JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Coarse-Grained Simulations of Heme Proteins: Validation and Study of Large Conformational Transitions

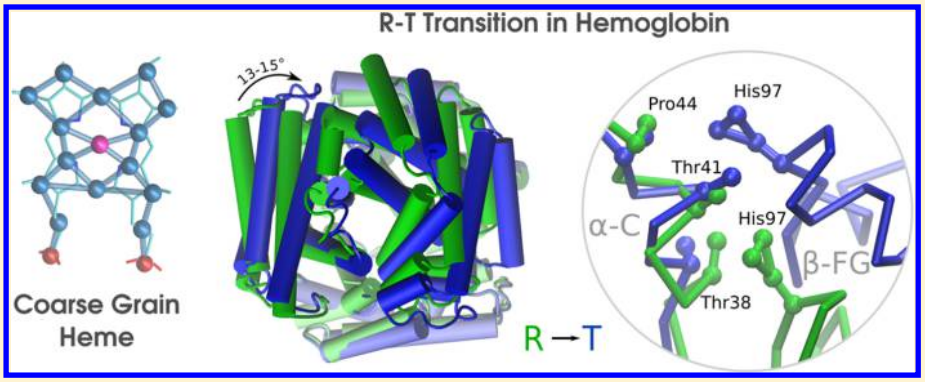
Claudia L. Ramírez,[†,‡] Ariel Petruk,[†] Mauro Bringas,[†] Dario A. Estrin,[†] Adrian E. Roitberg,[§] Marcelo A. Marti,[‡] and Luciana Capece*[,†]

[†]Dto. de Química Inorgánica, Analítica y Química Física, Fac. de Ciencias Exactas y Naturales, Univ. de Buenos Aires/INQUIMAE-CONICET, Buenos Aires, C1428EGA, Argentina

[‡]Dto. de Química Biologica Fac. de Ciencias Exactas y Naturales, Univ. de Buenos Aires/IQUIBICEN-CONICET, Buenos Aires, C1428EGA, Argentina

[§]Department of Chemistry, University of Florida, Gainesville, Florida 32611-7200, United States

Supporting Information



ABSTRACT: Heme proteins are ubiquitous in nature and perform many diverse functions in all kingdoms of life. Many of these functions are related to large-scale conformational transitions and allosteric processes. Sampling of these large conformational changes is computationally very challenging. In this context, coarse-grain simulations emerge as an efficient approach to explore the conformational landscape. In this work, we present a coarse-grained model of the heme group and thoroughly validate this model in different benchmark examples, which include the monomeric heme proteins myoglobin and neuroglobin and the tetrameric human hemoglobin where we evaluated the method's ability to explore conformational changes (as the formation of hexacoordinated species) and allosteric transitions (as the well-known R → T transition). The obtained results are compared with atomistic molecular dynamics simulations. Overall, the results indicate that this approach conserves the essential dynamical information on different allosteric processes.

INTRODUCTION

Heme proteins are a ubiquitous group of proteins that perform a wide variety of key biological functions, which include the sensing and transport of small molecules as in myoglobin (Mb)[1] and hemoglobin (Hb),[2] oxidation reactions (as in P450s),[3] and electron transport (as in cytochromes),[4] among others. They are also usually used as benchmark cases to test newly developed theoretical and experimental approaches.[5] These proteins contain a heme cofactor, a porphyrin macrocycle which possesses a central iron coordinated in its four equatorial coordination positions by the porphyrin. The two axial coordination positions can be occupied by protein residues like His or Cys, among others, and/or external ligands.

Atomistic computer simulations of heme proteins have allowed the study of many different phenomena in which they are involved, ranging from chemical reactivity and enzymatic catalysis inside the heme pocket to extended conformational changes that involve large protein regions. However, atomistic simulations, at least in their unbiased approach, are still limited to processes occurring in the nanoseconds to microseconds time scale and to systems of moderate size. Thus, typically systems of this size are trapped in local minima and unable to explore conformations separated from the initial configurations by barriers larger than a few kcal/mol. In this context, coarse-grained (CG) simulations emerge as a plausible tool to increase the sampling of the conformational space as well as the system size due to their much lower computational cost.[6]

In the recent years, many efforts have been done to develop protein CG models that provide an accurate representation of the protein dynamics in the native state.[7−9] Initially, the focus was on folding processes, with the Go model being the reference method for this type of study.[10] In this representation, each amino acid is replaced by one center or

Received: March 17, 2016
Published: June 6, 2016

DOI: 10.1021/acs.jctc.6b00278
J. Chem. Theory Comput. 2016, 12, 3390−3397

"bead" and the dynamics is based exclusively on the native contacts. However, very often it is desirable to retain information on the side chain positions and interactions between amino acids. For this purpose, specific CG force fields were developed in order to perform CG simulations in proteins, as the extension of the Martini Force Field to proteins,[8] the Sirah Force Field for proteins,[9] or the CG model for proteins developed by Hills et al.[7] In this last model, each residue is represented with one to five beads, depending on the complexity of the residue.[7] Parameters for this force field have been obtained using a bottom-up approach, in order to fit as best as possible the results obtained with atomistic simulations. Nonbonded interactions are represented using five different atom types, for which the associated 15 pairwise potentials have been adjusted using a force-matching approach. Bonded interactions have been parametrized in order to represent the Boltzmann distribution of the AA simulations, with the exception of the backbone torsions, which were taken from the PDB database. This CG model was successfully applied among other systems to the study of Trp cage folding and adenylate kinase conformational transitions[7] and the MsbA ABC transporter in a mixed DOPC/DOPE bilayer.[11]

Considering the above-mentioned ubiquity, the relevance of heme proteins, and the power of CG methodologies in this work, we developed a CG model of the heme group using a similar protocol to the one employed by Hills et al. Among the available CG models, the parametrization strategy followed by Hills et al. allowed us to obtain a CG model for the heme group in a straightforward way from atomistic simulations. This model allows performing CG simulations of heme proteins using the corresponding force field. Specifically, the heme model was tested in different heme proteins which include myoglobin (Mb),[5] neuroglobin (Ngb),[12] and hemoglobin (Hb),[2] where we studied several dynamic processes that cover different spatial and temporal scales. The results show a very good agreement with previous atomistic data, validating the heme group parametrization, and widening our capabilities for studying this amazing group of proteins.

