

# Revised CHARMM Force Field Parameters for Iron-Containing Cofactors of Photosystem II

Suliman Adam <sup>[a]</sup>, Michaela Knapp-Mohammady<sup>[b]</sup>, Jun Yi<sup>[c]</sup> and Ana-Nicoleta Bondar<sup>[a]</sup>

Photosystem II is a complex protein–cofactor machinery that splits water molecules into molecular oxygen, protons, and electrons. All-atom molecular dynamics simulations have the potential to contribute to our general understanding of how photosystem II works. To perform reliable all-atom simulations, we need accurate force field parameters for the cofactor molecules. We present here CHARMM bonded and non-bonded parameters for the iron-containing cofactors of photosystem II

that include a six-coordinated heme moiety coordinated by two histidine groups, and a non-heme iron complex coordinated by bicarbonate and four histidines. The force field parameters presented here give water interaction energies and geometries in good agreement with the quantum mechanical target data. © 2017 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.24918

## Introduction

Reliable description of the conformational dynamics of photosystem II is an essential step toward understanding how this membrane-embedded complex of proteins and cofactor molecules generates molecular oxygen by splitting water. Recent high-resolution crystal structures of photosystem II from *Thermosynechococcus vulcanus*<sup>[1,2]</sup> provide valuable information about intramolecular interactions potentially important for function, such as the presence of protein–water hydrogen-bonded networks that could be involved in proton transfer<sup>[1,3,4]</sup>; however, it remains unclear how motions and electrostatic interactions of the protein environment influence protonation of specific protein groups, or how protein groups and water molecules respond to changes in protonation. All-atom molecular dynamics simulations are valuable means to address questions pertaining to protonation-coupled protein dynamics.<sup>[5,6]</sup> In the case of photosystem II, a specific challenge is the requirement of reliable force field parameters to describe the geometry and non-bonded interactions of cofactor molecules. Here, we derived new Chemistry at Harvard Molecular Mechanics (CHARMM)<sup>[7]</sup> all-atom force field parameters for two iron-containing cofactors in photosystem II that include (i) a six-coordinated heme moiety with two histidines bound to iron, and (ii) a non-heme iron complex liganded with a bicarbonate ion and four histidine groups.

The crystal structure of the photosystem II dimer from *T. vulcanus* contains two heme b molecules bound to each monomer.<sup>[1]</sup> Each of the heme b molecules is coordinated by two histidine residues (Fig. 1a), and has additional interactions that depend on the immediate protein environment. In case of the heme b molecule located at the stroma side of the transmembrane region of the protein, the two heme b propionate groups interact with arginine side chains from the cytochrome b559 subunits  $\alpha$  (PsbE) and  $\beta$  (PsbF), respectively<sup>[8]</sup>; the coordinating histidine side chains (PsbE-H23 and PsbF-H24) are close to tyrosine or tryptophan side chains that provide stabilizing  $\pi$ -stacking interactions (PsbE-Y19 and PsbF-W20, respectively, Supporting Information Figure S1a). Experiments on mutant

photosystem II from *T. elongatus* indicate that phototropic growth is possible without heme at the stroma side<sup>[9]</sup>; nevertheless, mutation to Phe of the PsbE-Y19 group that is within one helical turn of PsbE-H23 (Supporting Information Figure S1a) influences the extent to which activity recovers after photoinhibition.<sup>[9]</sup> This observation highlights the importance of understanding the conformational dynamics close to the heme b moiety and the coordinating histidine groups.

The second heme molecule is located at the lumen side of the photosystem II monomer, where it binds to the extrinsic cytochrome c550 subunit PsbV (Fig. 1a). There, the histidine side chains contributed by cytochrome c550 appear to lack potentially stabilizing  $\pi$ -stacking interactions, although interactions with nearby protein groups could limit their rotational flexibility. The propionate groups of the heme b molecule are exposed to water (Fig. 1a). Differences in the protein environment of the heme iron complexes at the stroma (Supporting Information Figure S1a) versus the lumen side (Fig. 1a) are

[a] S. Adam, A.-N. Bondar

Theoretical Molecular Biophysics, Department of Physics, Freie Universität Berlin, Arnimallee 14, Berlin D-14195, Germany  
E-mail: nbondar@zedat.fu-berlin.de

[b] M. Knapp-Mohammady

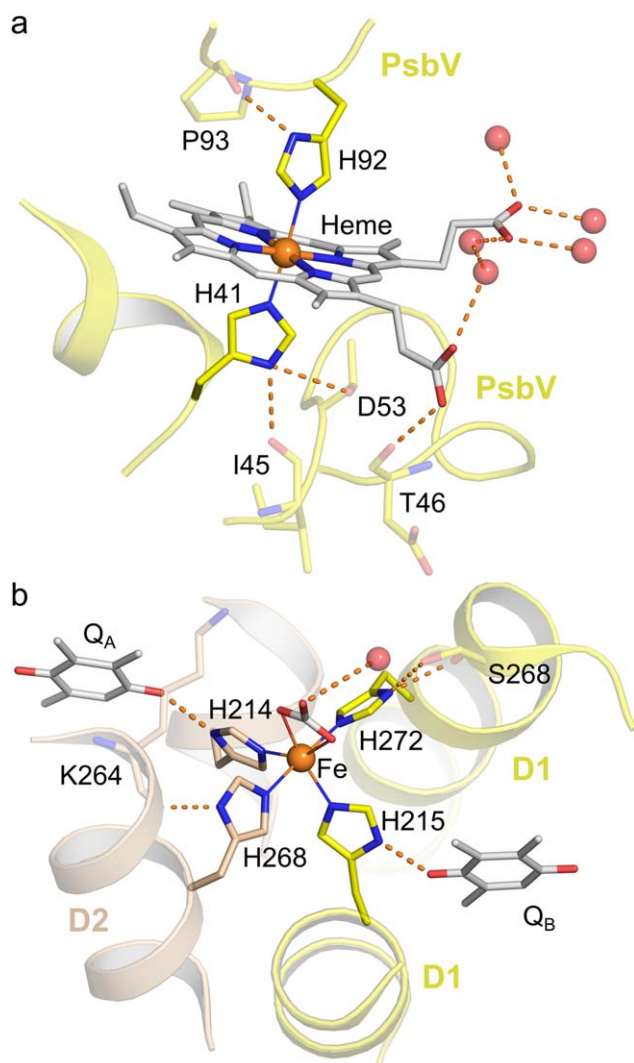
Division of Functional Genome Analysis, German Cancer Research Center (DKFZ) Heidelberg, Im Neuenheimer Feld 580, Heidelberg D-69120, Germany

[c] J. Yi

Department of Biological Engineering, School of Environmental and Biological Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China

Contract grant sponsor: German Research Foundation (DFG) Collaborative Research Center SFB 1078, Protein Dynamics in Protein Function (Project C4, to A.-N.B.); Contract grant sponsor: The North-German Supercomputing Alliance, HLRN; Contract grant number: bec00063 (to A.-N.B.); Contract grant sponsor: National Natural Science Foundation of China (to J.Y.); Contract grant numbers: 21271104 and 21671103; Contract grant sponsor: Excellence Initiative of the German Federal and State Governments provided via the Freie Universität Berlin (to A.-N.B.)

© 2017 Wiley Periodicals, Inc.



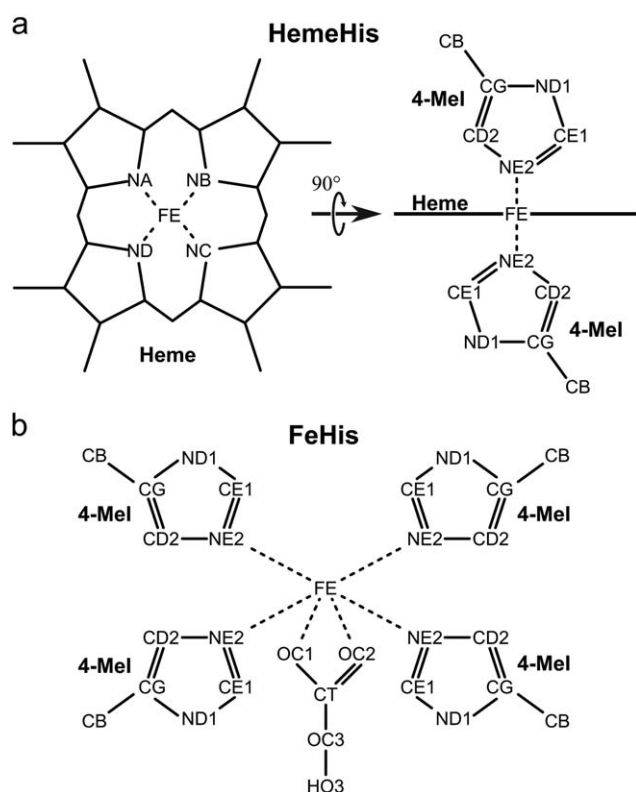
**Figure 1.** Heme and non-heme iron inside photosystem II. The thin solid lines indicate the coordination bonds involving heme and non-heme iron (orange spheres). Hydrogen bonds are shown as dashed lines. Water oxygens are shown as red spheres. a) The heme b (gray) at the lumen side of photosystem II forms coordination bonds with NE2 of H41 and H92 of subunit PsbV of photosystem II (both shown in yellow). b) The non-heme iron is bound to bicarbonate (gray), H214 and H268 of subunit D2 (wheat) and H215 and H272 of subunit D1 (yellow). H214 and H215 connect the non-heme iron–bicarbonate complex to the quinones  $Q_A$  and  $Q_B$  (gray) via a hydrogen-bonding network. The molecular graphics are based on monomer A, chain V (for HemeHis, panel a), and chains a and d (for FeHis, panel b) from the crystal structure (PDB ID: 3WU2<sup>[1]</sup>). Additional molecular graphics of the heme and non-heme iron complexes in photosystem II are given in Supporting Information Figure S1. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

likely to impact the dynamics of the heme b molecule and of the coordinating histidine side chains.

The non-heme iron is coordinated by four NE2 atoms from histidine side chains contributed by subunits D1 and D2, and by the carboxylate oxygen atoms of the bicarbonate molecule<sup>[1,10]</sup> (Figs. 1b and 2b). The hydroxyl group of the bicarbonate molecule hydrogen-bonds to a water molecule (Fig. 1b) and to two tyrosine side chains.<sup>[10]</sup> Two of the histidine side chains, H268 and H272, hydrogen-bond to nearby protein groups, whereas H214 and H215 hydrogen-bond to the quinone groups  $Q_A$  and  $Q_B$ , respectively (Fig. 1b). This network of

hydrogen bonds could be important for proton transfer at  $Q_B$ .<sup>[10]</sup> In addition to a potential role in proton transfer, we suggest that the complex composed of non-heme iron, bicarbonate, and coordinating histidine groups might be important for the local structure and dynamics of photosystem II. Indeed, this complex connects four transmembrane helices (Supporting Information Figure S1b), possibly coupling to the largely different orientations of these helices relative to the membrane normal.

