent molecules in all trajectories, and the average displacement (out–in conformation) of His64 (N_ϵ(H64)–O(CO)) is *ca.* 4 Å. For the $A_{1,2} \rightarrow A_3$ transition, Arg45 spends 33% of its time in solvent and its remaining time in the hydrophobic pocket. The (N_ϵ(H64)–N_η(R45)) distance traveled during this transition is *ca.* 2 Å. Since His64 moves *ca.* 2 Å more than Arg45, the $A_{1–3} \rightarrow A_0$ interconversion should depend more strongly on the nature of the solvent than the $A_{1,2} \rightarrow A_3$ transition, consistent with experimental results [112].

The proposed model avoids the need to invoke protonation changes between N_ϵ and N_δ which have been used repeatedly to account for the frequency shifts [94,97]. In the trajectories of this study, solvent approaches only one side of His64 (pointing away from the heme pocket), even after 1.2 ns simulation time. However, in order to aid in the transfer of a proton leading to a change in protonation from N_ϵ to N_δ, His64 would have to swing out of the heme pocket. Other groups have already pointed out that there is no plausible mechanism for a change in protonation on the time scale of a nanosecond which is the estimated interconversion time between $A_{1,2}$ and A_3 [119]. We based our choice of protonation at the N_δ atom of His64 on observations by neutron diffraction studies and NMR spectroscopy [82–84]. Recently, it was reported that His64 is primarily protonated at N_ϵ [121]. The current molecular dynamics simulations suggest that if His64 were protonated at N_ϵ then the proton would still be close enough to O(CO) to induce the frequency changes observed by IR spectroscopy. The important point is that as long as a proton from His64 is close to the ligand it can cause the red-shifted frequencies characteristic of the A-states, and at the same time its electrostatic effect on the CO ligand can be enhanced by interaction between its unprotonated nitrogen and Arg45. The in/N_ϵ conformation, as defined in this study, is not involved in the formation of the $A_{1–3}$ states.

4. Conclusion

Multiple short-time (400 ps) trajectories have been used to extend the timescale of molecular dynamics simulations to determine the molecular details of the spectroscopic A-states of MbCO. Thorough sampling of the rough potential energy surface of a protein such as MbCO is difficult, which results in incomplete sampling of the potential energy surface and therefore results in biases in computed molecular properties. The spectroscopic $A_0 \rightarrow A_{1–3}$ transition has been experimentally observed to be 10 μs at 300 K, which is too long to be observed by standard molecular dynamics simulations. From the total 5.6 ns of simulation time, two molecular motions involving the His64 and Arg45 sidechains were found to explain the spectroscopic states of MbCO. The molecular motions occurred only once over all trajectories, whereas the multiple trajectories sampled the A-states fully. Therefore, the method of multiple short trajectories has been found useful to examine the spectroscopic A-states of MbCO, and to extend molecular motion from 5.6 ns of total simulation time to experimental observations on the microsecond timescale.

Acknowledgements

We thank the Department of Defense (DAAH04-96-1-0311 and DAAG55-98-1-0067) for financial support to make this research possible.