COMPUTATIONAL METHODS

CG Representation of the Heme Group. In order to parametrize the heme group, we performed 20 ns of atomistic classical molecular dynamics simulation of the heme group. The system was inserted in a box of 1002 water molecules. Temperature was set at 300 K and pressure at 1 bar using the Berendsen thermostat and barostat, respectively.[13] A histidine residue was coordinated in one of the axial sites in order to properly describe the active site in the most typical coordination state. This allowed us also to parametrize the Fe−His bond. Atomistic simulations were run using heme parameters developed in our group and tested in a large list of heme proteins.[14,15]

CG beads positions were initially positioned mimicking the position of the beads in the CG amino acids with similar chemical structure, and the parameters for bonded interactions were obtained in order to reproduce the Boltzmann distributions observed in the atomistic simulations. Similarly to the parametrization performed by Hills et al., bonds were fitted to a harmonic potential, while angles required an up to three-order polynomial equation. Nonbonded interactions were represented using three of the five atom types already available in the CG force field. Figure 1 shows the classification of each bead in the heme CG model. The resulting model contains 17 CG beads: 1 bead representing the iron atom, 4 beads close to the pyrrolic nitrogen atoms, 9 beads in the porphyrin cycle, and 4 beads for the two propionate groups (Figure 1).

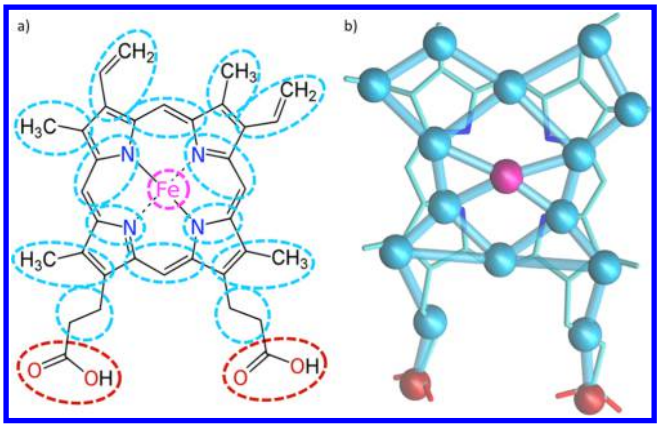


Figure 1. CG model of the heme group. A total of 17 beads have been used to represent the heme group in the coarse-grained model: 1 polar bead (magenta), 14 nonpolar beads (cyan), and 2 negative beads (red). (a) Atoms grouped in each CG bead. The position of each bead is located in the center of mass of each group of atoms circled with dashed lines. (b) Chemical structure of the heme group with the CG beads position and bonding scheme.

Details on the CG Simulations. CG simulations were performed using a locally modified version of the PMEMD module of AMBER11[16] for CPUs, where the CG force field developed by Voth et al. was implemented. Due to the absence of water molecules and to the reduction of the system size, it was not necessary at this stage to use parallel computing strategies. Future developments in the context of the Amber package include parallel and GPU CG simulations. Simulation conditions were similar to the ones recommended in the original work. A time step of 4 fs was used, and 0.5 μs of CG−MD simulations were run for pentacoordinated Mb (Mb-5c), hexacoordinated Mb (Mb-6c), pentacoordinated Ngb (Ngb-5c), and hexacoordinated Ngb (Ngb-6c). Here, 0.2 μs CG−MD simulations were run for human hemoglobin in the tense (T) and relaxed (R) states. Initial structures were taken from the crystal structures: 1VXD (for Mb),[17] 1OJ6 (for Ngb),[18] 2HHB (for HbT),[19] and 1HHO (for HbR).[20] These structures were converted to the CG representation, minimized, and slowly thermalized to the desired temperature. For each case, test equilibrium simulations were performed at 100, 150, and 200 K. In the case of Ngb and Mb, results at 150 K were more accurate in relation to stability and similarity to the atomistic simulations. (See SI for details on the results at different temperature conditions). This is also consistent with previous observations, which indicated the need to simulate at 0.6 T.[7] Thus, 150 K was the temperature of choice for all production simulations. For tetrameric hemoglobin, the temperature was set to 125 K to avoid distortions of the protein structure. Additionally, in order to avoid sliding movements of the heme group, a small positional restraint of 10 kcal/mol Å^2 was applied to the four polar heme CG beads, which lie in the proximity to the pyrrolic nitrogens. For comparative purposes, atomistic MD simulations in explicit solvent and the NPT ensemble were performed using the Amber99SB force field[21] and the PMEMD module of the Amber14 package, using standard simulation conditions as described in previous works.[22,23] Atomistic simulations were run up to 80 ns for

DOI: 10.1021/acs.jctc.6b00278
J. Chem. Theory Comput. 2016, 12, 3390−3397

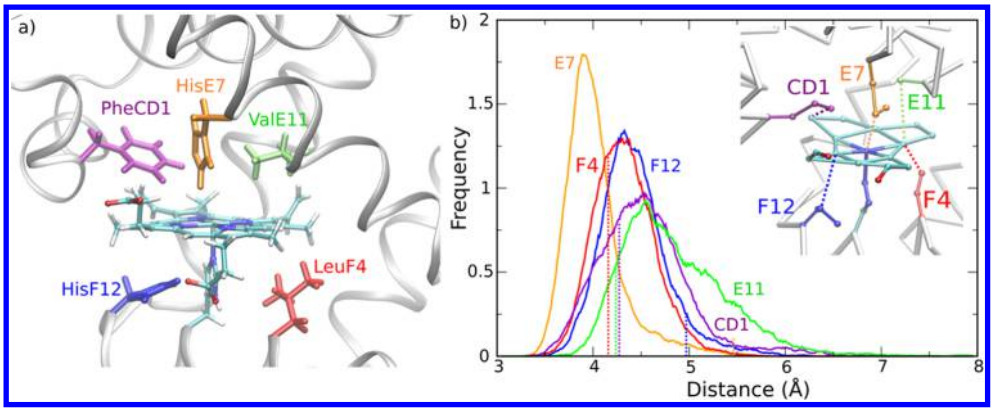


Figure 2. Analysis of the key interactions between the heme group and the protein matrix in myoglobin in the CG simulations. (a) Atomistic structure of myoglobin showing the residues that anchor the heme to the protein structure. (b) Distribution of the interaction distances along the CG−MD simulation. Inset: CG structure of Mb active site, with the measured distances showed as dotted lines. For comparison purposes, these distances were measured in myoglobin crystal structure[17] and converted to the CG representation yielding values in Å of 4.88 (HisF12), 4.27 (PheCD1), 4.16 (LeuF4), 5.49 (HisE7), and 4.26 (ValE11). These values are marked in panel b, in dotted vertical lines, using the same color code of the distribution lines.

the monomeric globins (Mb and Ngb) and 200 ns for tetrameric Hb.