Accurate description of the geometry and conformational dynamics of cofactor molecules is a key requirement for reliable molecular dynamics simulations of photosystem II. CHARMM<sup>[7,11,12]</sup> includes force field parameters for six-coordinated heme, allowing for a number of ligands, including a histidine side chain; issues pertaining to CHARMM force field parameters for heme and heme with various ligands have been addressed by others.<sup>[13–15]</sup> Recently, all-atom force field parameters of the heme iron complex of photosystem II were derived for the AMBER force field (Assisted Model Building with Energy Refinement).<sup>[16]</sup> This parametrization included Hartree–Fock (HF) computations in vacuum, and density functional theory (DFT) calculations with a continuum solvent representation.<sup>[16]</sup> Other parametrizations directly relevant to the heme molecule of photosystem II include the AMBER and CHARMM force field parameters for heme coordinated by



**Figure 2.** Chemical structures of the two complexes parametrized. a) Skeletal formula of the heme model compound of the HemeHis complex illustrating the central iron and the surrounding nitrogen of the heme plane with a side view of the full HemeHis complex. The NE2 atoms of the 4-methylimidazole (4-Mel) groups coordinate the heme iron (FE). b) The FeHis complex has iron coordinating 4-Mel and bicarbonate. The 4-Mel groups are  $\delta$ -protonated in both systems.

histidine and methionine side chains,<sup>[17]</sup> or the charge parametrization scheme for heme a with two coordinating histidine groups and for a five-coordinated heme molecule.<sup>[18]</sup>

Our aim is to describe the heme and non-heme iron complexes of photosystem II with CHARMM force field parameters that are consistent with the recommended protocol of the CHARMM General Force Field (CGenFF).<sup>[19]</sup> In this protocol, parameters for specific bonded and non-bonded interactions from the potential energy function of CHARMM are derived iteratively until certain convergence criteria are met. An important aspect of the CGenFF is that, to ensure internal consistency of the parameters, the atomic partial charges of the cofactor molecule are adjusted by fitting CHARMM water interaction energies against the corresponding values derived with HF/6-31G\*.<sup>[19]</sup>

For the heme b molecule coordinated by two histidine side chains (Fig. 1a), we relied heavily on the bonded and non-bonded parameters that have already existed in CHARMM. Our focus here was to derive bonded parameters involving the coordination bonds specific to heme b coordinated by two histidine moieties. In the case of the complex consisting of non-heme iron, bicarbonate, and four histidine side chains (Fig. 1b), we started from CHARMM parameters for the isolated bicarbonate and histidine moieties to derive bonded and non-bonded parameters specific to the complex.

## Methods

### CHARMM potential energy function

The potential energy function of CHARMM<sup>[11]</sup> contains terms for intramolecular and intermolecular interactions as follows:

$$\begin{aligned} U(\vec{R}) = & \sum_{\text{bonds}} K_b(b-b_0)^2 + \sum_{\text{angles}} K_\theta(\theta-\theta_0)^2 + \sum_{\text{dihedrals}} K_\phi(1+\cos(n\phi-\delta)) \\ & + \sum_{\text{impropers}} K_\omega(\omega-\omega_0)^2 \\ & + \sum_{\text{Urey-Bradley}} K_{\text{UB}}(S-S_0)^2 + \sum_{\text{non-bonded}} \left\{ \epsilon_{ij} \left[ \left( \frac{R_{ij}^{\text{min}}}{r_{ij}} \right)^{12} - 2 \left( \frac{R_{ij}^{\text{min}}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{\epsilon_1 r_{ij}} \right\} \end{aligned} \quad (1)$$

CHARMM describes bond stretching and valence angle bending by harmonic terms that include as parameters equilibrium values ( $b_0$ ,  $\theta_0$ ) and force constants ( $K_b$ ,  $\theta_b$ ). A harmonic potential with force constant  $K_\omega$  is also used to describe the improper dihedral angles  $\omega$ . The proper dihedral angle term is a sum of cosine functions of multiplicity  $n$  and phase shift  $\delta$ , the torsional energy barrier being described by the force constant  $K_\phi$ . The Urey-Bradley term can be used to treat the 1,3-non-bonded interactions of valence angles; it adds a harmonic potential with a force constant  $K_{\text{UB}}$  and equilibrium distance  $S_0$ . The non-bonded parameters comprise the partial charges  $q_i$  that give the Coulomb interactions between atoms separated by the distance  $r_{ij}$ , and parameters for the van der Waals interactions: The depth of the Lennard-Jones well  $\epsilon_{ij}$ , the minimum interaction radius  $R_{ij}^{\text{min}}$ , and the effective dielectric constant  $\epsilon_1$ . The CHARMM36 protein force field additionally includes an energy correction map (CMAP) term to improve the description of the backbone dihedrals in peptides.<sup>[20]</sup>

### Model systems and starting coordinates

For the starting coordinates of the complexes parametrized here, we used monomer A of the crystal structure of photosystem II (PDB ID: 3WU2<sup>[1]</sup>). To facilitate the quantum mechanical (QM) calculations and the parametrization procedure, we used two simplified model complexes as described below.

For the heme b molecule coordinated by two histidine side chains (Fig. 1a), we replaced the propionate side chains and the methyl groups of heme b with hydrogen atoms, and the two histidine residues were modeled as 4-methylimidazole (4-Mel, Fig. 2a). This complex composed of the heme b core and two 4-Mel groups is denoted here as HemeHis (Fig. 2a). Likewise, in the case of the complex consisting of the non-heme iron, four histidines and the bicarbonate ion (Fig. 1b), we described each histidine residue as 4-Mel, and denoted this complex as FeHis (Fig. 2b).

Coordinates for the hydrogen atoms were constructed with CHARMM.<sup>[7]</sup> The 4-Mel groups are single-protonated on the ND1 atoms. For completeness of the parametrization and to facilitate CHARMM computations of heme and non-heme iron compounds, we derived parameters for HemeHis and FeHis in both ferrous ( $\text{Fe}^{2+}$ ) and ferric ( $\text{Fe}^{3+}$ ) states. Ferrous and ferric HemeHis have a total charge of 0 and +1, respectively; whereas ferrous and ferric FeHis have a total charge of +1 and +2, respectively.

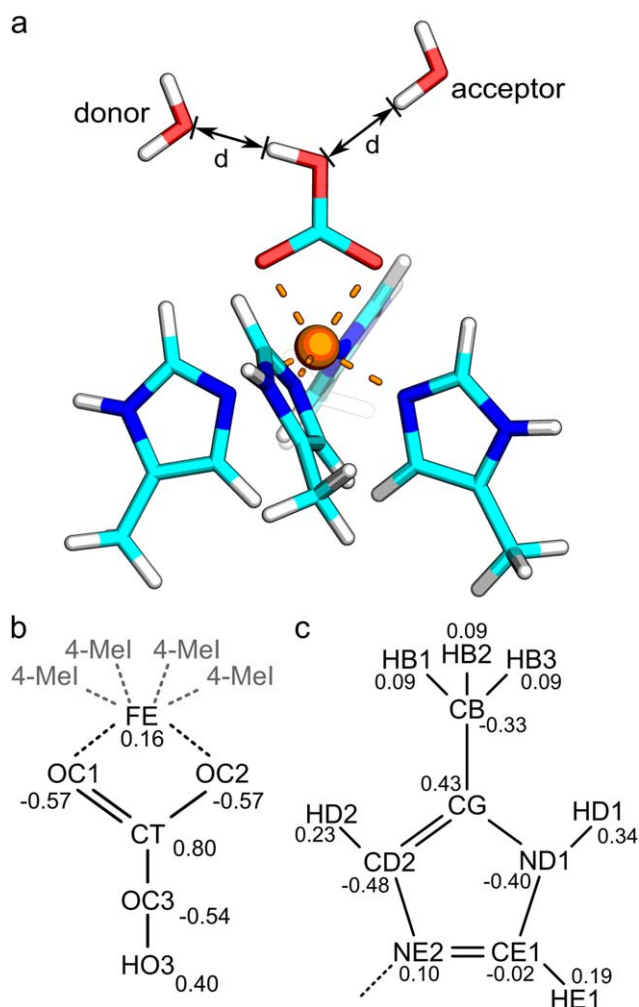
The total charge of the system is relevant for the protocol used to derive atomic partial charges. In case of molecules that have zero net charge, deriving the atomic partial charges based on water interaction energies involves scaling of the QM interaction energies by a factor of 1.16. Water interaction distances for the polar sites are offset by  $-0.2 \text{ \AA}$ .<sup>[19]</sup> Such a protocol was used recently for parametrizing neutral cofactor molecules of photosystem II.<sup>[21]</sup> Scaling of the interaction energies is not performed for compounds with non-zero net charge, such as FeHis and the ferric HemeHis complexes investigated here.

### QM geometry optimizations

QM-optimized geometries of the isolated model complexes serve as geometrical target data to validate the parameters. For the FeHis complex, we used second-order Møller-Plesset perturbation theory (MP2)<sup>[22]</sup> with a 6-31G\*<sup>[23-25]</sup> basis set. Because of computational cost, the ferrous and ferric HemeHis and the ferric FeHis complexes could not be optimized with MP2; consequently, we used the DFT hybrid functional Becke3, Lee-Yang-Parr (B3LYP)<sup>[26,27]</sup> for HemeHis, and unrestricted B3LYP (UB3LYP) for the ferric complexes. All QM calculations were performed with the Gaussian 09 software package<sup>[28]</sup> using default convergence criteria.

### Optimization of the partial atomic charges

The partial atomic charges are optimized such that the molecular mechanical (MM) water interaction energies  $E_{\text{MM}}$  and interaction distances  $d_{\text{MM}}$  reproduce the corresponding  $E_{\text{QM}}$  and  $d_{\text{MM}}$  values computed with HF/6-31G\* for all hydrogen



**Figure 3.** Optimization of atomic partial charges for FeHis. a) Illustration of two of the water interaction sites sampled during the parametrization of FeHis. A water molecule is constrained to a planar interaction geometry and optimized to obtain the interaction distance  $d$  and the interaction energy  $E$  for hydrogen bond donor and acceptor sites. Each water interaction site is sampled separately. b) Optimized partial charges of the bicarbonate and iron in FeHis. c) Optimized partial charges for 4-Mel. Because of symmetry, each of the four 4-Mel groups has the same set of atomic partial charges. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

bond donor and acceptor sites.<sup>[19]</sup> Two examples for possible water interactions sites of FeHis are given in Figure 3a.

