Exploring coupled redox and pH processes with a force field-based approach: applications to five different systems

Vinícius Wilian D. Cruzeiro [1], Gustavo Troiano Feliciano [2], Adrian E. Roitberg [1,]*

[1] Department of Chemistry, University of Florida, Gainesville, FL 32611, United States

[2] Departamento de Físico-Química, Instituto de Química, Universidade Estadual Paulista (Unesp), Araraquara, Brazil

ABSTRACT

Coupled redox and pH-driven processes are at the core of many important biological mechanisms. As the distribution of protonation and redox states in a system is associated with the pH and redox potential of the solution, having efficient computational tools that can simulate under these conditions become very important. Such tools have the potential to provide information that complement and drive experiments. In previous publications we have presented the implementation of the constant pH and redox potential molecular dynamics (C(pH,*E*)MD) method in AMBER and we have shown how multidimensional replica exchange can be used to significantly enhance the convergence efficiency of our simulations. In the current work, after an improvement in our C(pH,*E*)MD approach that allows a given residue to be simultaneously pH- and redox-active, we have employed our methodologies to study five different systems of interest in the literature. We present results for: capped tyrosine dipeptide, two maquette systems containing one pH- and redox-active tyrosine (α_3Y and peptide A), and two proteins that contain multiple heme groups (diheme cytochrome *c* from *Rhodobacter sphaeroides* and *Desulfovibrio vulgaris Hildenborough* cytochrome c_3). We show that our results can provide new insights into previous theoretical and experimental findings by using a fully force field-based and GPU-accelerated approach, which allows the simulations to be executed with high computational performance.

INTRODUCTION

Coupled redox and pH-driven processes are at the core of many important biological mechanisms. The protonation and redox states of titratable residues can affect the function and other properties like stability, ligand binding, catalysis, and absorption spectrum of proteins and other biomolecules [1,2]. Also, Proton-Coupled Electron Transfer (PCET) processes are important in the study of many biological processes such as respiration, nitrogen fixation, and photosynthesis [3–6].

Problems related to redox processes are commonly theoretically investigated using QM [7–9] or QM/MM [10–13] approaches. In a previous publication [14] we have presented the implementation of a fully force field-based approach in AMBER that allows the execution of constant pH and constant redox potential molecular dynamics (C(pH,*E*)MD). We have validated our methodologies with simulations of a small system that contains a single heme: NAcMP8 axially connected to a histidine-containing peptide. We have correctly predicted the behavior of the standard redox potential (E^o) of the heme as a function of pH and the pK_a of each pH-active residue as a function of the redox potential. We have also explored the use of Replica Exchange Molecular Dynamics (REMD) as an enhanced sampling technique and have observed a better converge efficiency in our C(pH,*E*)MD simulations [14]. In a follow-up publication [15] we have shown how the use of multidimensional Replica Exchange Molecular Dynamics (REMD) further improves the convergence efficiency in comparison to one-dimensional REMD simulations. Multidimensional REMD allows more information to be exchange across replicas, which enhances the sampling. It is important to emphasize that all our methodologies can be executed with high computational performance because they are available using AMBER's GPU-accelerated code.

In this work we use recent improvements we applied in our C(pH,*E*)MD methodology in order to allow protons and electrons to be exchanged simultaneously in a given titratable residue. In the first released version of our C(pH,*E*)MD code, as reported in reference [14], a given residue could either be pH-active or redox-active but not pH- and redox-active simultaneously.

The current work is devoted to applying our methodologies to systems of interest that were investigated experimentally in the literature. One of the systems we looked at is the diheme cytochrome *c* (DHC) from *Rhodobacter sphaeroides* (see Figure 1A). This protein contains two heme groups that are similar to the heme group from NAcMP8, with two histidines as axial ligands, and its standard redox potential values are known experimentally as -240 and -310 mV *vs*. NHE at pH 7.5 [16]. These values of E^o are obtained using experimental data and fittings based on the Nernst equation, however, they cannot be assigned one-to-one to specific heme groups that are present in the structure of DHC. As we will show, our methodologies can aid with this assignment.

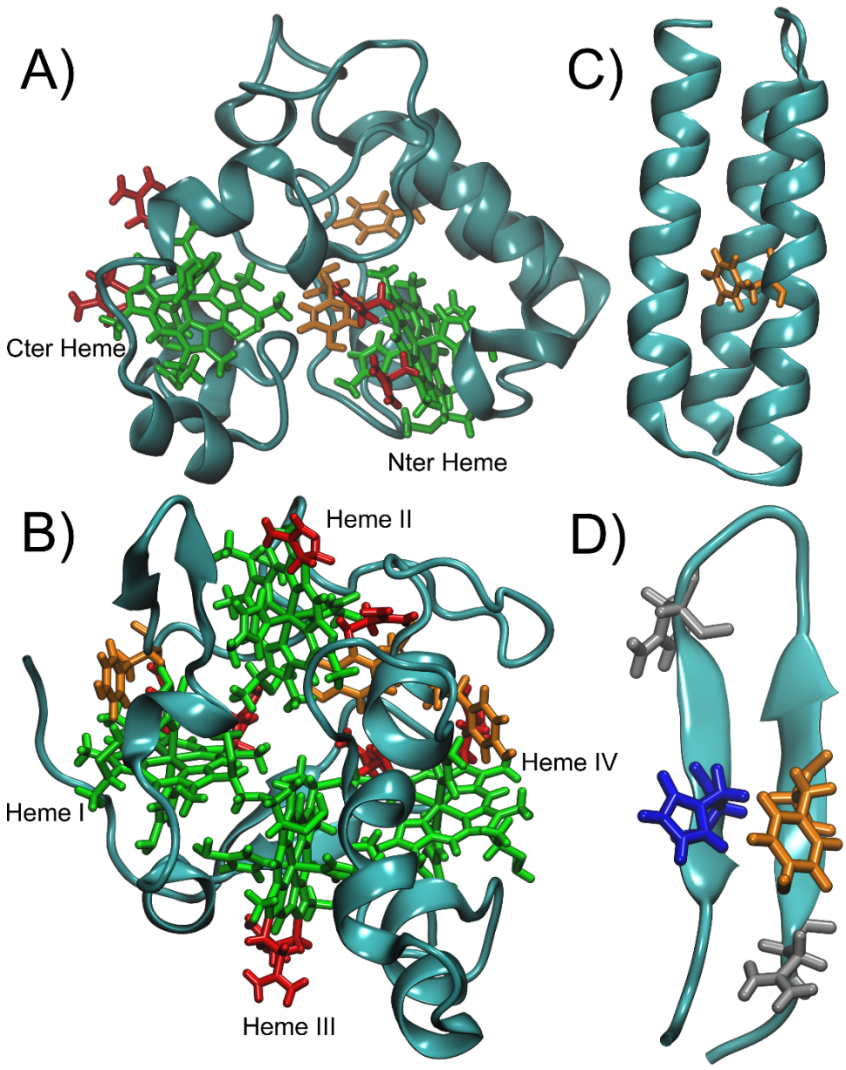

Figure 1. Structure of: A) diheme cytochrome *c* (DHC) from *Rhodobacter sphaeroides*; B) *Desulfovibrio vulgaris Hildenborough* cytochrome c_3 (DvHc_3); C) Alpha 3 Y (α_3Y); and D) Peptide A. The residues in green are the redox-active HEH residues (which contains the porphyrin ring and the side chains of the histidines and cysteines), the ones in red are the pH-active propionates, and the ones in orange are the pH- and redox-active tyrosines residues. For peptide A, a histidine is shown in blue, and two aspartates are shown in gray. All other pH-active residues are not explicitly shown in the figures. In our methodologies, hydrogens that have been

deprotonated are still present as dummy hydrogens with a charge of zero; this is why all four proton tautomeric positions are being displayed in the carboxyl groups.

Another system that we have studied is the *Desulfovibrio vulgaris Hildenborough* cytochrome c_3 (DvHc_3, see Figure 1B), a relatively small protein that contains four heme groups. Each heme group in this protein was uniquely associated to a different E^o value using a model and experimental data [17, 18], thus computational simulations can aid validating the assignment of the order in which the hemes are reduced. This ordering is of biological importance, among other things, because this protein acts as a mediator in the periplasmatic metabolism of H_2 in the sulfate-reducing bacteria [19–22]. The standard redox potential values obtained experimentally for the four hemes are -343, -302, -298 and -249 mV *vs*. NHE at pH 6.5 [18]. Previous theoretical studies in the literature [19, 23] have attempted to assign these standard redox potential values to specific heme groups of DvHc_3 but did not succeed.

In order to further evaluate our methodologies for describing coupled redox and pH processes, we have studied two different maquette systems that contain pH- and redox-active tyrosines. The first maquette system is the Alpha 3 Y (α_3Y, see Figure 1C) [24, 25]. This system contains three alpha helices (a total of 67 residues) and a single tyrosine that is buried (on position 32). The standard redox potential of this tyrosine has been measured experimentally. It is shifted down by approximately 193 mV at pH 13.4 in comparison to the standard redox potential of acetyl-L-tyrosinamide at the same pH [24].

The second maquette system is simply called peptide A (see Figure 1D) [26–28]. It is composed by a sequence of 18 amino acids, has a β-hairpin shape, and displays contacts between the side chains of a tyrosine and a histidine [26]. The standard redox potential of this system as a function of pH has been reported experimentally [26]. This system is an interesting target to be studied with our

methodologies because it displays a conformational change associated to electron and proton exchanges [27].

THEORY AND METHODOLOGIES

Constant pH and Redox Potential Molecular Dynamics (C(pH,*E*)MD)

A detailed description of our C(pH,*E*)MD approach has been presented in a previous publication [14]. Earlier attempts to describe pH and redox potential effects can be found in the literature [19, 29–35]. Our methodology makes use of Monte Carlo transitions between discrete protonation or redox states. Each state is represented by different atomic charge distributions. In our approach we perform molecular dynamics for a predetermined number of steps, then the simulation is halted, and afterward we attempt protonation and/or redox state changes. In our methodology the relation between proton and electron exchanges arises naturally because a successful protonation state change attempt will influence the next redox state change attempts in redox-active residues nearby and, likewise, a successful redox state change attempt will affect the following protonation state change attempts in neighboring pH-active residues.

The main equation used for protonation state change attempts is [14]:

$$\Delta G_{protonation} = k_b T(\ln 10)\left(pH - pK_{a,ref}\right) + \Delta G_{elec} - \Delta G_{elec,ref} \qquad (1)$$

And for redox state change attempts the equivalent equation is [14]:

$$\Delta G_{reduction} = vF\left(E - E^o_{ref}\right) + \Delta G_{elec} - \Delta G_{elec,ref} \qquad (2)$$

ΔG_{elec} is computed on each state change attempt and $\Delta G_{elec,ref}$ is a precomputed quantity used to reproduce $pK_{a,ref}$ or E^o_{ref} in the simulations for the reference compound. More information about these equations and their derivations can be found in reference [14].

