

PCCP

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Acidity Constants of Lumiflavin from First Principles Molecular Dynamics Simulations

Murat Kılıç Bernd Ensing*

June 9, 2014

Abstract

We have computed the free energy profiles of the deprotonation reactions of lumiflavin in the semiquinone and fully reduced oxidation states using constrained DFT-based molecular dynamics simulations. In the semiquinone state, the N5 nitrogen atom and the N1 nitrogen atom can become protonated. We find, in agreement with experiment, that the N5 site is the predominant proton acceptor, although the computed pK_a value is somewhat smaller than the experimental number. The computed pK_a for the N1 protonation in the fully reduced state is in good agreement with the experimental number. We employ two different, commonly used, reaction coordinates based on the distances between the proton and the donor and acceptor atoms. Further improvement of the accuracy of this type of pK_a calculations may require development of more advanced reaction coordinates that go beyond the description of only the first proton transfer step from a donor atom to a first solvation shell water molecule.

1 Introduction

Flavins such as flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN) are common cofactors that play a key role in a wide variety of enzyme catalysed oxidation and reduction reactions. They can accept up to two electrons in either a single step

*Van 't Hoff Institute for Molecular Science, University of Amsterdam, Science Park 904, 1098 XH, Amsterdam, The Netherlands, E-mail: B.Ensing@uva.nl, website: <http://molstim.chem.uva.nl>

or in two distinct one-electron transfer reactions. Flavins can also act as proton donors or acceptors. The electron and proton transfer reactions are coupled—the acidity constants depend on the redox state of the flavin and, vice versa, the reduction potentials are a function of the pH. Moreover, both the electron and proton affinities of the flavin are modulated by their protein or solvent environment, which underlies the versatile flavochemistry observed in nature.

The structural motif of flavins is a three-ring isoalloxazine moiety with a side group attached to the nitrogen (N10) atom in the middle ring (see Figure 1). Three other nitrogen atoms can accept or donate a proton depending on the oxidation state of the flavin. The acidity and redox properties of flavins have been subject of study since the 1930s¹ using potentiometric methods,² pulse radiolysis,^{3,4} NMR,⁵ and UV/Vis spectroscopy.^{3,5} The acidity constants (pK_a 's) and reduction potentials are therefore experimentally well-established, despite the complexity due to the intrinsic coupling of the redox and protonation reactions and the fact that the intermediate semiquinone radical is rather unstable with a formation constant of $K = [\text{FMN}_{\text{sq}}]^2 / ([\text{FMN}_{\text{ox}}][\text{FMN}_{\text{red}}]) \approx 10^{-3}$.⁶ Figure 2 illustrates schematically the different redox and protonation states of flavins. In the oxidised state, only the N3 atom is protonated with an associated pK_a of about 10.3. One-electron reduction of the neutral flavin leads to the anionic radical semiquinone state, or to the neutral semiquinone when simultaneously a proton is accepted at the N5 position. For riboflavin and FMN, the pK_a of the semiquinone is estimated to be about 8.3-8.55.^{2,3,6} The second one-electron reduction leads to the anionic hydroquinone state, or again a neutral state when accompanied with a second protonation, this time at the N1 atom. The pK_a of the riboflavin and FMN hydroquinone are estimated to be 6.25 and 6.72 respectively.²

Computational studies have the advantage that the electron and proton transfer processes can be decoupled by computing the reaction free energy of the proton (or electron) transfer reaction at a fixed oxidation (or protonation) state. Truhlar and coworkers have computed the redox and acidity properties of several flavins using density functional theory (DFT) and a continuum description of the aqueous solvent.^{7,8} For a series of lumiflavin compounds with different substituent at the C7 and C8 positions, they

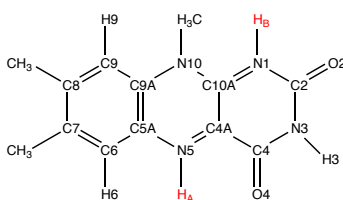


Figure 1: Structural motif of Lumiflavin. In the flavins most commonly found in nature, such as riboflavin (vitamine B2), FMN, and FAD, the methyl group attached to N10 is replaced by a longer side-chain. Depending on the oxidation state, the N1 and N5 nitrogen atoms can be protonated. The protons are labeled H_A and H_B .

found reduction potentials and pK_a values in good agreement with experiment, with the exception of the pK_a of 1.5 for the dimethyl species in the fully reduced state, which is much smaller than the experimental value of 6.5. DFT-based molecular dynamics (DFT-MD) simulations using explicit solvent require fewer empirical parameters such as the cavity radii and the dielectric constant, and more importantly, they allow for study of the role of the molecular environment in the charge transfer process. However, sampling of the solvent configurations is computationally demanding at the DFT-MD level of theory. And secondly, the calculation of the transfer free energy requires an advanced sampling method, such as free energy perturbation^{9,10} or constrained molecular dynamics.^{11–13}

We have recently studied¹⁴ the two one-electron reduction reactions of lumiflavin (LF) in aqueous solution using DFT-MD simulations. Lumiflavin is one of the smallest flavin members, in which the N10 side-chain is a methyl group as depicted in Figure 1. We used a free energy perturbation approach^{15,16} that relates the redox potential to the average energy required to add (or subtract) an electron to the system, *i.e.* the vertical gap energy. Computation of the pK_a is somewhat more involved because the proton, treated as a classical particle, has a position that should be chosen in an efficient way such that the configurations are physically relevant.^{10,17,18} We therefore use constrained molecular dynamics,^{11,19} to calculate the potential of mean force of the protonation reaction as function of a well-chosen reaction coordinate, from which the pK_a is obtained.²⁰