Water interaction energies  $E$  are computed<sup>[19]</sup> as the difference between the energy of the FeHis–water complex and the sum of energies computed, at the same level of theory (MM or QM), for the isolated FeHis complex and a water molecule,

$$E = E(\text{FeHis} + \text{water}) - E(\text{FeHis}) - E(\text{water}) \quad (2)$$

For acceptor interactions, the interaction distance  $d$  is the distance, measured for a geometry-optimized structure, between the water hydrogen atom and the FeHis heavy atom for which the interaction energy is computed (Fig. 3a). For donor sites, it is the distance between water oxygen and interacting FeHis hydrogen atom.

Following the CGenFF protocol,<sup>[19]</sup> water interaction energies for the possible water interaction sites were first

computed with HF/6–31G\* using the QM-optimized geometry of the complex and a TIP3P water molecule.<sup>[29]</sup> We utilized the Force Field Toolkit<sup>[30]</sup> to generate input files for water interactions. The complex was generated by placing a TIP3P water molecule in a linear interaction geometry with the hydrogen (donor) or heavy (acceptor) atom of the QM-optimized structure of the FeHis complex (see illustration in Fig. 3a). In the next step, a HF/6–31G\* optimization was performed for the distance between the two interacting atoms, and for the orientation of the water molecule relative to the FeHis complex. This optimization gives the HF/6–31G\* equilibrium distance  $d_{\text{QM}}$  and the interaction energy  $E_{\text{QM}}$  of the water molecule at that specific interaction site. This process was performed, with one water molecule at a time, for all accessible interaction sites of both the bicarbonate and the four histidine moieties.

The MM values for  $d_{\text{MM}}$  and  $E_{\text{MM}}$  were computed using the MM atomic partial charges of FeHis. The initial partial charges were the Merz–Singh–Kollman (MK) charges<sup>[31]</sup> calculated with HF/6–31G\* for the geometry optimized with MP2/6–31G\* for ferrous FeHis, B3LYP/6–31G\* for ferrous HemeHis, and UB3LYP/6–31G\* for the ferric systems.

For the MM calculations of water interaction energies at each interaction site considered, we again placed the respective TIP3P water molecule in an ideal interaction geometry, but used the MM-optimized FeHis geometry. To determine the MM interaction position  $d_{\text{MM}}$ , we constrained the orientation of the water molecule relative to FeHis to that from the MP2-optimized geometry, and scanned the distance between water and the FeHis interaction site in steps of 0.01 Å, until the energy minimum was reached.

We manually changed the partial charges of the FeHis complex and recalculated the water interaction energies. These computations were performed separately for ferrous and ferric FeHis. With the exception of the non-polar hydrogens HB1, HB2, and HB3 of the imidazole moieties (Fig. 3c), where the partial charge was kept at the standard CHARMM value for methyl hydrogen atoms of 0.09, the partial charges of the iron, bicarbonate, and 4-Mel atoms shown in Figures 3b and 3c were optimized until  $E_{\text{MM}}$  and  $E_{\text{QM}}$  agreed to within the recommended convergence criterion of 0.2 kcal/mol.<sup>[19]</sup>

The partial charge optimization of the HemeHis system was performed in the same manner, with the exception that partial atomic charges taken from the CHARMM36 protein force field were used as initial charge guesses for 4-Mel and ferrous heme.

### Refinement of bonds and angles for FeHis and HemeHis

We compared the equilibrium positions of the bonds and valence angles of the QM- and MM-optimized structures of HemeHis and FeHis. To improve the agreement between MM- and QM-optimized geometries, we refined the bonds and angles involving iron coordination bonds. For both complexes, we optimized the iron–histidine coordination bond FE–NE2, and the valence angles FE–NE2–CD2 and FE–NE2–CE1 (see Figs. 2a and 2b for the CHARMM atom labels). For HemeHis, we additionally modified bonds between the iron and the

coordinating nitrogen atoms NA, NB, NC, and ND of the heme plane (Fig. 2a). We refined the equilibrium positions until the deviation between MM and QM values converged to a value less than 0.03 Å for bonds, and 3° for valence angles.<sup>[19]</sup>

For the bicarbonate compound of FeHis, the starting parameters were taken from the carbonate ion parameters of the CHARMM36 protein force field. Here, we derived new bonds and angles parameters for the hydroxyl group.

For all ferric complexes, we used as starting point the bonded parameters optimized for the corresponding ferrous complexes, and then further tested and refined specific bond and angle parameters that had been previously refined for the ferrous complexes. That is, for both ferrous and ferric complexes, we refined only bonded parameters involving the iron ion, and kept the remaining bonded CHARMM parameters unchanged.

### Refinement of force constants of bonds and angles of HemeHis and FeHis

Comparison of structures geometry-optimized with QM versus MM indicated that a number of bond and angle parameters of ferrous HemeHis and FeHis required further refinement. Ideally, force constants for bonds and valence angles are optimized by analyzing the vibrational modes of model compounds. For molecules that are too large for QM vibrational analysis, such as the FeHis or HemeHis complexes studied here, CGenFF recommends potential energy scans (PES) of small perturbations of bonds and angles.<sup>[19]</sup> We used such an approach to optimize the force constants of the bonds and angles involving iron in FeHis and HemeHis.

In the first step, we used the QM-optimized geometries of ferrous HemeHis and FeHis to measure the equilibrium positions for bond lengths and the valence angles subject to refinement. In the second step, we performed PES separately for ferrous HemeHis and FeHis. To refine force constants for bonds, we started with an offset of  $-0.01$  Å from the equilibrium length, and conducted a three-step PES with a step size of  $0.01$  Å. For valence angles, we used an offset of  $-1^\circ$  and a step size of  $1^\circ$ . We adjusted the MM force constants until the agreement between the QM and MM PES could no longer be improved.

The refined force constants obtained for the ferrous systems were used as starting force constants for testing specific bonded force constants of ferric HemeHis and FeHis.

### Deriving parameters for dihedral angles of HemeHis

To derive parameters for describing heme b coordinated by two imidazole groups, we performed a PES for rigid and flexible dihedral angles defined as follows: Rigid dihedrals are dihedrals describing ring torsions, that is, the central two atoms of such dihedrals form part of a ring, and the dihedral angle is associated with a relatively high energy barrier; all other dihedral angles are denoted as flexible dihedral angles.

Starting from the position of the minimized structure, we used QM to derive the PES for rigid dihedrals, once with a step size of  $5^\circ$ , and once with a step size of  $-5^\circ$  for six steps

each. For all flexible dihedrals, we used  $5^\circ$  increments to perform an 18-step scan in each direction.

The PES was then computed with MM using the same step size and number of steps as in the corresponding QM PES described above. In optimizing the MM description of the dihedral angle, we aimed to reproduce the height of the energy barriers and the location of the energy minima. Reproducing the energy profile in the region of the energy minima is particularly important for molecular dynamics simulations because these are the regions that are sampled most often.<sup>[19]</sup> We adjusted the dihedral parameters by changing their multiplicity, phase shift, and force constant. Following the CHARMM protocol,<sup>[19]</sup> we used only  $0^\circ$  and  $180^\circ$  for the phase shift.

For both ferric and ferrous complexes, the parametrization steps described above were performed iteratively until self-consistency had been reached.

### New chemical atom types, and compatibility with the CHARMM protein force field

We introduced new chemical types for iron depending on its oxidation state and the system. Using these, new atom types allowed us to avoid overwriting parameters already present in CHARMM, and retain full CHARMM36-compatibility.

The parameters presented in here were implemented as an extension to the CHARMM36 protein force field.<sup>[11]</sup> All Lennard-Jones parameters were taken from standard CHARMM.

### MM molecular dynamics to test the HemeHis parameters

To test the performance of the ferrous HemeHis parameters derived here, we used as model system human brain neuroglobin, a relatively small soluble protein of 151 amino acid residues that contains a ferrous heme cofactor coordinated by two histidine residues. We took the starting coordinates of the protein and heme atoms from chain B of the crystal structure (PDB ID: 1OJ6<sup>[32]</sup>).

Hydrogen atoms were constructed with CHARMM.<sup>[7]</sup> The protein was placed at the center of the Cartesian coordinate system, and overlapped with a box of TIP3P water molecules. Water molecules of the solvent box that were overlapping with crystal structure coordinates were deleted. The simulation setup consisted of the protein, the heme molecule, the 42 crystal structure water molecules, 24,506 bulk water molecules, and six sodium ions added for charge neutrality in a simulation box of size  $90 \times 90 \times 90$  Å<sup>3</sup>. The system was assembled with CHARMM-GUI<sup>[33]</sup> Solvator.<sup>[34]</sup> We used the CHARMM36 protein force field,<sup>[11,20]</sup> the ion parameters of Roux and coworkers,<sup>[35]</sup> and the TIP3P water model.<sup>[29]</sup> For the heme and the two coordinating histidine side chains, we performed independent simulations using the two ferrous force field parameter sets refined here. For the non-bonded interactions, we used a switch distance of 10 Å, a cutoff distance of 12 Å and a pair list distance of 14 Å.

We used NAMD<sup>[36,37]</sup> with a Langevin dynamics scheme and a Nosé-Hoover Langevin piston<sup>[38,39]</sup> to perform simulations in the isobaric-isothermal ensemble (NPT) with isotropic cell fluctuations at a temperature of 300 K and a pressure of 1 bar,

**Table 1.** Interaction energies  $E$ , in kcal/mol, and distances  $d$ , in Å, computed with optimized CHARMM parameters for the ferrous FeHis complex.