Essential dynamics (ED) analysis was performed on the CG trajectories in order to compare the conformational space explored by the CG and the AA simulations.[24] ED involves diagonalization of the covariance matrices of atomic positions along the trajectory, yielding the eigenvectors that define the essential motions of the protein. This analysis was performed only for the Cα atoms (which represent the backbone atoms in the CG model). Terminal residues were excluded in order to avoid masking of the essential motions of the protein core by the high flexibility of terminal regions. ED analysis of combined trajectories was also performed to gain insight into the dynamics of the structural transitions involved. Finally, projection of the MD trajectories onto selected essential motions was performed to analyze the configurational space explored along the MD simulation. This type of analysis showed to be very useful in the characterization of the dynamics of globins and in particular the hexacoordination process.[22,23] Thus, performing the same analysis in the CG trajectories allows a more extensive comparison with the atomistic dynamics.

Finally, in order to further validate the heme CG model for the study of larger scale conformational transitions, we used steered molecular dynamics (SMD)[15] to force the transition between the R and T states in human hemoglobin. In SMD, the potential energy is modified by adding a time-dependent potential ($E(\vec{r}, t)$), which allows exploring a specific reaction coordinate. In this case, the potential was modified following eq 1:

$$E(\vec{r}, t) = k(\mathrm{RMSD}(\vec{r}, t) - \mathrm{RMSD}_o(\vec{r}, t))^2 \quad (1)$$

where RMSD(r_i) is the root-mean-square deviation of each frame with the reference structure, RMSD$_o$ is the reference value for the root-mean-square deviation, and $\vec{r}$ are the coordinates of the CG beads. The reaction coordinate ranges from 5.35 to 0.75 Å, being the maximum value of the RMSD between the initial and final conformation, and was explored in 70 ns using a 4 fs time step.

Source code for running CG−MD simulations in the PMEMD module of Amber11 and the herein developed heme CG parameters are available under request. Efforts are in place to include the CG code into Amber official releases.

RESULTS

Stability of Heme Protein Monomeric Structures in CG Representation and Comparison with All Atom Simulations. Myoglobin (Mb) is a well-known monomeric heme protein that belongs to the globin superfamily. This protein binds molecular oxygen and stores it in muscle tissues.[5] The heme group in absence of external ligands is pentacoordinated (5c) in the ferrous state.[17] As a first validation of the heme CG model, we performed a 0.5 μs CG−MD simulation of Mb in the 5c state. The structure remained stable along the simulation. However, as expected, RMSD values are higher in CG simulations than in atomistic simulations,[21] indicating a larger deviation from the crystal structure in the coarse grain simulations (Figure S1). The iron in the heme group of myoglobin is coordinated to a histidine residue located at the F helix (named HisF8) and is anchored to the protein matrix by several hydrophobic and aromatic interactions, which involve mainly residues PheCD1, HisE7, ValE11, HisF12, and LeuF4. As shown in Figure 2, these interactions were generally conserved in the CG simulations. The distance distributions are consistent with the observed values in the crystal structure (detailed in the caption of Figure 2) with the larger deviation observed for HisE7. This residue position is associated with the opening and closing of the HisE7 gate,[25] which is in turn highly related to the protonation position of the histidine residue.[26] The protonation site of the histidine residue is not distinguished in the force field developed by Voth et al.;[7] thus, consideration of this effect is not possible in the context of the present CG model. Overall, the results indicate that the heme CG model retains proper interactions with the protein matrix.

Neuroglobin (Ngb) is another monomeric heme protein that belongs to the globin family, showing a very similar tertiary structure to that of Mb. However, and opposite to Mb, crystallographic information indicates that in the absence of an external ligand, Ngb displays an hexacoordinated (6c) heme group.[18] Thus, in Ngb, HisE7 coordinates to the iron in the distal position, forming the archetypical 6c form. In previous works from our group by means of atomistic molecular