In this work, we study the proton transfer (PT) reactions at the N1 and N5 lu-

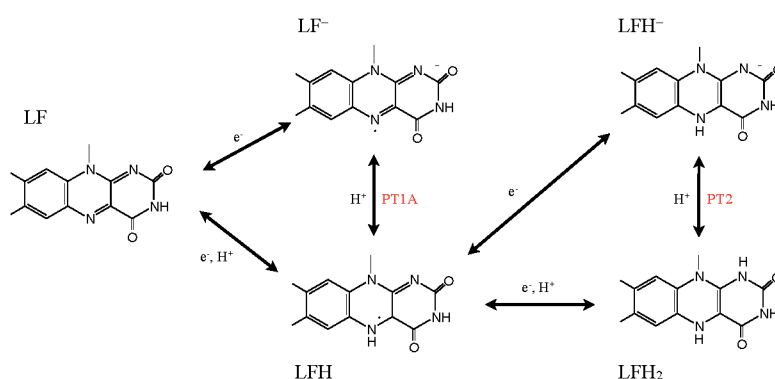


Figure 2: Lumiflavin (LF) can exist in several oxidation and protonation states. Electron and proton transfer reactions are indicated by arrows.

lumiflavin nitrogen atoms in the semiquinone and fully reduced oxidation states. In the semiquinone state, the protonation takes place at the N5 nitrogen, which we label PT1A as indicated in Figure 2. However, we can also measure the pK_a of the N1 site in the semiquinone state (labeled PT1B). The third reaction is the protonation of the N1 site in the fully reduced state (PT2), in which case the N5 site is considered already protonated.

In the following, we first briefly review the method to calculate the pK_a using constrained DFT–MD simulation and provide the details of our model. Next, the computed free energy profiles for the proton transfer reactions and the resulting pK_a values for the different flavin protonation sites are discussed. We have computed the free energy profiles twice using the two most commonly used reaction coordinates for proton transfer reactions. A suggestion is made on how these reaction coordinates can be further improved.

2 Methods

2.1 Calculation of pK_a from constrained dynamics

The acidity constant is related to the difference in the Helmholtz free energy, ΔA , between the protonated and deprotonated states of the molecule as:

$$\text{p}K_a = \frac{\Delta A}{k_B T \ln(10)}, \quad (1)$$

with k_B the Boltzmann constant and T the absolute temperature. We compute this free energy difference by performing a series of constrained molecular dynamics simulations, in which for each simulation the sampling of the molecular configurations is constrained to a certain fixed reaction coordinate value. The reaction coordinate is a geometric function of the atomic positions, $q(\mathbf{r}^N)$, that describes the progress of the protonation reaction from the reactant state to the product state. Integration of the mean force of constraint, *i.e.* the average force needed to maintain the reaction coordinate constraint during the simulation, results in a free energy profile as a function of the reaction coordinate:

$$\Delta A(q') = - \int_{q_0}^{q'} \langle f \rangle_q dq, \quad (2)$$

in which we choose $A(q_0) = 0$ and q_0 to be at the free energy minimum of the reactant state. The brackets denote that the force is an ensemble average and the subscript q indicates that the ensemble was constrained at the value q of the reaction coordinate, which is enforced every MD step using the method of Lagrange multipliers. For a distance constraint, the mean force of constraint is equal to the Lagrange multiplier. For more general coordinate types, the constraint force contains additional terms that unbiased the measured force to that of the unconstrained ensemble.¹⁹

The accuracy of the resulting free energy profile of the reaction depends crucially on how well the chosen reaction coordinate describes the reaction mechanism. For a proton transfer reaction, this choice of a reaction coordinate is complicated by the fact that after the proton has made its initial jump from the donor molecule to a nearby solvent water molecule to form a hydronium ion, another proton of this ion may make the next jump to a water molecule in the second coordination shell, and so forth. Previous work has therefore focussed mainly on the first step, the breaking of the initial donor-proton bond, thereby neglecting the changes in free energy that may occur when the charge-separation proceeds beyond the contact-ion pair distance. The good agreement of previous results with experiment so far indicates that the missing contributions are

small.

Here, we use two different types of reaction coordinates. The first coordinate is a function that estimates the coordination number, n_c , of the number of hydrogens within a distance d_0 of the donor nitrogen atom:

$$n_c = \sum_i \frac{1 - \left(\frac{d(\text{N}-\text{H}_i)}{d_0}\right)^n}{1 - \left(\frac{d(\text{N}-\text{H}_i)}{d_0}\right)^m}. \quad (3)$$

Here, the sum runs over all solvent hydrogen atoms in the system. The value of n_c is (close to) one in the protonated state and switches smoothly to zero as the distance $d(\text{N}-\text{H})$ increases. The smoothing parameters n and m were chosen to be respectively 8 and 16 and the cutoff radius d_0 was set to 1.3 Å. The coordination number coordinate is probably the most used reaction coordinate for this type of $\text{p}K_a$ calculations. One advantage of this coordinate is that it leaves the choice of the solvent molecule that accepts the proton free. A disadvantage of this coordinate is that it is difficult to simulate the reverse proton transfer from the solvent to the flavin molecule, because at very small values of n_c , the proton can escape into the bulk after which the coordinate loses its control on the proton position. The series of constrained MD simulations is therefore setup by starting from the protonated flavin state and stepwise decreasing n_c to generate initial conditions for the other simulations.

The second reaction coordinate type that we use includes also the distance of the proton to the accepting water molecule oxygen:

$$\Delta d = d(\text{N}-\text{H}) - d(\text{O}_w-\text{H}). \quad (4)$$

That is, Δd is the difference between the distance of the proton to the donating flavin nitrogen and the distance of the proton to the accepting water oxygen. Its value is negative in the protonated state, zero when the proton is exactly equidistant from the donor and acceptor atoms, and positive in the deprotonated state. The advantage of using this coordinate is that we can also generate a series of simulations starting from the deprotonated state. This deprotonated state was constructed from a simulation of

(deprotonated) lumiflavin in water by adding a proton to a water molecule that was hydrogen bonded to the flavin nitrogen atom. Subsequently, this system was equilibrated while maintaining the hydrogen bond using the Δd constraint and an additional coordination number restraint on the hydronium ion oxygen. This restraint worked as a repulsive harmonic wall to avoid that either of the other two hydronium hydrogens would escape to a nearby water molecule. The wall potential on this coordination number was zero as long as its value was larger than 1.6. After equilibration with this additional restraint, we performed the production simulations both without and with the restraint. The former results are the most interesting and are discussed hereafter, whereas the results obtained with the additional restraint on the proton accepting water molecule are provided in the supplementary information. By performing a constrained MD series starting from the equilibrated protonated state (forward series) as well as from the equilibrated deprotonated state (reverse series), we can assess whether omitted (solvent) degrees of freedom from the reaction coordinate cause hysteresis effects (see for an discussion of such hysteresis *e.g.* Ref.²¹).

