

# A molecular mechanics model for flavins

**Alexey Aleksandrov<sup>1\*</sup>**

<sup>1</sup>Laboratoire d'Optique et Biosciences (CNRS UMR7645, INSERM U1182), Ecole Polytechnique, IP Paris, 91128 Palaiseau, France

\*Corresponding author: [alexey.aleksandrov@polytechnique.edu](mailto:alexey.aleksandrov@polytechnique.edu)

Running title: A force field model for flavins

Keywords: flavin; riboflavin; flavin adenine nucleotide; flavin mononucleotide; flavoprotein; CHAMM; force field; molecular dynamics;

## **ABSTRACT**

Flavin containing molecules form a group of important cofactors that assist a wide range of enzymatic reactions. Flavins use the redox-active isoalloxazine system, which is capable of one- and two-electron transfer reactions and can exist in several protonation states. In this work, molecular mechanics force field parameters compatible with the CHARMM36 all-atom additive force field were derived for biologically important flavins, including riboflavin, flavin mononucleotide and flavin adenine dinucleotide. The model was developed for important protonation and redox states of the isoalloxazine group. The partial charges were derived using the CHARMM force field parametrization strategy, where quantum mechanics water-solute interactions are used to target optimization. In addition to monohydrate energies and geometries, electrostatic potential around the compound was used to provide additional restraints during the charge optimization. Taking into account the importance of flavin-containing molecules special attention was given to the quality of bonded terms. All bonded terms, including stiff terms and torsion angle parameters, were parametrized using exhaustive potential energy surface (PES) scans. In particular, the model reproduces well the butterfly motion of isoalloxazine in the oxidized and reduced forms as predicted by quantum mechanics in gas-phase. The model quality is illustrated by simulations of four flavoproteins. Overall, the presented molecular mechanics model will be of utility to model flavin cofactors in different redox states.

## INTRODUCTION

Flavoproteins participate in a wide range of biological functions and perform redox reactions and non-redox reactions, such as transferase, lyase, isomerase, and ligase reactions.<sup>1</sup> The flavoproteins have been extensively studied by experiments and represent one of the best-characterized families of proteins.<sup>2</sup> Flavoproteins have either flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD) as a cofactor, and more rarely both cofactors.<sup>3</sup> Apart from FAD and FMN, another biologically important flavin is riboflavin (vitamin B2) that serves as the precursor for FMN and FAD in almost all organisms.<sup>4</sup> FMN and FAD utilize in enzymatic reactions the redox-active isoalloxazine system, capable of one-electron and two electron transfer reactions, which can be coupled to the proton transfer. Flavins can also act in synergy with other cofactors such as iron–sulfur clusters, heme, molybdopterin, or thiamine diphosphate, increasing the diversity of reactions.<sup>1</sup> Time-resolved spectroscopy allowed gaining a wealth of information on the protein dynamics of flavoproteins, on time-scales directly accessible by Molecular Dynamics simulations.<sup>5-7</sup>

Molecular modeling techniques based on force fields have been applied to study different aspects of flavins in solvent<sup>8</sup> and in protein complexes.<sup>9-13</sup> The major requirement for such computer simulations is the existence of an adequate molecular mechanics force field model. Moreover, the underlying force field models determine the quality of the results of simulation studies relying on such models. A recent study compared the performance of different force field models including CHARMM36, Amber and OPLS in simulations of protein complexes with flavin mononucleotide.<sup>14</sup> It was found that the performance of the CHARMM36 force field was similar to OPLS in reproducing experimental 3J couplings, but worse in reproducing the effect of amino residue substitutions by alanine on FMN binding.<sup>14</sup> However, it should be noted that for the CHARMM simulations, atom types and force field parameters were assigned by the MATCH Web server.<sup>15</sup> Thus, parameters for similar compounds existing in the CGenFF force field were used, which were not necessarily optimal for flavin. Moreover, such “similar” compounds may not be available in CGenFF for the unique tricyclic heterocycle isoalloxazine structure of flavin.

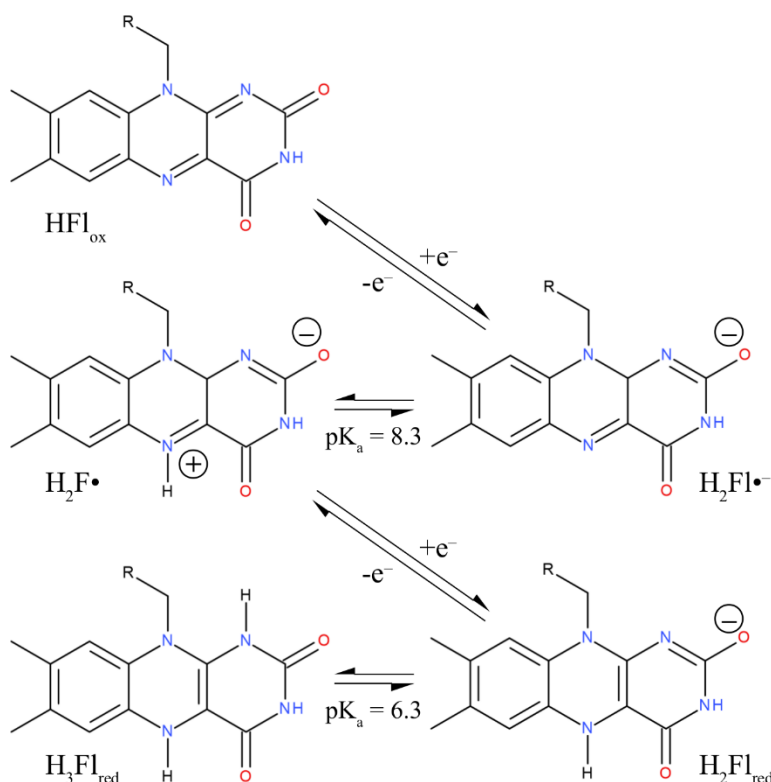
While the CHARMM force field model exists for a wide range of molecules including proteins<sup>16-18</sup> and nucleotides,<sup>19</sup> flavins, despite their importance, have been omitted in the force field development. The present study represents a systematic development of a force field model for flavins in important protonation and redox states. Using the standard CHARMM parametrization protocol, the force field is developed to be compatible with the CHARMM36 force field for proteins and nucleic acids,<sup>16-18</sup> and CGenFF force field for small molecules.<sup>20</sup> To summarize, the model developed in this work is suitable to investigate interactions of flavins with a wide range of flavoprotein and can be used as to parametrize other flavins, such as covalently linked to proteins.<sup>21-22</sup>

## RESULTS

### Set of parametrized molecules

We parametrized three redox forms of flavin shown in Figure 1. For each redox form of flavin the most important protonation states were considered at the physiological pH of 6.5. In particular, the flavin group in the

reduced form can be protonated and deprotonated on the N1 atom. The associated  $pK_a$  is 6.3 pH units suggesting that both forms coexist in solution at the physiological pH.<sup>23</sup> The  $pK_a$  of the N5 nitrogen of flavin is 8.3 pH units in the radical form.<sup>23</sup> When bound to a protein and in the absence of a proton donor, the anionic radical form can exist for a significant period of time after the oxidized form of flavin is reduced by the transfer of one electron.<sup>7</sup> Thus, this protonation state of the flavin radical was also parametrized in this work. We considered the flavin group in the following important cofactors: lumiflavin, riboflavin, flavin mononucleotide (FMN), and flavin adenine dinucleotide (FAD). The flavin group common to these molecules was first parametrized in lumiflavin, and then parameters were transferred to other molecules in accord with the development of the standard CHARMM force field.