DOI: 10.1021/acs.jctc.6b00278
J. Chem. Theory Comput. 2016, 12, 3390−3397

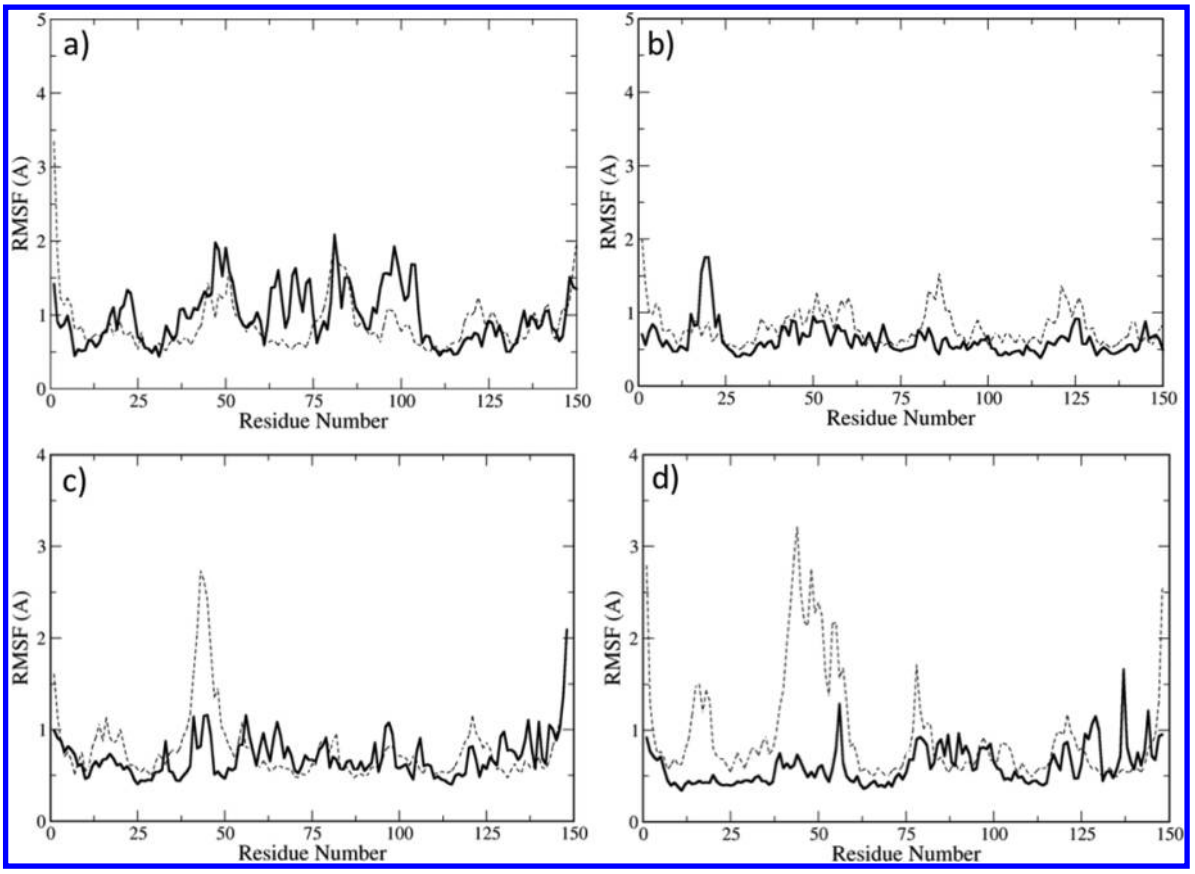


Figure 3. RMSF for (a) Mb-5c, (b) Mb-6c, (c) Ngb-5c, and (d) Ngb-6c. The data was calculated considering only the Cα atoms. Results for the CG and the atomistic simulations are shown in solid and dashed lines, respectively.

dynamics simulations, we observed that the main structural and dynamical differences between these two globins are located in a region called "CD region", which involves two short helixes (named C and D helixes) and their connecting loop (the CD loop).[22] To further assess the capabilities of the CG heme model, we performed 0.5 μs of CG−MD simulations for myoglobin and neuroglobin, both in the penta- and hexacoordinated states. As in the previous case, the structures remained stable along the simulations, as can be observed in the RMSD calculations shown in Figure S1. Protein flexibility in the CG model was analyzed calculating the root-mean-square fluctuations (RMSF) along the trajectory for each residue. Figure 3 compares the RMSF obtained in each simulation with the results obtained with the corresponding atomistic model. The CG results are in good agreement with the AA, especially for Mb, which is the less flexible protein. In Ngb, the CG and AA results are qualitatively consistent; however, the flexibility is significantly reduced in the CG simulations with respect to the AA simulations, especially in the CD region (residues 49 to 59 in Mb and 39 to 58 in Ngb). Interestingly, if the RMSF is calculated using the trajectory starting at 10 ns up to 80 ns (a similar time frame to the one used in the atomistic simulations) results are in better agreement (Figure S2). This might indicate that in the CG model after a certain simulation time the CD loop finds a local minimum and stays there for the rest of the simulation time, therefore reducing the average fluctuation in this region. This stabilization however requires very long time scales and thus is not observed in AA simulations. However, caution should be taken when comparing simulation times in atomistic and CG simulations. Although the applied CG force field was parametrized using atomistic simulations, this does not imply necessarily an exact conservation of the free energy surface. Indeed, due to the simplification of the system, it is almost impossible to completely conserve the kinetic barriers between minima. Thus, the time scales in atomistic and CG models are not directly comparable, making this last observation a qualitative implication without implying an exact determination of the time scale required to achieve an specific conformation in the CG model.

The 5c to 6c Transition in Ngb in the CG Model. The heme CG model was also tested analyzing the 6c ⇔ 5c transition in Ngb and Mb. We calculated the transition essential mode as explained in the computational methods for Ngb and Mb in the CG model and compared it with the results of the atomistic simulation. Figure 4 compares the projection of the principal mode obtained in this analysis (which corresponds to the transition mode) along the protein sequence for Ngb and Mb, following the procedure used in previous works to describe the 6c ⇔ 5c transition in Ngb and Mb.[22,23] It can be seen that qualitatively there is a good agreement between the CG and the AA models, especially in the case of Ngb, where the contribution of the CD region is significant. This contribution is also observed in the CG essential mode. In Mb, the transition mode is less specific and involves several regions of the protein, again as observed in the atomistic model.

CG−MD Simulations of Tetrameric Hemoglobin. Human hemoglobin (Hb) is a tetrameric heme protein displaying a well-known allosteric transition after binding oxygen to the heme, moving from the so-called T ("tense") to the R ("relaxed") states.[2] Hb structure is formed by two