2.2 Computational details

All electronic structure calculations and ab initio molecular dynamics simulations were performed using DFT with the Perdew-Burke-Ernzerhof (PBE)²² exchange-correlation functionals and Grimme's D3 Van der Waals correction²³ as implemented in the CP2K program (version 2.4).^{24,25} The CP2K program is based on a hybrid Gaussian and plane wave scheme, in which the wave functions are expanded using a Gaussian basis set, and an auxiliary basis of plane waves is employed to expand the density.²⁶ We used pseudopotentials of the Goedecker-Teter-Hutter (GTH) type, based on the parametrization of Hartwigsen-Goedecker-Hutter^{27,28} and adapted for the DFT. A split valence Gaussian basis set designed specifically for these pseudopotentials,²⁵ of double- ζ quality and one set of polarisation functions (DZVP), was employed for all atoms including hydrogen. The auxiliary plane wave basis expansion was cutoff at 300 Ry.

The system of aqueous LF contained one LF molecule and 102 water molecules in a periodic cubic unit cell with an edge of 15.148 Å. The DFT-based molecular dynamics

Proton transfer	PBE	B3LYP	M06-2X	M06-L	B3LYP ⁷	M06-L ⁷
PT1A: LFH \rightarrow LF ⁻ + H ⁺	338.0	339.4	336.0	339.0	334	335
PT1B: LFH \rightarrow LF ⁻ + H ⁺	327.8	330.3	327.4	329.7	-	-
PT2: LFH ₂ \rightarrow LFH ⁻ + H ⁺	335.4	338.2	336.0	337.6	332	333

Table 1: Zero Kelvin gas-phase energies (kcal/mol) of the three deprotonation reactions. The last two columns shows the results from ref⁷ using the B3LYP and M06-L density functionals.

simulations used the Born-Oppenheimer method with a time step of 0.5 fs in the canonical (NVT) ensemble. The CSVr (Canonical Sampling through Velocity Rescaling)²⁹ thermostat was used to maintain a constant temperature of $T = 300$ K. The systems were equilibrated for at least 5 ps, after which at least 5 ps of constraint simulation was performed at each reaction coordinate value parameter for analysis.

To assess the quality of the PBE functional, we performed a series of benchmark calculations of the proton transfer reactions in the gas-phase using the Gaussian-09 program.³⁰ The molecular geometries were computed at the PBE/6-31g(d,p) level, which were subsequently used to compare the deprotonation energies between different functionals using a 6-311G(2d,p) basis set. The results are compiled in Table 1, together with the B3LYP³¹⁻³³ and M06-L³⁴ results of Bhattacharyya *et al.*⁷ Comparing the results using the PBE functional, which is a first generation GGA functional, with the hybrid-GGA B3LYP, the meta-GGA M06-L, and the meta-hybrid-GGA M06-2X, we note small differences of 1-2 kcal/mol. Note in particular that the PBE difference between the PT1A and PT2 reactions is somewhat larger (2.6 kcal/mol) than that for the other functionals, B3LYP (1.2), M06-2X (0.0), and M06-L (1.4). Instead, the difference between PT1B and PT2 is very similar for all functionals, PBE (7.6), B3LYP (7.9), M06-2X (8.6), M06-L (7.9). The latter leads us to conclude that there is no indication that the PBE GGA functional shows a systematic deviation for the deprotonation reactions in the semiquinone radical state (PT1A and PT1B) compared to the closed shell state (PT2). For our flavin simulations in aqueous solution, we augment the PBE functional with the *a posteriori* Van der Waals correction by Grimme.

3 Results and discussion

Figure 3 summarises the main results from the constrained DFT–MD simulations that we performed to compute the pK_a values of lumiflavin. The different colours represent the three proton transfer reactions that we focussed on: the (de-) protonation in the semiquinone state at the N5 atom (black lines labeled PT1A) and at the N1 atom (red lines, PT1B), and the (de-) protonation in the fully reduced state at the N1 atom (green lines; PT2). The top panels show the free energy profiles as a function of the coordination number reaction coordinate (Eqn. 3) at the left side, and using the distance-difference coordinate (Eqn. 4) in the right panels. These profiles were obtained by integration of the measured mean force of constraint (see Eqn. 2), which are shown in the middle panels.

In the top-left panel, the free energy profiles show a minimum in the protonated state at a reaction coordinate value of $n_c = 0.9$ and increase when the proton is moved away from the flavin by decreasing n_c . The bottom panel shows the average $d(N-H)$ and $d(H-O_W)$ bond distances, which are close to respectively 1.1 Å and 1.8 Å in the protonated state as the proton forms a hydrogen bond with the nearby water molecule. For each of the three proton transfer reactions, a snapshot of this initial protonated state is shown in Figure 4 at the left.