**Figure 1.** Three oxidation forms of the flavin group parametrized in this work. The most important conjugate acid-bases of the flavin forms are shown.

### Charge optimization

The CHARMM atomic charges were derived targeting water-compound minimum interaction energies and geometries along with the dipole moment magnitude and orientation, and electrostatic potential. In what follows the charge optimization is discussed in detail only for lumiflavin in the oxidized and neutral reduced forms. The statistics for water-compound interactions for all compounds is given in Table 1. Empirical and *ab initio* interaction energies and distances are given in Table S1-S5 in the Supplementary Information. The Root Mean Square (RMSD) deviation for interaction energies with the optimal set of charges is  $0.34 \text{ kcal}\cdot\text{mol}^{-1}$  and  $0.42 \text{ kcal}\cdot\text{mol}^{-1}$  for lumiflavin in the oxidized and the neutral reduced forms respectively. With the initial charges provided by ParamChem the RMSD is  $2.70 \text{ kcal}\cdot\text{mol}^{-1}$  and  $0.61 \text{ kcal}\cdot\text{mol}^{-1}$  for the oxidized and neutral reduced

forms respectively. The RMS deviation for interaction distances is 0.20 and 0.17 Å with the optimal set of charges and 0.95 and 0.18 Å with the initial set for the oxidized and reduced forms of lumiflavin respectively. Overall, the optimization of charges improved water-compound interactions for all compounds.

**Table 1.** Statistics of water-compound interactions. All values are given for the optimal and initial sets of atomic charges.

Compound	energy (kcal·mol <sup>-1</sup> )		distance (Å)	
	RMSD	max  error	RMSD	max  error
lumiflavin/HFl <sub>ox</sub>	0.34/2.70	0.96/5.61	0.20/0.95	0.39/2.71
lumiflavin/H <sub>3</sub> Fl <sub>red</sub>	0.42/0.61	0.69/1.30	0.17/0.18	0.34/0.38
lumiflavin/H <sub>2</sub> Fl <sub>red</sub> <sup>-</sup>	0.24/1.60	0.54/2.73	0.27/0.28	0.89/0.93
lumiflavin/H <sub>2</sub> Fl <sup>-</sup>	0.81/3.82	2.98/9.73	0.31/0.29	0.56/0.57
lumiflavin/H <sub>2</sub> Fl <sup>-</sup>	0.49/4.21	1.18/5.88	0.18/0.37	0.28/1.45

Empirical and *ab initio* dipole moments and electrostatic potentials are given in Table 2. To be consistent with the protocol used to derive the CHARMM force field for proteins and small molecules the dipole moment was included only for neutral compounds. The resulting CHARMM dipole moment for lumiflavin in the oxidized form of 11.48 Debye is in good agreement with the overestimated QM dipole moment of 11.77 Debye demonstrating the expected overestimation. The dipole moment is also significantly improved relative to the initial estimate of 3.85 Debye with the ParamChem charges. The orientation of the dipole moment is also improved from 105.6° to 0.7° with the initial and optimal set of charges respectively. For the reduced form of lumiflavin the CHARMM dipole moment of 4.88 Debye with the optimal set of charges is smaller than the corrected QM dipole of 6.39 Debye and is not improved in comparison with 5.00 Debye computed with the ParamChem charges. However, the direction of the dipole moment is improved relative to the initial ParamChem dipole moment from 20.9° to 0.1° with the optimal set of charges.

**Table 2.** Empirical and *ab initio* dipole moments and electrostatic potentials. The corrected dipole moment was used as an additional restraint only for the neutral compounds during the atomic charge optimization.

Compound	dipole moment, (Debye)	<sup>a</sup> angle°	<sup>b</sup> RMSD esp
	<i>ab initio</i> /optimal/initial	optimal/initial	optimal/initial
lumiflavin/oxidized	11.77/11.48/3.85	0.7/105.6	3.27/15.87
lumiflavin/reduced	6.39/4.88/5.00	0.1/20.9	1.00/3.58
lumiflavin/ H <sub>2</sub> Fl <sub>red</sub> <sup>-</sup>	12.25/12.12/21.90	1.1/9.8	1.48/17.77
lumiflavin/H <sub>2</sub> Fl <sup>-</sup>	11.61/9.04/18.72	0.0/46.3	0.72/31.98
lumiflavin/HFl <sup>-</sup>	24.33/24.57/9.33	0.3/73.1	1.28/60.77

<sup>a</sup>angle between the *ab initio* and empirical dipole moment vectors; <sup>b</sup>RMS deviation in kcal·mol<sup>-1</sup>·Å<sup>-1</sup> between *ab initio* and empirical electrostatic potential.

The dipole moment for charged compounds was not targeted in the charge fitting as in the previous force

field determination works, since it depends on the system of coordinates.<sup>24</sup> In principle, if the molecule is considered in the same system of coordinate and in the same orientation, the *ab initio* and empirical dipole moments can be nevertheless compared. The dipole moment for charged compounds computed with the empirical and QM models is given in Table 2. In both *ab initio* and empirical calculations, the molecule was positioned in the same “standard” orientation used by the Gaussian software. Though it was not explicitly included in the charge fitting, the CHARMM model reproduces well the unscaled *ab initio* dipole moment for the charged molecules parametrized in this work. For example, in the case of the reduced deprotonated form of lumiflavin the QM and MM dipole moment with the optimal set of charges is 12.25 Debye and 12.12 Debye respectively. However, the dipole moment of 21.90 Debye is significantly overestimated with the initial ParamChem set of charges. The direction of the dipole moment is also more consistent with the optimal charge set with an angle between the QM and MM dipoles equal to 1.1°, in contrast to 9.8° obtained with the initial charges.

Electrostatic potential was included as an additional restraint during charge fitting to provide a better charge distribution in the model compound in accord with the previous study.<sup>25</sup> In all cases, including anionic model compounds, electrostatic potential is significantly improved relative to the initial values. For the oxidized form of lumiflavin the RMS deviation between MM and QM electrostatic potential is 3.27 and 15.87 kcal·mol<sup>-1</sup>·Å<sup>-1</sup> with the optimal and initial set of charges respectively. For the neutral reduced form, RMSD of electrostatic potential is 1.00 and 3.58 kcal·mol<sup>-1</sup>·Å<sup>-1</sup> with the optimal and initial set of charges respectively. Targeting the QM electrostatic potential was found particularly important for charge compounds, since the number of probe water interactions was fewer than for neutral compounds, due to the dominant contribution of the net charge to water interactions. Finally, to test the dependence of the results on the QM theory employed to optimize geometries and compute the dipole moment, geometry optimization and dipole moment calculations were also performed using the MP2 functional. The results are given in Table S10 in the Supplementary Information. Compound geometries optimized using the B3LYP and MP2 functionals are very similar, with the RMS deviation lower than 0.1 Å. Consequently, the dipole moments computed using the B3LYP and MP2 optimized structures also very similar, demonstrating that the B3LYP functional is appropriate to optimize non-bonded parameters of flavins.