Dealing with pH- and redox-active titratable residues

Our C(pH,*E*)MD methodology as reported in reference [14] and as originally released in AMBER 18 [36] could only deal with residues that are either pH- or redox-active, but not both at the same time. We will now report a recent improvement that we have applied in our C(pH,*E*)MD methodology that allows for a titratable residue to be simultaneously pH- and redox-active. This opens the way, for example, for the theoretical study of PCET processes in AMBER. In order to understand how this improvement works, we start considering the following thermodynamic cycle and employing equations 1 or 2 on each step:

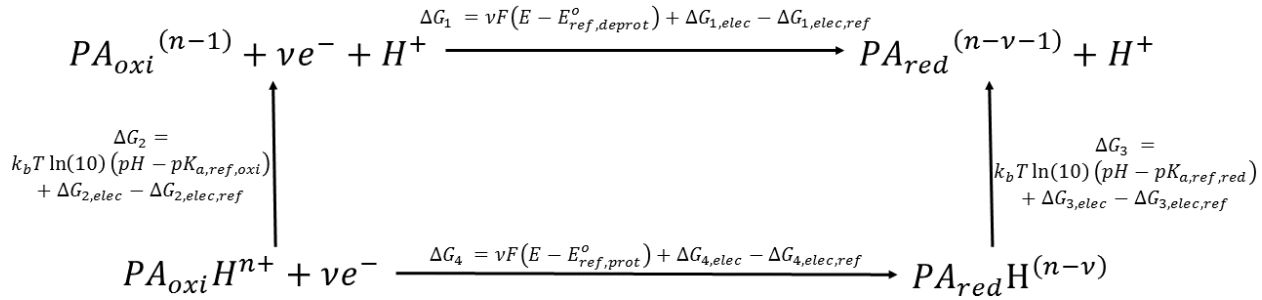

Figure 2. Thermodynamic cycle used in our most recent C(pH,*E*)MD methodology that allows residues that are both pH- and redox-active. *red* refers to the reduced state and *oxi* to the oxidized state.

Here, *A* is the pH- and redox-active residue and *P* is the rest of the protein. Because *A* can exchange both protons and electrons, there are two reference pK_a values and two reference E^o values associated with the same residue *A*, as can be seen from Figure 2. As a consequence of that, there are four $\Delta G_{elec,ref}$ values to be computed for this same residue. In our implementation it is possible to attempt switching from any of the four states shown in Figure 2 to any other, including switches between the diagonal states in the figure, which would represent a simultaneous exchange of

protons and electrons. This feature might be important for the description of some PCET processes where the simultaneous transfer of both a proton and an electron is important.

From Figure 2 we can discuss some more details about our most recent C(pH,*E*)MD methodology. We can see that when the current and the proposed states differ only by the number of protons, then equation 1 is used. When the states differ only by the number of electrons, then equation 2 is used. However, when the states differ by both the number of protons and electrons, then we must follow the thermodynamic cycle and ΔG is computed by combining equations 1 and 2.

Multidimensional Replica Exchange MD along the redox potential and pH dimensions

Detailed information about our multidimensional replica exchange MD along the redox potential and pH dimensions (*E*,pH-REMD) methodology can be found in our previous publication [15], where we show that *E*,pH-REMD has a better sampling convergence efficiency than either pH-REMD or *E*-REMD. In AMBER's multidimensional REMD module we perform sequential one-dimensional replica exchange attempts, thus only one type of information can be exchanged every time a replica exchange is attempted. Figure 3 illustrates how a *E*,pH-REMD simulation work in AMBER's multidimensional REMD module [15, 37]:

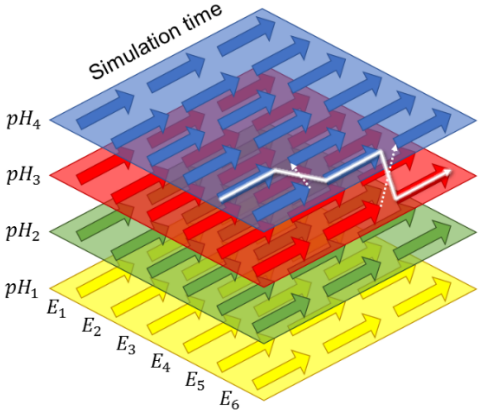

Figure 3. Schematic diagram of a *E*,pH-REMD simulation in AMBER. Each plane represents a different pH value, and the colored arrows are placed to visualize the different redox potential values. We see a possible path followed by a given replica in white.

As the figure shows, pH and redox potential exchange attempts are performed at different steps. It is also shown that redox potential exchange attempts are only done between replicas with the same pH and that pH exchange attempts are only done between replicas with the same redox potential. In the *E*,pH-REMD approach only the target *E* and pH values are exchanged, whereas the redox and protonation states exchanges are still dealt within each single replica through the C(pH,*E*)MD approach.

The Metropolis Monte Carlo probability of exchanging any two replicas in the pH dimension is [15, 38, 39]:

$$P_{i\to j} = \min\left\{1, e^{\ln(10)(pH_i - pH_j)\left(N_i^{H^+} - N_j^{H^+}\right)}\right\} \quad (3)$$

where pH_i and $N_i^{H^+}$ are, respectively, the pH and the total number of protons in replica i. And the probability of exchanging any two replicas in the redox potential dimension is [15]:

$$P_{i\to j} = \min\left\{1, e^{\frac{F}{k_b T}(E_i - E_j)\left(N_i^{e^-} - N_j^{e^-}\right)}\right\} \quad (4)$$

where F is the Faraday constant, and E_i and $N_i^{e^-}$ are, respectively, the redox potential and the total number of electrons in replica i.

Describing theoretically the pH-dependence of E^o and the redox potential-dependence of pK_a values

If we consider a system that contains only a single pH- and redox-active residue, it is possible to show by making thermodynamic considerations that [14]:

$$vF\left(E^o_{prot} - E^o_{deprot}\right) = k_b T \ln(10)\left(pK_{a,red} - pK_{a,oxi}\right) \quad (5)$$

$$E^o = E^o_{prot} + \frac{k_b T}{vF}\ln\left(\frac{10^{-pK_{a,red}} + 10^{-pH}}{10^{-pK_{a,oxi}} + 10^{-pH}}\right) \quad (6)$$

$$pK_a = pK_{a,red} + \log\left(\frac{e^{-vFE^o_{prot}/k_bT} + e^{-FE/k_bT}}{e^{-vFE^o_{deprot}/k_bT} + e^{-FE/k_bT}}\right) \quad (7)$$

where v is the number of electrons involved in the reduction, E^o_{prot} and E^o_{deprot} are the standard redox potential when, respectively, fully protonated or fully deprotonated, and $pK_{a,red}$ and $pK_{a,oxi}$ are the pK_a values when, respectively, fully reduced or fully oxidized. The detailed derivations for all equations are presented at the Supporting Information of a previous publication [14].

CALCULATION DETAILS

We performed all the simulations in this work using an in-house modified version of AMBER 18 [36]. This version contains the improvements we describe in this work that allow a given residue to be simultaneously pH- and redox-active. This improved C(pH,*E*)MD methodology will be part of the future AMBER 20 release.

Parametrization of the pH- and Redox-active Tyrosine residue (TYX)

Following the discussions made from Figure 2, we now present in Figure 4 the E^o and pK_a values associated with the tyrosine residue. These values are reported in references [40, 41]. We name the pH- and redox-active tyrosine residue as TYX in order to distinguish it from the regular TYR residue in AMBER.

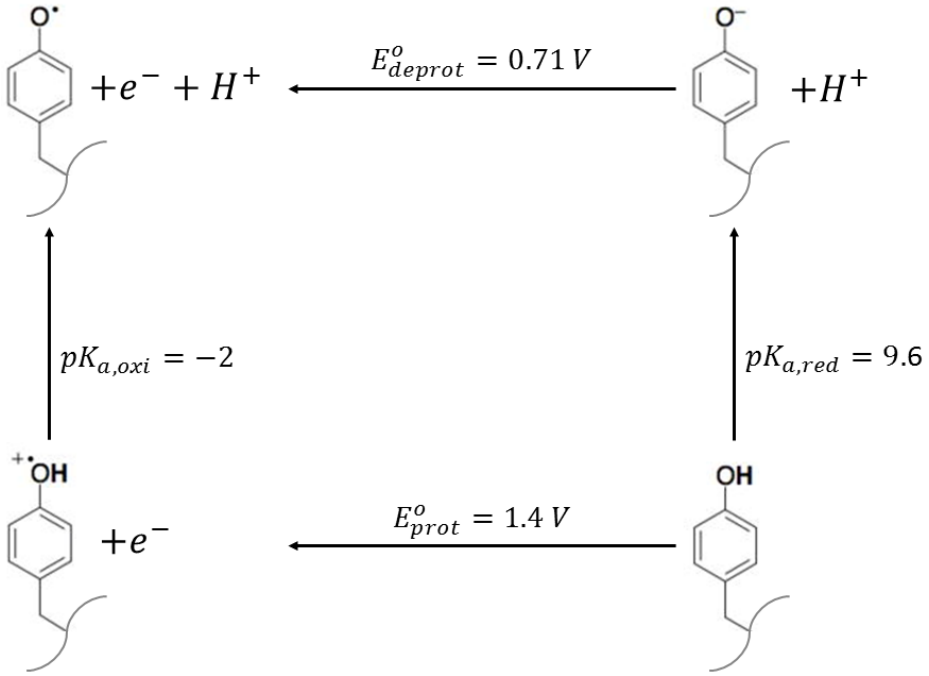

Figure 4. E^o and pK_a values associated with the tyrosine residue. These values are based on references [40, 41].

Reference [40] reports a value of 10.1 for p$K_{a,red}$, however, when using this value and all the others shown in Figure 4 the equation 5 is not satisfied. In order to satisfy this equation exactly we have modified p$K_{a,red}$ to be 9.6, which is in very close agreement with other pK_a measures for tyrosine [41]. The values shown in Figure 4 are in close agreement with the values obtained experimentally for acetyl-L-tyrosinamide [24].

We performed one individual CpHMD or C*E*MD simulation of the tyrosine dipeptide capped (ACE-TYX-NME) for each step shown in Figure 4. In each case we fitted the $\Delta G_{elec,ref}$ term from either equation 1 or equation 2 in order for the CpHMD or C*E*MD simulation to predict the respective E^o or pK_a value shown in Figure 4. Therefore, we ended up with four different $\Delta G_{elec,ref}$ values, where each one is associated to a given step in Figure 4 and its respective E^o or pK_a value.

The charges of each tyrosine state shown in Figure 4 were initially obtained with the Merz-Singh-Kollman scheme [42, 43] at the level HF/6-31G* using the Gaussian 09 program [44] and considering ACE-TYX-NME to be in vacuum. The results of this calculation were submitted to AMBER's Antechamber tool [36] for posterior RESP fittings. The RESP charges were adapted afterward so that the backbone charges are equal across all states. This is a common procedure in AMBER that aims at avoiding a residue sequence dependence when obtaining reference energies (see the AMBER18 manual [36]). Therefore, only the side chain has its charge distribution modified during a state change. Also, the total charge of each state remains an integer, in agreement with Figure 4. In Table 1 we present the charge distribution of each state, as well as the relative number of electrons and protons, and the total charge per state.