DOI: 10.1021/acs.jctc.6b00278
J. Chem. Theory Comput. 2016, 12, 3390−3397

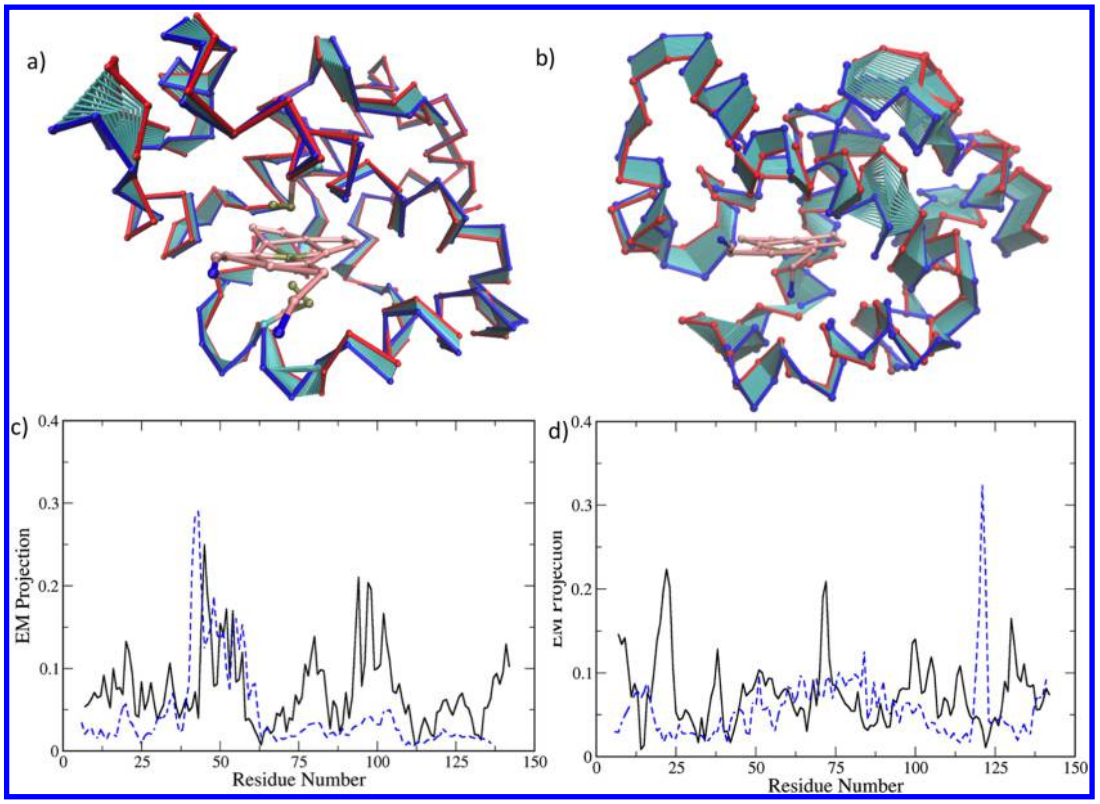


Figure 4. The 6c ⇔ 5c transition essential mode for myoglobin and neuroglobin in the CG representation. Panels (a) and (b) show superimposed structures along the 6c−5c transition essential mode for Ngb and Mb, respectively. Panels (c) and (d) show the projection of the mode along the protein sequence for Ngb and Mb, respectively. Results for CG are shown in black solid lines, and for AA blue dashed lines.

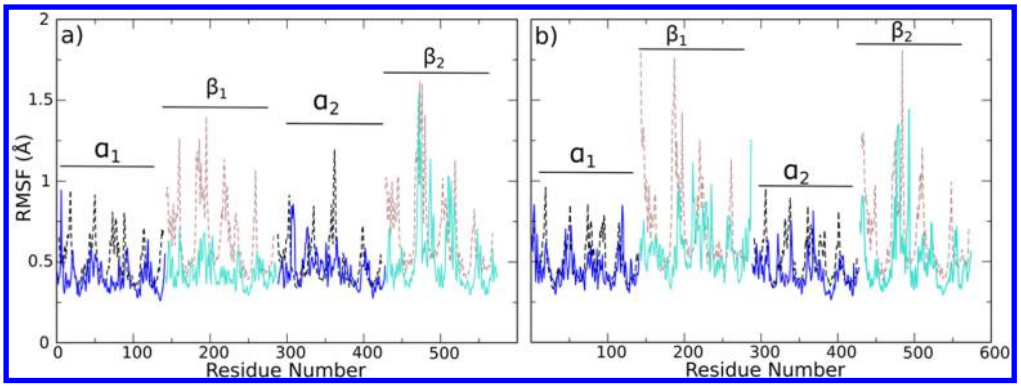


Figure 5. RMSF per residue in the CG (solid blue and cyan lines) and atomistic (dashed black and brown lines) MD simulations. Panels (a) and (b) show results for the R and T states, respectively. Here, α_1 and β_1 correspond to the α and β subunits of the first subunit, and α_2 and β_2 to the α and β subunits of the second subunit.

heterodimers, each one containing two subunits named α and β. In order to further validate the heme CG model for larger proteins and evaluate the ability of the model to describe allosteric transitions, we performed CG simulations of tetrameric human hemoglobin. Specifically, 0.2 μs equilibrium CG−MD simulations were performed starting from the crystal structures corresponding to the T and R states. In both cases, the structures remained stable, with an average RMSD of 4.4 Å in both R and T trajectories with respect to the initial structure, which is again acceptable in the context of the CG model (Figure S3). Additionally, the CG trajectories were compared with atomistic molecular dynamics simulations of tetrameric hemoglobin in both T and R states. RMSF calculations indicate that the fluctuations per residue observed in the CG trajectories are in very good agreement with the observed fluctuations in the atomistic simulation (Figure 5). Generally, the fluctuation patterns observed in atomistic simulations are conserved in the CG simulations. Interestingly, the larger fluctuations obtained in the β subunits with respect to the α subunits are also observed in CG simulations, showing the ability of the model to describe the flexibility difference between the alpha and beta subunits.