## Optimization of bonded terms

### *Stiff degrees of freedom*

In this work, all bonded terms including stiff terms were parametrized based on the reproduction of energies and geometries of potential energy surface (PES) scans. We first present results for the parametrization of stiff terms. Selected results are discussed in detail only for lumiflavin in the oxidized form and in the fully reduced neutral form. The results for empirical and *ab initio* structures and conformation energies are summarized in Table 3. To parametrize bonded terms of lumiflavin in the oxidized form a total of 324 optimizations of PES scans were performed. The agreement between QM and MM energies of PES scans for the oxidized and reduced forms of lumiflavin is illustrated in Figure S1 in the Supplementary Information. Overall, a good agreement

between QM and MM energies was achieved with the RMS deviation for energies is 0.11 kcal·mol<sup>-1</sup> and 0.50 kcal·mol<sup>-1</sup> for the oxidized and reduced forms respectively. The equilibrium values for bonds and angles for the structure optimized with the MM model are also in good agreement with the QM structure. Generally, the deviation for bonds is less than 0.02 Å; for covalent angles the RMSD is lower than 2.4° and for torsions the RMS angle deviation is in the range of 0.2 to 6.2°. The RMS deviation between *ab initio* and empirical geometries is 0.04 Å and 0.10 Å for lumiflavin in the oxidized and reduced forms respectively. Overall, with the optimized bonded parameters, the empirical model reproduces the QM equilibrium geometries very well.

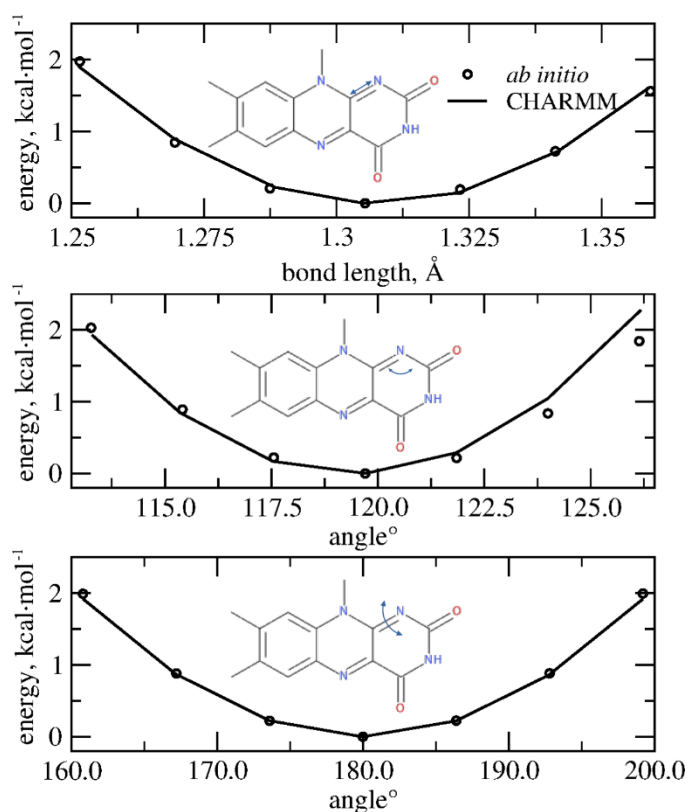
**Table 3.** Comparison between empirical and *ab initio* optimized geometries for equilibrium structures and energies of PES scans

compound	RMSD		
	<sup>a</sup> energy (kcal·mol <sup>-1</sup> )	<sup>b</sup> geometry (Å)	<sup>c</sup> bond (Å)/angle (°)/dihedral (°)
lumiflavin/HFl <sub>ox</sub>	325/0.11	0.05/0.04	0.02/2.0/0.2
lumiflavin/H <sub>3</sub> Fl <sub>red</sub>	244/0.50	0.19/0.10	0.02/1.9/6.2
lumiflavin/H <sub>2</sub> Fl <sub>red</sub> <sup>-</sup>	39/0.18	0.17/0.13	0.02/2.4/4.6
10-ethyl-flavin/ HFl <sub>ox</sub>	122/0.40	0.07/0.07	0.02/1.5/1.6
10-ethyl-flavin/ H <sub>3</sub> Fl <sub>red</sub>	125/0.33	0.10/0.06	0.02/1.6/2.4

<sup>a</sup>number of points on PES used to optimize bonded parameters and RMS deviation between QM and MM energies;

<sup>b</sup>RMS deviation between QM and MM optimized equilibrium structures for all atoms and heavy atoms of the compound; <sup>c</sup>RMS deviation for bond length, angles and dihedrals.

The agreement for energies is demonstrated in Figure 2 for selected degrees of freedom in the oxidized form of lumiflavin with the optimal set of bonded parameters. All energies are within 2.0 kcal·mol<sup>-1</sup> of the minimum energy, which we achieved by using an additional calculation to determine the range of values (Equation 1 and 2 in the Method section) for each bonded term. Overall, the model reproduces well the QM equilibrium conformations of the model compounds as well as QM energies of various deformations along parametrized degrees of freedom.



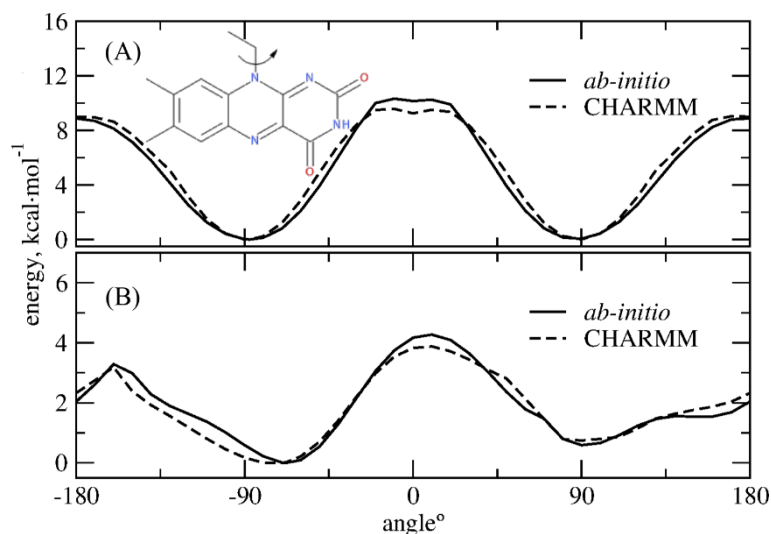
**Figure 2.** PES scans for selected degrees of freedom in lumiflavin in the oxidized form. The PES scan was performed for a bond (top), for a covalent angle (middle), and for a torsion (bottom). The corresponding term is shown by an arrow.