Table 1. Charge distributions for each of the four states of the pH- and redox-active tyrosine residue (TYX). The relative number of electrons and protons, and the total charge of each state is also shown.

Atom Name	State 1 (OH)	State 2 (•OH)	State 3 (O^-)	State 4 (•O)
	Atomic charges (a.u.)			
N	-0.4157	-0.4157	-0.4157	-0.4157
H	0.2719	0.2719	0.2719	0.2719
CA	-0.0014	-0.0014	-0.0014	-0.0014
HA	0.0876	0.0876	0.0876	0.0876
CB	-0.1163	-0.2911	-0.0500	-0.1941
HB2	0.0548	0.1123	0.0300	0.0931
HB3	0.0548	0.1123	0.0300	0.0931
CG	-0.0139	0.4479	-0.1786	0.0538
CD1	-0.1142	-0.1923	-0.1545	-0.1216
HD1	0.1615	0.2177	0.1418	0.1629
CE1	-0.3410	-0.1736	-0.4793	-0.3224
HE1	0.1911	0.2109	0.1318	0.1643
CZ	0.4198	0.5560	0.7197	0.6679
OH	-0.5278	-0.4376	-0.8026	-0.4519
HH	0.3621	0.4031	0.0000	0.0000
CE2	-0.3410	-0.1736	-0.4793	-0.3224

HE2	0.1911	0.2109	0.1318	0.1643
CD2	-0.1142	-0.1923	-0.1545	-0.1216
HD2	0.1615	0.2177	0.1418	0.1629
C	0.5973	0.5973	0.5973	0.5973
O	-0.5679	-0.5679	-0.5679	-0.5679
Relative number of electrons	1	1	0	0
Relative number of protons	1	0	1	0
Total charge	0	-1	1	0

The charges for states 1 and 3 (OH and O^-) are in good agreement with the charges used in AMBER for the protonated and deprotonated states, respectively, of the regular TYR residue in CpHMD simulations [45]. A good agreement is also observed against CHARMM charges for the same two states [27].

Parametrization of Diheme Cytochrome *c* (DHC) from *Rhodobacter sphaeroides*

There are two structures reported for the diheme cytochrome *c* (DHC) from *Rhodobacter sphaeroides* [16]: a native structure whose PDB entry is 2FWT, and a recombinant structure whose PDB entry is 2FW5. Neither of these structures contain the last residue and the first 11 residues of the amino acid sequence [16]. Also, the native structure lacks the residues 80 to 91. When comparing 2FWT to 2FW5, only minor changes are observed in the structures [16]. In this work we will consider only the recombinant structure because this is the more complete structure. The missing residues have been added after being modeled with EasyModeller 4.0 [46].

Figure 1A shows the structure of DHC. DHC contains two heme groups. Each heme is similar to the heme group in NAcMP8 (previously studied by us in references [14, 15]): each heme contains two histidines as axial ligands. Because of that, we have split each heme group as a HEH residue and two PRN residues and used the same E^o_{ref} and $\Delta G_{elec,ref}$ values (see equation 2) that we used for

NAcMP8 [14, 15]. HEH is the redox-active residue that has the porphyrin ring of the heme and parts of the histidines and cysteines (see reference [14] for more details) shown in green in Figure 1, and PRN is the pH-active propionate shown in red in Figure 1. DHC also contains two tyrosine residues, including one that is in between the two heme groups. We considered the two tyrosine residues as TYX (see previous subsection).

Following what was done in references [16, 47], we label each heme based on its proximity to either the N-terminus or C-terminus residues. The residues highlighted in Figure 1A are not the only titratable residues we considered. This system also contains seven aspartates, eight glutamates and three lysines that are pH-active titratable residues.

Parametrization of *Desulfovibrio vulgaris Hildenborough* Cytochrome c_3 (DvHc_3)

The structure of *Desulfovibrio vulgaris Hildenborough* cytochrome c_3 (DvHc_3) is reported in the literature with the PDB id 2CTH [48]. This protein contains four hemes with bis-histidinyl axial coordination, similarly to NAcMP8. The hemes were split into HEH and PRN residues. DvHc_3 contains three tyrosine residues that are near to three of the hemes. These tyrosines were assigned as TYX residues. In Figure 1B we present the structure of DvHc_3. We have used the same labeling for each heme, from I to IV, used by Machuqueiro and Baptista [19]. DvHc_3 contains additional pH-active titratable residues not shown in the figure: nine aspartates, four glutamates, and twenty lysines.

Parametrization of α_3Y

The structure of α_3Y has been reported with the PDB id 2MI7 [49] and is shown on Figure 1C. In addition to the pH- and redox-active tyrosine residue TYX highlighted in the figure, this system contains seventeen glutamates and seventeen lysines pH-active titratable residues that are not shown in the figure.

Parametrization of Peptide A

The structure of peptide A is shown in Figure 1D and has been provided to us by Bridgette Barry *et. al.* [28]. This is an averaged and minimized structure obtained by NMR spectroscopy [27, 28]. The figure highlights all titratable residues in this peptide: the pH- and redox-active tyrosine residue TYX, one pH-active histidine, and two pH-active aspartates.

Explicit Solvent Calculations

The procedure described here applies to the simulations we performed for α_3Y, peptide A, DHC, and DvHc_3. Each system was solvated in a truncated octahedron box with TIP3P waters with a buffer of 12 Å. Sodium or chloride ions were randomly added to neutralize the total charge of the initial configuration.

We then performed a minimization restraining all the backbone atoms with a 10 kcal/mol·Å^2 constant. The minimization started with 1000 steepest descent steps followed by 4000 conjugate gradient steps. Afterward, a heating simulation has been performed. This simulation was performed for 6 ns at constant volume with a backbone constraint of 1 kcal/mol·Å^2 and a 5 ps^{-1} friction frequency on the Langevin thermostat. During the initial 5 ns the target temperature of the simulation was linearly modified from 10 to 300 K.

An 8 ns equilibration simulation was then performed with a 1 bar constant pressure and a 300 K constant temperature. A backbone constraint of 0.1 kcal/mol·Å^2, a friction coefficient of 2 ps^{-1} for the Langevin thermostat and 5 ps relaxation time for the Berendsen barostat. The structure obtained from this simulation was submitted to *E*,pH-REMD simulations with no backbone restraints. The length of these simulations was different for α_3Y, peptide A, DHC, and DvHc_3. The length for each case is reported in the results section below. In all cases, part of the beginning of each *E*,pH-REMD simulation was considered to be equilibration and only the remaining steps were considered for production purposes and analyses.

In the production *E*,pH-REMD simulations the protonation and redox state change attempts were performed every 200 fs. Any successful attempt was followed by a 200 fs long solvent relaxation. The Exchange Attempt Frequency (EAF) used in the *E*,pH-REMD was of 2.5 ps^{-1}, that is, one exchange attempt every 400 fs. We constrained all the hydrogen bonds using SHAKE [50, 51] in all our simulations and we used a time step of 2 fs. We have employed the particle-mesh Ewald method [52, 53] to account for the long-range electrostatic interactions using a van der Waals cutoff and a direct space of 8 Å. We used the GB model proposed by Onufriev *et al*. [54] (*igb*=2 in AMBER) as the implicit solvent model during redox and protonation state change attempts.

RESULTS AND DISCUSSIONS

Tyrosine dipeptide

We start evaluating the behavior of a capped tyrosine dipeptide in water. For this system only explicit solvent C(pH,*E*)MD simulations were performed. Given the size of this peptide, convergence can still be reached with somewhat short simulations, so we did not need to use REMD calculations for this system. In order to study the pH-dependence of the standard redox

potential values of the tyrosine dipeptide we have considered pH values from -4 to 12 with an interval of 1 pH unit between nearest values. For each pH value, 6 redox potential values were considered in intervals of 48 mV: 3 values below the E^o predicted by the theoretical model at that pH (equation 6), and 3 values above. Figure 5A shows the E^o values versus pH that we obtained from the simulations, the expected curve from the theoretical model (shown in equation 6) with the parameters from Figure 4, and the experimental values obtained from reference [26]. For the evaluation of the redox potential-dependence of the pK_a values we have considered redox potentials between 0.65 and 1.46 V in intervals of 51 mV. For each redox potential 6 pH values were taken into consideration. The pH values were chosen to cover a range in which the pK_a value predicted by the theoretical model (equation 7) at the desired redox potential is included: we considered 3 pH values above and 3 values below the pK_a value predicted by the model, with an interval of 0.8 between nearest pH values. Figure 5B shows the pK_a of tyrosine as a function of redox potential obtained from the simulations and also with the theoretical model using equation 7 and the parameters in Figure 4. All the production C(pH,*E*)MD simulations executed to construct both Figures 5A and 5B were 15 ns long.

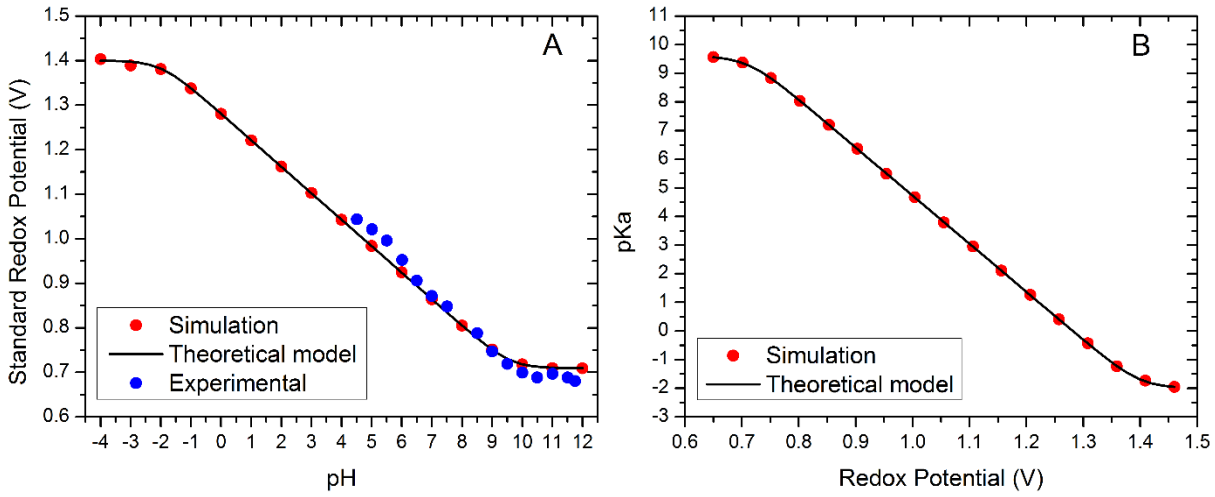

Figure 5. A) Standard redox potential as a function of pH, and B) pK_a as a function of redox potential for the pH- and redox-dependent tyrosine dipeptide capped. The points in red come from the simulations, the solid lines were obtained from the theoretical model, and the blue points are experimental results obtained from reference [26].