In order to analyze if this CG model is able to explore the transition between the T and R states, we performed steered molecular dynamics (SMD) simulations starting from either the T and R states, using as reaction coordinate the RMSD with respect to the target structure, as explained in the Computational Methods section. Several structural parameters were

DOI: 10.1021/acs.jctc.6b00278
J. Chem. Theory Comput. 2016, 12, 3390−3397

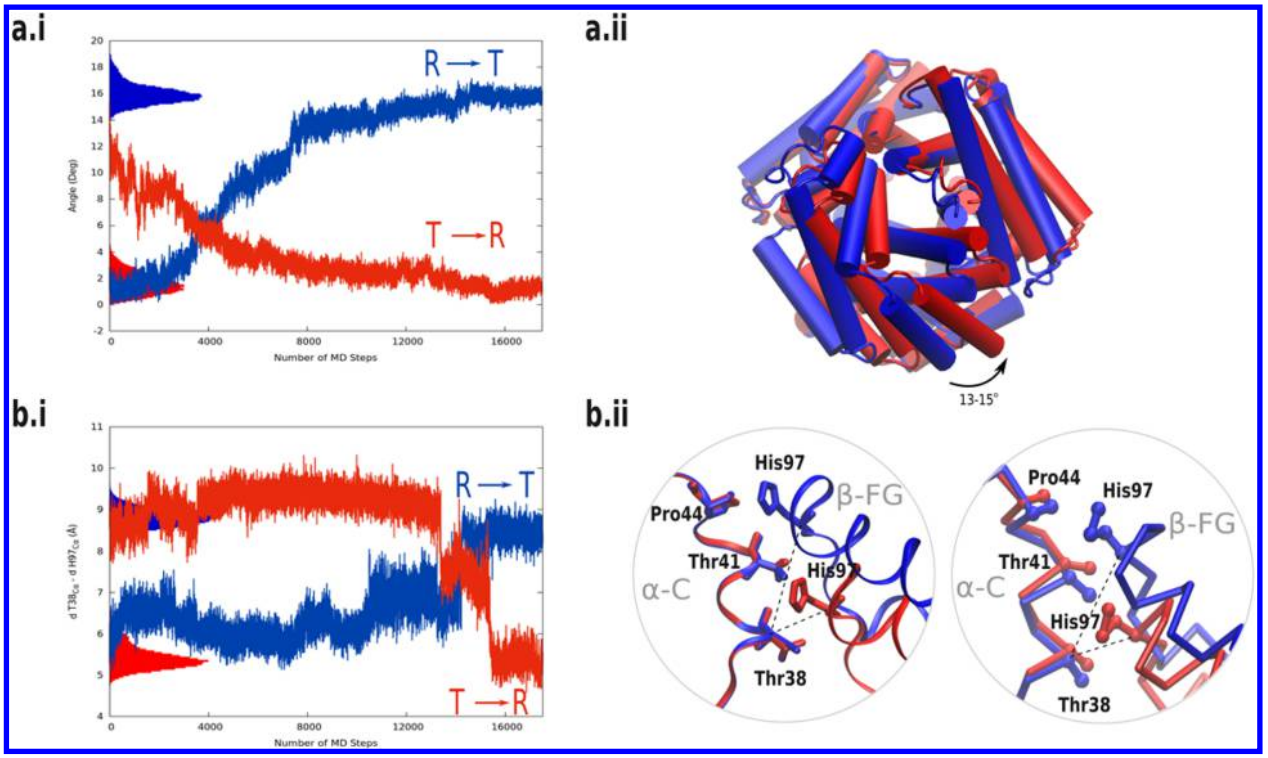


Figure 6. Analysis of the R−T transition in human hemoglobin. Transition from T → R and R → T are shown in red and blue lines, respectively. Distribution of the geometrical parameters along the equilibrium molecular dynamics for each state is shown, for reference, in red for the R state and in blue for the T state. (a) Rotation of the α_1/β_1 dimer with respect to the α_2/β_2, using as a reference the R state: (i) Evolution of the rotational angle and (ii) superimposed structures of the T and R state showing the angle rotation. (b) Interface switch in the T−R transition: (i) Evolution of the Thr38Cα−His97Cα distance along the transition and (ii) crystal (left) and coarse-grained structure (right) switch region. In panels (a.ii) and (b.ii), structures of R and T states are shown in red and blue, respectively.

analyzed along the SMD trajectory in order to characterize the path followed by the system in the CG trajectory and compare it with recent results obtained for this transition with atomistic methodologies.[27,28]

One of the most employed parameters to characterize the R−T quaternary transition is the rotation of the α_1/β_1 dimer with respect to the α_2/β_2 dimer, which shows a variation of 12°−15° between the two states[27,28] (Figure 6a). Figure 6a.i) shows the variation of this angle for the MD of R and T states (red and blue histograms, respectively) and also for the R−T transition (blue and red traces), using as a reference the R state. Clearly, while the equilibrium trajectories of the T and R states conserve their conformation, the trajectory obtained with SMD moves from the T to the R state and vice versa. This global structural parameter is accompanied as well by changes in the interface between the α and β subunits, the most important and well characterized is the His97 switch.[27] This change corresponds to a movement of the His97 side chain of the beta subunit, which is located between Thr38 and Thr41 of the α subunit in the R state and between Thr41 and Pro44 in the T state (Figure 6b). Figure 6b.I shows the distribution of the His97Cα−Thr38Cα distance in the R and T states (red and blue histograms, respectively) and the evolution of this distance along the transition trajectories. The results indicate that while in each state the position of the switch is conserved, in the transition between the T and R states, the SMD simulation is able to induce the change in the switch position. Finally, we calculated the transition essential mode between the R and T states as explained in the Computational Methods section. As in the previous analysis, we observe that while the equilibrium trajectories of the T and R states show sharp distributions of principal essential mode projection values, the SMD transitions cover the projection values between the T and R states (Figure S4). Altogether, these results indicate that for the reference heme protein human Hb, the coarse-grained model with the presently developed parameters for the heme group allows us not only to simulate the equilibrium T and R states but also to explore the transition between both allosteric states, correctly sampling both global quaternary rearrangements as well as local adjustments of the side chain positions.

DISCUSSION

In this work, we present a CG model of the heme group to be used in CG−MD simulations of heme proteins using the force field developed by Voth et al.[7] Parametrization of prosthetic groups and cofactors for CG models is not a trivial procedure since in contrast to parametrization for atomistic MD it is necessary to make an initial decision, for which there is not a unique answer: the position of the CG beads. Depending on the CG model, the position of the beads can be determined using similar assumptions than the ones made to determine the beads positions in the standard amino acids. However, often new criteria need to be applied and tested in order to define the CG representation for a specific chemical entity. Also, all the bonded and the nonbonded interactions need to be parametrized, which can be a complex problem. For instance, in the model by Voth et al., the coarse grain mapping reduces the transferability of the bonded parameters up to the point that for each amino acid side chain specific parameters are obtained. Thus, in this kind of model, it is required to develop bonded

DOI: 10.1021/acs.jctc.6b00278
J. Chem. Theory Comput. 2016, 12, 3390−3397

parameters for all the bonded interactions present in the prosthetic group or cofactor to be parametrized. Additionally, nonbonded interactions need to be described, which again requires the description of the new CG beads with the already available nonbonded parameters or the development of new interaction parameters. In the case of the heme group, we were able to assign the existing atom types for nonbonded parameters to all the CG beads but specific bonded parameters needed to be developed for all the bonded terms in the heme group.