### *Soft dihedral angles*

These terms include flexible dihedrals along which energy variations are within a range that is sampled in MD simulations, hence accurate treatment of these dihedral PES is important to describe adequately the conformational space of molecules. The parameters of these flexible dihedrals were determined based on points of PES scans. Here we illustrate the quality of the parameters that define the most important conformational changes. In particular, the torsion angle that defines rotations of 10-ethyl group in 10-ethyl-flavin was parametrized to reproduce the complete rotation of 360°. This torsion defines orientation of the flavin group relative to other groups, such as adenine in FAD. The results of the PES scan are shown in Figure 3. In the oxidized form, the 10-ethyl group has two symmetric minimum positions  $-90^\circ$  and  $90^\circ$  relative to the plane of the isoalloxazine group separated by barriers at  $0^\circ$  and  $180^\circ$ . In the protein crystal structures with FMN and FAD in the oxidized form (PDB access codes: 2XOD<sup>26</sup> and 5GV7<sup>27</sup>) this torsion angle is  $-83.1^\circ$  and  $92.3^\circ$  respectively, in agreement with the QM and MM PES scans in vacuum. In the reduced form, the torsion of the 10-ethyl group in vacuum has two asymmetric minima at  $-67.8^\circ$  and  $90.0^\circ$ . The PES near the minimum at  $90^\circ$  is shallow, suggesting that the reduced form of flavin is more flexible around this torsion in protein complexes than in the oxidized form. Indeed, in the protein crystal structures with FMN and FAD in the reduced form (PDB access codes: 4UTK<sup>28</sup> and 4U2S<sup>29</sup>) this

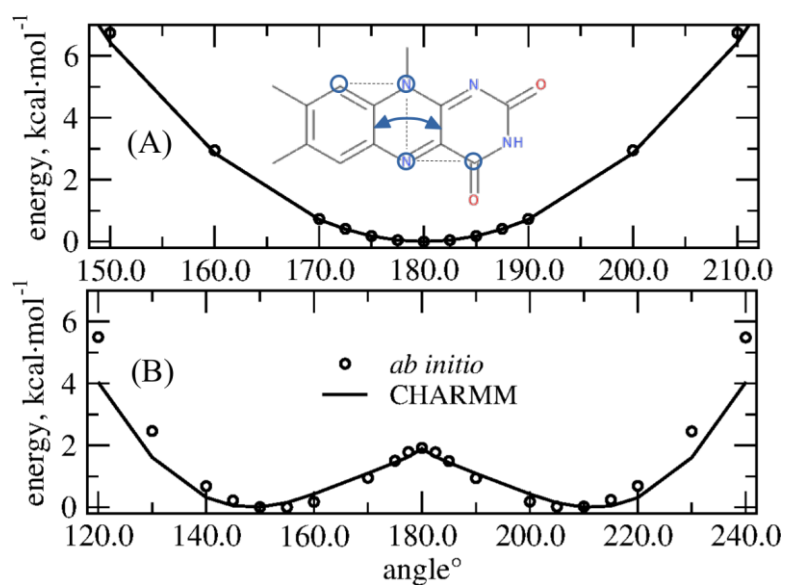


torsion is  $-66.2^\circ$  and  $102.9^\circ$ . The first value is close to  $-67.8^\circ$ , but the second torsion angle is different from  $90.0^\circ$ , in agreement with the shallow region of PES.



**Figure 3.** PES scan of the 10-ethyl group rotation in 10-ethyl-flavin in the oxidized form (A) and in the fully reduced form (B). *Ab initio* energies were computed at the B3LYP/6-31G(d) level. The arrow shows the dihedral angle.

Finally, we examine the bending (also called the butterfly) motion of the isoalloxazine group of flavin. The bending motion of flavin was measured by the torsion defined by atoms C4, N5, N10 and C9 of the isoalloxazine group. It is worth to note that the deformation along this degree of freedom is contributed by several dihedral energy terms and there is no particular dihedral angle term associated with this torsion. Thus, it illustrates the overall quality of the force field parameters. Figure 4 shows *ab initio* and empirical energies of PES scans along the torsion for the oxidized and fully reduced form of lumiflavin. The oxidized form is characterized by a single minimum at  $180.0^\circ$  and high energies for relatively small deformations. In contrast, the reduced form has two shallow minima connected by a small barrier, which is explained by the electron-rich  $\pi$  molecular orbital system of the fully reduced form.<sup>23</sup> The force field model reproduces well positions of the two minima in the reduced flavin. The empirical energy barrier of  $1.9 \text{ kcal}\cdot\text{mol}^{-1}$  is in good agreement with the QM energy barrier of  $1.8 \text{ kcal}\cdot\text{mol}^{-1}$ , which is important to reproduce structural transitions in simulations with the reduced form of flavin.



**Figure 4.** PES scan of the butterfly motion in lumiflavin in the oxidized form (A) and in the fully reduced form (B). *Ab initio* energies were computed at the B3LYP/6-31G(d) level. The torsion angle shown by the arrow is defined between atoms in circles.

### Molecular Dynamics simulations of protein complexes

To illustrate the quality of the model, Molecular Dynamics (MD) simulations of protein complexes with bound flavins were performed. The protein structures with a high-to-medium resolution were chosen for MD simulations. MD simulations were performed for 50 nanoseconds. The resulting RMS deviation between simulation and experimental structures is given in Table 4. Figure S2 in the Supplementary Information demonstrates the evolution of RMSD in the MD simulations. The RMS deviation for heavy atoms of flavin cofactors in all simulations is in the range of 0.29 Å to 0.53 Å and in all cases lower than the RMS deviation for unrestrained protein backbone atoms, demonstrating that the model performs as good as the standard CHARMM force field for proteins in protein simulations.

**Table 4.** Root Mean Square (RMSD) deviation in molecular dynamics simulations. Standard deviations are given in parenthesis.

cofactor	PDB ref. code	<sup>a</sup> Protein RMSD (Å)	<sup>b</sup> Cofactor RMSD (Å)
FMN <sub>ox</sub>	2XOD	0.53 (0.10)	0.45 (0.10)
FAD <sub>ox</sub>	5GV7	0.54 (0.05)	0.53 (0.11)
FMN <sub>red</sub>	4UTK	0.51 (0.10)	0.29 (0.05)
FAD <sub>red</sub>	4U2S	0.37 (0.02)	0.32 (0.05)

<sup>a</sup>RMSD was computed for unrestrained backbone atoms after superposition on the experimental structure; <sup>b</sup>RMSD was computed based on heavy atoms of the cofactor.

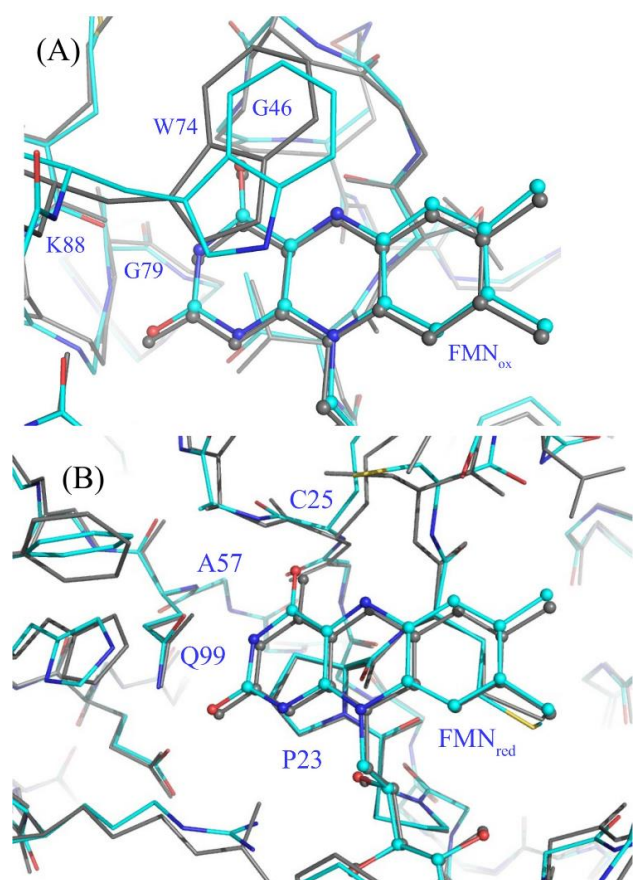
The distances between heavy atoms participating in hydrogen bonds are given in Table 5 for MD

