

Article

Analysis of fluctuation in the heme-binding pocket and heme distortion in hemoglobin and myoglobin

Hiroko X. Kondo ^{1,2,3,*}, and Yu Takano ²

¹ School of Regional Innovation and Social Design Engineering, Faculty of Engineering, Kitami Institute of Technology, 165 Koen-cho, Kitami, Hokkaido, 090-8507, Japan; h_kondo@mail.kitami-it.ac.jp

² Department of Biomedical Information Sciences, Graduate School of Information Sciences, Hiroshima City University, 3-4-1 Ozukahigashi Asaminamiku, Hiroshima, Hiroshima, 731-3194, Japan; ytakano@hiroshima-cu.ac.jp

³ Laboratory for Computational Molecular Design, RIKEN Center for Biosystems Dynamics Research, 6-2-3, Furuedai, Suita, Osaka, 565-0874, Japan

* Correspondence: h_kondo@mail.kitami-it.ac.jp; Tel.: +81-157-26-9401

Abstract: Heme is located in the active site of proteins and has diverse and important biological functions, such as electron transfer and oxygen transport and/or storage. The distortion of heme porphyrin is considered an important factor for the diverse functions of heme because it correlates with the physical properties of heme, such as oxygen affinity and redox potential. Therefore, clarification of the relationship between heme distortion and the protein environment is crucial in protein science. Here, we analyzed the fluctuation in heme distortion in the protein environment for hemoglobin and myoglobin using molecular dynamics (MD) simulations and quantum mechanical (QM) calculations. We also investigated the protein structures of hemoglobin and myoglobin stored in Protein Data Bank and found that heme is distorted along the doming mode, which correlates with its oxygen affinity, more prominently in the protein environment than in the isolated state, and the magnitude of distortion is different between hemoglobin and myoglobin. This tendency was also observed in the results of MD simulations and QM calculations. These results suggest that heme distortion is affected by its protein environment and fluctuates around its fitted conformation, leading to physical properties that are appropriate for protein functions.

Keywords: heme distortion; pocket rigidity; protein environment; hemoglobin; myoglobin; MD simulation; ONIOM

1. Introduction

Heme is a complex of iron and porphyrin that is present at the active site of heme proteins. Heme proteins, a group of proteins that bind heme(s) as a cofactor, perform diverse biological functions, such as oxygen binding (transport or storage)[1,2], electron transfer[3], redox reactions[4], microRNA processing[5], and transcriptional regulation [5,6], indicating that heme molecules have diverse functions. Although the functional mechanisms of individual heme proteins have been thoroughly studied, elucidating the mechanisms that regulate the diverse functions of heme is still challenging. To solve this issue, we have been investigating the structure–function relationship in heme proteins by focusing on the distortion of heme porphyrin. Heme distortion has recently been considered an important factor underlying the diverse functions of heme, because it correlates with the physical properties of heme, such as oxygen affinity and redox potential[8–11]. One of our interests is the structural regulation of heme by its host protein environment, and our recent study suggests a correlation between the amino acid composition of the heme-binding pocket and particular heme distortions[12]. According to a survey of the structures of heme proteins stored in the Protein Data Bank (PDB; <http://www.rcsb.org/>)[13], hemes in proteins with the same amino acid sequence have similar conformations.

Some studies have reported a correlation between the protein environment and heme distortion, suggesting that the protein environment of the heme-binding pocket is probably rigid and affects heme distortion. Li et al. demonstrated the stability of the heme-binding pocket by comparing protein structures in the apo (heme-unbound) and holo (heme-bound) states[14]. Most apo-holo pairs exhibit a small conformational change after heme binding (root mean square deviation of 1.03 Å or less). Furthermore, Sacquin-Mora et al. investigated the local flexibility of the active-site residues using Brownian dynamic simulations and showed that the amino acid residues around the heme-binding pocket must be tightly anchored to enable its biological function, except for the residues that are mobile according to their protein function[15]. We thus investigated the fluctuation of the heme-binding pocket and heme distortion in the host protein environment and the correlation between them.

We focused on two heme proteins, hemoglobin (Hb) and myoglobin (Mb). Although both have oxygen-binding activities, their functions are different. Whereas Mb is a monomeric protein that stores oxygen and facilitates its diffusion in muscle, Hb is a tetrameric protein that transports oxygen in the blood. Mb and Hb tertiary structures were determined by the earliest crystallographic studies[1,16], and their functional mechanisms have been thoroughly investigated. Even in recent years, some studies have focused on their dynamics. Aharoni and Tobi compared the dynamics of Mb and Hb using algorithms for alignment of Anisotropic Network Model modes of motion and indicated that the quaternary structure of Hb affects the intrinsic dynamics of each domain, leading to functional differences between Mb and Hb[17]. Tobi also carried out the clustering of globin family proteins based on their dynamics and succeeded in distinguishing the different states of Hb[18]. Bringas et al. elucidated a mechanism that regulates oxygen affinity through structural changes in Hb subunits using molecular dynamics (MD) simulations and quantum mechanics (QM)/molecular mechanics (MM) calculations [19]. Based on these meaningful results, it is important to elucidate the effect of these differences in protein dynamics on heme distortion.

In this study, we first estimated the effect of ligand(s) on heme distortion for an isolated heme via QM calculations and then compared heme distortion between an isolated heme and hemes in Hb and Mb based on PDB data. Next, to clarify the relationship between the protein environment and heme distortion, we investigated the fluctuation in the heme-binding pocket and heme distortion in its host protein for Hb and Mb using MD simulations and QM calculations. The results showed differences in flexibility of heme-binding pockets and in tendency of fluctuations in the doming distortion of heme between Hb and Mb, with the latter corresponding to the results obtained from the PDB survey. These differences might be related to differences in function between Hb and Mb.

2. Materials and Methods

2.1. Collation of structural data of hemoglobin and myoglobin

To compare the heme distortions in hemoglobin and myoglobin, first, we downloaded PDBx/mmCIF and fasta files of any kinds of heme proteins carrying heme *b* and heme *c* (Figure 1) with a structural resolution less than or equal to 2.0 Å from PDBj[20] and extracted structural information from them. The compound ID (_chem_comp.id) HEM, HEB, or HEC was used to search for heme proteins via SQL in the PDBj Mine relational database (<https://pd bj.org/mine>)[21]. Consequently, 3070 unique entries were selected from all PDB entries (as of December 22, 2021). For each entry, the structural data of heme without missing data of the atomic coordinates of the 25 heavy atoms composing the Fe-porphyrin skeleton, shown in Figure 1, were extracted. We obtained 3063 entries. The Biopython library[22] was used as a parser and structural data were collected using the MDTraj library[23]. Next, we carried out the clustering for amino acid sequences for collated entries using Cd-hit[24] with a clustering threshold of a sequence identity of 90 % and obtained the clusters of PDB chains of the α chain of hemoglobin, the β chain of

hemoglobin, or myoglobin. The cluster of the α chain of hemoglobin has 235 chains, that of the β chain of hemoglobin has 287 chains, and that of myoglobin has 279 chains. Heme porphyrin distortion was analyzed using normal-coordinate structural decomposition[25], which is a typical approach for evaluating heme distortion. Only the saddling, ruffling, and doming modes were considered in this study. The details of the calculation were described in a previous study[26].