As described in the subsection "Parametrization of the pH- and redox-active tyrosine residue (TYX)", our reference energies were only computed at the limiting cases where only one proton or one electron are exchanged, even though, as can be seen in Figure 5, both the pH-dependence of E^o and the redox potential-dependence of pK_a are correctly described in our simulations. We observe a good agreement with experimental results and see that the theoretical model is naturally satisfied in our simulations. It is interesting to point out that the solid curves shown in Figure 5 are not fittings from the results of the simulations; they are solely the direct application of equations 6 or 7 with data from Figure 4.

α_3Y

In order to evaluate the pH- and redox potential-dependence of α_3Y we have performed three independent *E*,pH-REMD simulations. In each *E*,pH-REMD simulation we have considered 10 pH values from 8.9 to 13.4 in intervals of 0.5, and 10 redox potentials from 360 to 630 mV in intervals of 30 mV, giving a total of 100 replicas considered. In each *E*,pH-REMD simulation the first 4.2 ns were considered as equilibration and only the following 19 ns were considered for analyses. Figure S1 in the Supporting Information shows the cumulative average of E^o at pH 13.4 which indicates that the 19 ns considered for production are sufficient for convergence. In Figure 6 we show the standard redox potential of α_3Y as a function of pH and compare it with the E^o values of the tyrosine dipeptide capped shown in Figure 5. The figure shows the average values of E^o for the three independent simulations of α_3Y as points and the standard deviations as error bars.

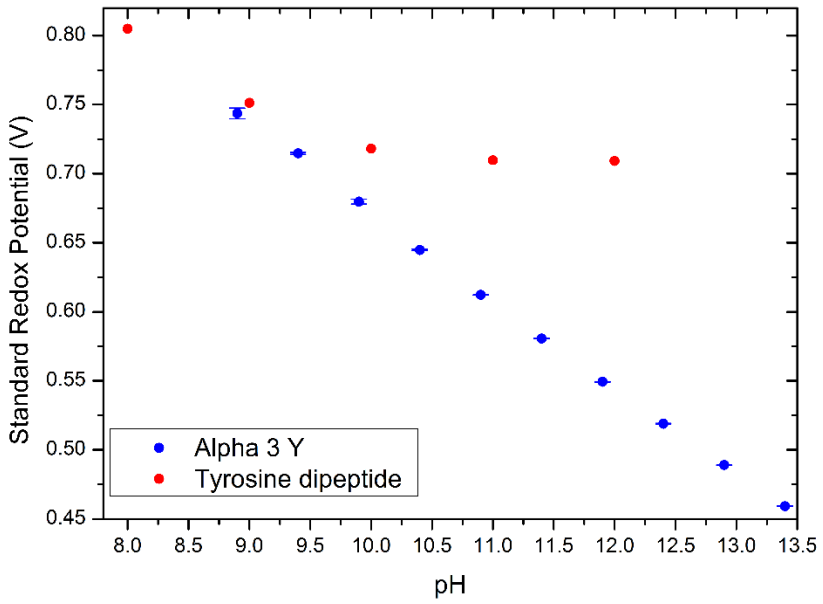

Figure 6. Standard redox potential as a function of pH for α_3Y and for capped tyrosine dipeptide. The points in blue for α_3Y represent the average values and the error bars represent the standard deviations from the three independent runs.

From Figure 6 we observe a good agreement between the three independent α_3Y simulations because of the small standard deviations. The larger error bar for the first point at pH 8.9 can be explained by our choice of redox potential values (360 to 630 mV) that does not cover well the E^o value of 744 mV at this pH.

Tommos *et al* [24] have shown experimentally that the E^o of α_3Y at pH 13.4 is shifted down by approximately 193 mV in comparison to the E^o of N-acetyl-L-tyrosinamide in this same pH. Our simulations predict a shift down of 251 mV at this pH in comparison to the tyrosine dipeptide capped, thus in good agreement with the experimental observation. The difference between our theoretical prediction and the experimental value is of 58 mV, which corresponds roughly to 1 pK_a unit at 300 K. This is usually taken as a small error in the type of calculations performed here.

Peptide A

For peptide A we have performed three independent *E*,pH-REMD simulations, each one 160 ns long. We have considered pH values from 5.0 to 11.5, in intervals of 0.5, and redox potential values from 640 mV to 970 mV, in intervals of 30 mV. This gives a total of 168 replicas considered. A previous publication has attempted to study a conformational change undergone by peptide A by performing simulations at constant protonation and redox states [27]. In our current work, we aim at evaluating the full pH- and redox potential-dependence of this conformational change.

In order to investigate the occurrence of conformational changes in our simulations for peptide A, in Figure 7A we show the average RMSD during the last 20 ns of one of the three independent *E*,pH-REMD simulations, for each target pH and redox potential values. The RMSD calculations were performed taking into consideration only backbone atoms. The reference structure for all the RMSD calculations is a folded structure taken from the trajectory at pH 5.0 and redox potential 0.64 V of the first independent *E*,pH-REMD simulation during the last 20 ns. This reference structure is shown in Figure S2 in the Supporting Information. In Figure 7B we show the average distance between the tyrosine and histidine during the last 20 ns for the same *E*,pH-REMD simulation. Only heavy atoms in the side chains were considered in order to compute the distance between the two residues. Figure 7C shows the average distance between the backbone atoms of the terminal residues, also during the last 20 ns of simulation. A version of Figure 7 for all three independent *E*,pH-REMD simulations is shown on Figure S3 at the Supporting Information.

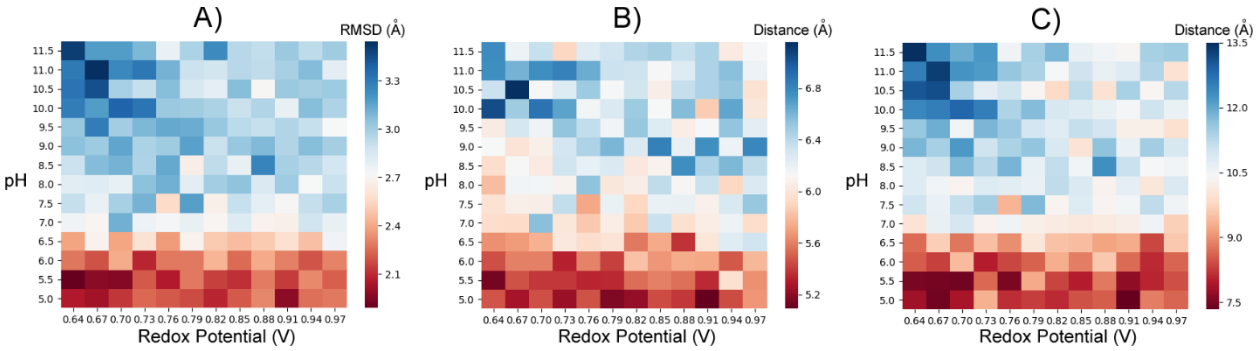

Figure 7. Evaluations of the last 20 ns of each of one of the three independent *E*,pH-REMD simulations performed for peptide A. Results are shown as a function of pH and redox potential. A) Average RMSD against a reference structure. B) Average distance between tyrosine and histidine. C) Average distance between the backbone atoms of the terminal residues.

Figure 7A shows that as the pH increases more distorted structures are obtained in comparison to our reference folded structure. The distortions become significantly more evident after pH 6.5. The redox potential also plays a role in the conformational change at higher pH values: for pH values of 10 or higher, we observe more distorted structures at lower redox potential values, however, this effect is not as significant as the pH effect. In Figure 7B and in Figure 7C we see that the conformational change investigated in Figure 7A correlates with an increase in distance between the tyrosine and the histidine, and with the terminal residues getting further apart. All the behaviors described here were reproduced in all three independent *E*,pH-REMD simulations (see Figure S3).

Figure 7A only shows us average RMSD values, but no information was given onto how much from the average value each structure deviates. In order to provide an insight on that effect, in Figure S4 in the Supporting Information we present the histogram of RMSD values from which the average values shown in Figure 7A were computed. Histograms are presented for pHs 5.0, 8.5 and 11.0 at the low and high redox potential limits. From Figure S4 we can clearly see that the probability of obtaining structures with higher RMSD values increases at larger pH values, at both redox potential limits, in agreement with the average behavior shown in Figure 7A.

In order to further investigate the conformational change undergone by peptide A, in Figure 8 we present the fraction of unfolded structures for each pH and redox potential during the last 20 ns of one of the three independent *E*,pH-REMD simulations. We arbitrarily considered a given structure to be unfolded if its RMSD against our reference folded structure is higher than 2.4 Å; we have observed the same trends shown in Figure 8 by using other choices of values around 2.4 Å. A version of Figure 8 for all three independent *E*,pH-REMD simulations is shown on Figure S5 at the Supporting Information.

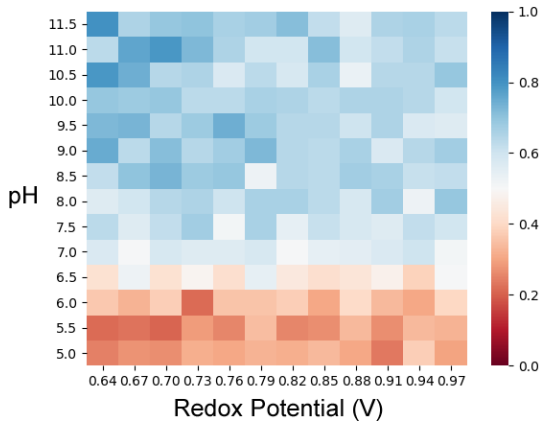

Figure 8. Fraction of unfolded structures as a function of pH and redox potential over the last 20 ns of each independent *E*,pH-REMD simulation performed for peptide A.

The analysis shown in Figure 8 indicates that unfolded structures are more likely at higher pH values, in agreement with Figure 7. Once at a high pH value, the figure shows unfolding is slightly more likely at low redox potential values.

The microscopic pK_a and E^o values of tyrosine and the microscopic pK_a of histidine should be different for folded and unfolded structures of peptide A. This is due to the significant change in the environment around these residues in the two cases. A previous study from our research group [55] indicates that fittings using equations 1 or 2 from our previous publication [15] should not be performed for predictions of pK_a or E^o values when such conformational change take place, but rather a complex model that takes the conformational change into consideration should have to be employed.

Another interesting point that is investigated from our simulations is what are the preferred states of the tyrosine and histidine during the simulations. In Figure 9 we show the percentage of each tyrosine and histidine state as a function of pH and redox potential. This figure was constructed with data from the last 20 ns of one of the independent *E*,pH-REMD simulations.