simulations of flavin mononucleotide FMN<sub>ox</sub> bound to the NrdI protein (PDB access code 2XOD). The distances for MD simulations with FAD<sub>ox</sub>, FMN<sub>red</sub>, FAD<sub>red</sub> are given in Table S6-S8. All distances observed in MD simulations are within the standard deviation of the corresponding distances observed in the experimental structures. Overall, the interactions between flavins and the proteins are well reproduced. The superposition of the experimental structures (PDB access codes 2XOD and 4UTK) with the average position of atoms from 50 ns MD simulations is shown in Figure 5. To compute the average position, 2,500 structures saved every 10 ps from 50 ns MD simulations were superimposed on the experimental structure based on the backbone atoms and were averaged. Overall, the model reproduces well the structure of FMN and FAD in the protein complexes in the oxidized and reduced forms.

**Table 5.** Selected distances (Å) in molecular dynamics simulations of flavin mononucleotide FMN<sub>ox</sub> in complex with the NrdI protein (PDB entry 2XOD)

<sup>a</sup> atom pair	X-Ray str.	MD simulation
NG46...O4 <sub>Fl</sub>	2.83	3.16 (0.27)
OM77...N3 <sub>Fl</sub>	3.05	2.98 (0.18)
NG79...O2 <sub>FL</sub>	2.84	3.05 (0.23)
NN71...O2 <sub>FL</sub>	2.92	3.34 (0.27)
NN71...N1 <sub>FL</sub>	3.23	3.48 (0.23)
OT42...O2' <sub>Fl</sub>	2.70	2.82 (0.15)

<sup>a</sup>protein atoms (left) are labeled by the amino acid to which they belong



**Figure 5.** The experimental structure (in gray) vs the average structure (in color) observed in MD simulations with (A) FMN<sub>ox</sub> in complex with the NrdI protein (PDB access code 2XOD); and (B) with FMN<sub>red</sub> in complex with xenobiotic reductase A (PDB access code 4UTK).

As an additional test, the most important soft dihedral angles parametrized in this work that define the geometry of the flavin group were further investigated. The torsion angles observed in MD simulations are given in Table 6. The angle between atoms C4, N5, N10, and C9 defining the butterfly motion of the isoalloxazine group is well reproduced in MD simulations with flavin in the oxidized and reduced forms. For example in MD simulations with FMN<sub>red</sub> this angle is 153.8° in the experimental structure 4UTK and 156.6±6.6° in MD simulations, which is within the standard deviation from the experimental value. The torsions between atoms C2', C1', N10, and C9A, and C1', N10, C10, and N1 defining the position of the flavin group relative to the rest of molecule are also reproduced within the standard deviation. Overall, the force field model reproduces well the torsions between atoms, which define the overall orientation of the isoalloxazine group relative to the rest of molecule.

**Table 6.** Flavin dihedrals observed in molecular dynamics simulations and experimental structures. Standard deviations are given in parenthesis.

---

C4-N5-N10-C9	C2'-C1'-N10-C9A	C1'-N10-C10-N1
--------------	-----------------	----------------

---

Cofactor	PDB ref. code	X-ray str.	MD simul.	X-ray str.	MD simul.	X-ray str.	MD simul.
FMN <sub>ox</sub>	2XOD	-175.2	-175.5 (3.4)	-83.1	-79.1 (7.5)	-2.5	0.5 (8.2)
FAD <sub>ox</sub>	5GV7	176.3	177.5 (4.7)	92.3	90.9 (6.4)	-77.4	-78.3 (4.8)
FMN <sub>red</sub>	4UTK	153.8	156.6 (6.6)	-66.2	-69.6 (7.6)	10.4	20.7 (11.4)
FAD <sub>red</sub>	4U2S	-164.1	-153.7 (6.3)	102.9	90.4 (7.5)	4.7	12.6 (8.9)

## Discussions

The present study represents a systematic development of a force field model for flavin cofactors, including riboflavin, flavin mononucleotide and flavin adenine dinucleotide in the most important protonation and redox states. The parametrization was performed to be consistent with the standard method used to develop the CHARMM additive force field. The model is, thus compatible with the other components in the CHARMM force field. The parameters were optimized targeting QM data and validated against experimental structural data. The initial guess for atomic charges was obtained from the CGenFF force field. The charges were then adjusted to reproduce interaction energies and distances of a large number of solute-water interactions, which was important to parametrize the large heterocycle isoalloxazine group. Targeting solute-water interactions was also important to maintain the balance between flavin interactions with solvent and protein residues. In addition, the model reproduces the magnitude and direction of the *ab initio* dipole moment as well electrostatic potential. Though the dipole moment for charged compounds was not targeted, the model reproduces the QM dipole moment for the anionic molecules. Including the QM electrostatic potential in the charge optimization, in accord with the previous study,<sup>25</sup> was found helpful to obtain a better charge distribution in the large isoalloxazine group.

A special emphasis was given to the quality of bonded parameters, with all bonded parameters, including soft torsions and stiff harmonic terms adjusted using computationally intensive PES scans. Initial parameters were provided by ParamChem, with parameters previously not existing in the CGenFF force field assigned values by analogy. The bonded parameters were then optimized to reproduce PES energies and geometries. The model reproduces the QM geometry for flavin in the oxidized and fully reduced and important protonation states. It is worth to note the quality of the model to reproduce PES for soft torsions. In particular, the force field model reproduces well the bending motion in the oxidized and reduced forms of flavin.

Model validation was based on molecular dynamics simulations of four protein complexes containing the flavin group as a cofactor. The result demonstrate that the model reproduces well interactions between the flavin group and proteins as well as the flavin conformation. Importantly, the force field model reproduces the flavin plasticity and conformational difference between the oxidized and reduced forms as observed in the experimental crystal structures.

## METHODS AND MATERIALS

### Parametrization Scheme

A force field of the additive all-atom CHARMM form was adopted in this work.<sup>30</sup> The CHARMM force field form is well described in other studies and reviews,<sup>30</sup> and hence is not discussed in detail here. The atom types were adopted from the CGenFF force field.<sup>20</sup> The ParamChem web server (<https://cgenff.paramchem.org>) was used to assign existing atomic types and to obtain initial guesses for the parameters of model compounds.<sup>31-32</sup> Parameters that were assigned zero “penalty” by ParamChem, i.e. already optimized in CGenFF, were not considered for optimization in the present study. Parameters of the Lennard-Jones potential were taken directly from the CGenFF force field and were not optimized. The atom names from the Protein Data Bank were used throughout this work. In particular, the atoms C4A and C5A on the isoalloxazine group in riboflavin and FMN were renamed in FAD to C4X and C5X respectively, to be unique and different from the C4A and C5A atoms of the adenine group.