References

[1] X. Wu, S. Wang, J. Chem. Phys. 110 (1999) 9401.
[2] H. Xu, B.J. Berne, J. Chem. Phys. 110 (1999) 10299.
[3] Z. Zhu, M.E. Tuckerman, S.O. Samuelson, G.J. Martyna, Phys. Rev. Lett. 88 (2002) 100201.
[4] A. Cooper, Proc. Natl. Acad. Sci. USA 73 (1976) 2740.
[5] R. Elber, M. Karplus, Science 235 (1987) 318.
[6] Y. Zhang, D. Kihara, J. Skolnick, Proteins Struct. Funct. Genet. 48 (2002) 192.
[7] R.K. Hart, R.V. Pappu, J.W. Ponder, J. Comp. Chem. 21 (2000) 531.
[8] L.S.D. Caves, C.S. Verma, Proteins Struct. Funct. Genet. 47 (2002) 25.
[9] J.A. Rahman, J.C. Tully, J. Chem. Phys. 116 (2002) 8750.
[10] M. Gruebele, Curr. Opin. Struct. Biol. 12 (2002) 161.
[11] L.S.D. Caves, J.D. Evanseck, M. Karplus, Protein Sci. 7 (1998) 649.
[12] D. Thirumalai, R.D. Mountain, T.R. Kirkpatrick, Phys. Rev. A 39 (1989) 3563.
[13] R. Zhou, B.J. Berne, J. Chem. Phys. 107 (1997) 9185.
[14] M. Karplus, A. McCammon, Nat. Struct. Biol. 9 (2002) 646.
[15] J.M. Yon, Cell. Mol. Life Sci. 53 (1997) 557.
[16] C.M. Dobson, A. Sali, M. Karplus, Angew. Chem. Int. Ed. Engl. 37 (1998) 868.
[17] A. Ciochetti, A.R. Bizzarri, S. Cannistraro, Biophys. Chem. 69 (1997) 185.
[18] H. Frauenfelder, S.G. Sligar, P. Wolynes, Science 254 (1991) 1598.
[19] M. Karplus, Biopolymers 68 (2003) 350.
[20] T. Hansson, C. Oostenbrink, W. Van Gunsteren, Curr. Opin. Struct. Biol. 12 (2002) 190.
[21] N. Wilding, D.P. Landau, Lecture Notes Phys. 605 (2003) 231.
[22] T.E. Cheatham III, B.R. Brooks, Theor. Chem. Accounts 99 (1998) 279.
[23] G. Moraitakis, A.G. Purkiss, J.M. Goodfellow, Reports Prog. Phys. 66 (2003) 383.
[24] B.J. Berne, J.E. Straub, Curr. Opin. Struc. Biol. 7 (1997) 181.
[25] J.B. Clarage, T. Romo, B.K. Andrews, B.M. Pettit, G.N. Phillips Jr., Proc. Natl. Acad. Sci. USA 92 (1995) 3288.
[26] A. Hodel, T. Simonson, R.O. Fox, A. Brunger, J. Phys. Chem. 97 (1993) 3409.
[27] D. Caspar, Structure 3 (1995) 327.
[28] P.H. Hunenbeger, A.E. Mark, W. Van Gunsteren, J. Mol. Biol. 252 (1995) 492.
[29] V. Daggett, M. Levitt, Ann. Rev. Biophys. Biomol. Struct. 22 (1993) 353.
[30] J.M. Troyer, F.E. Cohen, Proteins Struct. Funct. Genet. 23 (1997) 97.
[31] W. Van Gunsteren, P.H. Hunenbeger, A.E. Mark, P.E. Smith, I.G. Tironi, Comput. Phys. Commun. 91 (1995) 305.
[32] T.E. Cheatham, P.A. Kollman, J. Mol. Biol. 259 (1996) 434.
[33] I. Andricioaei, A.R. Dinner, M. Karplus, J. Chem. Phys. 118 (2003) 1074.
[34] B.G. Schulze, J.D. Evanseck, J. Am. Chem. Soc. 121 (1999) 6444.
[35] A. Elofsson, L. Nilsson, J. Mol. Biol. 233 (1993) 766.