The heme model developed in this work allows performing stable CG−MD simulations of different heme proteins using the force field of Voth et al. We demonstrate that this model can also describe conformational changes in heme proteins, with diverse complexity. In particular, we analyzed the hexacoordination process in monomeric globins myoglobin and neuroglobin and the R−T allosteric transition in tetrameric hemoglobin. In the first case, the process involves a rearrangement of a specific protein region (the CD region), which is observed in the essential dynamics analysis performed in the CG trajectories. Furthermore, the model is able to distinguish between a naturally hexacoordinated protein, as neuroglobin, and a naturally pentacoordinated protein, as myoglobin, showing the dynamical differences between the two proteins.

In the case of tetrameric hemoglobin, we not only observed stable equilibrium trajectories for the T and R states but also by means of steered molecular dynamics we were able to explore the complex R−T allosteric transition. Comparison with atomistic molecular dynamics simulations indicates that CG−MD simulations are able to explore a similar phase space with a significant reduction of the system size. In spite of the loss of atomistic detail, processes that involve local or global rearrangements can be described from a structural and a dynamical viewpoint but with a significant reduction of the computational cost.

The results presented in the present work for the Hb T-to-R transition show that it is possible to explore the underlying conformational change with a biased sampling scheme. Therefore, in principle, it is possible to obtain the corresponding free energy profiles and thus a direct link to experimental derived data such as equilibrium and kinetic constants. It is important to note, however, that the accuracy and thus the relevance of the comparison will depend on the capacity of the CG force field to correctly describe the underlying energy surface of the system. We envisage that our heme-developed parameters, together with other similar applications, will increase the applicability of the CG force field to a wider range of systems that should allow exploration of its applicability to obtain relevant thermodynamic parameters.

Additionally, the question arises on how are the time scales of CG simulations related to the time scales observed experimentally. CG simulations imply the reduction of the degrees of freedom of the system, and thus, a simplification of the free energy surface typically related to an increased softness of the surface curvature. It might imply as well, depending on the model, a modification of kinetic barriers. Thus, the time scale associated with a process does not necessarily need to be related with experimental time scales. In the case of our calculations on the R−T transition in hemoglobin, atomistic MD simulations indicate that spontaneous transitions are observed in the 10−100 ns time scale.[28] Experimental information obtained with Raman spectroscopy[29] and time resolved X-ray scattering[30] show that tertiary and quaternary transitions occur in the nanosecond to microsecond time scale. Therefore, we can consider that our simulation results are consistent with the time scale of the actual experimental transitions, although exact comparisons are obviously not possible from these simulations.

Given the ubiquity and relevance and their widespread use as a model system to study biological process at the molecular level, we expect the current development to provide a useful tool for those who want to study heme proteins as part of large heme protein complexes and/or their conformational transitions in the millisecond to microsecond time scale.

CONCLUSIONS

In this work we developed a heme coarse-grained model which can be used in the context of the CG force field of Voth et al. We demonstrate that this approach can be applied to the study of different conformational changes in heme proteins, retaining the key structural and dynamical information that govern the processes. The cases studied here involve from small protein regions (as the movement of the CD region in neuroglobin and its influence in globin hexacoordination) to large structural transitions as the transition from R to T states in human hemoglobin, opening the possibility to study the complex dynamic processes in heme proteins with coarse-grained simulations.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.6b00278.

RMSD analysis for Ngb, Mb, and Hb simulations, RMSF analysis for Ngb simulations using different time frames, and histograms of the projection of the first R−T transition essential mode in the MD−CG trajectories. (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: lula@qi.fcen.uba.ar.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported in part by grants from the Universidad de Buenos Aires (UBACYT project 20020120300025BA), Agencia Nacional de Promoción Científica y Tecnológica (project PICT 2012-2571), CONICET (PIP 112 201101 0085), Ministerio de Ciencia, Tecnología e Innovación Productiva (RAICES-Siembra), and National Institute of Health Grants (project GM086482). C.R. holds a CONICET Ph.D. Fellowship. M.A.M., D.A.E., and L.C. are members of CONICET. Additionally, this research is part of the Blue Waters Sustained-Petascale Computing Project, which is supported by the National Science Foundation (Award No. OCI 07-25070) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana−Champaign and its National Center for Supercomputing Applications. Research funding comes from the National Science Foundation (OCI-1440031). The authors thank the High-Performance Comput-

DOI: 10.1021/acs.jctc.6b00278
J. Chem. Theory Comput. 2016, 12, 3390−3397

ing Center at the University of Florida for providing further computational resources.