Decreasing the n_c reaction coordinate carries the proton from the flavin toward the nearest water molecule as seen from the increasing $d(N-H)$ distance and the decreasing $d(H-O_W)$ distance. This happens in a similar manner for all three proton transfer reactions, crossing the equidistant state at $n_c = 0.6$. The three free energy profiles show also a similar increase, but reach different final maxima. The last data point is drawn with an open circle to indicate that in that simulation a spontaneous second proton transfer was observed that carried the hydronium ion into the second coordination shell, after which we stopped the simulation. The free energy profiles should thus be interpreted as the work required to carry the proton from the flavin up to the reaction coordinate value at which the proton escapes barrierless into the solvent. This escape occurs earliest for PT1B at a free energy barrier of 6.2 kcal/mol and at the latest for

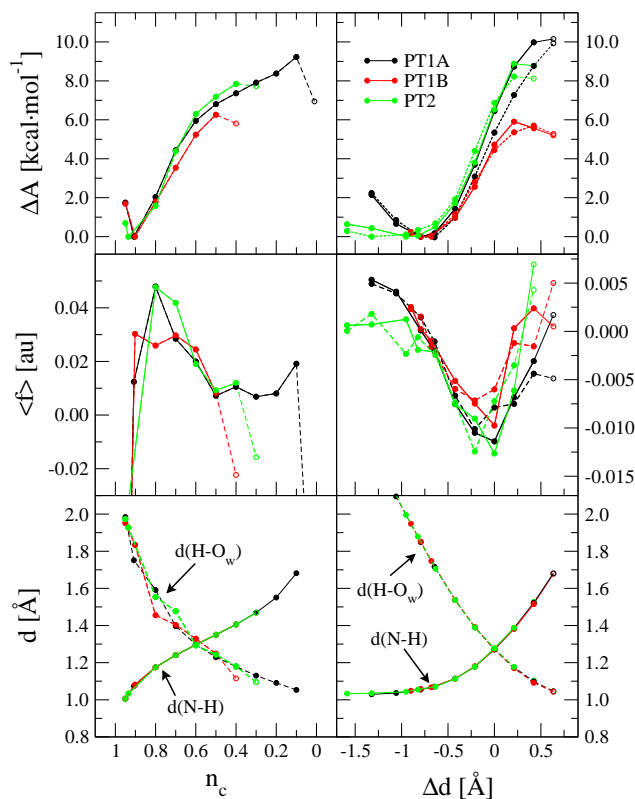


Figure 3: Top panels: free energy profiles for the three protonation reactions computed using the n_c reaction coordinate (left) and the Δd coordinate (right panels). The dashed lines in the right panels results from the backward reactions and allow for assessment of hysteresis effects. Middle panels: average force of constraint. Open circles connected by dashed lines denote the runs in which the proton escapes the control of the reaction coordinate. Bottom panels: distance between the proton and the lumiflavin N1 or N5 nitrogen atom and distance between the proton and the accepting water molecule oxygen atom.

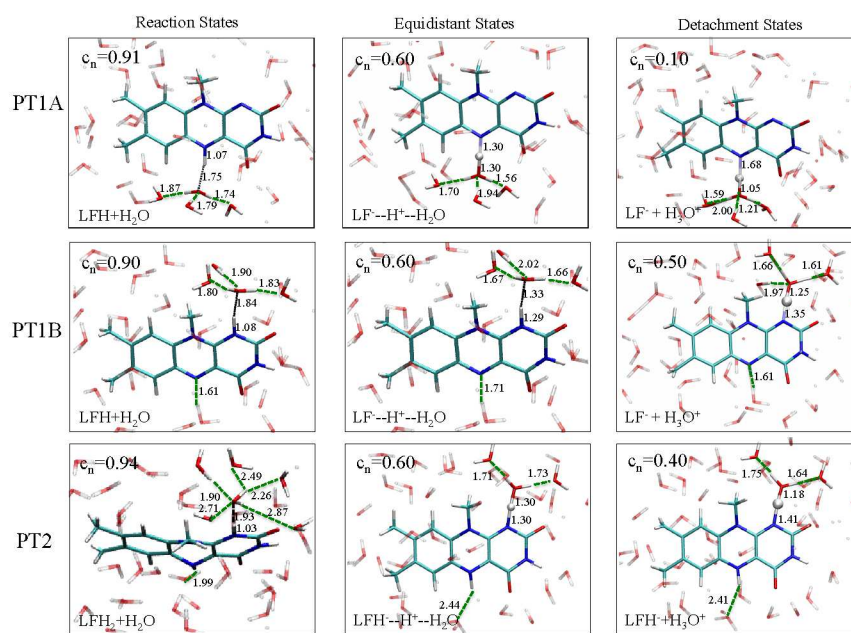


Figure 4: Representative snapshots from the constrained simulations of the PT1A, PT1B and PT2 reactions, at the protonated reactant state (left panels), the intermediate equidistant state (middle column), and the state just before the proton escapes to the second solvation shell.

Proton transfer	ΔA		pK_a		
	n_c	Δd	n_c	Δd	exp.
PT1A: LFH \rightarrow LF ⁻ + H ⁺	9.1	9.4	6.7	6.9	8.5
PT1B: LFH \rightarrow LF ⁻ + H ⁺	6.2	5.6	4.5	4.1	-
PT2: LFH ₂ \rightarrow LFH ⁻ + H ⁺	7.8	8.5	5.7	6.2	6.5

Table 2: Deprotonation free energies, ΔA , (kcal/mol) and pK_a values computed for the N5 nitrogen (PT1A) and N1 nitrogen (PT1B) sites in the semiquinone state and the N1 nitrogen (PT2) site in the fully reduced state, computed using two different reaction coordinates (n_c and Δd). Experimental pK_a values were taken from Ref.³⁵

PT1A at 9.1 kcal/mol. These free energies, together with the pK_a values computed using Eqn. 1, are compiled in Table 2.

The free energy profiles as a function of the distance–difference reaction coordinate, shown at the right in Figure 3, show the same trend for the escape barriers of the three proton transfer reactions. As explained in the method section, with this reaction coordinate, we are also able to perform the series of constraint simulations starting from the deprotonated state, which results in the profiles shown by the dotted lines. The hysteresis between the forward and reverse reaction profiles is most likely due to small differences in the solvent reorganisation that is not controlled by these simple reaction coordinates. The effect is the largest for the PT1A reaction, resulting in an overestimated deprotonation free energy of 10.0 kcal/mol for the forward reaction compared to an underestimated protonation free energy of 8.8 kcal/mol in the reverse direction.