### Determination of the intermolecular force field parameters

The intermolecular energy is contributed from Coulomb and Lennard–Jones terms. Consistent with the development of the CHARMM force field, atomic charges were optimized targeting interactions between the model compound and individual water molecules, and the dipole moment of the model compound. Quantum mechanical (QM) electrostatic potential was also included as additional target data in the charge fitting similar to the other work.<sup>25</sup> The charge optimization was performed on the compound structures optimized with the B3LYP functional<sup>33</sup> and 6-31G(d) basis set.<sup>34</sup> Gaussian09 (Gaussian, Inc, Wallingford CT, 2009) was used for all QM calculation.

Atoms of the model compound that can participate in hydrogen bonds were probed by individual water molecules placed in idealized linear orientations.<sup>16</sup> Different rotations of the water molecule were considered around the interaction axis: for polar atoms, the complex was calculated every 45° or 90° of the water probe rotation, and only one or two orientations for non-polar atoms. Each water-compound structure was then optimized by varying the interaction distance, to find the energy minimum for the water position. The B3LYP/6-31G(d) geometry was used for the compound and TIP3P model geometry for the water molecule. The single angle defining the orientation of the water molecule was held fixed during the optimization. Following the same protocol that was used to parametrize the CHARMM force field, calculations were done at the HF/6-31G(d) level.<sup>16, 20</sup> The interaction energy was calculated for the optimal distance. In accord with the standard CHARMM parametrization protocol,<sup>16</sup> the *ab initio* interaction energies were scaled by an empirical factor of 1.16 only for neutral polar compounds and the HF/6-31G(d) minimum interaction distance was corrected by subtracting 0.2 Å for all polar interactions involving neutral compounds. To parametrize radical states of flavin, water interaction calculations were performed with the B3LYP potential and unrestricted open-shell wavefunctions using the same 6-31G\* basis set. No correction of QM interaction energies and distances was made for the radical forms of flavin.

The molecular dipole moment, which is defined by the charge distribution, was used to provide additional

target data for the optimization of the atomic charges. The dipole moment was used only for the neutral compounds in the charge fitting; for charged molecules, the dipole moment was not targeted as in the previous works.<sup>24</sup> The dipole moment was calculated in vacuum at the HF 6-31G(d) level using the B3LYP optimized conformation. As in the previous studies, to account the molecular polarizability implicitly, the MM optimization targeted dipole moments overestimated by 30% with respect to the QM values for polar compounds.<sup>20</sup> During the charge optimization, both the magnitude and direction of the dipole moment were targeted.<sup>24</sup>

Since we consider only probe water interactions with a few proton donors and acceptors, QM water interaction data may not be sufficient to define partial charges on all atoms for large compounds, such as isoalloxazine parametrized in this work. This is especially problematic for charged molecules, for which the dipole moment is not targeted, and not all proton acceptors and donors can be probed due to the dominant contribution of the net charge to the water electrostatic interaction. For this reason, we included electrostatic potential to provide additional target data in charge fitting.<sup>25</sup> To remain consistent with the standard protocols used for the CHARMM and AMBER force fields, the QM electrostatic potential calculations were performed at the HF/6-31G\* level.<sup>25</sup> At each iteration during the charge optimization, the Root Mean Square deviation between QM and MM electrostatic potential was evaluated and added with the corresponding weight to the target function.

The charge optimization was performed with the C++ program that was used to parametrize a large number of modified nucleotides in the previous work.<sup>24</sup> The following terms were included with different weights in the target function: the RMS deviation between empirical and *ab initio* minimum interaction energies, the RMS deviation between *ab initio* and empirical minimum interaction distances, the absolute difference between the norms of the empirical and *ab initio* dipole moments, the angle between the empirical and *ab initio* dipole moments, the RMS deviation between *ab initio* and empirical electrostatic potential and a term associated with restraints on the charges. The latter term was introduced to prevent large deviations from the starting guess for charges. Charges of symmetrical atoms had identical values during the charge optimization. The initial partial charges were obtained from the ParamChem online server. Charges that were already optimized in the CGenFF force field, for example for adenosine, were not modified in this work. Finally, charges of aliphatic groups were not optimized, in accord with the standard CHARMM method with all methyl groups having a charge of  $0.09e$  on protons. The LJ parameters were not considered to optimization.

### **Optimization of flexible dihedral parameters**

Dihedrals within a molecule can be considered in two classes, stiff and rotatable. Terms associated with stiff dihedrals characterized by a single minimum and high energy for small deformations. Rotatable dihedrals are characterized by a shallow energy surface, and thus may undergo large fluctuations during simulations and may have several minima. Since the molecule can perform large conformational motions along these degrees of freedom, the treatment of this class of dihedral terms is crucial. To parametrize these terms potential energy surface (PES) scans were performed on the torsions, in which torsion angles were scanned in the range from  $-180^\circ$  to  $180^\circ$  in  $10^\circ$  increments. Each conformation for the MM calculations was extracted from the QM scan and minimized

with a harmonic restraint force constant of  $5 \cdot 10^4$  kcal·mol·radian<sup>-2</sup> on the target torsion. The MM dihedral parameters were optimized to achieve a minimum deviation between the QM and MM surfaces only in the lower energy regions, PES points with the QM energy of more than 7 kcal·mol above the minimum energy were not considered.

### Determination of bonded harmonic energy terms

Parameters for the bonded terms described by harmonic potentials, i.e. bonds, angles, Urey-Bradley, and improper dihedrals, as well as for the stiff dihedrals were optimized using the following protocol. Only parameters not optimized in the CGenFF force field, i.e. with the non-zero “penalty”, were optimized. The initial guess was provided by the ParamChem online server. In this work, we used the same method used to parametrize the terms associated with soft dihedrals in CHARMM described in the previous section. In particular, we perform an adiabatic PES scan for each degrees freedom that has adjustable parameters in the force field. It is worth to note that a similar method is also used in the original CHARMM force field to determine force constants by three-point potential energy scans, when the assignment of contributions of the internal coordinates to the vibrations is ambiguous.<sup>24, 35</sup> This method, in contrast to the original method based on the Normal Mode Analysis,<sup>35</sup> departs from the local minimum, and is more computationally expensive. For example, to parametrize lumiflavin in the oxidized form in this study, 324 optimizations were performed of lumiflavin deformed along the parametrized degrees of freedom.