Interaction site	$E_{QM}$	$E_{MM}$	$\Delta E$	$d_{QM}$	$d_{MM}$	$\Delta d$
HB-OHH	-3.41	-3.01	0.40	2.54	2.68	0.14
HD1-OHH	-9.98	-9.90	0.08	1.95	1.85	-0.10
HE1-OHH	-3.49	-3.45	0.04	2.30	2.45	0.15
OC1/2-HOH	-7.42	-7.46	-0.04	1.94	1.74	-0.20
OC3-HOH	-2.91	-2.85	0.06	2.23	1.98	-0.25
HO3-OHH	-8.31	-8.07	0.24	1.86	1.82	-0.04
Average deviation			0.13			-0.05
Av. absolute dev.			0.14			0.15

Atoms of the interaction sites are labeled according to the CHARMM nomenclature from Figure 2b.

and with a Langevin damping coefficient of  $5.0 \text{ ps}^{-1}$ . We used SHAKE<sup>[40]</sup> to constrain covalent bonds involving hydrogen atoms. Heating was performed with velocity rescaling in the canonical ensemble (NVT). During heating and the first 2 ns of equilibration, we used an integration time step of 1 fs. For the rest of the simulation, we used the reversible multiple time step integration scheme<sup>[41,42]</sup> with steps of 1 fs for the bonded forces, 2 fs for short-range non-bonded, and 4 fs for long-range non-bonded interactions. The simulation was prolonged to 100 ns, and coordinate sets were saved every 10 ps. Unless specified otherwise, average distances and histograms were computed using 5000 equally spaced coordinate snapshots from the last 50 ns of the simulation.

All molecular graphics and skeletal formulae were prepared using PyMol<sup>[43]</sup> and BKchem,<sup>[44]</sup> respectively. QtiPlot<sup>[45]</sup> was used to generate data plots.

## Results and Discussion

We derived a new set of revised, CHARMM-compatible parameters for HemeHis and FeHis. The parametrized bonds and angles of the final MM-optimized structures are within  $0.01 \text{ \AA}$  and  $1.8^\circ$  (HemeHis), and within  $0.03 \text{ \AA}$  and  $0.8^\circ$  (FeHis) of their QM counterparts. The water interaction energies of ferrous and ferric HemeHis agree to within 0.4 kcal/mol with QM (Table 2, Supporting Information Table S11). The atomic partial charges optimized here for ferrous and ferric FeHis give water interaction energies that agree to within 0.5 kcal/mol with the HF/6-31G\* values (Table 1, Supporting Information Table S10).

**Table 2.** Interaction energies  $E$ , in kcal/mol, for the ferrous HemeHis complex computed with optimized CHARMM parameters using parameter Set 1 and Set 2.

Interaction site	$E_{QM}$	$E_{Set 1}$	$\Delta E$	$E_{Set 2}$	$\Delta E$
HB-OHH	-1.35	-1.06	0.29	-1.24	0.11
HD1-OHH	-6.91	-6.89	0.01	-5.02	1.88
H <sup>Heme methine</sup> -OHH	-1.02	-1.24	-0.22	-1.27	-0.25
H <sup>Heme methyl</sup> -OHH	-0.82	-0.59	0.23	-0.61	0.20
Average deviation			0.08		0.49
Av. absolute dev.			0.19		0.61

Atoms of the interaction sites are labeled according to the CHARMM nomenclature from Figure 2a.

In the Supporting Information, we provide topology and parameter files for the HemeHis and FeHis force field parameters reported here. To facilitate CHARMM force field computations of systems with non-heme iron compounds, we have also included a set of parameters for ferrous and ferric iron coordinated by four histidines, that is, FeHis without bicarbonate. These additional parameters were derived with the protocol used for FeHis and are documented by data provided in Supporting Information Fig. S8 (partial charges) and Supporting Information Tables S13–S15 (water interaction energies and geometry parameters).

In what follows, we discuss main observations from revising selected force field parameters for HemeHis and FeHis, and the performance of the HemeHis parameters in describing the dynamics of neuroglobin.

### Water interactions of FeHis and HemeHis

We computed water interaction energies with HF and with CHARMM to derive atomic partial charges for both systems. The MM atomic partial charges were adjusted such that the interaction distance  $d_{MM}$  and the interaction energy  $E_{MM}$  for each interaction site were within good agreement with the corresponding QM values  $d_{QM}$  and  $E_{QM}$ . For ferrous FeHis, this procedure led to the final partial charges shown in Figures 3b and 3c. The partial charges of the other systems studied here are given in Supporting Information Figures S5–S8.

The optimized water interaction energies and interaction distances for all FeHis interaction sites studied here are summarized in Table 1. The interaction distances  $d_{MM}$  and  $d_{QM}$  agree to within  $0.04$ – $0.25 \text{ \AA}$  (Table 1). Qualitatively, we obtained good agreement of water interaction distances for donor and acceptor sites most likely to interact with water: the histidine side chains, and the hydroxyl group of bicarbonate (Table 1). The QM and MM interaction energies agree on average to within 0.4 kcal/mol for all sites that could be sampled. The partial charges of the ferrous FeHis complex are briefly described below.

The optimized parameter set for FeHis has atomic partial charges of bicarbonate atoms different from those used for bicarbonate in CGenFF; for example, in our final set, the atoms OC1, OC2, and CT, which are close to the iron, have a partial charge of  $-0.57$  and  $0.80$  for the oxygen and the carbon atoms, respectively. In CGenFF, the corresponding atoms have partial charges  $-0.76$  and  $0.69$ . These differences are explained by the electrostatic interactions of the bicarbonate compound with other groups of the FeHis system. Likewise, the partial charges of the four histidine moieties differ from standard CHARMM charges. The CHARMM charge of unprotonated NE2 in single histidine is  $-0.70$ ; in FeHis, where NE2 coordinates iron, the optimized charge on NE2 is  $0.10$ . Because of the geometry of FeHis (Fig. 1b), the CD2 site could not be accessed directly; however, by adjusting the CD2–HD2 and ND1–HD1 dipole moment as well as the charges on CG and CB, we obtained an overall good description of the FeHis water interaction energies (Table 1, Supporting Information Table S9).

**Table 3.** Comparison of QM and MM geometries, using the initial guess and final optimized force field parameters.

Parameter	HemeHis (Fe <sup>2+</sup> )			FeHis (Fe <sup>2+</sup> )		
	QM	MM <sub>initial</sub>	MM <sub>final</sub>	QM	MM <sub>initial</sub>	MM <sub>final</sub>
	Bonds (Å)			Bonds (Å)		
FE–N <sub>Heme</sub>	2.01	1.96	2.01	–	–	–
FE–NE2	2.01	2.08	2.01	1.98	2.23	1.98
FE–OC1/2	–	–	–	2.04	1.72	2.04
CT–OC1/2	–	–	–	1.27	1.30	1.26
CT–OC3	–	–	–	1.35	1.33	1.35
	Angles (°)			Angles (°)		
FE–NE2–CD2	128	131	128	129	136	129
FE–NE2–CE1	126	121	127	125	119	125
FE–OC1/2–CT	–	–	–	86	77	86
OC1–FE–OC2	–	–	–	66	83	66
CT–OC3–HO3	–	–	–	106	106	106

Average bonds and angles are given for: 4-Mel; the heme plane nitrogens NA, NB, NC, and ND (denoted as N<sub>Heme</sub>); and the OC1 and OC2 bonds and angles. Atoms of the interaction sites are labeled according to the CHARMM nomenclature from Figure 2.

The iron of FeHis is buried deep inside the complex and could not be accessed for a direct calculation of its water interaction energies. We thus chose to keep the partial charge of the iron ion at the starting MK charge value of +0.16. That the charge of iron in FeHis is smaller than the +0.24 value in six-coordinated heme from CHARMM could be explained by the presence of the negatively charged bicarbonate molecule in FeHis.

Using the CHARMM36 protein force field,<sup>[20]</sup> ferrous HemeHis can be modeled by applying twice the patch that connects heme to a single histidine.<sup>[13,46]</sup> This approach leaves the partial charges of the HemeHis compounds at the standard CHARMM values. The original CHARMM partial charges yield a good description of the non-polar interaction sites, but underestimate the HD1 interaction by 1.88 kcal/mol (Table 2). We adjusted the MM partial charges on 4-Mel and obtained agreement to within 0.01 kcal/mol with the QM value for this interaction site. This set of HemeHis parameters is denoted here as Set 1.

In addition to the ferrous HemeHis parameter Set 1 with optimized partial charges and bonded parameters, we derived a second ferrous HemeHis parameter set where only the bonded parameters were optimized, while the partial charges were kept at their original CHARMM values. This set of HemeHis parameters is denoted here as Set 2.

For the ferric complexes, we derived new partial charges starting from MK calculations.

We note that for some of the buried sites close to the iron the water interaction energies and distances could not be calculated.

### Bond and angle parameters for HemeHis and FeHis

Starting bonded parameters for HemeHis were taken from the CHARMM36 protein force field and its extension for heme with a single histidine.<sup>[13,46]</sup> Comparison of the MM- and QM-optimized geometries of HemeHis indicated that the initial

MM parameters underestimated the strength of the interaction between iron and the NE2 atoms of the coordinating 4-Mel groups. The distance between the iron and the 4-Mel nitrogen NE2 is 2.01 Å and 1.99 Å in the QM-optimized structure of ferrous and ferric HemeHis, respectively, as compared to 2.08 Å and 2.07 Å for the MM-optimized geometry when using the standard CHARMM parameters for ferrous heme with one coordinating histidine group (Table 3, Supporting Information Table S12).

For FeHis, the initial guess for the parameters for the bond between iron and the bicarbonate oxygens OC1 and OC2 (Fig. 2b) was taken from the bond between heme iron and carbon dioxide found in the heme patch of the CHARMM36 protein force field. As for HemeHis, the bond between non-heme iron and 4-Mel nitrogen NE2 was based on the CHARMM parameters for heme with a single histidine.