Although we cannot estimate the amount of hysteresis in the case of the coordination number reaction coordinate, we should of course expect a similar overestimate for the deprotonation free energy, due to the solvent reorganisation lagging behind. Nevertheless, these free energies are smaller (in the cases of PT1A and PT2) than those computed using the Δd coordinate, even though we could correct the latter numbers for the hysteresis effect by taking the average of the forward and reverse estimates (see Table 2). The main difference between the n_c and Δd reaction coordinates is that n_c is effectively only a function of the distance between the proton and the donating nitrogen atom, whereas Δd incorporates also the distance between the proton and the accepting water oxygen. Not controlling the $d(\text{H} - \text{O}_w)$ with n_c leads to larger variation in the

average distance of the accepting water to the leaving proton as is seen in the bottom-left panel of Figure 3 by comparing the three $d(\text{H} - \text{O}_W)$ curves at the early stage of the deprotonation reaction (*i.e.* at $n_c \approx 0.9 - 0.7$). Note how a further away water oxygen (*e.g.* the red dot at $n_c = 0.9$, or the black and green dots at $n_c = 0.8$, or the green dot at $n_c = 0.7$) corresponds to a larger average force of constraint shown in the middle-left panel. Note also that in the n_c case the largest absolute force is measured at the early stages of the deprotonation, so that the steepest part of the free energy curve is around $n_c = 0.8$, whereas for the Δd coordinate the largest constraint force is seen near the equidistant state, so that the ΔA curve is the steepest around $\Delta d = 0$. A third difference seen in the behaviour of the n_c and Δd reaction coordinates is seen in the last stage of the deprotonation reaction. Using the n_c coordinate, the average constraint force remains positive, *i.e.* pushing the system back toward the protonated state, until the proton escapes into the solution. Instead, when using the Δd coordinate, the constraint force tends back to zero, which results in a free energy curve that shows a maximum, as expected when reaching a transition state barrier. This indicates that in the latter case, the proton escape occurs close to the actual free energy barrier, whereas in the case of the n_c coordinate the proton may escape prematurely depending on a random fluctuation of the distance between the proton and the accepting water molecule.

The final state of the system is different for each of the three proton transfer reactions, irrespective of the reaction coordinate used in the constraint simulations. For PT1A, the hydronium ion remains close to the N5 site, simply hopping back and forth between first and second shell water molecules through an Eigen-Zundel-Eigen mechanism. In one simulation, using the n_c coordinate constraint at 0.1, the proton was seen to be accepted by the lumiflavin O4 atom for most of the simulation. For PT1B, in all simulations in which the proton escapes (*i.e.* at $n_c \leq 0.4$ or $\Delta d \geq 0.6 \text{ \AA}$) the hydronium ion did not remain close to the N1 donor atom for more than a picosecond, but moved via several Grothuss-like hops along a wire of hydrogen bonds in the water solvent to terminate at the lumiflavin N5 atom (or that of a periodic copy of the lumiflavin molecule). And finally, for PT2 the hydronium ion would also travel away from the N1 atom but in this case remain far away from the lumiflavin molecule, at a distance

apparently only limited by the size of the system box.

The pK_a values computed from the deprotonation free energy barriers using the two different reaction coordinates are in fair agreement with each other (see Table 2). Compared to the experimental numbers, we see good agreement for the PT2 case, but a significant underestimation for the PT1A pK_a . Similar to previous pK_a calculations using the same approach,^{13,36–38} we estimate the accuracy of the employed methodology to be within 1 pK_a unit. The discrepancy for the PT1A pK_a between our computed value and the experimental number is most likely due to a systematic underestimation of the free energy barrier due to the limited range of the employed reaction coordinates. As described above, in the PT1A case the hydronium ion remains near the N5 protonation site when the maximum range of either the n_c or Δd reaction coordinate is reached. This indicates that the actual free energy barrier to separate the hydronium ion from the semiquinone anion is larger than we can compute with these types of reaction coordinates.

An interesting approach to compensate for the missing part of the free energy profile beyond the transition state is based on the reversible work theorem,³⁹ which relates the proton–conjugate base radial distribution function, $g(d)$, to a free energy profile as

$$g(d) = \exp[-\Delta A(d)/k_B T], \quad (5)$$

in which $\Delta A(d)$ is the free energy (or average work) to bring the proton from infinity to a distance d from the conjugate base. The dissociation constant can then be expressed as

$$K_c^{-1} = c_0 \int_0^{D_c} dd 4\pi d^2 \exp[-\Delta A(d)/k_B T], \quad (6)$$

with c_0 the standard concentration and D_c a suitable cut-off distance to distinguish between the covalently bonded state and the dissociated state. However, due to the limited box size, L , that can be used in DFT–MD simulations, the free energy profile from infinity until the maximum dissociation distance, $L/2$, remains unknown. Moreover, in practice, we only have a free energy profile of the bonded state up to the point that the proton escapes the control of the reaction coordinate. Davies *et al.* circumvent these

