

Calculating Chemically Accurate Redox Potentials for Engineered Flavoproteins from Classical Molecular Dynamics Free Energy Simulations[†]

Benedict M. Sattelle and Michael J. Sutcliffe*

School of Chemical Engineering and Analytical Science, University of Manchester, Manchester Interdisciplinary Biocentre, 131 Princess Street, Manchester M1 7DN, United Kingdom

Received: May 1, 2008; Revised Manuscript Received: August 13, 2008

The tricyclic isoalloxazine nucleus of the redox cofactors flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) acts as an electron sink in life-sustaining biological electron transfer (eT). The functional diversity of flavin-containing proteins (flavoproteins) transcends that of free flavins. A large body of experimental evidence attributes natural control of flavoprotein-mediated eT to tuning of the thermodynamic driving force by the protein environment. Understanding and engineering such modulation by the protein environment of the flavin redox potential (ΔE°) is valuable in biotechnology and device design. In this study we employed classical molecular dynamics free energy simulations (MDFES), within a thermodynamic integration (TI) formalism, to calculate the change in FMN first reduction potential ($\Delta\Delta E^\circ_{\text{ox/sq}}$) imparted by 6 flavoprotein active site mutations. The combined performance of the AMBER ff03 (protein) and GAFF (cofactor) force fields was benchmarked against experimental data for mutations close to the isoalloxazine *re*- and *si*-faces that perturb the wild-type $\Delta E^\circ_{\text{ox/sq}}$ value in *Anabaena* flavodoxin. The classical alchemical approach used in this study overestimates the magnitude of ΔE° values, in common with other studies. Nevertheless, chemically accurate $\Delta\Delta E^\circ$ values—calculated to within 1 kcal mol⁻¹ of the experimental value—were obtained for five of the six mutations studied. We have shown that this approach is practical for quantitative *in silico* screening of the effect of mutations on the first reduction potential where experimental values and structural data are available for the wild-type flavoprotein. This approach promises to be useful as an integral part of future interdisciplinary strategies to engineer desired thermodynamic properties in flavoproteins of biotechnological interest.

Introduction

The flavoprotein cofactors flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) mediate life-sustaining biological electron transfer (eT). This ubiquity is attributed to the versatile electronic structure of isoalloxazine, which may adopt neutral or anionic forms of oxidized, one- and two-electron reduced species. Over the past 40 years a sustained and systematic program of mutagenesis, combined with functional studies, has shown that the eT thermodynamic driving force (ΔG°) of Marcus theory,¹ upon which the feasibility and outcomes of flavoprotein eT reactions are critically dependent, is modulated by the protein active site microenvironment to which flavin cofactors are typically noncovalently bound.² This is reflected in the wide range of redox potentials (ΔE°) observed for this class of eT proteins. Engineering flavoprotein redox potentials is of practical interest; current and prospective biotechnological applications include biosensors,^{3–5} biocatalysis,^{6,7} bioremediation^{8–10} and bioelectronics.^{11,12}

Early attempts to establish structural determinants of flavoprotein redox potentials focused on flavin chemical modification and reconstitution into the apoprotein.¹³ This approach has since been employed to engineer reversal of redox chemistry in the archetypal flavoprotein old yellow enzyme (OYE).¹⁴ However, this is not universally applicable as the altered wild-type enzyme function can result in, for example, perturbation of either flavin chemistry or substrate binding affinity.¹⁵ Most

information pertaining to flavin redox potential stabilization by the apoprotein derives from traditional site-directed mutagenesis and functional assays in the form of reduction potential measurements. *In silico* screening of active site mutations prior to mutagenesis would provide a useful tool to guide the current experimentally led engineering strategy for imparting desired thermodynamic properties in flavoproteins of biotechnological interest.