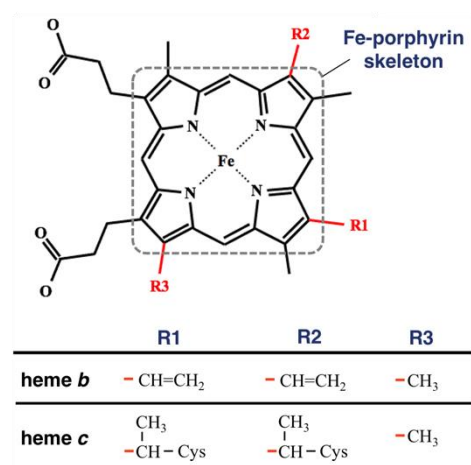


Figure 1. Structure of heme *b* and heme *c*. The red lines in the bottom table correspond to the bonds in the structure of the heme molecule. The atoms composing the Fe-porphyrin skeleton are surrounded by a square with a dotted line.

2.2. Setup of simulation systems

In this study, we considered six systems, deoxy-Mb (dMb), oxy-Mb (oMb), apo-Mb (apoMb), deoxy-Hb (dHb), oxy-Hb (oHb), and apo-Hb (apoHb). The initial structures of the simulations were prepared using the atomic coordinates determined by X-ray crystallographic studies (PDB ID: 1A3N[27] for dHb and apoHb (human), 2DN1[28] for oHb (human), 1MBD (sperm whale) for dMb and apoMb, and 1MBO[29] (sperm whale) for oMb). The initial structures of an apo protein were prepared by removing the heme molecule(s) from its complex structure in the deoxy state. All ionizable side chains were configured in their ionized states at pH 7.0, as calculated using the H++ server[30], except for 146His in oHb. Because protonation of the C-terminal histidines of Hb β chains, namely 146His (HIP), is an important factor for stabilizing the quaternary T state[31,32], only the N ϵ atom of 146His was protonated in the oxygenated state. Each system was solvated with a 150 mM NaCl aqueous solution in a rectangular box. The distance between the solute and box was set to 1.0 nm.

2.3. Simulation details

All simulations were conducted using GROMACS software, version 2019.5[33]. The Amber ff14SB force field[34] was applied to normal amino acids and ions, and the TIP3P model[35] was applied to water molecules. For heme molecules, we used the force field according to Bringas et al.[19]. The periodic boundary condition was applied, and electrostatic interactions were treated using the Particle Mesh Ewald method[36]. The cutoff distances for both the Ewald real space and van der Waals truncation were 1.2 nm. The neighbor list was updated every 20 steps for MD simulations. The simulation protocols were the same for all systems.

After the energy minimizations on the whole system, we gradually heated the system from 0.5 K to 300 K for 120 ps (50 K/20 ps) using a V-rescale thermostat[37]. Subsequently, two equilibration processes with the position restraints of the heavy atoms of the proteins

were applied to relax the hydrogen atoms as follows: 100 ps NVT (constant volume and constant temperature) simulation with force constants of 250 kJ mol⁻¹ nm⁻² for backbone atoms and 50 kJ mol⁻¹ nm⁻² for sidechain atoms and 200 ps NPT (constant pressure and constant temperature) simulation with force constants of 50 kJ mol⁻¹ nm⁻² for backbone atoms. We performed each 200 ns production run under NPT conditions, with a time step of 2 fs. The temperature and pressure were maintained at 300 K and 1 bar with a Nosé–Hoover thermostat[38–40] and Parrinello–Rahman barostat[41,42], respectively. All bonds involving hydrogen atoms were constrained using the LINCS algorithm[43]. For the simulation of Mb (dMb and oMb), a distance restraint was applied between the Fe atom of heme and the nitrogen atom (NE2) of the ligand histidine with a harmonic potential with a force constant of 1000 kJ mol⁻¹ nm⁻² because the coordination between heme and the ligand histidine could not be maintained without restraints. The reference distance was set as the coordination distance in the X-ray structure as follows: 0.2096 nm for dMb and 0.2065 nm for oMb. Two simulations with different initial velocities were conducted for each system.

2.4. Trajectory analysis of MD simulations

Snapshots were sampled every 20 ps for each trajectory. In the trajectory analyses, we defined "pocket residues" as amino acid residues for which heavy atoms are within 4.5 Å of the Fe-porphyrin skeleton of the heme molecule in the energy-minimized structure (reference structure in the following section). The structural stability of the whole structure and pocket residues was estimated using root mean square deviation (RMSD) values from the reference structure. RMSD values were computed for C α atoms after superimposing each region (whole structure or pocket residues) onto the reference structure. The root mean square fluctuation (RMSF) values were calculated for each heavy atom after least-squares fitting of the C α atoms of the whole protein and averaged over the atoms in each residue. The MDTraj library[23] was used for the analyses.

For the analysis of volume and clustering of the heme-binding pocket, snapshots were sampled every 5 ns for each protein subunit. The pocket volume was computed using POVME 3.0 [44]. The sampled structures (heme–protein complex) were superposed by Fe and the four nitrogen atoms in heme. The volume was calculated using the following parameters: the center of the mean coordinates of the Fe atom of heme as the center of an inclusion sphere and 8.5 Å as the radius of an inclusion sphere. Clustering of the heme-binding pocket was also performed with the results obtained from the volume calculation using POVME 3.0, and the representative structures of each cluster were obtained.

2.4. Model construction for QM calculations

In this study, we used three types of models as follows: the QM model of isolated heme (1), our own N-layered integrated molecular orbital and molecular mechanics (ONIOM) model of heme complexed with its host protein (2), and the QM model of heme and adjacent residues (3). (1) A model of isolated heme with or without ligands consists of heme or heme and its ligand(s). The initial coordinates were extracted from the X-ray structures of deoxygenated and oxygenated hemoglobin (PDB IDs: 1A3N and 2DN1, respectively). (2) The ONIOM model includes heme and its host protein subunit, for example, an Hb chain. The high layer consists of heme and its ligand(s), and the C α atom of the ligand, histidine, is replaced with a hydrogen atom. (3) The QM model of heme and adjacent residues is composed of heme and amino acid residues, for which heavy atoms are within 4.5 Å of the Fe-porphyrin skeleton of heme. The backbone atoms of the amino acid residues were removed and the C α atoms were replaced with hydrogen atoms. X-ray structures (PDB ID: 1MBD, 1MBO, and 1A3N (chain A)) were used as the initial structures.