Since the PES scans were performed by varying one stiff degree of freedom, the potential energy can become very high, even with small deformations. Such high-energy points on PES are not sampled during typical MD simulations and should not be considered in the optimization of boded parameters. To use geometries of the compound within a range of energies that can be reached during MD simulations we propose and use the following protocol. First, using initial values of distortions, force constants of energy terms are estimated. The initial values of the distortion,  $\Delta x$  are 0.06 Å, 4.0°, 25.0° and 25.0° for the bond length, valence angle, dihedral and improper angle respectively. The compound geometry is optimized with a chosen internal coordinate constrained to  $x = x_0 - \Delta x$ , where  $x_0$  is the equilibrium value. The optimization is performed at the B3LYP/6-31G(d) level. The initial force constant of each energy term being parametrized is estimated using the following equation:

$$k = 2 (E(x) - E_0) / \Delta x^2, \quad [\text{Eq. 1}]$$

where  $E_0$  and  $E(x)$  is the minimum energy and energy of the deformed structure. Using this force constant, the new range  $\Delta x'$  of  $x$  values was estimated with the maximum energy of  $|E(x) - E_0| < \Delta E_{max}$  relative to the minimum energy  $E_0$ :

$$\Delta x' = \sqrt{2 \Delta E_{max} / k}, \quad [\text{Eq. 2}]$$

To optimize each bonded term we use seven points on PES equally spaced in the range of  $x \in [x_0 - \Delta x', x_0 + \Delta x']$ , including the minimum energy structure at  $x = x_0$ . In this work 2.0 kcal·mol<sup>-1</sup> was used for  $\Delta E_{max}$  in Equation 2. All PES scans were performed at the B3LYP/6-31G(d) level.

Once points on PES were computed, the equilibrium values of the MM parameters and force constants



were adjusted simultaneously using a C++ program based on the Powell and Amoeba minimization algorithms from Numerical Recipes.<sup>36</sup> Each conformation for the MM calculations was extracted from the QM scan and minimized with a harmonic restraint force constant of  $5 \cdot 10^4$  kcal·mol·Å<sup>-2</sup> or  $5 \cdot 10^4$  kcal·mol·radian<sup>-2</sup> on the target bond and angle respectively. At each optimization iteration of bonded parameters, PES scans were performed with CHARMM using a new set of MM parameters. The target function included RMS deviation between *ab initio* and empirical PES energies and geometries; and restraints to the initial set of parameters provided by the ParamChem server. The MM parameters were adjusted until the RMSD between the MM and the optimized QM geometries and energies could not be reduced significantly.

### **Molecular Dynamics simulations**

To evaluate the quality of the force field model of flavins, Molecular Dynamics (MD) simulations of protein complexes were performed. The protein complexes used in this work are summarized in Table S9. The crystal structures with a high to medium resolution were obtained from the Protein Data Bank (PDB). The simulation setup was similar to previous studies.<sup>37-38</sup> In brief, the simulations included protein residues within a 24 Å sphere, centered on the N5 atom of the bound FAD or FMN cofactor in the crystal structures. Protonation states of histidines were assigned by visual inspection and ideal stereochemistry; protonation states of other residues were assigned using PROPKA.<sup>39-40</sup> Protein atoms between 20 and 24 Å from the sphere's center were harmonically restrained to their experimentally determined positions. In the case of XenA, simulations also included residues and crystal waters from the second monomer of the dimeric structure within 24 Å of the N5 atom of the flavin group.<sup>28</sup>

In addition to crystal waters, a 75 Å cubic box of water was overlaid, and waters overlapping the protein, ligands and crystal water molecules were removed. Periodic boundary conditions were assumed; i.e. the entire 75-Å box was replicated periodically in all directions. All long range electrostatic interactions were computed efficiently by the particle mesh Ewald method,<sup>41</sup> and the appropriate number of potassium counterions were included to render the system electrically neutral. A smooth switching function was used to truncate all van der Waals interactions at the distance of 11 Å. Long range electrostatic forces were evaluated every 4 steps, while short-range non-bonded interactions were computed at each step. MD simulations were performed at constant room temperature and pressure, after 200 ps of thermalization. Constant pressure was maintained using the Berendsen pressure bath coupling<sup>42</sup> with the relaxation of 500 fs the compressibility parameter of liquid water. Constant temperature was maintained by simulating temperature coupling to a heat bath with a room temperature by correcting forces as implemented in the NAMD program.<sup>43</sup> The CHARMM36 force field was used for the protein<sup>44-45</sup> and the TIP3P model for water.<sup>16, 46-47</sup> The flavin cofactors were modeled using the force field model specifically developed in this work. Calculations were done with the NAMD program.<sup>43</sup> MD simulations of the protein complexes were continued for 50 nanoseconds.

### **Supporting Information**

Tables with water-compound interactions; selected distances observed in molecular dynamics simulations

of flavins in protein complexes; experimental protein structures used for Molecular Dynamics simulations. Figure with empirical vs *ab initio* energies of lumiflavin conformations; Figure with RMS deviation observed in MD simulations of the protein complexes; the developed force field model in the CHARMM format.

## Acknowledgements

We thank Prof. Hannu Myllykallio and Dr. Marten Vos for helpful discussions. This work was supported by grant ANR-18-CE44-0002. This work was performed using HPC resources from GENCI-CINES (Grant 2018-A0040710436).