In this study we validate the use of classical molecular dynamics free energy simulations (MDFES), using a combination of the AMBER ff03¹⁶ (protein) and GAFF¹⁷ (cofactor) force fields; such an *in silico* approach has, to the best of our knowledge, not been used previously as a tool with which to guide rational modification of flavin redox potentials in flavin-containing proteins and synthetic biologically inspired devices. To facilitate a systematic study, our simulations were based on the wild-type and 6 mutant forms, W57A, W57F, W57Y, Y94A, Y94F and Y94W (both W57 and Y94 contribute to stabilization of all three redox states;¹⁸ their position is illustrated in Figure 1), of long-chain *Anabaena* (a genus of filamentous cyanobacteria) flavodoxin for which experimentally observed changes in reduction potential were available.¹⁸ Additionally, the relatively small size (168 amino acid residues) and the availability of a crystal structure of wild-type enzyme rendered this a practical choice for study using MDFES. Flavodoxins contain a single noncovalently bound FMN molecule. Their functional role as eT proteins has been exploited in diverse biotechnological applications including biosensors and mixed protein/metal nanostructures,^{19,20} and they have been employed alongside

[†] Part of the "Sason S. Shaik Festschrift".

* To whom correspondence should be addressed. E-mail: mike.sutcliffe@manchester.ac.uk.

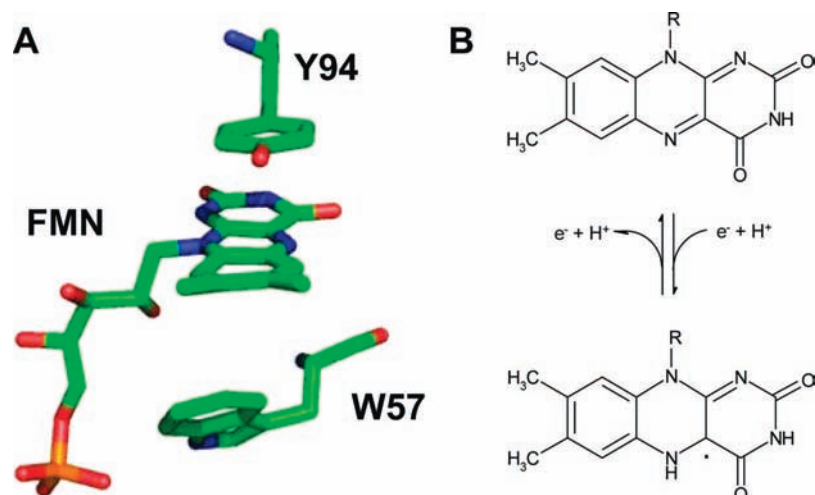


Figure 1. (A) *Anabaena* flavodoxin active site residues that mediate the first reduction potential of FMN via through-space interaction with the tricyclic isoalloxazine moiety (based on PDB accession code 1FLV). (B) Modeled neutral forms of oxidized (top) and one-electron reduced semiquinone (bottom) isoalloxazine. FMN: R = CH₂(CHOH)₄-phosphate. FAD: R = CH₂(CHOH)₄-pyrophosphate adenosine.

cytochromes as components of synthetic multidomain eT systems to explore their potential use in bioelectronics.¹² Flavodoxins from pathogenic bacteria are potential drug targets.^{21,22} They exhibit extremely low semiquinone(sq)/hydroquinone(hq) (E_1) redox potentials, the lowest measured is that of *Azotobacter vinelandii* flavodoxin at -522 mV;²³ this is attributed to protein–flavin aromatic interactions and the abundance of negatively charged residues in the active site, which destabilize flavin–protein interactions when the FMN is in the anionic hydroquinone form.^{24,25}