2.5. Estimation of heme porphyrin distortion based on QM calculations

We performed geometry optimizations on models (1) to (3) in the aforementioned section using the PBE0 hybrid function[45] with the 6-31G(d) basis set[46–48]. For model (1), the coordinates of all atoms were optimized, and the protein environment or water

solvent was included using the polarizable continuum model (PCM)[49–51] with a dielectric constant of 4.0 or 78.4, respectively. For model (2), only the coordinates of heme and its ligand(s) were optimized (backbone atoms except the C α atom of the ligand histidine were fixed). An AMBER force field was used for the low (molecular mechanics) layer. For model (3), only the coordinates of the heme, oxygen atoms in the O $_2$ molecule, and hydrogen atoms in the amino acid residues were optimized. The six dihedral angles involved in the propionates were fixed during optimization, and the protein environment was included using PCM with a dielectric constant of 4.0. All QM calculations were performed using the Gaussian16 program package[52].

3. Results and Discussion

3.1. Distortion of heme porphyrin in isolated heme

Bikiel et al.[8] analyzed the distortion of the Fe-porphyrin molecule (heme without side-chains) in the 5-coordinate state (mono His-coordinated) and 6-coordinate state (His-O $_2$ -coordinated) through DFT calculations and showed that the doming distortion tends to decrease from the 5-coordinate state to the 6-coordinate state. In this study, to examine the effect of ligand(s) on its distortions in heme *b* (the chemical structure is presented in Figure 1), we computed the porphyrin distortion of heme *b* in the mono His-coordinate and His-O $_2$ coordinate states. The coordinates of all atoms were optimized (no atom was fixed) by PBE0/6-31G(d) with the PCM with dielectric constants of 4.0 and 78.4 to incorporate the electrostatic protein environment and water solvent, respectively. The porphyrin distortions along the saddling, ruffling, and doming modes in the optimized structures are listed in Table 1. The equilibrium structure of isolated heme *b* was nearly the same as that of the Fe-porphyrin molecule, although it was slightly distorted along the saddling and ruffling modes (only in the result with a dielectric constant of 78.4; a distortion of 0 corresponds to the equilibrium structure of the Fe-porphyrin molecule). The magnitude of the ruffling distortion tends to increase from the 5-coordinate state to the 6-coordinate state, whereas that of the doming distortion has the opposite tendency. These tendencies are consistent with a previous study that analyzed the distortion of Fe-porphyrin molecules, by Bikiel et al.[8]. The difference in the dielectric constant as a parameter of the PCM did not significantly affect the equilibrium structure of heme *b*.

Table 1. Porphyrin distortions in heme *b*.

System	Dielectric constant	Saddling	Ruffling	Doming
heme <i>b</i>	4.0	-0.133	-0.001	-0.004
His-coordinated heme <i>b</i>	4.0	-0.150	-0.040	-0.115
His-O $_2$ -coordinated heme <i>b</i>	4.0	-0.065	-0.642	0.069
heme <i>b</i>	78.4	-0.136	0.105	-0.004
His-coordinated heme <i>b</i>	78.4	0.005	0.034	-0.138
His-O $_2$ -coordinated heme <i>b</i>	78.4	-0.043	-0.393	0.016

3.2. Difference in heme distortions between Hb and Mb homologs in the oxy-state and deoxy-state

To examine the effect of the protein pocket environment on heme distortion, we collated the structural data of Hb and Mb from PDB and analyzed the distribution of heme distortions with 5-coordinate (mono-His coordinated) heme and 6-coordinate heme, for which ligands are His-O $_2$ (oxygen molecule) or His-CO (carbon monoxide). The homologous protein groups of the Hb α/β chain and Mb were obtained by clustering the amino

acid sequences of heme proteins containing heme *b* or heme *c* with a threshold of 90 % sequence similarity. Figure 2 shows the heme distortion along the saddling, ruffling, and doming modes in the Hb α/β chain and Mb. Whereas no significant difference was observed between the 5-coordinated and 6-coordinated hemes in the saddling and ruffling distortions, except for the Hb β chain, the doming distortion (absolute values) tended to decrease from the 5-coordinate state to the 6-coordinate state. This corresponds to the tendency observed in isolated heme *b* in Section 3.1.

In addition, a difference was observed in the doming distribution between the Hb α/β chain and Mb, even though they have similar functions (oxygen binding) and structure. The mean values of the doming distortion in each subunit in the His-coordinated and His-O₂-coordinated states were -0.633 (HIS) and -0.354 (HIS-OXY) for the Hb α chain, -0.467 (HIS) and -0.264 (HIS-OXY) for the Hb β chain, and -0.321 (HIS) and -0.221 (HIS-OXY) for Mb. The change in doming distortion from the deoxygenated state to the oxygenated state in Mb was smaller than that in Hb. Since doming distortion is known to correlate with the oxygen affinity of heme, these differences in doming distortions could be involved in their protein function. Our previous study[12] suggested a correlation between the protein environment and heme distortion along saddling, ruffling, and doming modes. Considering these points, a protein environment might control the heme distortion, as it is appropriate to protein function. In the following, we focus on the doming distortion of heme molecules in Hb and Mb.

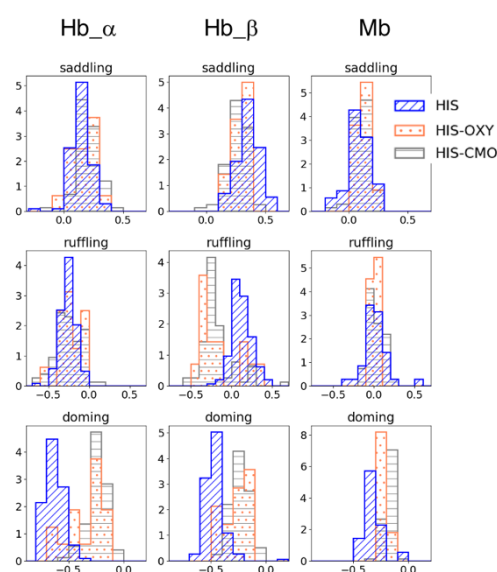


Figure 2. Porphyrin distortions of hemes in Hb subunits (the left two columns) and Mb (the right column). The blue with diagonal line represents the distortion in mono His-coordinated heme (HIS), the coral with dots; in His-O₂-coordinated heme (HIS-OXY), and the gray with horizontal stripes; in His-CO-coordinated heme (HIS-CMO). The first, second, and third rows show the saddling, ruffling, and doming distortions, respectively.

3.3. Stability of protein whole structure and conformation of heme-binding pocket

We computed the RMSD values of C α atoms for the whole protein structure and pocket residues to analyze the difference in structural stability between the deoxy-heme-bound, heme-unbound, and oxy-heme-bound proteins. Mean values and standard deviations are presented in Table 2. RMSD values were averaged over all snapshots in the two trajectories for each system. For the Hb trajectories, RMSD values were averaged for each

subunit (α or β chain). The entire protein structure was very stable in the deoxy-heme-bound and oxy-heme-bound proteins, and no significant difference was observed between them. In the apo (heme-unbound) protein, the RMSD values were larger than those in the holo (deoxy-heme-bound) state, but less than 2.0 Å. Compared to the RMSD values of the pocket residues (Table 2b), the structural stability was high, particularly in the heme-binding pocket in the holo proteins (deoxy and oxy), and was low in the heme-binding pocket in the apo protein, suggesting that heme binding stabilizes the pocket conformation of its host protein. In the comparison between apoHb and apoMb, the standard deviation of RMSD was much larger, especially in the heme-binding pocket in Mb, than that in the Hb α/β chains. This indicated that the pocket conformation of Mb is more flexible than that of Hb.