REFERENCES

(1) Gros, G.; Wittenberg, B. A.; Jue, T. Myoglobin's Old and New Clothes: From Molecular Structure to Function in Living Cells. *J. Exp. Biol.* **2010**, *213*, 2713−2725.
(2) Perutz, M. F.; Wilkinson, A. J.; Paoli, M.; Dodson, G. G. The Stereochemical Mechanism of the Cooperative Effects in Hemoglobin Revisited. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 1−34.
(3) Danielson, P. B. The Cytochrome P450 Superfamily: Biochemistry, Evolution and Drug Metabolism in Humans. *Curr. Drug Metab.* **2002**, *3*, 561−597.
(4) Margoliash, E. Primary Structure and Evolution of Cytochrome C. *Proc. Natl. Acad. Sci. U. S. A.* **1963**, *50*, 672−679.
(5) Frauenfelder, H.; McMahon, B. H.; Fenimore, P. W. Myoglobin: The Hydrogen Atom of Biology and a Paradigm of Complexity. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 8615−8617.
(6) Saunders, M. G.; Voth, G. A. Coarse-Graining Methods for Computational Biology. *Annu. Rev. Biophys.* **2013**, *42*, 73−93.
(7) Hills, R. D.; Lu, L.; Voth, G. A. Multiscale Coarse-Graining of the Protein Energy Landscape. *PLoS Comput. Biol.* **2010**, *6*, e1000827.
(8) Monticelli, L.; Kandasamy, S. K.; Periole, X.; Larson, R. G.; Tieleman, D. P.; Marrink, S.-J. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.* **2008**, *4*, 819−834.
(9) Darré, L.; Machado, M. R.; Brandner, A. F.; González, H. C.; Ferreira, S.; Pantano, S. SIRAH: A Structurally Unbiased Coarse-Grained Force Field for Proteins with Aqueous Solvation and Long-Range Electrostatics. *J. Chem. Theory Comput.* **2015**, *11*, 723−739.
(10) Abe, H.; Go, N. Noninteracting Local-Structure Model of Folding and Unfolding Transition in Globular Proteins. II. Application to Two-Dimensional Lattice Proteins. *Biopolymers* **1981**, *20*, 1013−1031.
(11) Ward, A. B.; Guvench, O.; Hills, R. D. Coarse Grain Lipid−protein Molecular Interactions and Diffusion with MsbA Flippase. *Proteins: Struct., Funct., Genet.* **2012**, *80*, 2178−2190.
(12) Pesce, A.; Bolognesi, M.; Bocedi, A.; Ascenzi, P.; Dewilde, S.; Moens, L.; Hankeln, T.; Burmester, T. Neuroglobin and Cytoglobin. Fresh Blood for the Vertebrate Globin Family. *EMBO Rep.* **2002**, *3*, 1146−1151.
(13) Berendsen, H.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684.
(14) Capece, L.; Boechi, L.; Perissinotti, L. L.; Arroyo-Mañez, P.; Bikiel, D. E.; Smulevich, G.; Marti, M. A.; Estrin, D. A. Small Ligand-Globin Interactions: Reviewing Lessons Derived from Computer Simulation. *Biochim. Biophys. Acta, Proteins Proteomics* **2013**, *1834*, 1722−1738.
(15) Marti, M. A.; Capece, L.; Bidon-Chanal, A.; Crespo, A.; Guallar, V.; Luque, F. J.; Estrin, D. A. Nitric Oxide Reactivity with Globins as Investigated through Computer Simulation. *Methods Enzymol.* **2008**, *437*, 477−498.
(16) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. AMBER 11, University of California, San Francisco, 2010.
(17) Yang, F.; Phillips, G. N., Jr. Crystal Structures of CO-, Deoxy- and Met-Myoglobins at Various pH Values. *J. Mol. Biol.* **1996**, *256*, 762−774.
(18) Pesce, A.; Dewilde, S.; Nardini, M.; Moens, L.; Ascenzi, P.; Hankeln, T.; Burmester, T.; Bolognesi, M. Human Brain Neuroglobin Structure Reveals a Distinct Mode of Controlling Oxygen Affinity. *Structure* **2003**, *11*, 1087−1095.
(19) Fermi, G.; Perutz, M. F.; Shaanan, B.; Fourme, R. The Crystal Structure of Human Deoxyhaemoglobin at 1.74 A Resolution. *J. Mol. Biol.* **1984**, *175*, 159−174.
(20) Shaanan, B. Structure of Human Oxyhaemoglobin at 2.1 A Resolution. *J. Mol. Biol.* **1983**, *171*, 31−59.
(21) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber ff99SB Protein Force Field. *Proteins: Struct., Funct., Genet.* **2010**, *78*, 1950−1958.
(22) Capece, L.; Marti, M. A.; Bidon-Chanal, A.; Nadra, A.; Luque, F. J.; Estrin, D. A. High Pressure Reveals Structural Determinants for Globin Hexacoordination: Neuroglobin and Myoglobin Cases. *Proteins: Struct., Funct., Genet.* **2009**, *75*, 885−894.
(23) Boron, I.; Capece, L.; Pennacchietti, F.; Wetzler, D. E.; Bruno, S.; Abbruzzetti, S.; Chisari, L.; Luque, F. J.; Viappiani, C.; Marti, M. A.; Estrin, D. A.; Nadra, A. D. Engineered Chimeras Reveal the Structural Basis of Hexacoordination in Globins: A Case Study of Neuroglobin and Myoglobin. *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850*, 169−177.
(24) Amadei, A.; Linssen, A. B.; Berendsen, H. J. Essential Dynamics of Proteins. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 412−425.
(25) Perutz, M. F.; Mathews, F. S. An X-Ray Study of Azide Methaemoglobin. *J. Mol. Biol.* **1966**, *21*, 199−202.
(26) Boechi, L.; Arrar, M.; Martí, M. A.; Olson, J. S.; Roitberg, A. E.; Estrin, D. A. Hydrophobic Effect Drives Oxygen Uptake in Myoglobin via Histidine E7. *J. Biol. Chem.* **2013**, *288*, 6754−6762.
(27) Fischer, S.; Olsen, K. W.; Nam, K.; Karplus, M. Unsuspected Pathway of the Allosteric Transition in Hemoglobin. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 5608−5613.
(28) Hub, J. S.; Kubitzki, M. B.; de Groot, B. L. Spontaneous Quaternary and Tertiary T-R Transitions of Human Hemoglobin in Molecular Dynamics Simulation. *PLoS Comput. Biol.* **2010**, *6*, e1000774.
(29) Jones, E. M.; Balakrishnan, G.; Spiro, T. G. Heme Reactivity Is Uncoupled from Quaternary Structure in Gel-Encapsulated Hemoglobin: A Resonance Raman Spectroscopic Study. *J. Am. Chem. Soc.* **2012**, *134*, 3461−3471.
(30) Cammarata, M.; Levantino, M.; Wulff, M.; Cupane, A. Unveiling the Timescale of the R-T Transition in Human Hemoglobin. *J. Mol. Biol.* **2010**, *400*, 951−962.

DOI: 10.1021/acs.jctc.6b00278
J. Chem. Theory Comput. 2016, 12, 3390−3397