Methods

Preparation of the System. The model of explicitly solvated wild-type *Anabaena* flavodoxin was constructed from its crystallographic coordinates (PDB accession code 1FLV) as follows. Hydrogen atoms were added using AMBER 9²⁶ with the AMBER ff03¹⁶ force field. The FMN parameters were assigned using the general AMBER force field (GAFF).¹⁷ The isoalloxazine hydrogen atoms were appended manually according to the redox state (Figure 1B). The protein and bound cofactor were centered within a box of pre-equilibrated TIP3P water molecules. Periodic box dimensions for the wild-type flavodoxin were approximately $60 \times 60 \times 60$ Å, and a minimum distance of 10 Å was maintained between the protein exterior and the solvent box edge. Flavodoxins are highly acidic proteins; to facilitate convergence of the nonbonded Coulomb potential, the appropriate number of neutralizing Na⁺ counterions were automatically placed within the system using AMBER 9.²⁶ This produced a configuration for wild-type *Anabaena* flavodoxin containing 17158 atoms, comprising 2641 protein/cofactor/counterion atoms and 14517 water atoms. To generate starting structures for the simulations involving the six mutants W57A, W57F, W57Y, Y94A, Y94F and Y94W, *in silico* mutations were performed using the Rotamer Explorer tool²⁷ within the Molecular Operating Environment (MOE). For all mutations (except those to alanine) a number of rotamers were possible, energy minimization for each mutant with each rotamer identified the lowest energy conformation and improved the starting structures by removing unfavorable contacts. The lowest energy rotamers were used for subsequent MDFES.

Calculation of First Reduction Potentials. Classical MD-FES were used within a TI formalism to calculate the FMN first reduction potential ($\Delta E_{\text{ox/sq}}^\circ$). The single-topology TI

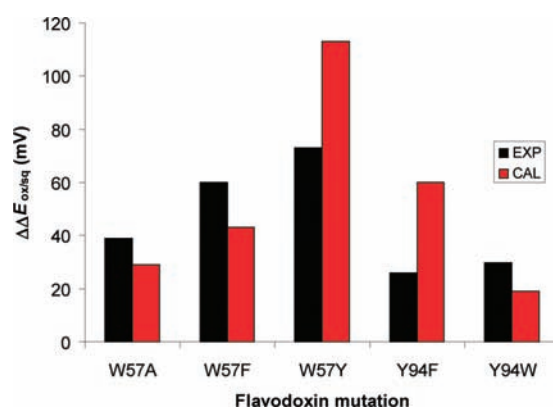


Figure 2. Calculated (CAL) and experimental (EXP) $\Delta\Delta E_{\text{ox/sq}}^\circ$ values for the mutant flavodoxin reductive half-reactions. Experimental values from ref 18.

implementation²⁸ employed systematic stepwise mutation of a classical molecular mechanics potential describing one molecule of noncovalently protein bound FMN from eT reactant to eT product states (Figure 1B) in a series of 14 production molecular dynamics runs corresponding to discrete values of a coupling parameter λ between 0 and 1 (the values of λ used are given in Figure 3A). Modeling the oxidative process and negating the computed Gibbs free energy for eT, to obtain the value for the $1\text{H}^+/1\text{e}^-$ reductive process, removed instabilities associated with the appearance of the isoalloxazine pyrimidine N5(H) atom when modeling the reductive half-reaction. The oxidized and semiquinone isoalloxazine charge distributions were generated from Merz–Singh–Kollman population analysis²⁹ at the B3LYP/6-31G(d) level of theory^{30,31} using Gaussian 03.³² As described in the AMBER 9²⁶ manual, the transformation is given by the mixed function

$$U(\lambda) = (1 - \lambda)U_0 + \lambda U_1 \quad (1)$$

where $U(\lambda)$ is the system potential energy evaluated at a particular value of the coupling parameter and U_0 and U_1 correspond to the potential energies of the initial and perturbed Hamiltonians, respectively. This linear mixing of the Hamiltonians means that

$$\frac{\partial U(\lambda)}{\partial \lambda} = U_1 - U_0 \quad (2)$$

The cumulative ensemble average of the partial derivatives of the potential energy with respect to the coupling parameter λ