Table 2. The mean values and standard deviation of root mean square deviation (RMSD) values of the whole protein structure (a) and heme-binding pocket (b).

(a) RMSD values of the entire protein structure.

System	RMSD (deoxy) [Å]	RMSD (apo) [Å]	RMSD (oxy) [Å]
Hb α chain	0.716 ± 0.084	1.079 ± 0.217	0.716 ± 0.080
Hb β chain	0.998 ± 0.179	1.702 ± 0.197	0.989 ± 0.119
Mb	1.076 ± 0.145	1.902 ± 0.587	0.941 ± 0.196

(b) RMSD values of the heme-binding pocket

System	RMSD (deoxy) [Å]	RMSD (apo) [Å]	RMSD (oxy) [Å]
Hb α chain	0.541 ± 0.090	1.276 ± 0.310	0.537 ± 0.074
Hb β chain	0.705 ± 0.179	1.901 ± 0.248	0.664 ± 0.101
Mb	0.726 ± 0.141	2.045 ± 0.750	0.559 ± 0.103

3.4. Fluctuations in amino acid residues in the heme-binding pocket

We computed the RMSF values of amino acid residues in the heme-binding pocket to estimate the fluctuation of each residue, including the side chain atoms. The RMSF values for each residue were averaged over two trajectories and plotted for each chain in the graph shown in Figure 3. The black arrows in Figure 3 represent pocket residues. The mean values and standard deviations of the RMSF pocket residues are shown in Table 3. The fluctuation in the pocket residues was very small in the holo (heme-bound) state (less than 0.9 Å) and became slightly larger in the apo state, at approximately 1.0 Å. The increase in the fluctuation amplitude was remarkable in the Hb α chain.

In addition to the fluctuations in each residue, we computed the time course of pocket volumes every 5 ns (Figure 4). No large difference was observed in pocket volumes from the initial state to the last state for all systems. The mean values and standard deviations of the pocket volumes for each subunit averaged over two trajectories (and the same subunits for Hb) were 467.427 ± 31.661 Å³ for dHb α chains, 458.805 ± 31.892 Å³ for dHb β chains, 444.067 ± 26.814 Å³ for oHb α chains, 472.110 ± 33.784 Å³ for oHb β chains, 585.037 ± 55.945 Å³ for dMb, and 510.512 ± 38.262 Å³ for oMb. The mean values averaged over each trajectory are presented in Table S1. Mb had a larger absolute value of the pocket volume and fluctuation amplitude than Hb. Because the variances in RMSD values of the pocket residues in Mb in the holo (deoxy- and oxy-heme-bound) states were almost the same or smaller than those in the Hb β chain, this fluctuation of the pocket volume would be due to the fluctuations of the side chains of the pocket residues.

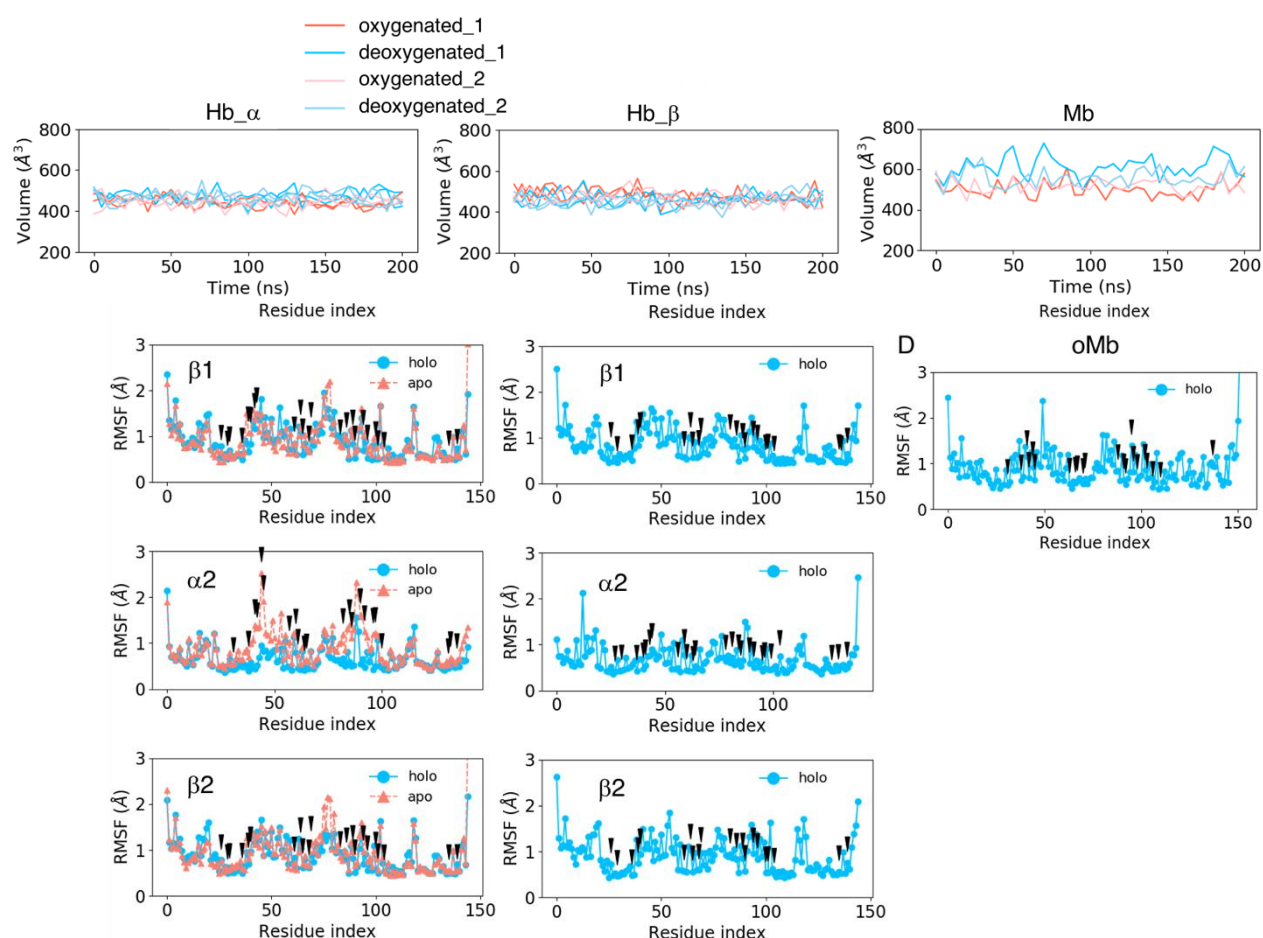


Figure 3. The root mean square deviation (RMSF) values of each residue averaged over two trajectories for dHb (A), oHb (B), dMb (C), and oMb (D). The sky-blue lines represent the holo (heme-bound) proteins, and the coral lines are apo proteins. The black arrows represent the pocket residues.