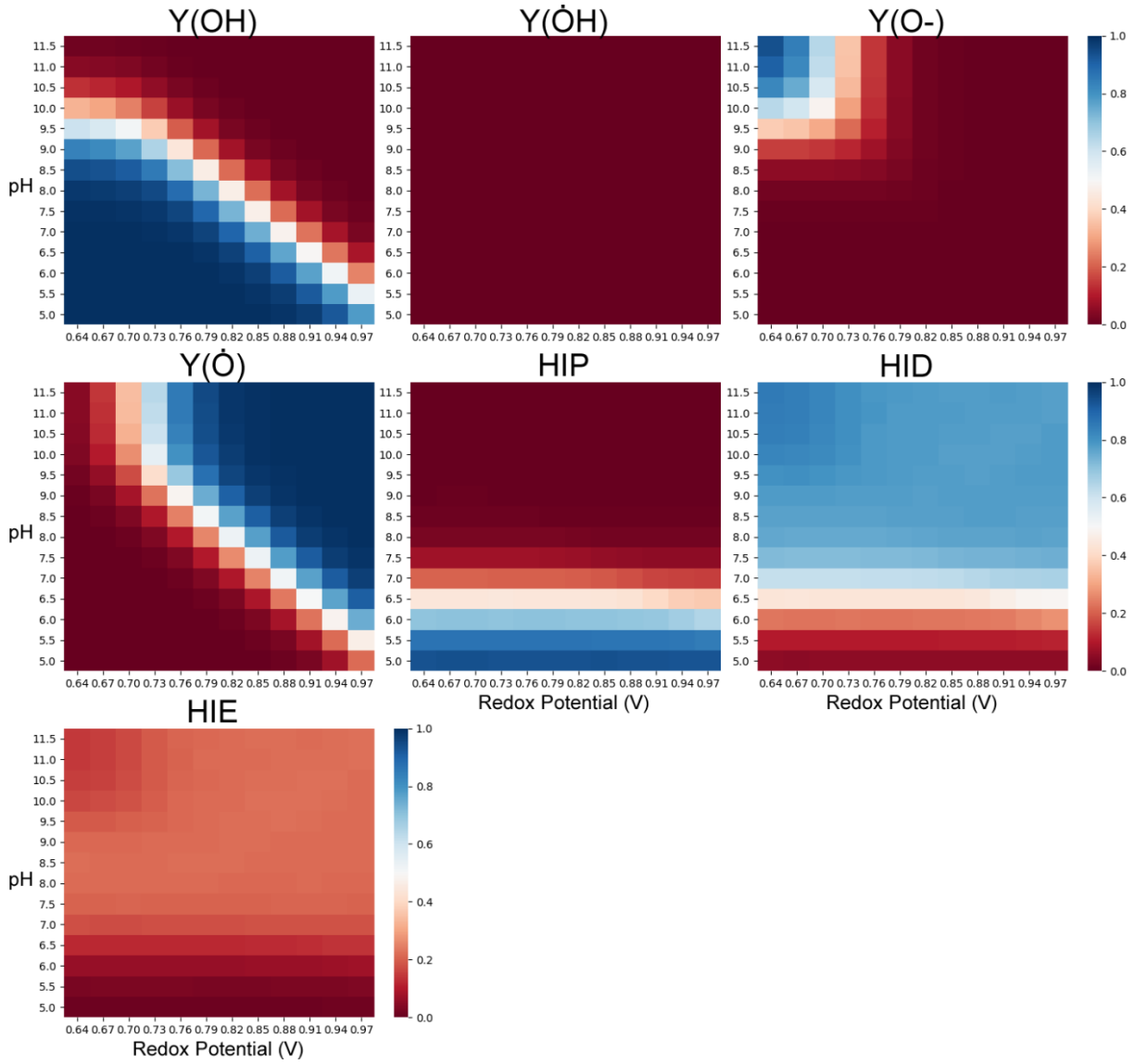

Figure 9. Fraction of each state of tyrosine (Y) and histidine (H) as a function of pH and redox potential over the last 20 ns of one of the three independent *E*,pH-REMD simulations performed for peptide A. The labeling for the four states of the tyrosine shown here is detailed in Figure 4.

As can be seen in Figure 9, the •OH state of the tyrosine is not visited at any value of redox potential and pH. From the figure we can see that at the regions with lower RMSD values in Figure 7A, corresponding to pH ≤ 6.5 and any redox potential, the state of the tyrosine is mostly OH and the state of histidine is mostly the doubly protonated state (HIP). At high pH values the histidine is

mostly in the singly protonated at the delta position state (HID), and the tyrosine is either in the •O state at the high redox potential values or in the O^- state at low redox potential values. It is interesting to notice that the tyrosine assuming a negative charge (O^- state) affects the tautomeric equilibrium between the HID and HIE states of the histidine.

When comparing Figure 9 with Figures 7A and 8 we can clearly see how the states of each residue correlate with the unfolding process: we observe that the deprotonation of the histidine is the most responsible for the appearance of more unfolded structures for pHs ≥ 6.5, and then when the tyrosine goes to its O^- state this slightly contributes to further unfold the structures.

Sibert *et al*. [26] have reported the experimental values of the standard redox potential of peptide A as a function of pH (see Figure 7 of reference [26]). The experimental results indicate two inflection points, observed around pH 7.0 and 8.6. The authors associate these inflection points with, respectively, the p$K_{a,red}$ and p$K_{a,oxi}$ values of histidine. The conclusions we reached from our *E*,pH-REMD simulations allows us to hypothesize that this behavior observed experimentally could be associated to the unfolding of peptide A, which according to our simulations is associated with the deprotonation of the histidine.

Hwang *et al*. [27] indicated that the tyrosine radical state (•O) is more likely to yield unfolded structures in peptide A than the states OH and O^- tyrosine states by doing a few regular MD simulations, that is, at constant protonation and redox states. In our work, with the use of an efficient replica exchange technique and by performing simulations that allow both protonation and redox states to be sampled, we have actually observed that the deprotonation of the histidine plays the most important role in the unfolding process. We have concluded that the influence of the tyrosine state is minor in the unfolding process. For the record, it is important to mention that

the tyrosine charges used by Hwang *et al*. [27] match closely with ours for the OH and O^- states, however, significant discrepancies are observed for the •O state.

Diheme Cytochrome *c* (DHC) from *Rhodobacter sphaeroides*

For DHC we have also performed three independent *E*,pH-REMD simulations. We have selected pH values from 3.5 to 10.0, with intervals of 0.5, and redox potential values from -450 to -180 mV in intervals of 30 mV, totalizing 140 replicas considered. Each *E*,pH-REMD simulation was 57 ns long, where the first 9.5 ns were considered as equilibration and only the next 47.5 ns were considered for production purposes. At Figure S6 in the Supporting Information we present the cumulative average of E^o of both DHC hemes at pH 7.5, which indicates that the 47.5 ns considered for production are enough for convergence. In Figure 10 we show the standard redox potential as a function of pH for both Nter and Cter hemes (Figure 1A also shows the labeling of each DHC heme). The figure shows the average values of E^o for the three independent *E*,pH-REMD simulations as points and the respective standard deviations as error bars.

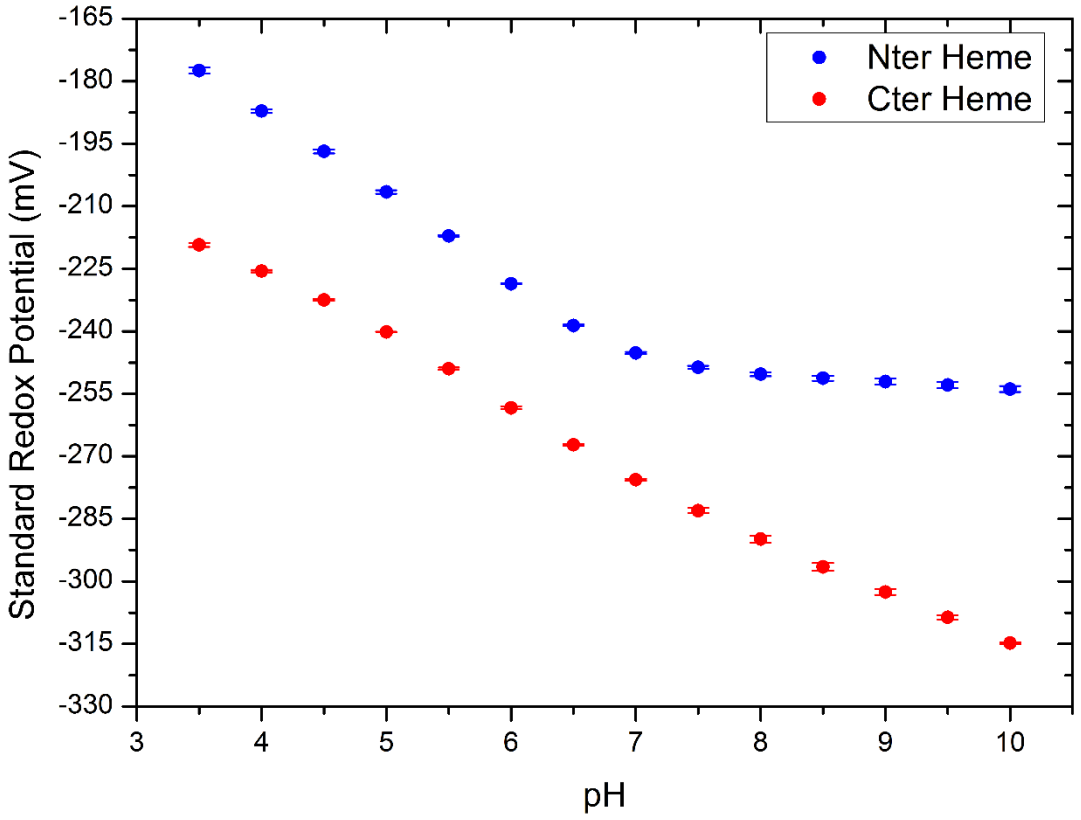

Figure 10. Standard redox potential as a function of pH for each heme of DHC. The points represent the average values and the error bars represent the standard deviations from the three independent *E*,pH-REMD runs.

Figure 10 shows that the E^o of the Cter heme is lower than the E^o of the Nter heme at all pH values. The figure also shows that above pH 7.0 the E^o of the Nter heme almost does not change whereas the E^o for the Cter heme continues to decrease. This behavior can be explained by the pK_a values of the titratable residues around each heme and will be discussed below. Table 2 shows the pK_a values of all pH-active residues at the low and high redox potential limits.

Table 2. pK_a values of all pH-active residues of DHC at the low (E = -450 mV) and high (E = -180 mV) redox potential limits. The table shows average and standard deviation values obtained from the three independent E,pH-REMD simulations. PRN 127 and 128 are the propionates from Cter heme, and PRN 130 and 131 are the propionates from Nter heme. pK_a values could not be fitted for the tyrosines (TYX) because the fraction of protonated species was essentially 100% in all pH values.