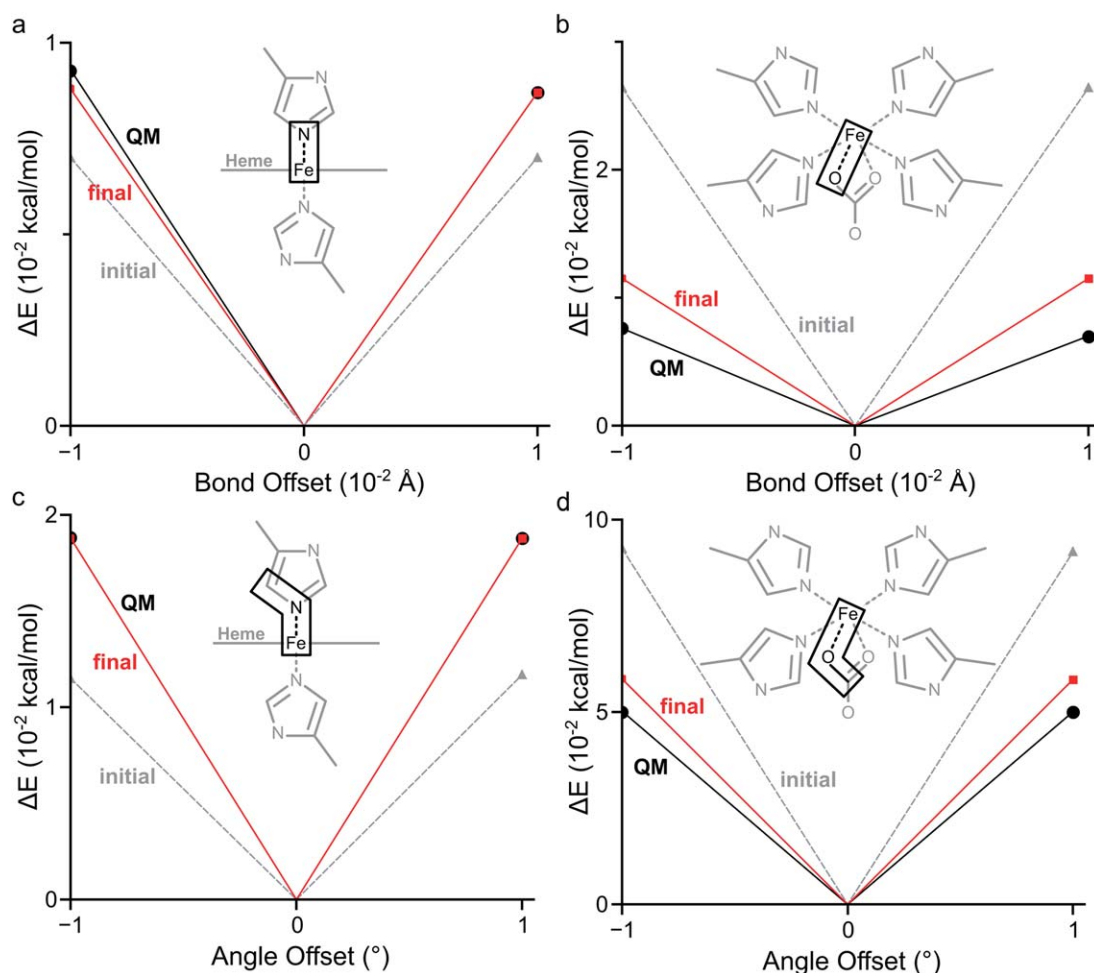
Comparison of the QM-optimized structure of ferrous and ferric FeHis with that from MM optimizations indicated that the length of the coordination bond between iron and the coordinating nitrogen atoms was overestimated. In ferrous FeHis, the starting MM parameters yielded a bond length of 2.23 Å, as compared to the 1.98 Å value obtained with MP2. In ferric FeHis, UB3LYP gives a bond length of 1.99 Å as compared to 2.24 Å with the initial guess parameters; in the crystal structure (PDB ID: 3WU2, monomer A<sup>[1]</sup>), the distances between the coordinated histidine NE2 atoms and iron range from 2.10 to 2.29 Å.

We used PES computations to optimize the force constant for the FE–NE2 coordination bonds in FeHis and HemeHis (Fig. 4a). This led us to increase the force constant from the standard CHARMM value of 65 kcal/mol/rad<sup>[2]</sup> to 80 and 140 for HemeHis and FeHis, respectively (Table 4).

The force constants optimized for the ferrous systems gave a good description of the ferric systems, and we only optimized the equilibrium values of the bond and angle parameters (Supporting Information Table S16).

With the initial guess, the PES analysis of the iron–bicarbonate coordination bond in ferrous FeHis indicated that the energy for stretching the FE–OC1 and FE–OC2 bonds by 0.01 Å was overestimated by ~260% and ~190%, respectively. We reduced the force constant for bond stretching by ~60%, and obtained agreement between the MP2 and MM bond stretching energies to within  $3 \times 10^{-3}$  kcal/mol (Fig. 4b, Table 4). Likewise, the PES for the valence bond angle involving FE and the bicarbonate oxygen atoms indicated an overestimation of the energy for a 1° angle bending by ~180% (Fig. 4d), and we reduced the associated force constants by 5 kcal/mol/rad<sup>[2]</sup>.

The revised parameters give good agreement for FeHis and HemeHis structures geometry-optimized with MM versus QM, using MP2 for ferrous FeHis, B3LYP for ferrous HemeHis (Table 3), and UB3LYP for the ferric complexes (Supporting Information Table S12). For ferrous FeHis, valence angles involving NE2 coordination bonds are the same with QM and revised MM. We note, however, that the lengths of these coordination bonds in the QM-optimized model systems are shorter than in the starting crystal structure (PDB ID: 3WU2). For ferrous



**Figure 4.** Three-point PES to refine force constants for bonds and valence angles. a) Optimization of the force constant of the iron–4-Me1 bond in ferrous HemeHis. Initial MM results are shown in dashed gray. The final optimized MM energy differences and the target QM data are indicated by black circles and red squares, respectively. b) PES for the bond between bicarbonate and non-heme iron in ferrous FeHis. c) PES for the iron–4-Me1 angle in ferrous HemeHis. d) PES for the angle involving the non-heme iron and bicarbonate in ferrous FeHis. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

HemeHis, there is improved agreement between QM and MM values involving NE2 coordination bonds (Table 3).

#### Optimization of selected dihedral angles for HemeHis and FeHis

We used the standard CHARMM parameters to compute a starting PES for the heme–histidine torsion about the dihedral angle CE1–NE2–FE–NA of ferrous HemeHis (Fig. 5a). Comparison of

this starting MM profile with the one computed with QM revealed that, at the 45°-twisted geometry, the MM energy profile yields an energy barrier of 0.72 kcal/mol, as compared to the 1.21 kcal/mol value calculated with QM (Fig. 5b). Moreover, the local minimum of the twisted geometry was situated at 80° in the MM computation, as compared to 90° in the QM profile, and was energetically less favorable relative to the untwisted starting geometry (0.49 kcal/mol and 0.35 kcal/mol for MM and QM, respectively, see Fig. 5b). To improve the agreement between the MM and QM energy profiles for the dihedral angle torsion, we increased the force constant of the term with multiplicity 4 from 0.05 kcal/mol/rad<sup>2</sup> to 0.14 kcal/mol/rad<sup>2</sup>, and we introduced two additional multiplicities (force constant 0.07 kcal/mol/rad<sup>2</sup>, multiplicity 2; force constant 0.04 kcal/mol/rad<sup>2</sup>, multiplicity 3). This adjustment allowed us to reproduce the QM barrier height and energy difference between structures at a torsion of 0° and ±90°.

For CE1–NE2–FE–NA in ferric HemeHis, QM calculations showed an increased barrier height of 1.92 kcal/mol (Supporting Information Figure S2d). We performed MM PES of the dihedral, and optimized its force constants and multiplicities. The optimized parameters give an excellent description of the

**Table 4.** Comparison of the initial guess, taken from standard CHARMM, for force constants of specific bonded parameters with values optimized here.

Parameter	Initial	HemeHis	FeHis
		Bonds (kcal/mol/Å <sup>2</sup> )	
FE–NE2	65	80	140
OC1/2–FE	250	–	100
		Angles (kcal/mol/rad <sup>2</sup> )	
FE–NE2–CD2	30	25	25
FE–NE2–CE1	30	25	25

Atoms of the interaction sites are labeled according to the CHARMM nomenclature from Figure 2.



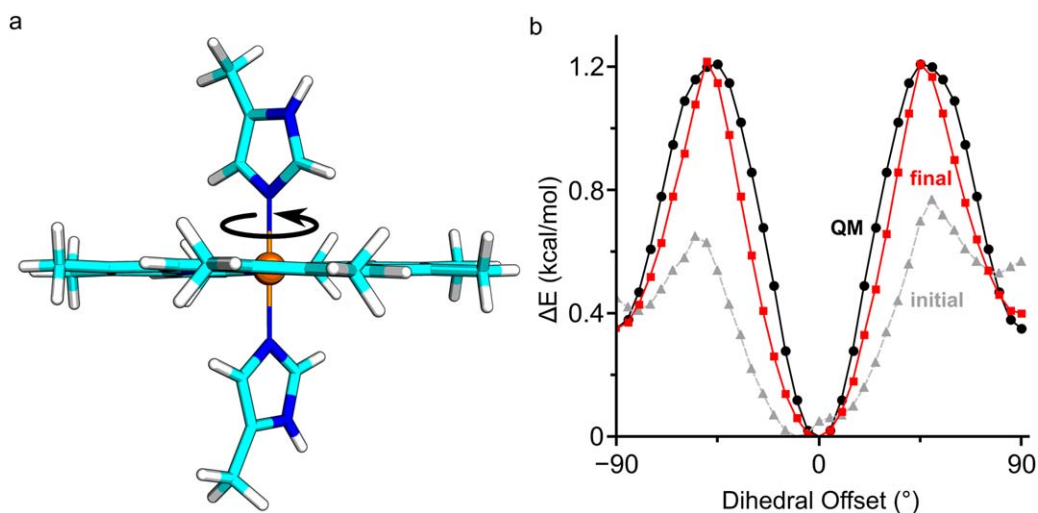


Figure 5. Optimization of dihedral parameters for ferrous HemeHis. a) Structure of HemeHis optimized with B3LYP. The arrow indicates the torsional axis about which the PES is performed. b) Energy profiles for PES computed with QM (black circles), initial CHARMM parameters (gray triangles), and with the optimized parameters presented here (red squares). Note the marked improvement of the agreement between the final MM energy profile and the target QM data. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

torsional potential of the heme–histidine torsion in ferric HemeHis, with an MM barrier height of 1.93 kcal/mol (Supporting Information Figure S2d).

In FeHis, the dihedral angle OC1–CT–OC3–HO3 describes the motion of the hydroxyl group of the bicarbonate, and the dihedral FE–OC1–CT–OC3 describes the rotation about the coordination bond between iron and bicarbonate. To derive parameters for these two important dihedral angles, we used MP2 (UB3LYP for ferric FeHis) to calculate the energy required to twist each of the two dihedral angles. As reference value, we used the energy of the geometry-optimized FeHis structure without any constraints. Starting from this reference structure, for each of the two dihedral angles, we performed PES using a step size of 5°, with maximum twists of +90° (clockwise twist)

and –90° (counterclockwise twist, see Fig. 6a and Supporting Information Figure S2). We then used the QM PES of each dihedral angle to manually adjust CHARMM dihedral angle parameters until the MM PES could no longer be further refined.