Table 3. The mean values and standard deviations of root mean square deviation (RMSF) values of the pocket residues.

System	RMSF (deoxy) [Å]	RMSF (apo) [Å]	RMSF (oxy) [Å]
Hb α chain	0.558 ± 0.180	0.913 ± 0.461	0.670 ± 0.454
Hb β chain	0.692 ± 0.172	0.780 ± 0.209	0.699 ± 0.179
Mb	0.849 ± 0.308	1.045 ± 0.357	0.720 ± 0.224

Figure 4. Time courses of volumes of heme-binding pocket for each subunit of Hb (the left two panels) and Mb (the right panel). The coral and pink lines represent the trajectories of proteins in the oxygenated state, and the sky-blue and light blue lines are those in the deoxygenated state.

3.5. Estimation of the effect of protein conformations on heme distortion

Although our previous study suggested a correlation between heme distortion and the composition of amino acid residues around heme, protein dynamics must also affect the conformation (distortion) of heme. We estimated the doming distortion in representative pocket conformations in our simulation trajectories. First, we clustered the

conformations of the heme-binding pocket to extract representative structures. Snapshots were extracted every 5 ns from each trajectory (41 frames per trajectory), and snapshots from two trajectories of the same system were combined (82 frames per system). These snapshots were divided into four clusters using a hierarchical clustering method, and four representative samples of the heme-binding pocket were selected using POVME 3.0. Table 4a lists the number of samples in each cluster. Most samples belonged to cluster 1 or cluster 2, implying that the distribution did not have characteristic clusters.

Next, we optimized the representative conformations of each cluster obtained using POVME 3.0 software by using the ONIOM method (2nd model in Section 2.4) and computed the doming distortion of the optimized structures (the validity of the estimation of doming distortion by ONIOM is discussed in the next section). For two of the 40 clusters, the optimization of a representative structure did not converge, and another conformation in the same cluster was used. For one of these clusters, optimization did not converge in the alternative structure (cluster 4 of oHb β 1). The calculated doming distortions are presented in Table 4b. The mean values over the clusters were the weighted means obtained by multiplying the sample number; for example, the mean for dHb α 1 was calculated as $(-1.058 \times 71 - 0.825 \times 7 - 0.962 \times 2 - 0.689 \times 2) / 82$. Comparing the doming distortions in Hb with those in Mb, we could see that heme molecules tend to be distorted along the doming mode more largely in Hb than in Mb, and the difference in the doming distortion between the oxygen-unbound (mono His-coordinated) state and oxygen-bound (His-O₂-coordinated) state in Hb (Hb α chain) was larger than that in Mb (Hb β chain). These corresponded to the observations from the PDB data mentioned in Section 3.2. The magnitude of distortion was distributed over a wide range in each system, for example, -1.136 to -0.352 for the dHb α chain and -0.760 to -0.129 for dMb. The physical property of proteins appears as their average.

Table 4. The number of members in each cluster (a) and the doming distortions estimated for the representative structures of each cluster (b).

(a)

System	Cluster1	Cluster2	Cluster3	Cluster4
dHb α 1	71	7	2	2
dHb α 2	77	3	1	1
dHb β 1	54	21	6	1
dHb β 2	66	8	6	2
oHb α 1	72	5	4	1
oHb α 2	72	5	3	2
oHb β 1	69	10	2	1
oHb β 2	71	5	3	3
dMb	38	27	11	6
oMb	59	17	5	1

(b)

System	Cluster1	Cluster2	Cluster3	Cluster4	Mean (clusters)	Mean (subunits)
dHb α 1	-1.058	-0.825	-0.962	-0.689	-1.027	
dHb α 2	-1.136	-0.961	-0.352	-0.852	-1.116	
dHb β 1	-0.904	-0.815	-0.828	-0.647	-0.873	-0.980
dHb β 2	-0.905	-0.865	-0.981	-0.763	-0.903	
oHb α 1	-0.517	-0.665	-0.408	-0.451	-0.520	-0.590

oHb α 2	-0.430	-0.329	-0.371	0.026	-0.410	
oHb β 1	-0.817	-0.485	-0.446	-	-0.758	
oHb β 2	-0.702	-0.476	-0.520	-0.455	-0.672	
dMb	-0.760	-0.364	-0.431	-0.129	-0.540	-
oMb	-0.422	-0.422	0.212	-0.559	-0.385	-

We also analyzed heme distortion in the simulation trajectories. Snapshots were sampled every 20 ps, and histograms were generated for the doming distortion (Figure 5). As shown in Table 5, the doming distortion of each system qualitatively corresponded to the results obtained from the QM calculations. Hemes tend to be less distorted along the doming mode in dMb than dHb under the same force field of the heme molecule, suggesting an effect of the protein environment on heme distortion. As mentioned in the Methods section, we constrained the distance between heme and its ligand histidine in the dMb and oMb simulations because coordination could not be maintained without a constraint. This might be caused by the difference in heme conformation between Mb and Hb; the force field of heme used in this study was optimized for heme in hemoglobin and was unsuitable for simulations of heme in Mb. The doming distribution in the unconstrained dMb trajectories is shown in Figure 5B. The distribution was bimodal and showed an inversion of the distortion. It is possible that the conformation of hemes is affected by its binding pocket environment and is important for maintaining coordination with its host protein.

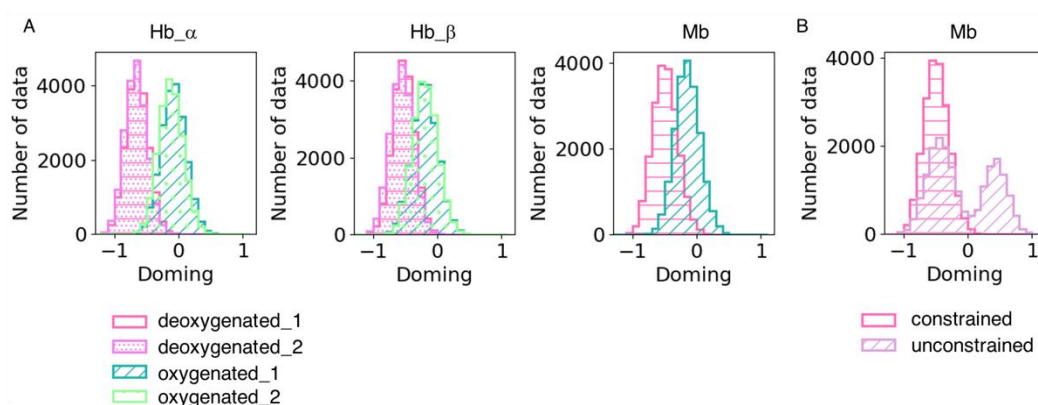


Figure 5. (A) Doming distortions in each subunit in the simulation trajectories. Two trajectories of the same system were combined. The hot pink with horizontal stripe and the violet with dots represent the deoxygenated state, and the sage green with a diagonal line and pale green with rough dots are the oxygenated state. (B) Comparison of doming distortions in the simulation trajectories with and without constraint between heme and its ligand His. The hot pink with a horizontal stripe and the plum with a diagonal line represent the dMb trajectories with and without constraint, respectively.