[36] M.L. Carlson, R.M. Regan, Q.H. Gibson, Biochemistry 35 (1996) 1125.
[37] P. Auffinger, E. Westhof, Biophys. J. 71 (1996) 940.
[38] J.E. Straub, A.B. Rashkin, D. Thirumalai, J. Am. Chem. Soc. 116 (1994) 2049.
[39] C.D. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Science 220 (1983) 671.
[40] B.G. Schulze, H. Grübmuller, J.D. Evanseck, J. Am. Chem. Soc. 122 (2000) 8700.
[41] H. Grübmuller, Phys. Rev. E 52 (1995) 2893.
[42] V. Hornak, C. Simmerling, Proteins Struct. Funct. Genet. 51 (2003) 577.
[43] A.-S. Wilkinson, P.K. Bryant, S.O. Meroueh, M.G.P. Page, S. Mobashery, C.W. Wharton, Biochemistry 42 (2003) 1950.
[44] R. Elber, M. Karplus, J. Am. Chem. Soc. 112 (1990) 9161.
[45] G.A. Worth, F. Nardi, R.C. Wade, J. Phys. Chem. B 102 (1998) 6260.
[46] V. Daggett, Curr. Opin. Struct. Biol. 10 (2000) 160.
[47] A.A. Gorfe, P. Ferrara, A. Caflisch, D.N. Marti, H.R. Bosshard, I. Jelesarov, Proteins Struct. Funct. Genet. 46 (2002) 41.
[48] E.C. Sherer, J. Phys. Chem. B 106 (2002) 5075.
[49] B.A. Springer, S.G. Sligar, J.S. Olson, G.N. Phillips Jr., Chem. Rev. 94 (1994) 699.
[50] J.T. Sage, P.M. Champion, in: Comprehensive Supramolecular Chemistry, vol. 5. Elsevier B.V., New York, 1996, p. 171.
[51] R.H. Austin, K.W. Beeson, L. Eisenstein, H. Frauenfelder, I.C. Gunsalus, Biochemistry 14 (1975) 5355.
[52] R.H. Austin, K.W. Beeson, L. Eisenstein, H. Frauenfelder, I.C. Gunsalus, V.P. Marshall, Science 181 (1973) 541.
[53] W.S. Caughey, J.O. Alben, S. McCoy, S.H. Boyer, S. Carache, P. Hathaway, Biochemistry 8 (1969) 59.
[54] W.S. Caughey, H. Shimada, M.G. Choc, M.P. Tucker, Proc. Natl. Acad. Sci. USA 78 (1981) 2903.
[55] N. Alberding, R.H. Austin, S.S. Chan, L. Eisenstein, H. Frauenfelder, I.C. Gunsalus, T.M. Nordlund, J. Chem. Phys. 65 (1976) 4701.
[56] G. Toth, R.F. Murphy, S. Lovas, J. Am. Chem. Soc. 123 (2001) 11782.
[57] I. Yeh, G. Hummer, J. Am. Chem. Soc. 124 (2002) 6563.
[58] Y. Wang, K. Kuczera, J. Phys. Chem. 100 (1996) 2555.
[59] A.J. Charlton, E. Haslam, M.P. Williamson, J. Am. Chem. Soc. 124 (2002) 9899.
[60] R. Dvorsky, J. Sevcik, L.S.D. Caves, R.E. Hubbard, C.S. Verma, J. Phys. Chem. B 104 (2000) 10387.
[61] D. Vitkup, D. Ringe, M. Karplus, G. Petsko, Proteins Struct. Funct. Genet. 46 (2002) 345.
[62] P. Barthe, C. Roumestand, H. Demene, L. Chiche, J. Comp. Chem. 23 (2002) 1577.
[63] S.L. Kazmirski, V. Daggett, J. Mol. Biol. 277 (1998) 487.
[64] C.A. Baysal, Biophys. J. 83 (2002) 699.
[65] D. Alonso, V. Daggett, Protein Sci. 7 (1998) 860.
[66] B. Schiott, T.C. Bruice, J. Am. Chem. Soc. 124 (2002) 14558.
[67] P. Doruker, A.R. Atilgan, I. Bahar, Proteins Struct. Funct. Genet. 40 (2000) 512.
[68] A. Li, V. Daggett, J. Mol. Biol. 257 (1996) 412.
[69] C.A. Marhefka, B.M. Moore, T.C. Bishop, L. Kirkovsky, A. Mukherjee, J.T. Dalton, D.D. Miller, J. Med. Chem. 44 (2001) 1729.
[70] J. Tsai, M. Levitt, D. Baker, J. Mol. Biol. 291 (1999) 215.
[71] K. Csaszar, N. Spackova, R. Stefl, J. Sponer, N.B. Leontis, J. Mol. Biol. 313 (2001) 1073.
[72] J.D. Williams, K. Hall, Biophys. J. 76 (1999) 3192.
[73] P. Auffinger, S. Louise-May, E. Westhof, Biophys. J. 76 (1999) 50.
[74] P. Auffinger, S. Louise-May, E. Westhof, J. Am. Chem. Soc. 117 (1995) 6720.
[75] A. Lahiri, L. Nilsson, Biophys. J. 79 (2000) 2276.
[76] J. Sarzynska, T. Kulinski, L. Nilsson, Biophys. J. 79 (2000) 1213.
[77] J.P. Duneau, S. Crouzy, N. Garnier, Y. Chapron, M. Genest, Biophys. Chem. 76 (1999) 35.
[78] R.J. Law, L.R. Forrest, K.M. Ranatunga, P. La Rocca, D.P. Tieleman, M. Sansom, Proteins Struct. Funct. Genet. 39 (2000) 47.
[79] M.J. Scheider, S.E. Feller, J. Phys. Chem. B 105 (2001) 1331.
[80] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, H. Shindyalov, P.E. Bourne, Nucleic Acids Res. 28 (2000) 235.