Figure 6a shows the improvement of the description of the energetics close to the minimum position of the bicarbonate dihedral angle OC1–CT–OC3–HO3 in ferrous FeHis. For relatively minor twists of  $\leq 10^\circ$ , which cost  $\sim 1$  kcal/mol and can be easily sampled at room temperature, the standard CHARMM parameters give a good description of the MP2 profile (Fig. 6a). Larger twists of  $\sim 45^\circ$  require  $\sim 3$  kcal/mol with the guess parameters, as compared to  $\sim 5$  kcal/mol with MP2 or the revised parameters presented here (Fig. 6a). Likewise,

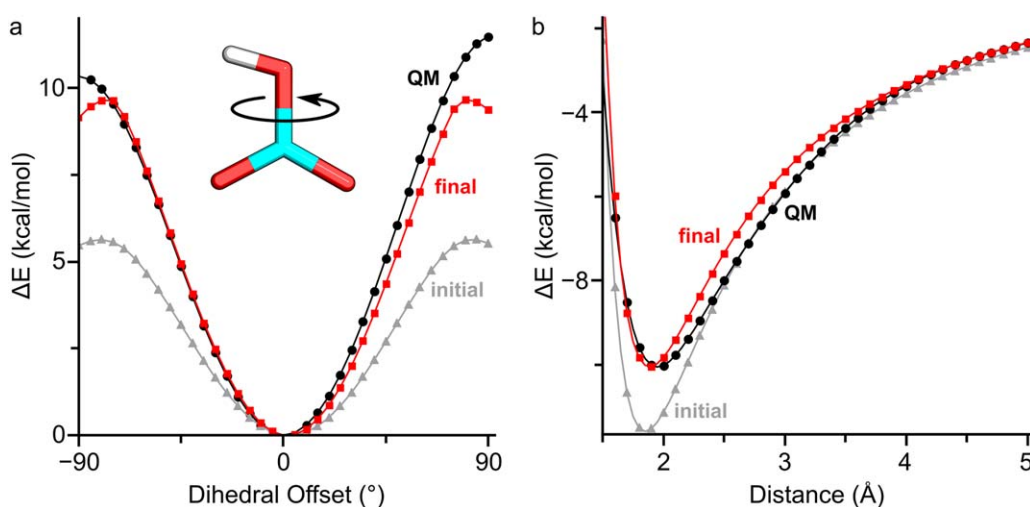
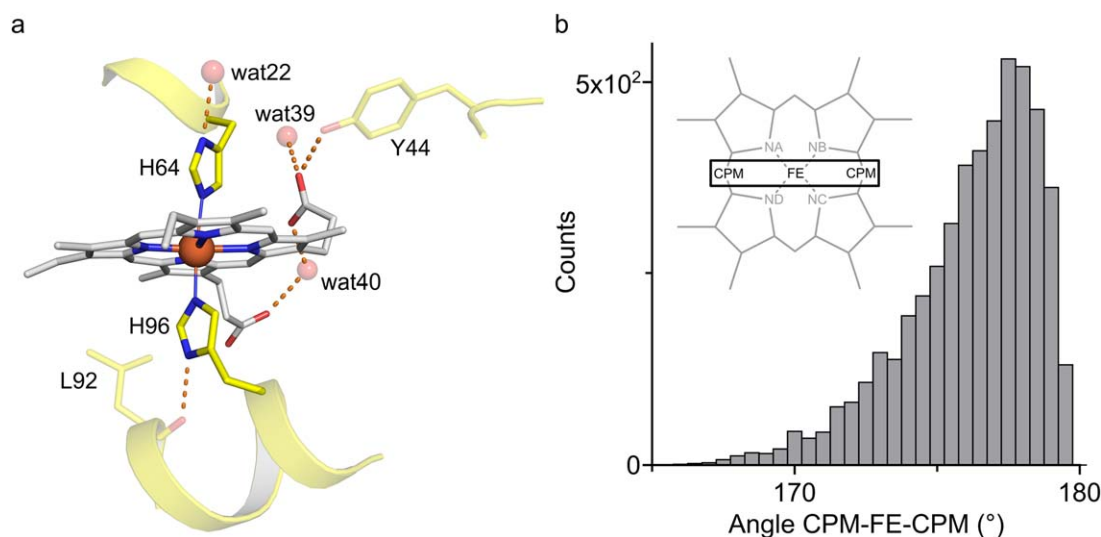


Figure 6. Dihedral PES of the bicarbonate in ferrous FeHis, and water interaction energies of ferrous FeHis. a) Energy profiles for PES for the torsion of the hydroxyl group of bicarbonate (see arrow) computed with MP2 (QM, black circles), initial CHARMM parameters (gray triangles), and optimized CHARMM parameters (red squares). The dihedral offset, in degrees, gives the value of the dihedral angle relative to its value of  $-2.4^\circ$  in the MP2-minimized structure. b) Water interaction energies for the HD1 donor site for one of the 4-Mel groups of FeHis computed with HF/6–31G\* (black circles), with the initial values (gray triangles), and with the optimized CHARMM force field parameters (black circles). [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**Figure 7.** Close view of the heme–histidine complex in neuroglobin, and planarity of the heme plane. a) Molecular graphics showing the heme b, coordinating histidine side chains, and nearby hydrogen-bonding groups in the starting crystal structure (PDB ID: 1OJ6, chain B<sup>[32]</sup>). Water molecule wat40 bridges the heme propionate groups via hydrogen bonds with distances of 2.5 and 2.8 Å. The hydrogen bond between the backbone oxygen of L92 and ND1 of H96 was conserved throughout our simulation, and had a value of  $2.9 \pm 0.2$  Å. b) Histogram of the out-of-plane bending of the heme iron in neuroglobin in sim1. Throughout most of the simulation the heme stayed close to planarity, with an average angle CPM–Fe–CPM of  $176 \pm 2^\circ$  during the last 50 ns of the simulation. [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

compared to the MP2 PES, torsional freedom of the dihedral FE–OC1–CT–OC3 in ferrous FeHis was overestimated when using standard CHARMM parameters (Supporting Information Figure S2). The revised bonded and non-bonded parameters presented here markedly improve the torsional energy profile for FE–OC1–CT–OC3 (Supporting Information Figure S2a).

Following the same protocol as described above for ferrous FeHis, we optimized the torsional parameters for the dihedrals OC1–CT–OC3–HO3 and FE–OC1–CT–OC3 of ferric FeHis, using UB3LYP computations. The PES of these dihedral angles show a marked improvement of the torsional properties of bicarbonate in ferric FeHis (Supporting Information Figures S2b and S2c).

We note that using the same chemical types for all four 4-Mel histidine groups of FeHis limits the ability of the force field to account for asymmetry in the orientation of histidine side chains in both FeHis complexes as observed in the QM-optimized structure.

#### Testing the revised HemeHis parameters with molecular dynamics of neuroglobin

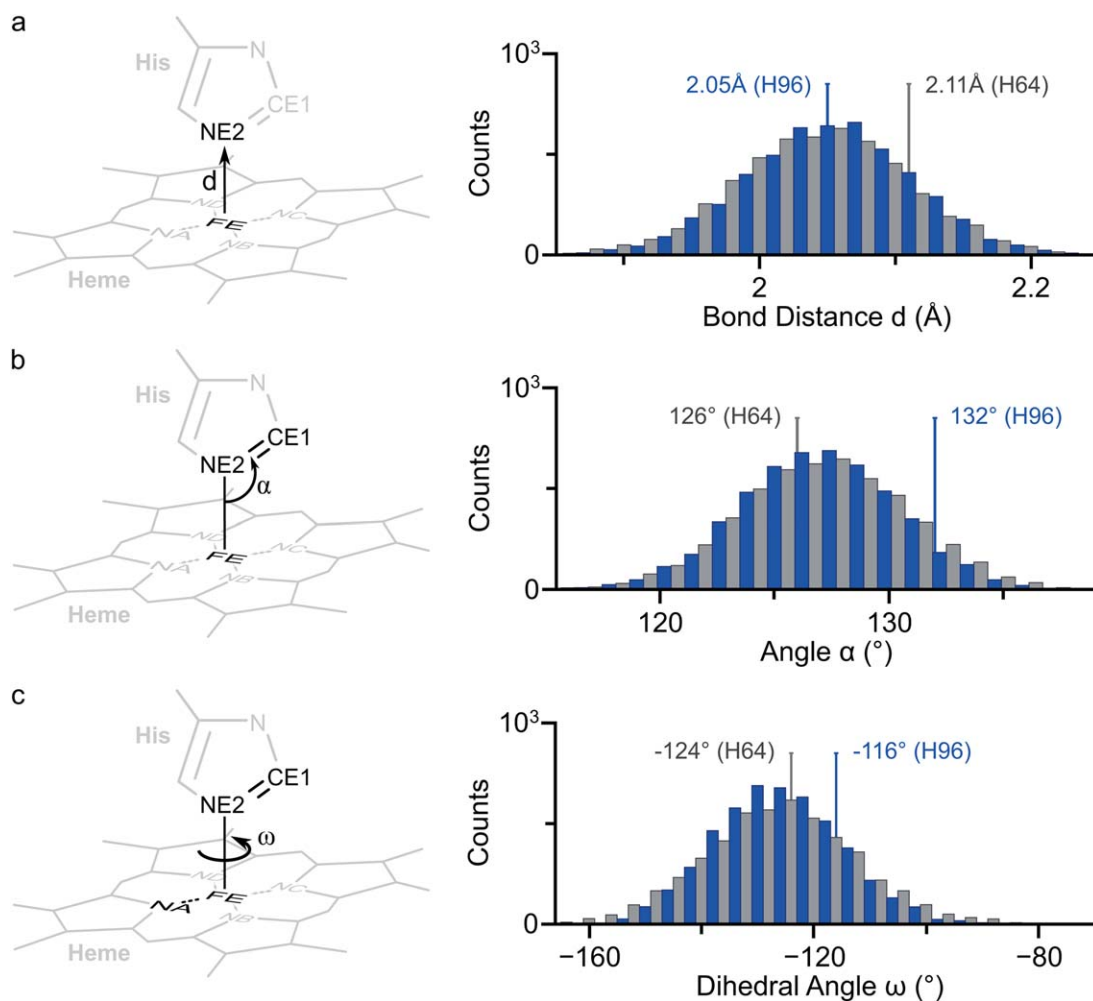
The accuracy of the ferrous HemeHis parameters was further evaluated by performing four independent all-atom molecular dynamics simulations of neuroglobin in a box of water molecules. Sim1 used the ferrous HemeHis parameter Set 2, and sim2 used parameter Set 1. Sim3 and sim4 are repeat simulations using Set 2 and Set 1, respectively. Heme b of neuroglobin is coordinated by H64 and H96 (Fig. 7a).