Table 5. The mean values and standard deviation (std) of the doming distortions for each subunit in each simulation trajectory. The mean values over four subunits were also calculated for dHb and oHb (right column).

System	Mean \pm std	Mean (subunits)
dHb α 1	-0.645 ± 0.173	-0.601
dHb α 2	-0.666 ± 0.173	
dHb β 1	-0.530 ± 0.171	
dHb β 2	-0.561 ± 0.178	
oHb α 1	-0.085 ± 0.192	-0.150
oHb α 2	-0.125 ± 0.188	
oHb β 1	-0.195 ± 0.198	
oHb β 2	-0.194 ± 0.194	
dMb	-0.491 ± 0.199	-
oMb	-0.148 ± 0.194	-

3.5. Validation for an estimation of doming distortion with the ONIOM method

In the ONIOM model used in Section 3.4, the high layer consists only of heme and its ligands, and all atoms in the low layer are fixed during optimization, allowing an efficient estimation of the doming distortion of the heme molecule. Figure 6 shows the correlation between the doming distortions observed in the X-ray structures of Hb and Mb and the distortions estimated by the ONIOM model (purple line in Figure 6, $R^2 = 0.80$). In addition, we compared the results of the ONIOM model with those from the calculations of larger QM model systems. The latter model system consists of heme and adjacent amino acid residues for which the heavy atoms are within 4.5 Å from the Fe-porphyrin skeleton of the heme molecule. Only the heme and hydrogen atoms were optimized. The results are shown in Figure 6 as points and lines with salmon color (we examined only three systems because this process is time-consuming). The doming values calculated using the former ONIOM method corresponded well with those obtained from the latter large QM model (regression coefficient = 0.9, $R^2 = 1.0$). From these results, we determined that the estimation of doming distortion using the ONIOM model was appropriate.

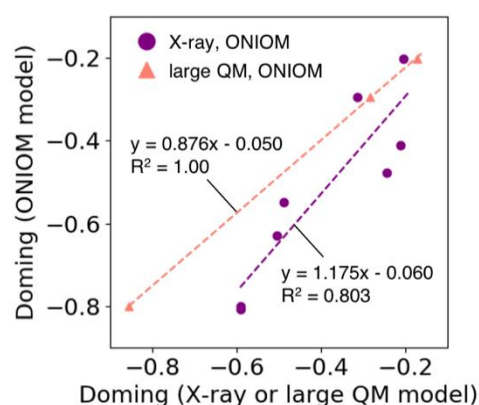


Figure 6. Plot of the doming distortion calculated based on the ONIOM model versus that based on the X-ray structure (filled circle colored in purple) and the large QM model (filled triangle colored in salmon). The horizontal axis represents the doming values calculated from X-ray structures or computed based on the large QM model. The vertical axis shows the doming values calculated based on the ONIOM model. The dotted lines represent the regression line for each group (line colors correspond to the colors used in the scatter plot).

4. Conclusions

In this study, we investigated fluctuations in the conformation of the heme-binding pocket and heme distortion for Hb and Mb to elucidate the effect of the protein environment on the distortion of heme porphyrin. First, to estimate the effect of ligand(s) on heme distortion, we optimized the model systems of isolated heme by performing QM calculations. The calculated distortion of the optimized structures showed a tendency similar to that of the Fe-porphyrin molecule reported by Bikiel et al. [8]. We compared the heme distortion between isolated heme and hemes in Hb and Mb from PDB data. This comparison indicated that the protein environment distorts heme in the doming mode more than the isolated ligand(s), especially for proteins in the deoxygenated state. The difference in doming distortion between the deoxygenated and oxygenated states in Hb was larger than that in Mb. We focused on doming distortion because it correlates with the oxygen affinity of heme.

Next, we computed the fluctuation of the heme-binding pocket and heme distortion in its host protein for Hb and Mb using MD simulations and QM calculations to clarify the relationship between the protein environment and heme distortion. The trajectory analysis of the simulations suggested that the protein environment is more flexible for Mb than for Hb. Clustering was then performed for the structures of each protein subunit sampled based on an MD simulation, and the doming distortion was estimated for the representative conformation of each cluster using the ONIOM model. The results showed the same tendency as the survey of the PDB data; specifically, heme tended to be distorted along the doming mode more largely in the dHb than in the dMb, and the difference in the distortion between the deoxygenated and oxygenated states was larger in Hb than in Mb and was different even between the Hb α chain and β chain. These differences might be related to differences in function between Hb and Mb. A previous study by Bringas et al.[19] showed a detailed difference in the mechanism underlying the affinity change between Hb α and β chains. It is reasonable that a difference in the doming distortion of heme was observed between Hb α and β chains.

We also calculated the doming distortion of heme from the MD trajectories and showed that the classical MD simulation can reproduce the distribution of heme distortion obtained by QM calculations. Finally, we validated the ONIOM model used in this study by comparing the doming values obtained from the X-ray structure and those calculated using our ONIOM model. A high correlation was observed between them. Although this method should be examined for other proteins, it could be possible to refine the heme structures obtained from X-ray crystallographic studies using our method.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Table S1: Mean values and standard deviation of the pocket volume in each subunit in each trajectory.