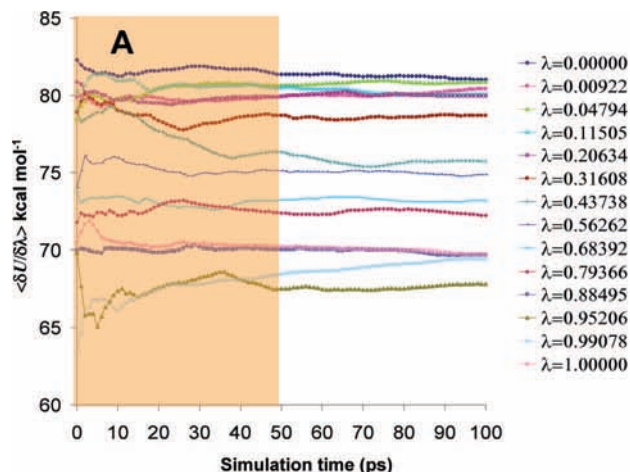
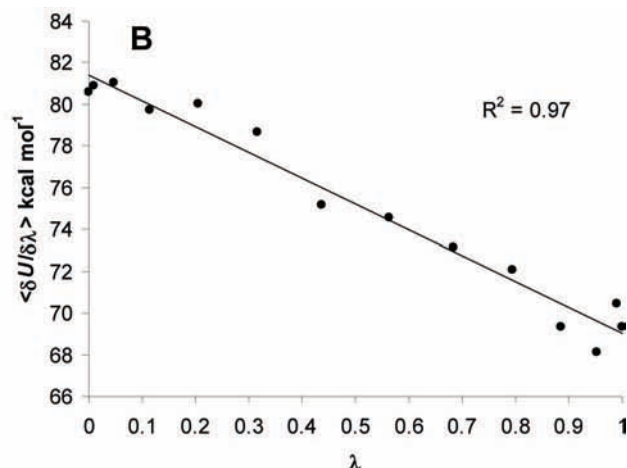


Figure 3. (A) Convergence of the free energy derivative $\langle \delta U / \delta \lambda \rangle$ during 100 ps equilibration MDFES for wild-type *Anabaena* flavodoxin. Data from the first 50 ps (shaded region) and values at $\lambda = 0$ and $\lambda = 1$ were disregarded for calculation of ΔG° . (B) $\langle \delta U / \delta \lambda \rangle$ vs λ for wild-type *Anabaena* flavodoxin.



was then integrated from 0 to 1 (eq 3) to give the Gibbs free energy difference between the related systems:²⁸

$$\Delta G = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (3)$$

The integral in eq 3 was computed using a 14-point Gaussian quadrature. Flavodoxin reduction potentials were calculated from the computed Gibbs free energy via the Nernst relationship, which effectively links the observed macroscopic ΔE° to the microscopic and calculable eT Gibbs free energy ΔG° . Because the oxidative (rather than reductive) reactions were simulated, the computed Gibbs free energies were negated to give values for the reductive processes. Addition of the SHE (+4.28 V per electron) and accounting³³ for the pH at which the experiments were done (pH 7 in all cases) rendered the calculated redox potentials comparable with experimental values.

Molecular dynamics simulations were performed as follows. The 7 explicitly solvated flavodoxins (i.e., wild-type and 6 mutant forms) were allowed to evolve unrestrained from their initial energy minimized configurations during 100 ps of self-guided Langevin dynamics.³⁴ The coupling parameter for TI, λ , was set to its initial value of 1, representative of the FMN isoalloxazine electrostatic and van der Waals parameters in the neutral semiquinone state. The systems were heated from 0 to 300 K during this time. Except for the perturbed FMN isoalloxazine moiety, the SHAKE algorithm was applied to bonds involving all hydrogen atoms. All simulations employed a 1 fs time step. Energetic and structural data were collected each picosecond. Default particle-mesh Ewald evaluation of long-range nonbonded electrostatic contributions to the energy were truncated at 9 Å. Long range van der Waals interactions were estimated by the default AMBER 926 continuum model.³⁵ Equilibration was confirmed by inspection of the conservation and variance (<0.4% of the average during last 50 ps of thermalization) of total energy. Following thermalization and equilibration, production MDFES comprised 14 runs of 100 ps within the NPT ensemble at the respective value of λ . Standard abscissas and weights were taken from the AMBER 926 manual. The structural stability of each system during production MDFES was verified by calculating the time averaged root-mean-square deviation (rmsd) of backbone C α , nitrogen and oxygen atoms from their positions in the energy minimized starting configuration during the 1400 ps of production MDFES. The low computed standard deviation of these values, on

average $\pm 11\%$ of the corresponding average, was indicative of structural equilibration. Final structures were analyzed to verify no loss of secondary or tertiary protein structure. Each simulation (i.e., each wild-type/mutant flavodoxin) was run for 1.5 ns.