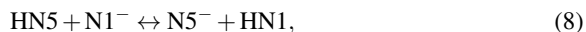
issues by deriving an expression for K_a in terms of an D_c dependent dissociation fraction, $\alpha(D_c)$. By fitting the $\text{p}K_a$ for water dissociation to the experimental value, they find a cut-off $D_c = 1.22 \text{ \AA}$ to be optimal.³⁶ Ivanov *et al.* also use the water dissociation as a reference, but simply take the ratio of the dissociation constants,

$$\frac{K_d(\text{acid})}{K_d(\text{H}_2\text{O})} = \frac{\int_0^{D_c} dd d^2 \exp[-\Delta A_{\text{acid}}(d)/k_B T]}{\int_0^{D_c} dd d^2 \exp[-\Delta A_{\text{H}_2\text{O}}(d)/k_B T]} \quad (7)$$

which is supposedly less sensitive to the cut-off radius; they use $D_c = 1.35 \text{ \AA}$.⁴⁰ We cannot straightforwardly employ this equation for our flavin case, because we have not computed the $K_d(\text{H}_2\text{O})$ at the same DFT level of theory. However, the same approach can be used to compare different deprotonation reactions, for example the PT1A reaction versus the PT2 reaction as reference. We therefore remap our $\Delta A(\Delta d)$ profiles to $\Delta A(d)$ functions using the bottom-right panel of Figure 3 and employ Equation 7. However, the result for the PT1A $\text{p}K_a$ is that it is almost equal or even less than the (reference) PT2 value, depending on the D_c value used, which is much worse than our results based on the dissociation barriers. This can be rationalised by noting that the PT2 ΔA curve is largely equal or somewhat higher than the PT1A curve until very close to the dissociation barrier at $d = 1.4 \text{ \AA}$ (see top panels in Figure 3). In other words, this approach is based on the assumption that the shape of the free energy profile between $d = 0$ and D_c is always and in the same manner representative for the overall free energy difference between bonded and dissociated states. Clearly, this assumption does not hold in our case.

For the PT1B, we find a value of 4.1 using the Δd coordinate. To our knowledge, the experimental $\text{p}K_a$ of the N1 site in the semiquinone state is unknown (Land and Swallow measured 2.3 and 8.3 for the semiquinone $\text{p}K_a$'s using pulse radiolysis, however here the first number refers most likely to the N1 $\text{p}K_a$ while the N5 site is protonated³). Similar to the PT2 case, our simulations captured the PT1B deprotonated including the escape of the proton into the solvent, so that we expect a similarly good accuracy. This allows us to estimate the relative propensity, K , for the protonation of the N1 site versus

the N5 site in the semiquinone state,



as

$$K = \frac{[\text{N5}^-][\text{HN1}]}{[\text{HN5}][\text{N1}^-]} = \frac{[\text{H}^+][\text{N5}^-]}{[\text{HN5}]} \cdot \frac{[\text{HN1}]}{[\text{H}^+][\text{N1}^-]} = 10^{-8.5}/10^{-4.1} = 6.3 \cdot 10^{-5}, \quad (9)$$

in which we use the experimental number of the $\text{p}K_a$ of PT1A.

4 Conclusions

Using the method of constrained DFT-based molecular dynamics simulation, we have studied the (de-) protonation of the prototypical flavin named lumiflavin in aqueous solution. In the semiquinone state, the N1 and the N5 nitrogen atoms can act as proton acceptors, whereas in the fully reduced oxidation state, only the N1 site is available for protonation.

We employed two different reaction coordinates in the constrained simulations to compute the deprotonation free energy profiles, from which the $\text{p}K_a$ values were obtained. The first reaction coordinate was functionalised as the coordination number of the number of hydrogens around the nitrogen atom, and thus effectively only considered the distance between the proton and the donor atom. The second reaction coordinate included also the distance from the proton to the accepting water oxygen atom, by taking the difference between the two distances. The latter reaction coordinate allowed us to estimate systematic errors due to omitted solvent degrees of freedom in these simple reaction coordinates, by measuring the hysteresis in the free energy profiles between the forward deprotonation and the backward protonation reactions. This hysteresis effect was found to be relatively small. Secondly, we noted that omission of the distance of the proton to the accepting water molecule in the first reaction coordinate leads to larger fluctuations in the measured force of constraint and free energy profile.

In the semiquinone state, the N5 nitrogen is the predominant protonation site as expected, however the computed pK_a of 6.9 (using the distance difference reaction coordinate) was smaller than the experimental number of 8.5. We attribute this to the fact that the current reaction coordinates do not control the deprotonation reaction beyond the second coordination shell of the donor molecule. This problem did not play a role in the deprotonation of the N1 site in the semiquinone and fully reduced oxidation states. In the first case the proton spontaneously diffused through the solvent toward the N5 site, whereas in the fully reduced case the hydronium ion remained as far from the lumiflavin as possible in the limited simulation system. The pK_a of the reduced lumiflavin was found to be 6.2 in good agreement with the experimental number of 6.5.

Supporting Information

The Supporting Information contains a comparison of the free energy profiles of the three proton transfer reactions using the Δd coordinate with and without an additional restraint on the proton accepting water molecule and radial distribution functions of the solvent environment around the donating lumiflavin site and the accepting water molecule.

Acknowledgement

The work was financially supported by NWO (The Netherlands Organisation for Scientific Research) through a Vidi grant of BE.

References

- [1] L. Michaelis, M. P. Schubert and C. V. Smythe, *J. Biol. Chem.*, 1936, **116**, 587–607.
- [2] R. Draper and L. Ingraham, *Arch. Biochem. Biophys.*, 1969, **125**, 802.
- [3] E. Land and A. Swallow, *Biochemistry*, 1969, **8**, 2117–2125.