Residue	pK_a (low E)	pK_a (high E)	ΔpK_a
ASP 6	3.28±0.08	3.19±0.05	0.09±0.13
GLU 12	6.27±0.01	6.14±0.01	0.13±0.02
TYX 20	>10.0	>10.0	N/A
ASP 36	5.70±0.01	5.57±0.01	0.13±0.02
GLU 43	6.14±0.00	5.90±0.01	0.24±0.01
ASP 44	3.68±0.01	3.56±0.00	0.12±0.01
ASP 48	3.18±0.00	3.16±0.01	0.02±0.01
GLU 49	3.93±0.02	3.92±0.01	0.01±0.03
GLU 56	4.72±0.01	4.62±0.01	0.10±0.02
TYX 58	>10.0	>10.0	N/A
ASP 65	4.01±0.02	3.98±0.01	0.04±0.03
ASP 80	4.73±0.02	4.69±0.01	0.04±0.03
GLU 87	4.21±0.04	4.16±0.03	0.06±0.07
LYS 92	10.06±0.01	9.93±0.00	0.13±0.01
LYS 94	9.45±0.01	9.35±0.01	0.10±0.02
ASP 97	3.61±0.02	3.54±0.04	0.07±0.06
GLU 98	5.73±0.02	5.36±0.02	0.37±0.03
GLU 105	4.20±0.01	4.13±0.01	0.06±0.02
LYS 106	10.03±0.01	9.91±0.02	0.13±0.03
GLU 121	4.95±0.03	4.85±0.03	0.10±0.06
PRN 127	7.21±0.03	6.94±0.04	0.28±0.07
PRN 128	8.30±0.03	7.83±0.06	0.46±0.08
PRN 130	2.94±0.08	2.74±0.10	0.20±0.18
PRN 131	5.74±0.08	5.44±0.07	0.30±0.15

As expected, all the pK_a values at low E shown in Table 2 are larger than the pK_a values at high E. When the hemes are in their reduced state, they are more negatively charged, and this makes it easier for a protonation to take place in a given pH-active titratable residue nearby, increasing then the pK_a of this residue. Looking at the ΔpK_a values in the table we observe that some residues

almost didn't change their pK_a values, which means that they are not interacting significantly with the hemes and that a change in redox state of the hemes does not affect their pK_a values. Equation 5 shows us that a ΔpK_a value is related to how much a given residue contributes to the pH-dependence of the redox-active residues nearby. From Table 2 we observe that the heme propionates and the residues GLU 98 and GLU 43 are the ones that contribute the most to the pH-dependence of the E^o values of the hemes. This can be justified by the fact that GLU 98 is very near to Cter heme and its PRN 128, and because GLU 43 is also very near to Nter heme.

Table 2 shows that the pK_a values of the two hemes are clearly different: the pK_a values of the propionates that belong to the Cter heme (PRN 127 and 128) are higher than the pK_a values of the propionates that belong to the Nter heme (PRN 130 and 130). This can be explained by the fact that the propionates of Cter heme are more solvent exposed than the propionates of Nter heme. The pK_a values of the Cter heme's propionates are close to the pK_a values obtained for NAcMP8 (~6.7) reported in our previous publication [14]. The propionates of NAcMP8 are completely solvent exposed. The pK_a values of the propionates in NAcMP8 differ only by approximately 0.15 pH unit. As one can observe, this difference is significantly larger for the Cter heme (~ 1.0 pH unit) and specially for Nter heme (~ 2.7 pH units). This indicates that the environment is different around each propionate. The pK_a shifts we observe in comparison to the values in NAcMP8 allows us to guess the charge environment surrounding each propionate (a negative environment must shift the pK_a up and a positive environment should shift it down). An analysis of the trajectories in our simulations can identify which residues are influencing the pK_a shifts in each propionate.

The fact that the E^o of the Nter heme is almost constant after pH 7.0 means that no pH-active residues near the Nter heme are changing their protonation state anymore. However, the E^o of Cter heme keeps decreasing after pH 7.0. This happens because the Cter heme's propionates have pK_a

values around 7.0 or 8.0 and because all the lysines (that have pK_a values around 9.5 or 10.0) are near the Cter heme.

Gibson *et al*. [16] reported the experimental standard redox potential values of DHC to be -240 and -310 mV *vs*. NHE at pH 7.5. In our simulations we obtained a E^o of -248.6±0.3 mV for Nter heme and of -283.0±0.6 mV for Cter heme. Even though it is clear that in our simulations the E^o of Cter heme is always smaller than E^o of Nter heme, it is important to emphasize that the accuracy of the ΔE^o between the hemes is pH-dependent. For this reason the accuracy of ΔE^o also depends on the accuracy of the pK_a values we predicted. Please notice that our ΔE^o at, for example, pH 9.5 agrees much better with the experimental ΔE^o. Previous CpHMD publications using AMBER [45] have already shown to under or overestimate pK_a values by 1 unit or more, and such inaccuracy would certainly affect our ΔE^o predictions.

Our *E*,pH-REMD simulations clearly assign the experimental E^o values of -240 and -310 mV respectively to the Nter heme and the Cter heme. This conclusion is in disagreement with a previous theoretical study performed by Daidone *et al*. [47] in which the opposite assignment was concluded. Regardless of any limitations that the theoretical approach used by Daidone *et al*. might have, it is important to notice that in their simulations the protonation state of all pH-active residues was kept constant, however, as our *E*,pH-REMD simulations show, many pH-active residues titratable at pH 7.5, including the heme propionates, and this effect is certainly important for the proper description of the E^o values.

Desulfovibrio vulgaris Hildenborough **Cytochrome c_3 (DvHc_3)**

Three independent *E*,pH-REMD simulations were performed for DvHc_3. We considered pH values from 3.5 to 10.0 with an interval of 0.5, and redox potential values from -450 to -180 mV with an

interval of 30 mV, which gives a total of 140 replicas. Each *E*,pH-REMD simulation was 64.5 ns long. The initial 10 ns were considered as equilibration and only the following 60 ns were considered for analyses. In Figure S7 at the Supporting Information we show the cumulative average of E^o at pH 6.5 for each heme. This figure indicates that the length of our production *E*,pH-REMD simulations is enough for E^o convergence. In Figure 11 we show the E^o of all four hemes of DvHc_3 as a function of pH (please see Figure 1A to refer to the labeling of each heme). Figure 11 shows the average values of E^o for the three independent *E*,pH-REMD simulations as points and the respective standard deviations as error bars.

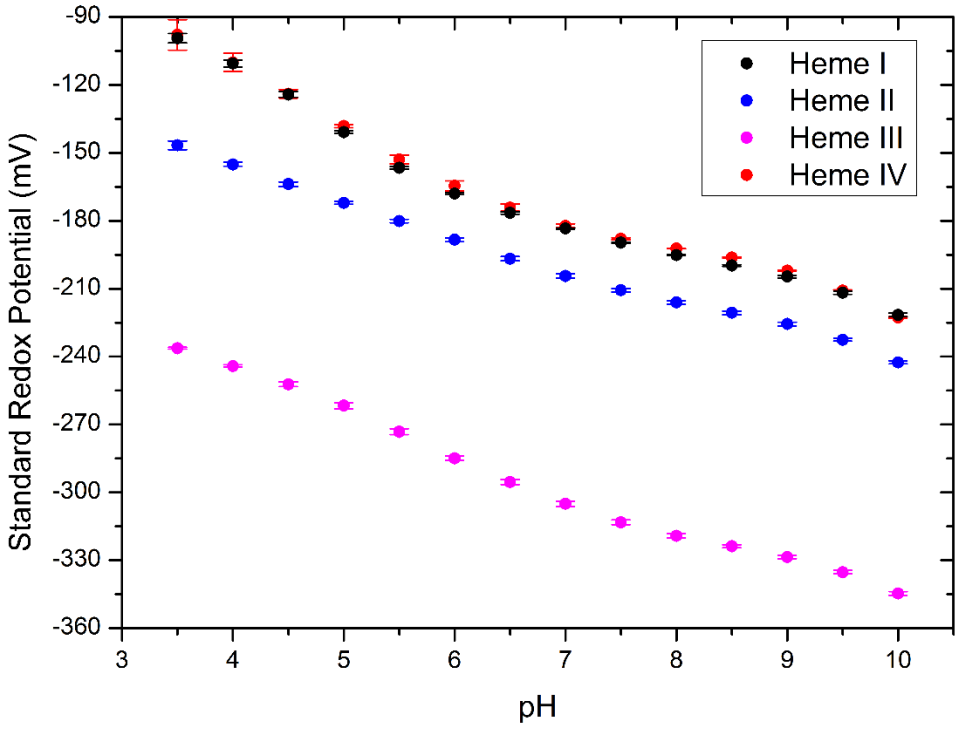

Figure 11. Standard redox potential as a function of pH for each heme of DvHc_3. The points represent the average values and the error bars represent the standard deviations from the three independent *E*,pH-REMD runs.

Figure 11 shows that the E^o of heme III is smaller than the E^o of heme II, that the E^o of heme II is smaller than the E^o of heme I, and that the E^o of heme I is slightly below or essentially equal to the E^o of heme IV. This is true for all pH values evaluated.

In Table 3 below we show the pK_a values of all pH-active residues of DvHc_3 at the low and high redox potential limits considered.

Table 3. pK_a values of all pH-active residues of DvHc_3 at the low (E = -450 mV) and high (E = -180 mV) redox potential limits. The table shows average and standard deviations values obtained from the three independent *E*,pH-REMD simulations. PRN 109 and 110 are the propionates from heme IV, PRN 112 and 113 are from heme II, PRN 115 and 116 are from heme I, and PRN 118 and 119 are from heme III. pK_a values could not be fitted for the tyrosines (TYX) because the fraction of protonated species was essentially 100% in all pH values.

Residue	pK_a (low E)	pK_a (high E)	ΔpK_a
LYS 3	10.27±0.01	10.12±0.01	0.14±0.02
ASP 7	3.61±0.03	3.52±0.03	0.09±0.06
LYS 10	10.38±0.01	10.21±0.01	0.17±0.02
GLU 12	4.20±0.01	4.17±0.01	0.03±0.03
LYS 15	10.64±0.01	10.49±0.02	0.15±0.04
LYS 26	10.34±0.03	10.10±0.01	0.24±0.04
LYS 29	10.18±0.01	10.03±0.01	0.15±0.02
ASP 32	4.71±0.09	4.61±0.06	0.10±0.15
LYS 40	10.16±0.03	10.10±0.03	0.06±0.06
GLU 41	5.45±0.05	5.32±0.05	0.13±0.09
ASP 42	2.63±0.07	2.63±0.09	0.00±0.16
TYX 43	>10.0	>10.0	N/A
LYS 45	9.97±0.01	10.19±0.03	-0.23±0.04
ASP 53	3.55±0.04	3.50±0.05	0.05±0.10
ASP 56	2.65±0.06	2.65±0.04	0.00±0.10