In the starting crystal structure (PDB ID: 1OJ6, chain B<sup>[32]</sup>), both propionate groups of the heme hydrogen-bond to water molecules, and one of the propionates has an additional hydrogen bond to Y44 (Fig. 7a). H64 and H96 hydrogen-bond to water and protein backbone, respectively (Fig. 7a).

On the  $\sim 100$  ns timescale of our simulations, the  $\alpha$ -helical segments of neuroglobin remain largely stable, as indicated by the  $C_\alpha$  root-mean-square deviation computed relative to the starting crystal structure (Supporting Information Figure S3). Importantly, the average distance between the  $C_\alpha$  atoms of H64 and H96 during the last 50 ns of the simulation,  $12.5\text{--}12.7 \pm 0.2$  Å, is very close to the  $12.2 \pm 0.1$  Å average distance computed for the four neuroglobin chains in the starting crystal structure<sup>[32]</sup> (Supporting Information Figure S4).

To characterize the dynamics of the heme and its two coordinating histidine moieties (Fig. 7a), we use here the CPM–Fe–CPM angle as a measurement of the local planarity of the heme (Fig. 7b), and structural parameters that describe the orientation of the histidine groups relative to the heme plane: the distance  $d$  between the heme iron and NE2 atoms of H64 and H96, the angle  $\alpha$  describing the bend of H64/H96 relative to the heme plane, and the dihedral angle  $\omega$  that gives the twist of the histidine imidazole group relative to the heme (Fig. 8 and Table 5). As reference values for these three structural parameters, we use information from the starting crystal structure of neuroglobin (PDB ID: 1OJ6<sup>[32]</sup>), and from other neuroglobin crystal structures that contain the heme cofactor with two coordinating histidine groups (see Table 6).

As illustrated in Figures 7 and 8 and Table 5, both sets of the revised parameters of HemeHis allow reliable description of the geometry of the heme and of the heme–histidine interactions. For both H64 and H96, the average FE–NE2 distance agrees well with the crystal structure (Fig. 8a) and with QM (Table 3). The average angle of the heme CPM–Fe–CPM angle,  $176 \pm 2^\circ$  (Fig. 7b), indicates that the heme plane is slightly bent. This observation is compatible with data from Autenrieth et al.,<sup>[17]</sup> based on their B3LYP computations on heme models and inspection of heme-containing proteins.



**Figure 8.** Testing the HemeHis parameters in classical MD simulations of neuroglobin. The bond lengths, valence angles and dihedral angles, measured for sim1, are illustrated together with their corresponding histogram. Histograms show the distribution for the length of the coordination bond FE–NE2 (a), the angle FE–NE2–CE1 (b), and the dihedral NA–FE–NE2–CE1 (c) sampled during the last 50 ns of the simulation. The gray and blue bars correspond to data sampled from H64 and H96, respectively. The numbers indicated for H64 and H96 are measured from the starting crystal structure coordinates (PDB ID: 1OJ6, chain B<sup>[32]</sup>). The distributions agree well with the crystal structure data (shown within the histograms). [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

Inspection of the average values for  $d$ ,  $\alpha$ , and  $\omega$  (Fig. 8) indicates that, overall, the orientation of H64 relative to the heme plane is well described by the revised force field

parameters presented here (Table 5). The average lengths of the FE–NE2 coordination bond for H64 and H96 computed from simulations are in excellent agreement with the crystal structure (Fig. 8a and Table 5). Likewise, the average angle CE1–NE2–FE for H64 is  $127 \pm 4^\circ$  in our simulations, as compared to  $126 \pm 1^\circ$  in the four protein chains of the crystal structure (Table 5).

**Table 5.** Using the ferrous HemeHis parameters to probe the dynamics of neuroglobin.

Parameter	His	Crystal	Sim1	Sim2
Bond distance $d$ (Å)				
FE–NE2	H64	$2.11 \pm 0.05$	$2.06 \pm 0.06$	$2.05 \pm 0.06$
	H96	$2.05 \pm 0.04$	$2.05 \pm 0.06$	$2.05 \pm 0.06$
Angles ( $^\circ$ )				
CE1–NE2–FE ( $\alpha$ )	H64	$126 \pm 1$	$128 \pm 4$	$127 \pm 4$
	H96	$132 \pm 2$	$127 \pm 3$	$126 \pm 3$
CD2–NE2–FE	H64	$125 \pm 1$	$127 \pm 4$	$128 \pm 3$
	H96	$119 \pm 2$	$128 \pm 3$	$128 \pm 3$
Dihedral $\omega$ ( $^\circ$ )				
CE1–NE2–FE–NA	H64	$-124 \pm 4$	$-124 \pm 14$	$-117 \pm 12$
	H96	$-116 \pm 4$	$-127 \pm 11$	$-121 \pm 12$

“Crystal” refers to values measured in the crystal structure (PDB ID: 1OJ6<sup>[32]</sup>). For H64 and H96, we report the geometry parameters illustrated in Figure 8.

**Table 6.** Heme planarity described by the angle CPM–FE–CPM in crystal structures of neuroglobin.

System	Planarity ( $^\circ$ )	Resolution (Å)
Sim1	$176 \pm 2$	–
Sim2	$177 \pm 2$	–
Sim3	$176 \pm 2$	–
Sim4	$177 \pm 2$	–
1OJ6 <sup>[32]</sup>	$176 \pm 1$	1.95
1Q1F <sup>[47]</sup>	$173 \pm 5$	1.50
2VRY <sup>[48]</sup>	$175 \pm 4$	1.87
4B4Y <sup>[49]</sup>	$170 \pm 2$	2.30
4MPM <sup>[50]</sup>	$176 \pm 2$	1.74

The  $\omega$  dihedral angle that characterizes the orientation of the imidazole group relative to the plane of the heme is well described for both H64 and H96 (Fig. 8c and Table 5). In case of H64, the difference between the angle CD2–NE2–FE sampled in simulations and the crystal structure average is within the error of the simulations (Table 5). The average value of the angle CD2–NE2–FE of H96 is  $128 \pm 3^\circ$  in our simulations, as compared to  $119 \pm 2^\circ$  in the crystal structure of neuroglobin (Table 5).

The hydrogen bond present in the crystal structure between the backbone oxygen of L92 and H96 ND1 was persistently sampled throughout all simulations. In the simulations using HemeHis parameter Set 2 (sim1, sim3), the hydrogen bond was present  $\sim 80\%$  of the time, whereas Set 1 (sim2, sim4) gave a higher hydrogen-bonding frequency of  $\sim 90\%$ . During the last 50 ns, the average hydrogen bond distance is  $2.9 \pm 0.2 \text{ \AA}$  in sim1 and sim3, and  $2.8 \pm 0.1 \text{ \AA}$  in sim2 and sim4. These values are close to the  $2.7 \pm 0.0 \text{ \AA}$  distance measured in the crystal structure. Although both sets, Set 1 and Set 2, allow good description of hydrogen-bonding at the ND1 site of H96, Set 1 appears to give slightly better agreement with the crystal structure. As anticipated, based on the crystal structure (Fig. 7a), both heme propionate groups are always hydrogen-bonding to water in all simulations. We note that Y44 was bound to heme only during 23% of our trajectory, with this hydrogen bond breaking and reforming throughout the simulation.

The crystal structure shows wat22 to be within hydrogen-bonding distance of H65 (Fig. 7a). When using Set 2, where the ND1 water interaction is weaker than the target QM value (Table 2), wat22 initially hydrogen-bonds with H64, but leaves the interaction site after 27 ns and 78 ns in sim1 and sim3, respectively, with no new water molecule entering. The charge-optimized parameter Set 1 used in sim2 and sim4 correctly reproduces the QM interaction energy at the ND1 site (Table 2), and during these two simulations, wat22 hydrogen-bonds to H64  $\sim 97\%$  of the time and does not leave the ND1 site.

## Conclusions

We derived CHARMM all-atom force field parameters for the two iron-containing cofactors of photosystem II, HemeHis, and FeHis (Fig. 1) in their ferrous form. To facilitate CHARMM computations of heme and non-heme iron compounds, we further optimized partial charges and bonded parameters for ferric HemeHis and FeHis, and for ferrous and ferric iron coordinated by four histidines.

The HemeHis complex, in which two histidine moieties coordinate the heme molecule (Figs. 2a and 2b), is found, for example in neuroglobin<sup>[48]</sup> (Fig. 7a) or in cytochromes.<sup>[51,52]</sup> The parameters presented here facilitate reliable all-atom simulations of proteins that contain heme and non-heme iron complexes.

To describe the HemeHis complex with CHARMM, we used as starting point the existing parameters for heme coordinated by one histidine group. We revised bonded interactions that

describe the geometry involving the coordination bonds: force constants and equilibrium values for the bond lengths and for angle bending (Figs. 4a and 4c, Table 4), and the torsional potential for rotation of the imidazole group relative to the heme plane (Fig. 5a). For each of these parameters, we achieved improved agreement between the CHARMM energy profiles and the target QM data (Figs. 4a, 4c, and 5b).

Tests on the dynamics of neuroglobin using the new force field parameters for HemeHis indicate that, regardless of the parameter set used (Set 1 or Set 2), we obtain a reliable description of the geometry of the heme, and of the orientation of the histidine side chains relative to the heme plane (Figs. 7a and 8, Table 5). Adjusting the histidine partial charges (Set 1), however, improved water hydrogen-bonding close to H64.

For specific protein systems, data from protein dynamics could be used to further adjust selected parameters, such as to account for the influence that the protein environment can have on the dynamics of the histidine side chains. That dynamics of the heme iron environment are influenced by protein interactions of the heme has been suggested based on work on cytochrome c and myoglobin.<sup>[53]</sup>

The FeHis model of the non-heme iron complex from photosystem II includes the iron ion,<sup>[54–57]</sup> a negatively charged bicarbonate moiety, and four histidine side chains (Figs. 1b and 2b). To our knowledge, standard CHARMM lacks specific force field parameters for this cluster, and we derived new parameters to describe bonded and non-bonded interactions.

For bonded interactions of FeHis, we kept the standard CHARMM parameters for intramolecular interactions of histidine and bicarbonate groups, and derived parameters for the coordination bonds involving iron, imidazole groups, and bicarbonate (Figs. 4b and 4d and Table 4). As summarized in Table 3, we obtained excellent description of these bonds and valence angles. For the torsion of the bicarbonate CT–OC3 bond, the new parameters give very good agreement with the QM energy profile, correcting the shallow energy profile obtained with the starting MM parameters (Fig. 6a).