- [4] R. Anderson, *Biochim. Biophys. Acta*, 1983, **722**, 158–162.
- [5] S. Ghisla, P. Macheroux and F. Müller, Flavins and flavoproteins, proceedings of the 10th internat. symposium, 1991, pp. 27–32.
- [6] S. G. Mayhew, *Eur. J. Biochem.*, 1999, **265**, 698–702.
- [7] S. Bhattacharyya, M. Stankovich, D. G. Truhlar and J. Gao, *J. Phys. Chem. A*, 2007, **111**, 5729–5742.
- [8] M. North, S. Bhattacharyya and D. G. Truhlar, *J. Phys. Chem. B*, 2010, **114**, 14907–14915.
- [9] R. W. Zwanzig, *J. Chem. Phys.*, 1954, **22**, 1420.
- [10] M. Sulpizi and M. Sprik, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5238.
- [11] E. A. Carter, G. Ciccotti, J. T. Hynes and R. Kapral, *Chem. Phys. Lett.*, 1989, **156**, 472–477.
- [12] B. Trout and M. Parrinello, *Chem. Phys. Lett*, 1998, **288**, 343.
- [13] M. Sprik, *Chem. Phys.*, 2000, **258**, 139–150.
- [14] M. Kılıç and B. Ensing, *J. Chem. Theory Comput.*, 2013, **9**, 3889–3899.
- [15] A. Warshel, *J. Phys. Chem.*, 1982, **86**, 2218–2224.
- [16] J. Blumberger, L. Bernasconi, I. Tavernelli, R. Vuilleumier and M. Sprik, *J. Am. Chem. Soc.*, 2004, **126**, 3928–3938.
- [17] J. Cheng, M. Sulpizi and M. Sprik, *J. Phys. Chem.*, 2009, **131**, 154504.
- [18] J. Cheng and M. Sprik, *J. Chem. Theory Comput.*, 2010, **6**, 880–889.
- [19] W. K. Den Otter and W. J. Briels, *J. Chem. Phys.*, 1998, **109**, 4139.
- [20] M. Sprik, *Faraday Discuss.*, 1998, **110**, 437–445.
- [21] B. Ensing, E. J. Meijer, P. E. Blöchl and E. J. Baerends, *J. Phys. Chem. A*, 2001, **105**, 3300.

- [22] J. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
- [23] S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.
- [24] CP2K developers group, <http://cp2k.berlios.de>.
- [25] J. Vandevondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, *Comput. Phys. Comm.*, 2005, **167**, 103–128.
- [26] G. Lippert, J. Hutter and M. Parrinello, *Mol. Phys.*, 1997, **92**, 477–487.
- [27] S. Goedecker, M. Teter and J. Hutter, *Phys. Rev. B*, 1996, **54**, 1703–1710.
- [28] C. Hartwigsen, S. Goedecker and J. Hutter, *Phys. Rev. B*, 1998, **58**, 3641–3662.
- [29] G. Bussi, D. Donadio and M. Parrinello, *J. Phys. Chem. A*, 2007, **126**, 014101.
- [30] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ā. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *GaussianĀĹij09 Revision D.01*, Gaussian Inc. Wallingford CT 2009.
- [31] A. Becke, *Phys. Rev. A*, 1988, **38**, 3098–3100.
- [32] C. Lee, W. Yang and R. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.

- [33] P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623.
- [34] Y. Zhao and D. G. Truhlar, *J. Chem. Phys.*, 2006, **125**, 194101.
- [35] S. Ghisla and V. Massey, *Biochemical Journal*, 1986, **239**, 1–12.
- [36] J. Davies, N. L. Doltsinis, A. Kirby, C. D. Roussev and M. Sprik, *J. Am. Chem. Soc.*, 2002, **124**, 6594.
- [37] N. L. Doltsinis and M. Sprik, *Phys. Chem. Chem. Phys.*, 2010, **5**, 2612–2618.
- [38] L. Bernasconi, E. Baerends and M. Sprik, *J. Phys. Chem. B*, 2006, **110**, 11444.
- [39] D. Chandler, *Introduction to modern statistical mechanics*, Oxford University Press, New York, 1987.
- [40] I. Ivanov, B. Chen, S. Raugei and M. L. Klein, *J. Phys. Chem. B*, 2006, **110**, 16365–6371.

5 Supplementary Information

5.1 Radial distribution functions

To quantify the solvent structure around the proton donating nitrogen atom of lumiflavin and the accepting first solvation shell water oxygen, we have computed the radial distribution functions (RDF) of the hydrogens around these nitrogen atoms (labeled N1 and N5) and oxygen atom (O_w), see also figure 4 in the paper. The graphs are shown in figure 5 for the deprotonation of the N5 site in the semiquinone state (labeled PT1A, top panels), the N1 deprotonation in the semiquinone state (PT1B, middle panels), and the N1 deprotonation in the fully reduced state (PT2, bottom panels). The different colours show the RDFs at different stages of the proton transfer reactions, with in black the initial flavin protonated state, in red an intermediate state in which the proton is (almost) equidistant from the donor and the acceptor atoms, and in green the final stage at which the second proton transfer, from first to second coordination shell, takes place.

The first steep peaks in the left-hand-side panels are due to the position of the transferring proton, which moves to the right (*i.e.* larger N-H distance) as the reaction proceeds. The smaller second peak is mainly due to the two hydrogens of the accepting water molecule, but overlaps with contributions due to other nearby hydrogens, in particular from a hydrogen that is donated to the accepting O_w by another first shell water molecule (see also figure 4 in the paper) and from one or two water molecules that flank the flavin near the donating N atom. As a result from this overlap, the integrals of the RDFs shown by dashed lines do not show a clear step for the number of hydrogens contributing to the second peak. For all three proton transfer reactions, the peak moves to the left and becomes somewhat more pronounced as the transfer proceeds, which is due to the shortening of the N- O_w distance.

At the accepting water oxygen (right panels), the first peaks are due to the two water hydrogens and the proton. The peak due to the latter moves to the left (*i.e.* shorter O_w -H distance) as the proton transfer occurs. The integral indicates a fourth hydrogen (note the right-hand-axis), which is due to a second shell water molecule that donates an H-bond to the O_w , except in the later stages of the PT2 reaction (bottom

right panel). Comparing the solvent structure of the accepting water molecule between the three proton transfer reactions, we find that in the PT2 case the water molecule is always somewhat under-coordinated as also illustrated in figure 4 in the paper. Why this under-coordination only occurs in the fully reduced state is not clear.

5.2 Free energy profiles using the addition H_3O^+ restraint

As described in the paper, we equilibrated the systems also in the product state, in which the flavin has donated its proton to the nearest water molecule, by using, in addition to the Δd constraint, a restraint that avoids that the two hydrogens, other than the transferred proton, leave the H_3O^+ ion. For this restraint, we used a coordination number variable of the number of hydrogens (other than the proton) within a cutoff distance to the oxygen, subject to a steep wall potential that would become effective if the coordination number would drop below 1.6 when one of the hydrogens would attempt to leave.