LYS 57	10.20±0.01	9.97±0.01	0.23±0.02
LYS 58	10.67±0.01	10.61±0.02	0.06±0.03
ASP 59	2.80±0.09	2.80±0.10	0.00±0.19
LYS 60	10.53±0.01	10.41±0.03	0.12±0.04
LYS 63	10.41±0.01	10.14±0.01	0.27±0.03
TYX 65	>10.0	>10.0	N/A
TYX 66	>10.0	>10.0	N/A
HIS 67	4.61±0.05	4.51±0.09	0.10±0.13
ASP 71	3.05±0.03	2.96±0.04	0.10±0.07
LYS 72	9.95±0.01	9.78±0.01	0.17±0.01
LYS 75	10.78±0.01	10.56±0.01	0.22±0.03
LYS 77	10.98±0.02	10.57±0.02	0.41±0.04
GLU 85	4.79±0.04	4.61±0.02	0.19±0.06
ASP 90	3.20±0.01	3.14±0.00	0.05±0.02
LYS 93	10.31±0.04	10.15±0.04	0.15±0.08
LYS 94	9.98±0.01	9.91±0.01	0.07±0.01
LYS 95	9.70±0.01	9.57±0.01	0.13±0.02
ASP 96	1.92±0.21	1.49±0.21	0.43±0.42
LYS 101	10.18±0.02	9.97±0.01	0.21±0.03
LYS 102	10.70±0.03	10.61±0.03	0.09±0.06
LYS 104	11.48±0.07	11.10±0.04	0.38±0.11
GLU 107	4.97±0.01	4.88±0.01	0.09±0.01
PRN 109	5.99±0.02	5.76±0.02	0.23±0.04
PRN 110	6.33±0.15	5.81±0.16	0.51±0.31
PRN 112	4.75±0.02	4.63±0.04	0.13±0.06
PRN 113	6.04±0.03	5.91±0.04	0.12±0.07
PRN 115	6.15±0.02	5.99±0.02	0.16±0.04
PRN 116	8.44±0.05	7.60±0.05	0.85±0.09
PRN 118	6.38±0.04	5.84±0.04	0.54±0.08
PRN 119	6.47±0.03	6.21±0.04	0.25±0.06

Similar to what was shown in Table 2 for DHC, in Table 3 we also see larger pK_a values at low E in comparison to the pK_a values at high E, where the only exception is LYS 45. This lysine is located at a flexible loop of DvHc_3, and, as we discussed in the results for peptide A, when conformational changes take place we would need to devise a new model that can account for that. Therefore, equation 7, that predicts larger pK_a value at higher values of redox potential, cannot be applied when conformational changes take place.

Table 3 shows that the propionates have different pK_a values for each heme. As can be seen from Figure 1B, heme III is the heme with its propionates most solvent exposed. Both propionates of heme III are essentially solvent exposed, and this is why their pK_a values are close to the pK_a values of the propionates of NAcMP8 (~6.7) presented in our previous publication [14].

By feeding experimental data into a theoretical model that assumes a single protonation site to describe the whole pH-dependence of the system, Turner *et al*. [18] presented values for the E^o of each heme of DvHc_3 as a function of pH (see Figure 6 in reference [18]). At pH 6.5 these values are -343, -302, -298 and -249 mV *vs*. NHE [18], and are assigned to, respectively, hemes III, II, I and IV [17–19]. At pH 6.5 we obtain the following E^o values from our *E*,pH-REMD simulations: -295.5±1.1 mV for heme III, -196.8±0.9 mV for heme II, -176.5±0.5 mV for heme I, and -174.2±1.6 mV for heme IV. This ordering that we obtained theoretically for which heme gets oxidized first (from low *E* to high *E*) matches exactly with the experimental ordering. Also, this ordering does not match with any of orderings computationally predicted by Machuqueiro *et al*. [19, 23] for DvHc_3 and by Ullmann [29] for a similar protein (cytochrome c_3 from *Desulfovibrio vulgaris Miyazaki*).

When we compare our E^o predictions (-296, -197, -177, and -174 mV) with the experimental ones (-343, -302, -298 and -249 mV) we do not observe a good match. When we compare our Figure 11 with Figure 6 from Turner *et al*. [18], we can clearly see that considering more pH-active residues does play an important role in describing the pH-dependence of DvHc_3, which may affect the E^o values of the hemes and consequently the assignment of the ordering in which the hemes get oxidized. One of the experimental data from which Turner *et al*. [18] fitted their theoretical model are measures of the total fraction of reduced species of DvHc_3 as a function of redox potential for two different pH values. Even though their theoretical model uniquely assigns E^o values, it is possible to devise a different set of E^o values for the four hemes in such a way that the experimental

fraction of reduced species versus redox potential would also be reproduced. Just as an example of this [29], different publications have reached different assignments for cytochrome c_3 from *Desulfovibrio vulgaris Miyazaki* [56–58], a system similar to DvHc_3.

Even though having in mind the limitations pointed out in the previous paragraph about the theoretical model from which the experimental E^o values were extracted, it is important to evaluate how far our E^o predictions are from these experimental values. Our predictions overestimate the E^o by 47 mV for heme III, 105 mV for heme II, 121 mV for heme I, and 75 mV for heme IV. These differences correspond, respectively, to 0.8, 1.8, 2.0, and 1.3 pH units. Errors of this order have been observed before in previous CpHMD publications [45].

CONCLUSIONS

In this work, we employed an improvement we performed in our C(pH,*E*)MD approach that allows a given residue to be simultaneously pH- and redox-active. Our C(pH,*E*)MD approach is fully force field-based and can be performed with AMBER's GPU-accelerated engine. We have studied five different systems: capped tyrosine dipeptide, α_3Y, peptide A, DHC, and DvHc_3. These are systems of relevant interest in the literature for which experimental results are available. For the tyrosine dipeptide, even though our reference energies were computed only for the cases in which only one proton or only one electron is exchanged, we were able to correctly predict the full pH-dependence of E^o and the full redox potential-dependence of pK_a. This description involves coupled proton and electron exchanges. The results from our simulations are in good agreement with the expected values from a theoretical model and with experimental results. For α_3Y our simulations correctly predict, in agreement with experimental results, a shift down of E^o in comparison to the tyrosine dipeptide capped at pH 13.4. For peptide A we have described the full pH- and redox potential-dependence of its conformational change, which provides enhanced information in comparison to a previous publication that studied this conformational change using simulations with constant protonation and redox states. From our *E*,pH-REMD simulations for peptide A we conclude that the deprotonation of the histidine (from doubly to singly protonated) plays an important role in the appearance of conformational changes. Once this deprotonation happens (for pH values 6.5 and higher), slightly more distorted conformations are observed when the tyrosine is in the O^- state (observed at low redox potential values) in comparison to when it is in the •O state (observed at high redox potential values). We conclude then that the histidine is the most responsible residue for the unfolding of peptide A, and the tyrosine only plays a minor role in this process. For DHC and DvHc_3 we have been able to provide insights into the ordering in

which the hemes are reduced by predicting one E^o value for each heme. The ordering that we have obtained theoretically is in good agreement with the experimental ordering. Also, we see that our methodology is capable of pointing out which pH-active residues interact more strongly with the redox-active residues, therefore, showing which residues are the most responsible for the pH-dependence of the E^o values. All the findings reported in this work contribute to further validate our methodology for the study of coupled redox and protonation processes.

AUTHOR INFORMATION

Corresponding Author

* E-mail: roitberg@ufl.edu

Notes

The authors declare no competing financial interest.

Supporting Information

See Supporting Information to access additional figures. These figures include cumulative averages and standard deviations of E^o values obtained in the simulations, an image of the reference structure used in RMSD calculations of peptide A, and more detailed figures for the three independent simulations performed for peptide A. This information is available free of charge via the Internet at https://pubs.acs.org/

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from CAPES (Brazil). G.T.F. acknowledges the Capes/Fulbright visiting scholar program for the grant that partially funded this project. The authors are thankful for the technical assistance provided by Daniel Roe, Kavindri Ranasinghe,

and Joe Garcia. This project was performed using computational resources from the Oak Ridge National Laboratory (ORNL) and the Texas Advanced Computing Center (TACC).

REFERENCES

[1] Creighton, T. E. *Proteins : Structures and Molecular Properties*; Freeman, 2013.

[2] Fersht, A. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; Freeman, 1998.

[3] Weinberg, D. R.; Gagliardi, C. J.; Hull, J. F.; Murphy, C. F.; Kent, C. A.; Westlake, B. C.; Paul, A.; Ess, D. H.; McCafferty, D. G.; Meyer, T. J. Proton-Coupled Electron Transfer. *Chem. Rev.*, **2012**, *112* (7), 4016–4093.

[4] Huynh, M. H. V.; Meyer, T. J. Proton-Coupled Electron Transfer. *Chem. Rev.*, **2007**, *107* (11), 5004–5064.

[5] Hammesschiffer, S.; Hatcher, E.; Ishikita, H.; Skone, J.; Soudackov, A. Theoretical Studies of Proton-Coupled Electron Transfer: Models and Concepts Relevant to Bioenergetics. *Coord. Chem. Rev.*, **2008**, *252* (3–4), 384–394.

[6] Costentin, C. Electrochemical Approach to the Mechanistic Study of Proton-Coupled Electron Transfer. *Chem. Rev.*, **2008**, *108* (7), 2145–2179.

[7] Landman, M.; Pretorius, R.; Buitendach, B. E.; van Rooyen, P. H.; Conradie, J. Synthesis, Structure, and Electrochemistry of Fischer Alkoxy- and Aminocarbene Complexes of Tungsten: The Use of DFT To Predict and Understand Oxidation and Reduction Potentials. *Organometallics*, **2013**, *32* (19), 5491–5503.

[8] Chun, H.-J.; Apaja, V.; Clayborne, A.; Honkala, K.; Greeley, J. Atomistic Insights into Nitrogen-Cycle Electrochemistry: A Combined DFT and Kinetic Monte Carlo Analysis of NO Electrochemical Reduction on Pt(100). *ACS Catal.*, **2017**, *7* (6), 3869–3882.

[9] Jensen, J. H.; Li, H. Calculation of Reduction Potential and p K A. In *Encyclopedia of Inorganic Chemistry*; John Wiley & Sons, Ltd: Chichester, UK, 2009.

[10] Dohm, S.; Spohr, E.; Korth, M. Developing Adaptive QM/MM Computer Simulations for Electrochemistry. *J. Comput. Chem.*, **2017**, *38* (1), 51–58.

[11] Wang, L.-P.; Van Voorhis, T. A Polarizable QM/MM Explicit Solvent Model for Computational Electrochemistry in Water. *J. Chem. Theory Comput.*, **2012**, *8* (2), 610–617.

[12] Blumberger, J. Free Energies for Biological Electron Transfer from QM/MM Calculation: Method, Application and Critical Assessment. *Phys. Chem. Chem. Phys.*, **2008**, *10* (37), 5651.

[13] Li, G.; Zhang, X.; Cui, Q. Free Energy Perturbation Calculations with Combined QM/MM Potentials Complications, Simplifications, and Applications to Redox Potential Calculations. *J. Phys. Chem. B*, **2003**, *107* (33), 8643–8653.

[14] Cruzeiro, V. W. D.; Amaral, M. S.; Roitberg, A. E. Redox Potential Replica Exchange Molecular Dynamics at Constant pH in AMBER: Implementation and Validation. *J. Chem. Phys.*, **2018**, *149* (7), 072338.