Atomic partial charges are particularly important for all-atom molecular dynamics simulations. To derive atomic partial charges for HemeHis and FeHis, we used HF/6–31G\* to compute water interaction energies and distances for all accessible unique sites of the complex (Fig. 3). We then adjusted the CHARMM atomic partial charges to reproduce the QM water interaction energies and distances. When using the optimized set of atomic partial charges (Fig. 3), ferrous FeHis water interaction energies and distances computed with MM versus QM agree to within 0.40 kcal/mol and 0.25 Å, respectively (Table 1).

In the protein environment, the ND1 atoms of FeHis hydrogen bond to quinone or protein groups (Fig. 1b). At room temperature in a flexible protein environment, these interactions can be dynamic. The good agreement between the MM and QM data for water interaction energies of the FeHis ND1 atoms (Fig. 6b) is encouraging, and suggests that water and protein interactions will likely be well described during molecular dynamics simulations (Fig. 6b).

We anticipate that the CHARMM topology and parameter files provided here will be useful for all-atom simulations of photosystem II and other proteins that have heme or non-heme iron complexes. Observations on the geometry and molecular interactions from such simulations can provide information to further improve the current force-field parameters. Of particular interest will be to include additional atom chemical types and account for asymmetry in the specific protein environment of the cofactor molecules.

## Acknowledgments


We acknowledge the cluster of the *Zentraleinrichtung für Datenverarbeitung (ZEDAT)* of the *Freie Universität Berlin* for computing time for QM computations, and *Jens Dreger* for excellent technical support with the local computing cluster of the *Department of Physics, Freie Universität Berlin*. *M.K.-M.* acknowledges computing resources from the *German Cancer Research Center Heidelberg, the DKFZ*. We thank the *Center for International Cooperation of the Freie Universität Berlin* for supporting the research visit of *J.Y.* *NAMD* was developed by the *Theoretical and Computational Biophysics Group in the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign*. We acknowledge the usage of the *CGenFF* tutorial from the webpage of the *Alex MacKerell group at the University of Maryland* to generate CHARMM input files for the parametrization.

## Authors' Contributions

S.A. and M.K.-M. carried out QM calculations to derive QM target data and to sample interaction energies for water interaction sites. S.A. performed the parameter optimization and the neuroglobin simulation, and carried out the data acquisition and statistical analysis. J.Y. contributed to the design and to interpretation of data on neuroglobin and heme structure and dynamics. A.-N.B. initiated and supervised the entire project. S.A. and A.-N.B. wrote the manuscript with input from all coauthors. All authors have read and accepted the manuscript.

**Keywords:** CHARMM force field · heme · non-heme iron · parametrization · photosystem II

How to cite this article: S. Adam, M. Knapp-Mohammady, J. Yi, A.-N. Bondar. *J. Comput. Chem.* **2018**, *39*, 7–20. DOI: 10.1002/jcc.24918

 Additional Supporting Information may be found in the online version of this article.

- [1] Y. Umena, K. Kawakami, J.-R. Shen, N. Kamiya, *Nature* **2011**, *473*, 55.
- [2] M. Suga, F. Akita, K. Hirata, G. Ueno, H. Murakami, Y. Nakajima, T. Shimizu, K. Yamashita, M. Yamamoto, H. Ago, J.-R. Shen, *Nature* **2015**, *517*, 99.
- [3] A.-N. Bondar, H. Dau, *Biochim. Biophys. Acta* **2012**, *1817*, 1177.
- [4] S. Lorch, S. Capponi, F. Pieront, A.-N. Bondar, *J. Phys. Chem. B* **2015**, *119*, 12172.
- [5] M. Andersson, A.-N. Bondar, J. A. Freites, D. J. Tobias, H. R. Kaback, S. H. White, *Structure* **2012**, *20*, 1893.
- [6] F. Guerra, A.-N. Bondar, *J. Membr. Biol.* **2015**, *248*, 443.
- [7] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187.
- [8] H.-A. Chu, Y.-F. Chiu, *Front. Plant Sci.* **2016**, *6*, 1261.
- [9] M. Sugiura, M. Nakamura, K. Koyama, A. Boussac, *Biochim. Biophys. Acta* **2015**, *1847*, 276.
- [10] F. Müh, A. Zouni, *Photosynth. Res.* **2013**, *116*, 295.
- [11] A. D. MacKerell, Jr., D. Bashford, M. Bellot, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. I. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorcikiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586.
- [12] B. R. Brooks, C. L. I. Brooks, A. D. MacKerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caffisch, L. Caves, Q. Cui, A. D. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, M. Karplus, *J. Comput. Chem.* **2009**, *30*, 1545.
- [13] K. Kuczera, J. Kuriyan, M. Karplus, *J. Mol. Biol.* **1990**, *213*, 351.
- [14] R. J. Loncharich, B. R. Brooks, *J. Mol. Biol.* **1990**, *215*, 430.
- [15] M. Meuwly, O. M. Becker, R. Stote, M. Karplus, *Biophys. Chem.* **2002**, *98*, 183.
- [16] L. Zhang, D.-A. Silva, Y. Yan, X. Huang, *J. Comput. Chem.* **2012**, *33*, 1969.
- [17] F. Autenrieth, E. Tajkhorshid, J. Baudry, Z. Luthey-Schulten, *J. Comput. Chem.* **2004**, *25*, 1613.
- [18] M. P. Johansson, V. R. Kaila, L. Laakkonen, *J. Comput. Chem.* **2007**, *29*, 753.
- [19] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. D. MacKerell Jr, *J. Comput. Chem.* **2010**, *31*, 671.
- [20] A. D. MacKerell, Jr., M. Feig, C. L. I. Brooks, *J. Comput. Chem.* **2004**, *25*, 1400.
- [21] F. Guerra, S. Adam, A.-N. Bondar, *J. Mol. Graph. Model.* **2015**, *58*, 30.
- [22] M. Head-Gordon, J. A. Pople, M. Frisch, *J. Chem. Phys. Lett.* **1988**, *153*, 503.
- [23] P. C. Hariharan, J. A. Pople, *Theor. Chim. Acta* **1973**, *28*, 213.
- [24] W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.* **1972**, *56*, 2257.
- [25] V. A. Rassolov, J. A. Pople, M. A. Ratner, T. L. Windus, *J. Chem. Phys.* **1998**, *109*, 1223.
- [26] K. Kim, K. D. Jordan, *J. Phys. Chem.* **1994**, *98*, 10089.
- [27] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623.
- [28] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian 09, Revision D.01, Gaussian, Inc., Wallingford CT, **2013**.
- [29] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.
- [30] C. G. Mayne, J. Saam, K. Schulten, E. Tajkhorshid, J. C. Gumbart, *J. Comput. Chem.* **2013**, *34*, 2757.
- [31] U. C. Singh, P. A. Kollman, *J. Comput. Chem.* **1984**, *5*, 129.
- [32] A. Pesce, S. Dewilde, M. Nardini, L. Moens, P. Ascenzi, T. Hankeln, T. Burmester, M. Bolognesi, *Structure* **2003**, *11*, 1087.
- [33] S. Jo, T. Kim, V. G. Iyer, W. Im, *J. Comput. Chem.* **2008**, *29*, 1859.
- [34] S. Jo, T. Kim, V.G. Iyer, W. Im, CHARMM-GUI Solvator, **2008**.
- [35] D. Beglov, B. Roux, *J. Chem. Phys.* **1994**, *100*, 9050.

- [36] L. Kalé, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan, K. Schulten, *J. Comput. Phys.* **1999**, *151*, 283.
- [37] J. C. Phillips, B. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten, *J. Comput. Chem.* **2005**, *26*, 1781.
- [38] S. E. Feller, Y. Zhang, R. W. Pastor, B. Brooks, *J. Chem. Phys.* **1995**, *103*, 4613.
- [39] G. J. Martyna, D. J. Tobias, M. L. Klein, *J. Chem. Phys.* **1994**, *101*, 4177.
- [40] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327.
- [41] H. Grubmüller, H. Heller, A. Windemuth, K. Schulten, *Mol. Simul.* **1991**, *6*, 121.
- [42] M. Tuckermann, B. J. Berne, G. J. Martyna, *J. Chem. Phys.* **1992**, *97*, 1990.
- [43] The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC, **2017**.
- [44] B. Kosata, BKChem Version 0.13, **2009**.
- [45] I. Vasilief, QtiPlot Version 0.9, **2011**.
- [46] J. W. Petrich, J. C. Lambry, K. Kuczera, M. Karplus, C. Poyart, J. L. Martin, *Biochemistry* **2002**, *30*, 3975.
- [47] B. Vallone, K. Nienhaus, M. Brunori, G. U. Nienhaus, *Proteins: Struct. Funct. Bioinformatics* **2004**, *56*, 85.
- [48] A. Arcovito, T. Moschetti, P. D'Angelo, G. Mancini, B. Vallone, M. Brunori, S. Della Longa, *Arch. Biochem. Biophys.* **2008**, *475*, 7.
- [49] C. Lechauve, M. Jager, L. Laguerre, L. Kiger, G. Correc, C. Leroux, S. Vinogradov, M. Czizek, M. C. Marden, X. Bailly, *J. Biol. Chem.* **2013**, *288*, 6957.
- [50] B. G. Guimaraes, D. Hamdane, C. Lechauve, M. C. Marden, B. Golinelli-Pimpaneau, *Acta Crystallogr. D* **2014**, *70*, 1005.
- [51] E. A. Berry, F. A. Walker, *J. Biol. Inorg. Chem.* **2008**, *13*, 481.
- [52] F. A. Walker, *Chem. Rev.* **2004**, *104*, 589.
- [53] B. M. Leu, Y. Zhang, L. Bu, J. E. Straub, J. Zhao, W. Sturhahn, E. E. Alp, J. T. Sage, *Biophys. J.* **2008**, *95*, 5874.
- [54] J. P. McEvoy, G. W. Brudvig, *Biochemistry* **2008**, *47*, 13394.
- [55] A. Remy, K. Gerwert, *Nat. Struct. Biol.* **2003**, *10*, 637.
- [56] P. Chernev, I. Zaharieva, H. Dau, M. Haumann, *J. Biol. Chem.* **2011**, *286*, 5368.
- [57] C. Berthomieu, R. Hienerwadel, *Biochemistry* **2001**, *40*, 4044.

---

Received: 2 March 2017

Revised: 5 July 2017

Accepted: 3 August 2017

Published online on 28 August 2017