[81] J. Kuriyan, S. Wilz, M. Karplus, G.A. Petsko, J. Mol. Biol. 192 (1986) 133.
[82] X. Cheng, B.P. Schoenborn, Acta Cryst. B46 (1990) 195.
[83] X. Cheng, B.P. Schoenborn, J. Mol. Biol. 220 (1991) 381.
[84] K. Osapay, Y. Theriault, P.E. Wright, D.A. Case, J. Mol. Biol. 244 (1994) 183.
[85] A. Brunger, M. Karplus, Proteins Struct. Funct. Genet. 4 (1988) 148.
[86] A.D. MacKerell Jr., D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, J. Gao, H. Guo, S. Ha, D. Joseph, 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. Smith, R. Stote, J.E. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, M. Karplus, J. Phys. Chem. 102 (1998) 3586.
[87] M. Eichinger, H. Grübmuller, H. Heller, P. Tavan, J. Comp. Chem. 18 (1997) 1729.
[88] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, J. Chem. Phys. 79 (1983) 926.
[89] R.J. Loncharich, B.R. Brooks, J. Mol. Biol. 215 (1990) 439.
[90] C.R. Brooks, M. Karplus, B.M. Pettitt, Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics, Wiley, New York, 1988.
[91] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, J.D. States, S. Swaminathan, M. Karplus, J. Comp. Chem. 4 (1983) 187.
[92] A. Brunger, X-PLOR, Version 2.1, Yale University, 1992.
[93] M. Levitt, J. Mol. Biol. 168 (1983) 595.
[94] E. Oldfield, K. Guo, J.D. Augspurger, C.E. Dykstra, J. Am. Chem. Soc. 113 (1991) 7537.
[95] K.D. Park, K. Guo, F. Adebodun, M.L. Chiu, S.G. Sligar, E. Oldfield, Biochemistry 30 (1991) 2333.
[96] D.P. Braunstein, K. Chu, K.D. Egeberg, H. Frauenfelder, J.R. Mourant, G.U. Nienhaus, P. Ormos, S.G. Sligar, B.A. Springer, R.D. Young, Biophys. J. 65 (1993) 2447.
[97] G.B. Ray, X.-Y. Li, J.A. Ibers, J.L. Sessler, T.G. Spiro, J. Am. Chem. Soc. 116 (1994) 162.
[98] T. Li, M.L. Quillin, G.N. Phillips Jr., J.S. Olson, Biochemistry 33 (1994) 1433.
[99] M. Lim, T.A. Jackson, P.A. Anfinrud, Science 269 (1995) 962.
[100] A.M. Navarro, M. Maldonado, J. Gonzalez-Lagoa, R. Lopez-Mejia, J. Lopez-Garriga, J.L. Colon, Inorg. Chim. Acta 243 (1996) 161.
[101] M. Unno, J.F. Christina, J.S. Olson, J.T. Sage, P.M. Champion, J. Am. Chem. Soc. 120 (1998) 2670.
[102] M.T. McMahon, A.C. deDios, N. Godbout, R. Salzmann, D.D. Laws, H. Le, R.H. Havlin, E. Oldfield, J. Am. Chem. Soc. 120 (1998) 4784.
[103] J.S. Olson, G.N. Phillips Jr., J. Biol. Inorg. Chem. 2 (1997) 544.
[104] J.F. Christina, M. Unno, P.M. Champion, Biochemistry 36 (1997) 11198.
[105] P. Jewsbury, T. Kitagawa, Biophys. J. 67 (1994) 2236.
[106] T.G. Spiro, P.M. Kozlowski, J. Biol. Inorg. Chem. 2 (1997) 516.
[107] T.G. Spiro, P.M. Kozlowski, J. Am. Chem. Soc. 120 (1998) 4524.
[108] F. Yang, G.N. Phillips Jr., J. Mol. Biol. 256 (1996) 762.
[109] J. Ma, S. Huo, J.E. Straub, J. Am. Chem. Soc. 119 (1997) 2541.
[110] N. Agmon, J.J. Hopfield, J. Chem. Phys. 79 (1983) 2042.
[111] P. Jewsbury, S. Yamamoto, T. Minato, M. Saito, T. Kitagawa, J. Phys. Chem. 99 (1995) 12677.
[112] W.D. Tian, J.T. Sage, P.M. Champion, E. Chien, S.G. Silgar, Biochemistry 35 (1996) 3487.
[113] D. Morikis, P.M. Champion, B.A. Springer, S.G. Sligar, Biochemistry 28 (1989) 4791.
[114] M.F. Perutz, F.S. Mathews, J. Mol. Biol. 21 (1966) 199.

[115] D.A. Case, M. Karplus, J. Mol. Biol. 132 (1979) 343.
[116] J. Kottalam, D.A. Case, J. Am. Chem. Soc. 110 (1988) 7690.
[117] L. Zhu, J.T. Sage, A.A. Rigos, D. Morikis, P.M. Champion, J. Mol. Biol. 224 (1992) 207.
[118] J. Ramsden, T.G. Spiro, Biochemistry 28 (1989) 3125.
[119] J.B. Johnson, D.C. Lamb, H. Frauenfelder, J.D. Muller, B. McMahon, G.U. Nienhaus, R.D. Young, Biophys. J. 71 (1996) 1563.
[120] J.A. McCammon, S.C. Harvey, Dynamics of Proteins and Nucleic Acids, Cambridge University Press, Cambridge, 1987.
[121] S. Bhattacharya, S.F. Sukits, K.L. MacLaughlin, J.T. Lecomte, Biophys. J. 73 (1997) 3230.