[15] Cruzeiro, V. W. D.; Roitberg, A. E. Multidimensional Replica Exchange Simulations for Efficient Constant pH and Redox Potential Molecular Dynamics. *J. Chem. Theory Comput.*, **2019**, *15* (2), 871–881.

[16] Gibson, H. R.; Mowat, C. G.; Miles, C. S.; Li, B.-R.; Leys, D.; Reid, G. A.; Chapman, S. K. Structural and Functional Studies on DHC, the Diheme Cytochrome c from Rhodobacter Sphaeroides , and Its Interaction with SHP, the Sphaeroides Heme Protein. *Biochemistry*, **2006**, *45* (20), 6363–6371.

[17] Paquete, C. M.; Turner, D. L.; Louro, R. O.; Xavier, A. V.; Catarino, T. Thermodynamic

and Kinetic Characterisation of Individual Haems in Multicentre Cytochromes C3. *Biochim. Biophys. Acta - Bioenerg.*, **2007**, *1767* (9), 1169–1179.

[18] Turner, D. L.; Salgueiro, C. A.; Catarino, T.; Legall, J.; Xavier, A. V. NMR Studies of Cooperativity in the Tetrahaem Cytochrome C3 from Desulfovibrio Vulgaris. *Eur. J. Biochem.*, **1996**, *241* (3), 723–731.

[19] Machuqueiro, M.; Baptista, A. M. Molecular Dynamics at Constant PH and Reduction Potential: Application to Cytochrome C3. *J. Am. Chem. Soc.*, **2009**, *131* (35), 12586–12594.

[20] Coutinho, I. B.; Xavier, A. V. [9] Tetraheme Cytochromes. *Methods Enzymol.*, **1994**, *243*, 119–140.

[21] Pereira, I. A. C.; Teixeira, M.; Xavier, A. V. Hemeproteins in Anaerobes. In *Bioinorganic Chemistry*; Springer Berlin Heidelberg, 1998; pp 65–89.

[22] Pereira, I. A. C.; Romão, C. V.; Xavier, A. V.; LeGall, J.; Teixeira, M. Electron Transfer between Hydrogenases and Mono- and Multiheme Cytochromes in Desulfovibrio Ssp. *JBIC J. Biol. Inorg. Chem.*, **1998**, *3* (5), 494–498.

[23] Henriques, J.; Costa, P. J.; Calhorda, M. J.; Machuqueiro, M. Charge Parametrization of the DvH-C3 Heme Group: Validation Using Constant-(PH,E) Molecular Dynamics Simulations. *J. Phys. Chem. B*, **2013**, *117* (1), 70–82.

[24] Tommos, C.; Skalicky, J. J.; Pilloud, D. L.; Wand, A. J.; Dutton, P. L. De Novo Proteins as Models of Radical Enzymes. *Biochemistry*, **1999**, *38* (29), 9495–9507.

[25] Martínez-Rivera, M. C.; Berry, B. W.; Valentine, K. G.; Westerlund, K.; Hay, S.; Tommos, C. Electrochemical and Structural Properties of a Protein System Designed To Generate Tyrosine Pourbaix Diagrams. *J. Am. Chem. Soc.*, **2011**, *133* (44), 17786–17795.

[26] Sibert, R.; Josowicz, M.; Porcelli, F.; Veglia, G.; Range, K.; Barry, B. A. Proton-Coupled

Electron Transfer in a Biomimetic Peptide as a Model of Enzyme Regulatory Mechanisms. *J. Am. Chem. Soc.*, **2007**, *129* (14), 4393–4400.

[27] Hwang, H.; McCaslin, T. G.; Hazel, A.; Pagba, C. V.; Nevin, C. M.; Pavlova, A.; Barry, B. A.; Gumbart, J. C. Redox-Driven Conformational Dynamics in a Photosystem-II-Inspired β-Hairpin Maquette Determined through Spectroscopy and Simulation. *J. Phys. Chem. B*, **2017**, *121* (15), 3536–3545.

[28] Pagba, C. V; McCaslin, T. G.; Veglia, G.; Porcelli, F.; Yohannan, J.; Guo, Z.; McDaniel, M.; Barry, B. A. A Tyrosine–Tryptophan Dyad and Radical-Based Charge Transfer in a Ribonucleotide Reductase-Inspired Maquette. *Nat. Commun.*, **2015**, *6*, 10010.

[29] Ullmann, G. M. The Coupling of Protonation and Reduction in Proteins with Multiple Redox Centers: Theory, Computational Method, and Application to Cytochrome C3. *J. Phys. Chem. B*, **2000**, *104* (26), 6293–6301.

[30] Teixeira, V. H.; Soares, C. M.; Baptista, A. M. Studies of the Reduction and Protonation Behavior of Tetraheme Cytochromes Using Atomic Detail. *J. Biol. Inorg. Chem.*, **2002**, *7* (1–2), 200–216.

[31] Gunner, M. R.; Alexov, E. A Pragmatic Approach to Structure Based Calculation of Coupled Proton and Electron Transfer in Proteins. *Biochim. Biophys. Acta - Bioenerg.*, **2000**, *1458* (1), 63–87.

[32] Gunner, M. R.; Mao, J.; Song, Y.; Kim, J. Factors Influencing the Energetics of Electron and Proton Transfers in Proteins. What Can Be Learned from Calculations. *Biochim. Biophys. Acta*, **2006**, *1757* (8), 942–968.

[33] Ullmann, G. M.; Knapp, E.-W. Electrostatic Models for Computing Protonation and Redox Equilibria in Proteins. *Eur. Biophys. J.*, **1999**, *28* (7), 533–551.

[34] Moutevelis, E.; Warwicker, J. Prediction of pKa and Redox Properties in the Thioredoxin Superfamily. *Protein Sci.*, **2004**, *13* (10), 2744–2752.

[35] Cheng, J.; Liu, X.; VandeVondele, J.; Sulpizi, M.; Sprik, M. Redox Potentials and Acidity Constants from Density Functional Theory Based Molecular Dynamics. *Acc. Chem. Res.*, **2014**, *47* (12), 3522–3529.

[36] Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E. I.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. AMBER 18. University of California: San Francisco 2018.

[37] Bergonzo, C.; Henriksen, N. M.; Roe, D. R.; Swails, J. M.; Roitberg, A. E.; Cheatham, T. E. Multidimensional Replica Exchange Molecular Dynamics Yields a Converged Ensemble of an RNA Tetranucleotide. *J. Chem. Theory Comput.*, **2014**, *10* (1), 492–499.

[38] Itoh, S. G.; Damjanović, A.; Brooks, B. R. PH Replica-Exchange Method Based on Discrete Protonation States. *Proteins Struct. Funct. Bioinforma.*, **2011**, *79* (12), 3420–3436.

[39] Swails, J. M.; Roitberg, A. E. Enhancing Conformation and Protonation State Sampling of Hen Egg White Lysozyme Using PH Replica Exchange Molecular Dynamics. *J. Chem. Theory Comput.*, **2012**, *8* (11), 4393–4404.

[40] Warren, J. J.; Winkler, J. R.; Gray, H. B. Redox Properties of Tyrosine and Related

Molecules. *FEBS Lett.*, **2012**, *586* (5), 596–602.

[41] Harris, T. K.; Turner, G. J. Structural Basis of Perturbed PKa Values of Catalytic Groups in Enzyme Active Sites. *IUBMB Life (International Union Biochem. Mol. Biol. Life)*, **2002**, *53* (2), 85–98.

[42] Besler, B. H.; Merz, K. M.; Kollman, P. A. Atomic Charges Derived from Semiempirical Methods. *J. Comput. Chem.*, **1990**, *11* (4), 431–439.

[43] Singh, U. C.; Kollman, P. A. An Approach to Computing Electrostatic Charges for Molecules. *J. Comput. Chem.*, **1984**, *5* (2), 129–145.

[44] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V; Cioslowski, J.; and D. J. Fox. Gaussian 09 Revision A.02. Wallingford 2009.

[45] Swails, J. M.; York, D. M.; Roitberg, A. E. Constant PH Replica Exchange Molecular Dynamics in Explicit Solvent Using Discrete Protonation States: Implementation, Testing, and Validation. *J. Chem. Theory Comput.*, **2014**, *10* (3), 1341–1352.

[46] Kuntal, B. K.; Aparoy, P.; Reddanna, P. EasyModeller: A Graphical Interface to MODELLER. *BMC Res. Notes*, **2010**, *3* (1), 226.

[47] Daidone, I.; Paltrinieri, L.; Amadei, A.; Battistuzzi, G.; Sola, M.; Borsari, M.; Bortolotti, C. A. Unambiguous Assignment of Reduction Potentials in Diheme Cytochromes. *J. Phys. Chem. B*, **2014**, *118* (27), 7554–7560.

[48] Simões, P.; Matias, P. M.; Morais, J.; Wilson, K.; Dauter, Z.; Carrondo, M. A. Refinement of the Three-Dimensional Structures of Cytochrome C3 from Desulfovibrio Vulgaris Hildenborough at 1.67 Å Resolution and from Desulfovibrio Desulfuricans ATCC 27774 at 1.6 Å Resolution. *Inorganica Chim. Acta*, **1998**, *273* (1–2), 213–224.

[49] Glover, S. D.; Jorge, C.; Liang, L.; Valentine, K. G.; Hammarström, L.; Tommos, C. Photochemical Tyrosine Oxidation in the Structurally Well-Defined α 3 Y Protein: Proton-Coupled Electron Transfer and a Long-Lived Tyrosine Radical. *J. Am. Chem. Soc.*, **2014**, *136* (40), 14039–14051.

[50] Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. . Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.*, **1977**, *23* (3), 327–341.

[51] Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.*, **1992**, *13* (8), 952–962.

[52] Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N Log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.*, **1993**, *98* (12), 10089–10092.

[53] 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.

[54] Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-Scale

Conformational Changes with a Modified Generalized Born Model. *Proteins Struct. Funct. Bioinforma.*, **2004**, *55* (2), 383–394.

[55] Di Russo, N. V; Estrin, D. A.; Martí, M. A.; Roitberg, A. E. pH-Dependent Conformational Changes in Proteins and Their Effect on Experimental PK(a)s: The Case of Nitrophorin 4. *PLoS Comput. Biol.*, **2012**, *8* (11), e1002761.

[56] Fan, K.; Akutsu, H.; Kyogoku, Y.; Niki, K. Estimation of Microscopic Redox Potentials of a Tetraheme Protein, Cytochrome c_3 of Desulfovibrio Vulgaris, Miyazaki F and Partial Assignments of Heme Groups. *Biochemistry*, **1990**, *29* (9), 2257–2263.

[57] Gayda, J. P.; Yagi, T.; Benosman, H.; Bertrand, P. EPR Redox Study of Cytochrome c_3 from Desulfovibrio Vulgaris Miyazaki. *FEBS Lett.*, **1987**, *217* (1), 57–61.

[58] Park, J.-S.; Kano, K.; Niki, K.; Akutsu, H. Full Assignment of Heme Redox Potentials of Cytochrome c_3 of D. Vulgaris Miyazaki F by 1 H-NMR. *FEBS Lett.*, **1991**, *285* (1), 149–151.