Author Contributions: Conceptualization, H.X.K. and Y.T.; methodology, H.X.K. and Y.T.; software, H.X.K.; investigation, H.X.K.; resources, H.X.K.; data curation, H.X.K.; writing—original draft preparation, H.X.K.; writing—review and editing, Y.T.; visualization, H.X.K.; supervision, Y.T.; project administration, H.X.K.; funding acquisition, H.X.K. and Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a Grant for Basic Science Research Projects from the Sumitomo Foundation. We are grateful to the Ministry of Education, Culture, Sports, Science and Technology (MEXT) for a Grant-in-Aid for Scientific Research on Transformative Research Areas (A) "Hyper-Ordered Structures Science," grant number No. 20H05883, and to the Japan Society for the Promotion of Science (JSPS) for Grants-in-Aid for Scientific Research (C), grant number JP19K06589, and Grants-in-Aid for Scientific Research (B), grant number JP19H02752.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The computations were performed at the Research Center for Computational Science, Okazaki, Japan and at the RIKEN Advanced Center for Computing and Communication (ACCC). This work was performed in part under the Collaborative Research Program of Institute for Protein Research, Osaka University, CR-15-05, CR-16-05, CR-17-05, CR-18-05, CR-19-05, and CR-20-05.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Perutz, M. F.; Rossmann, M. G.; Cullis, A. F.; Muirhead, H.; Will, G.; North, A. C. T. Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis. *Nature* **1960**, *185* (4711), 416–422. <https://doi.org/10.1038/185416a0>.
- Molecular Theory of Solvation*; Hirata, F., Ed.; Understanding Chemical Reactivity; Kluwer Academic Publishers: Dordrecht, 2004; Vol. 24. <https://doi.org/10.1007/1-4020-2590-4>.
- Poulos, T. L. The Janus Nature of Heme. *Nat. Prod. Rep.* **2007**, *24* (3), 504–510. <https://doi.org/10.1039/B604195G>.
- Shaik, S.; Kumar, D.; de Visser, S. P.; Altun, A.; Thiel, W. Theoretical Perspective on the Structure and Mechanism of Cytochrome P450 Enzymes. *Chem. Rev.* **2005**, *105* (6), 2279–2328. <https://doi.org/10.1021/cr030722j>.
- Faller, M.; Matsunaga, M.; Yin, S.; Loo, J. A.; Guo, F. Heme Is Involved in MicroRNA Processing. *Nat. Struct. Mol. Biol.* **2007**. <https://doi.org/10.1038/nsmb1182>.
- Sun, J.; Hoshino, H.; Takaku, K.; Nakajima, O.; Muto, A.; Suzuki, H.; Tashiro, S.; Takahashi, S.; Shibahara, S.; Alam, J.; Taketo, M. M.; Yamamoto, M.; Igarashi, K. Hemoprotein Bach1 Regulates Enhancer Availability of Heme Oxygenase-1 Gene. *EMBO J.* **2002**, *21* (19), 5216–5224. <https://doi.org/10.1093/emboj/cdf516>.
- Chapman, S. K.; Daff, S.; Munro, A. W. Heme: The Most Versatile Redox Centre in Biology?; 1997; pp 39–70. https://doi.org/10.1007/3-540-62870-3_2.
- Bikiel, D. E.; Forti, F.; Boechi, L.; Nardini, M.; Luque, F. J.; Martí, M. A.; Estrin, D. A. Role of Heme Distortion on Oxygen Affinity in Heme Proteins: The Protoglobin Case. *J. Phys. Chem. B* **2010**, *114* (25), 8536–8543. <https://doi.org/10.1021/jp102135p>.
- Imada, Y.; Nakamura, H.; Takano, Y. Density Functional Study of Porphyrin Distortion Effects on Redox Potential of Heme. *J. Comput. Chem.* **2018**, *39* (3), 143–150. <https://doi.org/10.1002/jcc.25058>.
- Kanematsu, Y.; Kondo, H. X.; Imada, Y.; Takano, Y. Statistical and Quantum-Chemical Analysis of the Effect of Heme Porphyrin Distortion in Heme Proteins: Differences between Oxidoreductases and Oxygen Carrier Proteins. *Chem. Phys. Lett.* **2018**, *710*, 108–112. <https://doi.org/10.1016/j.cplett.2018.08.071>.
- Sun, Y.; Benabbas, A.; Zeng, W.; Kleingardner, J. G.; Bren, K. L.; Champion, P. M. Investigations of Heme Distortion, Low-Frequency Vibrational Excitations, and Electron Transfer in Cytochrome C. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111* (18), 6570–6575. <https://doi.org/10.1073/pnas.1322274111>.
- Kondo, H. X.; Fujii, M.; Tanioka, T.; Kanematsu, Y.; Yoshida, T.; Takano, Y. Global Analysis of Heme Proteins Elucidates the Correlation between Heme Distortion and Heme-Binding Pocket. *Authorea* **2021**. <https://doi.org/10.22541/au.162860462.26233360/v1>.
- Berman, H. M. The Protein Data Bank / Biopython. *Presentation* **2000**, *28* (1), 235–242. <https://doi.org/10.1093/nar/28.1.235>.
- Li, T.; Bonkovsky, H. L.; Guo, J. Structural Analysis of Heme Proteins: Implications for Design and Prediction. *BMC Struct. Biol.* **2011**, *11* (1), 13. <https://doi.org/10.1186/1472-6807-11-13>.
- Sacquin-Mora, S.; Lavery, R. Investigating the Local Flexibility of Functional Residues in Hemoproteins. *Biophys. J.* **2006**, *90* (8), 2706–2717. <https://doi.org/10.1529/biophysj.105.074997>.
- Kendrew, J. C.; Dickerson, R. E.; Strandberg, B. E.; Hart, R. G.; Davies, D. R.; Phillips, D. C.; Shore, V. C. Structure of Myoglobin: A Three-Dimensional Fourier Synthesis at 2 Å. Resolution. *Nature* **1960**, *185* (4711), 422–427. <https://doi.org/10.1038/185422a0>.
- Aharoni, R.; Tobi, D. Dynamical Comparison between Myoglobin and Hemoglobin. *Proteins* **2018**, *86* (11), 1176–1183. <https://doi.org/10.1002/PROT.25598>.
- Tobi, D. Dynamics Based Clustering of Globin Family Members. *PLoS One* **2018**, *13* (12), e0208465. <https://doi.org/10.1371/JOURNAL.PONE.0208465>.
- Bringas, M.; Petruk, A. A.; Estrin, D. A.; Capece, L.; Martí, M. A. Tertiary and Quaternary Structural Basis of Oxygen Affinity in Human Hemoglobin as Revealed by Multiscale Simulations. *Sci. Reports* **2017**, *7* (1), 1–10. <https://doi.org/10.1038/s41598-017-11259-0>.
- Kinjo, A. R.; Suzuki, H.; Yamashita, R.; Ikegawa, Y.; Kudou, T.; Igarashi, R.; Kengaku, Y.; Cho, H.; Standley, D. M.; Nakagawa, A.; Nakamura, H. Protein Data Bank Japan (PDBj): Maintaining a Structural Data Archive and Resource Description Framework Format. *Nucleic Acids Res.* **2012**, *40* (D1). <https://doi.org/10.1093/nar/gkr811>.