We have considered whether it would be a good idea to also perform the production runs with this restraint applied, in an attempt to overcome the problem that the free energy profiles can only be computed up to the point that the proton jumps away into the bulk water. Unfortunately, this approach does not work. The main reason is that the transport of the proton, after it reached the first solvent water and formed a hydronium ion, does not proceed as an intact H_3O^+ ion, but rather hops from the first water to next via an intermediate H_5O_2^+ complex (aka the Grotthus mechanisms), so that the first water maintains its H-bond with the flavin donor atom. As the restraint enforces the hydronium ion to remain as an intact H_3O^+ ion, the H-bond with the flavin is stretched when the Δd reaction coordinate is further increased. This results in a negative force (see the middle right panel in Figure 6, and therefore a too high free energy profile (see the top panel).

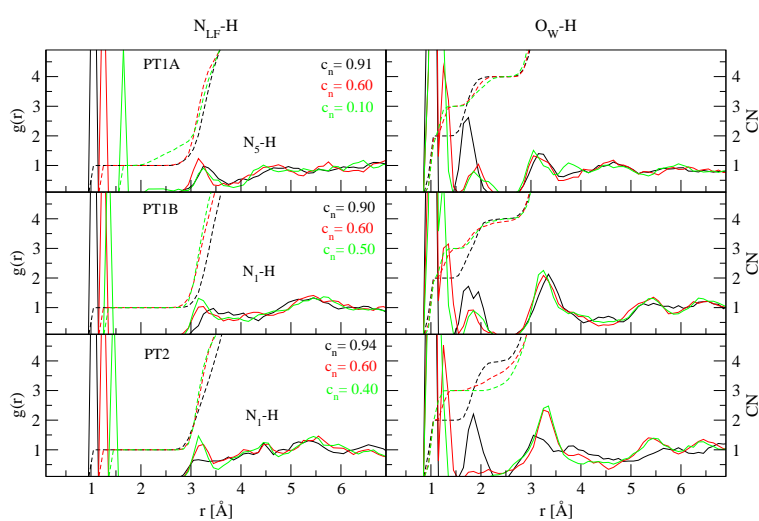


Figure 5: Radial distribution functions of the number of water solvent hydrogen atoms around the proton donating flavin nitrogen atoms (left panels) and around the proton accepting water oxygen atom (right panels), at different stages of the (de-) protonation reactions. Top-to-bottom: N5 deprotonation in the flavin semiquinone state (PT1A), N1 deprotonation in the semiquinone state (PT1B), and N1 deprotonation in the fully reduced state. The dashed lines show the integration of the RDFs from which the coordination numbers can be read, using the right-hand-side axis.

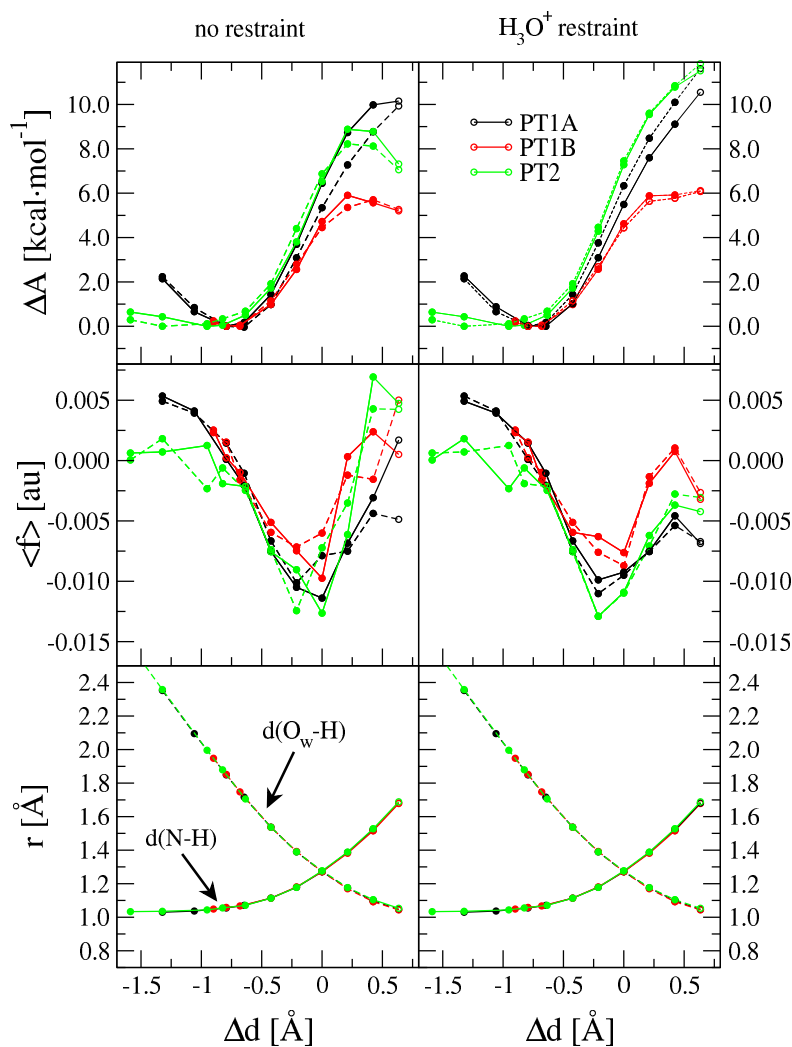


Figure 6: Top panels: free energy profiles for the three protonation reactions computed using the Δd coordinate and an additional H_3O^+ restraint (right panels) and without this restraint (left panels). Note that the latter is thus the same as shown in figure 4 in the paper. Middle panels: average force of constraint. Bottom panels: distance between the proton and the lumiflavin N1 or N5 nitrogen atom and distance between the proton and the accepting water molecule oxygen atom.

DFT-based molecular dynamics simulations predict the acidity of lumiflavin in different redox states.