## References

1. Macheroux, P.; Kappes, B.; Ealick, S. E., Flavogenomics-a genomic and structural view of flavin-dependent proteins. *FEBS J.* **2011**, *278* (15), 2625-34.
2. De Colibus, L.; Mattevi, A., New frontiers in structural flavoenzymology. *Curr. Opin. Struct. Biol.* **2006**, *16* (6), 722-8.
3. Hefti, M. H.; Vervoort, J.; van Berkel, W. J., Deflavination and reconstitution of flavoproteins. *Eur. J. Biochem.* **2003**, *270* (21), 4227-42.
4. Fischer, M.; Bacher, A., Biosynthesis of vitamin B2: Structure and mechanism of riboflavin synthase. *Arch. Biochem. Biophys.* **2008**, *474* (2), 252-65.
5. Sundstrom, V., Femtobiology. *Annu. Rev. Phys. Chem.* **2008**, *59*, 53-77.
6. Laptinok, S. P.; Bouzahir-Sima, L.; Lambry, J. C.; Myllykallio, H.; Liebl, U.; Vos, M. H., Ultrafast real-time visualization of active site flexibility of flavoenzyme thymidylate synthase ThyX. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (22), 8924-9.
7. Nag, L.; Sournia, P.; Myllykallio, H.; Liebl, U.; Vos, M. H., Identification of the TyrOH(\*+) Radical Cation in the Flavoenzyme TrmFO. *J. Am. Chem. Soc.* **2017**, *139* (33), 11500-11505.
8. van den Berg, P. A. W.; Feenstra, K. A.; Mark, A. E.; Berendsen, H. J. C.; Visser, A. J. W. G., Dynamic Conformations of Flavin Adenine Dinucleotide: Simulated Molecular Dynamics of the Flavin Cofactor Related to the Time-Resolved Fluorescence Characteristics. *The Journal of Physical Chemistry B* **2002**, *106* (34), 8858-8869.
9. Sattelle, B. M.; Sutcliffe, M. J., Calculating Chemically Accurate Redox Potentials for Engineered Flavoproteins from Classical Molecular Dynamics Free Energy Simulations. *J. Phys. Chem. A* **2008**, *112* (50), 13053-13057.
10. Schneider, C.; Suhnel, J., A molecular dynamics simulation of the flavin mononucleotide-RNA aptamer complex. *Biopolymers* **1999**, *50* (3), 287-302.
11. Dittrich, M.; Freddolino, P. L.; Schulten, K., When Light Falls in LOV: A Quantum Mechanical/Molecular Mechanical Study of Photoexcitation in Phot-LOV1 of *Chlamydomonas reinhardtii*. *The Journal of Physical Chemistry B* **2005**, *109* (26), 13006-13013.
12. Freddolino, P. L.; Gardner, K. H.; Schulten, K., Signaling mechanisms of LOV domains: new insights from molecular dynamics studies. *Photochemical & photobiological sciences: Official journal of the European Photochemistry Association and the European Society for Photobiology* **2013**, *12* (7), 1158-70.
13. Lans, I.; Seco, J.; Serrano, A.; Burbano, R.; Cossio, P.; Daza, M. C.; Medina, M., The Dimer-of-Trimers Assembly Prevents Catalysis at the Transferase Site of Prokaryotic FAD Synthase. *Biophys. J.* **2018**, *115* (6), 988-995.
14. Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L., Performance of Protein-Ligand Force Fields for the Flavodoxin-Flavin Mononucleotide System. *J. Phys. Chem. Lett.* **2016**, *7* (15), 3032-6.
15. Yesselman, J. D.; Price, D. J.; Knight, J. L.; Brooks, C. L., 3rd, MATCH: an atom-typing toolset for molecular mechanics force fields. *J. Comput. Chem.* **2012**, *33* (2), 189-202.
16. Mackerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Smith, J.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M., An all-atom empirical potential for molecular modelling and dynamics study of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586-3616.
17. MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., 3rd, Improved treatment of the protein backbone in empirical force fields. *J. Am. Chem. Soc.* **2004**, *126* (3), 698-9.
18. Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E.; Mittal, J.; Feig, M.; Mackerell, A. D., Jr., Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone  $\phi$ ,  $\psi$  and side-chain  $\chi(1)$  and  $\chi(2)$  dihedral angles. *J. Chem. Theory Comput.* **2012**, *8* (9), 3257-3273.
19. Foloppe, N.; MacKerell, A. D., All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *J. Comput. Chem.* **2000**, *21* (2), 86-104.
20. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comp Chem* **2010**, *31* (4), 671-690.
21. Mewies, M.; McIntire, W. S.; Scrutton, N. S., Covalent attachment of flavin adenine dinucleotide (FAD) and flavin mononucleotide (FMN) to enzymes: the current state of affairs. *Protein Sci.* **1998**, *7* (1), 7-20.
22. Heuts, D. P.; Scrutton, N. S.; McIntire, W. S.; Fraaije, M. W., What's in a covalent bond? On the role and formation of covalently bound flavin cofactors. *FEBS J.* **2009**, *276* (13), 3405-27.
23. Kyte, J., *Mechanism in Protein Chemistry*. Garland Science: 1995
24. Xu, Y.; Vanommeslaeghe, K.; Aleksandrov, A.; MacKerell, A. D.; Lennart Nilsson, J. L., Additive CHARMM force field for

- naturally occurring modified ribonucleotides. *J. Comput. Chem.* **2016**, *37* (10), 896-912.
25. Huang, L.; Roux, B., Automated Force Field Parameterization for Nonpolarizable and Polarizable Atomic Models Based on Ab Initio Target Data. *J. Chem. Theory Comput.* **2013**, *9* (8), 3543-3556.
26. Johansson, R.; Torrents, E.; Lundin, D.; Sprenger, J.; Sahlin, M.; Sjoberg, B. M.; Logan, D. T., High-resolution crystal structures of the flavoprotein NrdI in oxidized and reduced states--an unusual flavodoxin. *Structural biology. FEBS J.* **2010**, *277* (20), 4265-77.
27. Takaba, K.; Takeda, K.; Kosugi, M.; Tamada, T.; Miki, K., Distribution of valence electrons of the flavin cofactor in NADH-cytochrome b5 reductase. *Sci. Rep.* **2017**, *7*, 43162.
28. Werther, T.; Wahlefeld, S.; Salewski, J.; Kuhlmann, U.; Zebger, I.; Hildebrandt, P.; Dobbek, H., Redox-dependent substrate-cofactor interactions in the Michaelis-complex of a flavin-dependent oxidoreductase. *Nat. Commun.* **2017**, *8*, 16084.
29. Golden, E.; Karton, A.; Vrieling, A., High-resolution structures of cholesterol oxidase in the reduced state provide insights into redox stabilization. *Acta Crystallogr. D Biol. Crystallogr.* **2014**, *70* (Pt 12), 3155-66.
30. Vanommeslaeghe, K.; MacKerell, A. D., Jr., CHARMM additive and polarizable force fields for biophysics and computer-aided drug design. *Biochim. Biophys. Acta* **2015**, *1850* (5), 861-71.
31. Vanommeslaeghe, K.; MacKerell, A. D., Jr., Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *Journal of chemical information and modeling* **2012**, *52* (12), 3144-54.
32. Vanommeslaeghe, K.; Raman, E. P.; MacKerell, A. D., Jr., Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *Journal of chemical information and modeling* **2012**, *52* (12), 3155-68.
33. Becke, A. D., Density functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648.
34. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A., Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **1980**, *72* (1), 650-654.
35. Vanommeslaeghe, K.; Yang, M.; MacKerell, A. D., Jr., Robustness in the fitting of molecular mechanics parameters. *J. Comput. Chem.* **2015**, *36* (14), 1083-101.
36. Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P., *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press: 2007; p 1256.
37. Aleksandrov, A.; Simonson, T., Molecular Dynamics Simulations of the 30S Ribosomal Subunit Reveal a Preferred Tetracycline Binding Site. *J Am Chem Soc.* **2008**, *130* (4), 1114-1115.
38. Aleksandrov, A.; Schuldt, L.; Hinrichs, W.; Simonson, T., Tet repressor induction by tetracycline: a molecular dynamics, continuum electrostatics, and crystallographic study. *J Mol Biol.* **2008**, *378* (4), 896-910.
39. Olsson, M. H.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H., PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **2011**, *7* (2), 525-37.
40. Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A., PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **2004**, *32* (Web Server issue), W665-7.
41. Darden, T., Treatment of Long-Range Forces and Potential. In *Computational Biochemistry & Biophysics*, Becker, O. M.; MacKerell, A. D., Jr; Roux, B.; Watanabe, M., Eds. Marcel Dekker, NY.: 2001.
42. Berendsen, H.; Postma, J.; van Gunsteren, W.; DiNola, A.; Haak, J., Molecular dynamics with coupling to an external bath. . *J Chem Phys.* **1984**, *811*, 3684-3690.
43. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26* (16), 1781-802.
44. Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D., Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone  $\phi$ ,  $\psi$  and Side-Chain  $\chi_1$  and  $\chi_2$  Dihedral Angles. *J. Chem. Theory Comput.* **2012**, *8* (9), 3257-3273.
45. Huang, J.; MacKerell, A. D., Jr., CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *J. Comput. Chem.* **2013**, *34* (25), 2135-45.
46. Jorgensen, W.; Chandrasekar, J.; Madura, J.; Impey, R.; Klein, M., Comparison of simple potential functions for simulating liquid water. . *J. Chem. Phys.* **1983**, *79*, 926-935.
47. Neria, E.; Fischer, S.; Karplus, M., Simulation of activation free energies in molecular systems. *J1 Chem1 Phys.* **1996**, *105*, 1902-1921.