21. Kinjo, A. R.; Yamashita, R.; Nakamura, H. PDBj Mine: Design and Implementation of Relational Database Interface for Protein Data Bank Japan. *Database* **2010**, 2010, baq021–baq021. <https://doi.org/10.1093/database/baq021>.
22. Cock, P. J. A.; Antao, T.; Chang, J. T.; Chapman, B. A.; Cox, C. J.; Dalke, A.; Friedberg, I.; Hamelryck, T.; Kauff, F.; Wilczynski, B.; De Hoon, M. J. L. Biopython: Freely Available Python Tools for Computational Molecular Biology and Bioinformatics. *Bioinformatics* **2009**, 25 (11), 1422–1423. <https://doi.org/10.1093/bioinformatics/btp163>.
23. McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L. P.; Lane, T. J.; Pande, V. S. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **2015**, 109 (8), 1528–1532. <https://doi.org/10.1016/j.bpj.2015.08.015>.
24. Li, W.; Godzik, A. Cd-Hit: A Fast Program for Clustering and Comparing Large Sets of Protein or Nucleotide Sequences. *Bioinformatics* **2006**, 22 (13), 1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>.
25. Jentzen, W.; Song, X. Z.; Shelnutt, J. A. Structural Characterization of Synthetic and Protein-Bound Porphyrins in Terms of the Lowest-Frequency Normal Coordinates of the Macrocycle. *J. Phys. Chem. B* **1997**, 101 (9), 1684–1699. <https://doi.org/10.1021/jp963142h>.
26. Kondo, H. X.; Kanematsu, Y.; Masumoto, G.; Takano, Y. PyDISH: Database and Analysis Tools for Heme Porphyrin Distortion in Heme Proteins. *Database* **2020**, 2020, baaa066. <https://doi.org/10.1093/database/baaa066>.
27. Tame, J. R. H.; Vallone, B. The Structures of Deoxy Human Haemoglobin and the Mutant Hb Tyr α 42His at 120 K. *urn:issn:0907-4449* **2000**, 56 (7), 805–811. <https://doi.org/10.1107/S0907444900006387>.
28. Park, S. Y.; Yokoyama, T.; Shibayama, N.; Shiro, Y.; Tame, J. R. H. 1.25 Å Resolution Crystal Structures of Human Haemoglobin in the Oxy, Deoxy and Carbonmonoxy Forms. *J. Mol. Biol.* **2006**, 360 (3), 690–701. <https://doi.org/10.1016/J.JMB.2006.05.036>.
29. Phillips, S. E. V. Structure and Refinement of Oxymyoglobin at 1.6 Å Resolution. *J. Mol. Biol.* **1980**, 142 (4), 531–554. [https://doi.org/10.1016/0022-2836\(80\)90262-4](https://doi.org/10.1016/0022-2836(80)90262-4).
30. Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: A Server for Estimating PKas and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* **2005**, 33 (Web Server issue), W368. <https://doi.org/10.1093/NAR/GKI464>.
31. Shih, T.; Jones, R. T.; Bonaventura, J.; Bonaventura, C.; Schneider, R. G. Involvement of His HC3 (146) Beta in the Bohr Effect of Human Hemoglobin. Studies of Native and N-Ethylmaleimide-Treated Hemoglobin A and Hemoglobin Cowtown (Beta 146 His Replaced by Leu). *J. Biol. Chem.* **1984**, 259 (2), 967–974. [https://doi.org/10.1016/S0021-9258\(17\)43552-6](https://doi.org/10.1016/S0021-9258(17)43552-6).
32. 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 (5), e1000774. <https://doi.org/10.1371/JOURNAL.PCBI.1000774>.
33. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, 1–2, 19–25. <https://doi.org/10.1016/J.SOFTX.2015.06.001>.
34. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, 11 (8), 3696–3713. https://doi.org/10.1021/ACS.JCTC.5B00255/SUPPL_FILE/CT5B00255_SI_001.PDF.
35. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, 79 (2), 926–935. <https://doi.org/10.1063/1.445869>.
36. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, 103 (19), 8577–8593. <https://doi.org/10.1063/1.470117>.
37. Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J. Chem. Phys.* **2007**, 126 (1), 014101. <https://doi.org/10.1063/1.2408420>.
38. Nosé, S. A Unified Formulation of the Constant Temperature Molecular Dynamics Methods. *J. Chem. Phys.* **1984**, 81 (1), 511–519. <https://doi.org/10.1063/1.447334>.
39. Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, 31 (3), 1695–1697. <https://doi.org/10.1103/PhysRevA.31.1695>.
40. Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nosé-Hoover Chains: The Canonical Ensemble via Continuous Dynamics. *J. Chem. Phys.* **1992**, 97 (4), 2635–2643. <https://doi.org/10.1063/1.463940>.
41. Parrinello, M.; Rahman, A. Crystal Structure and Pair Potentials: A Molecular-Dynamics Study. *Phys. Rev. Lett.* **1980**, 45 (14), 1196–1199. <https://doi.org/10.1103/PhysRevLett.45.1196>.
42. Parrinello, M.; Rahman, A. Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method. *J. Appl. Phys.* **1981**, 52 (12), 7182–7190. <https://doi.org/10.1063/1.328693>.
43. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, 18 (12), 1463–1472. [https://doi.org/10.1002/\(SICI\)1096-987X\(199709\)18:12<1463::AID-JCC4>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H).
44. Wagner, J. R.; Sørensen, J.; Hensley, N.; Wong, C.; Zhu, C.; Perison, T.; Amaro, R. E. POVME 3.0: Software for Mapping Binding Pocket Flexibility. *J. Chem. Theory Comput.* **2017**, 13 (9), 4584–4592. <https://doi.org/10.1021/acs.jctc.7b00500>.
45. Adamo, C.; Barone, V. Toward Reliable Density Functional Methods without Adjustable Parameters: The PBE0 Model. *J. Chem. Phys.* **1999**, 110 (13), 6158–6170. <https://doi.org/10.1063/1.478522>.
46. Ditchfield, R.; Hehre, W. J.; Pople, J. A. Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1971**, 54 (2), 724–728. <https://doi.org/10.1063/1.1674902>.
47. Hariharan, P. C.; Pople, J. A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theor.*

- Chim. Acta* **1973**, 28 (3), 213–222. <https://doi.org/10.1007/BF00533485>.
48. Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. 6-31G* Basis Set for Atoms K through Zn. *J. Chem. Phys.* **1998**, 109 (4), 1223–1229. <https://doi.org/10.1063/1.476673>.
 49. Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic Interaction of a Solute with a Continuum. A Direct Utilization of AB Initio Molecular Potentials for the Prediction of Solvent Effects. *Chem. Phys.* **1981**, 55 (1), 117–129. [https://doi.org/10.1016/0301-0104\(81\)85090-2](https://doi.org/10.1016/0301-0104(81)85090-2).
 50. Miertuš, S.; Tomasi, J. Approximate Evaluations of the Electrostatic Free Energy and Internal Energy Changes in Solution Processes. *Chem. Phys.* **1982**, 65 (2), 239–245. [https://doi.org/10.1016/0301-0104\(82\)85072-6](https://doi.org/10.1016/0301-0104(82)85072-6).
 51. Pascual-Ahuir, J. L.; Silla, E.; Tuñón, I. GEPOL: An Improved Description of Molecular Surfaces. III. A New Algorithm for the Computation of a Solvent-excluding Surface. *J. Comput. Chem.* **1994**, 15 (10), 1127–1138. <https://doi.org/10.1002/JCC.540151009>.
 52. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16, Revision C.01*; Gaussian, Inc., Wallingford CT, 2016.