Results and Discussion

Choice of Computational Method. Computation of chemically accurate redox potentials, i.e., to within ~ 1 kcal mol⁻¹ (~ 50 mV) of experimental values,³⁶ remains intractable for flavoproteins due to the computational cost associated with the high level of electronic structure theory required for calculation of Gibbs free energies for eT. Indeed, it is generally accepted that for all computational methods except those allowing the most complete description of electronic structure, relative values of calculated molecular properties are more reliable than absolute values. Additionally, experimental redox potentials for wild-type flavoproteins are almost always available because functional characterization is often the first experiment to be performed once a flavoprotein has been identified, cloned and expressed. Thus, computation of absolute redox potentials is typically not required as values calculated relative to wild-type are equally insightful for engineering purposes. Free energy perturbation (FEP) and thermodynamic integration (TI),²⁸ within MM or combined semiempirical QM/MM formalisms, are typically employed to calculate solution phase ΔG° and corresponding ΔE° values in biomolecular redox systems. Currently, ΔE° values are calculable only to within hundreds of mV at these levels of theory. Despite this inherent caveat, it was anticipated that classical MDFES calculations,²⁸ based on relative redox potentials, were likely to prove useful in the context of flavoprotein engineering. Because engineered proteins are of potential interest, the computed redox potential of each *in silico* mutant can be calculated relative to the *in silico* wild-type value, with the aim of achieving improved accuracy on typically unreliable absolute redox potentials at the computationally efficient classical molecular mechanics level of theory. The performance of the method, a measure of its future applicability for prediction of reduction potentials for engineered flavoproteins, was tested on a data set comprising 6 flavoprotein mutations for which experimentally observed changes in reduction potential were available.

Calculated Reduction Potentials. The calculated first reduction potential for the wild-type *Anabaena* flavodoxin was -600 mV (-75 kcal mol⁻¹), Table 1. This calculated result is

TABLE 1: Calculated and Experimental *Anabaena* Flavodoxin $\Delta E^{\circ}_{\text{ox/sq}}$ and $\Delta\Delta E^{\circ}_{\text{ox/sq}}$ Values

	$\Delta E^{\circ}_{\text{ox/sq}}{}^{\text{EXP}}$ (mV) ^a	$\Delta E^{\circ}_{\text{ox/sq}}{}^{\text{CAL}}$ (mV) ^a	$\Delta\Delta E^{\circ}_{\text{ox/sq}}{}^{\text{EXP}}$ (mV) ^d	$\Delta\Delta E^{\circ}_{\text{ox/sq}}{}^{\text{CAL}}$ (mV) ^d	ABS DEV (mV) ^b
Wild-type	-212 ± 5	-600 ± 17			
W57A	-173 ± 5	-571 ± 19	39 ± 7	29 ± 25	10 ± 26
W57F	-152 ± 5	-557 ± 15	60 ± 7	43 ± 23	17 ± 24
W57Y	-139 ± 5	-487 ± 14	73 ± 7	113 ± 22	39 ± 23
Y94A	-203 ± 5	-631 ± 13	9 ± 7 ^c	-32 ± 21 ^c	41 ± 22
Y94F	-186 ± 5	-540 ± 19	26 ± 7	60 ± 25	34 ± 26
Y94W	-182 ± 5	-581 ± 14	30 ± 7	19 ± 22	11 ± 23

^a Experimental (EXP)¹⁸ and calculated (CAL) are relative the SHE. ^b ABSDEV = $|\Delta\Delta E^{\circ}_{\text{ox/sq}}{}^{\text{CAL}} - \Delta\Delta E^{\circ}_{\text{ox/sq}}{}^{\text{EXP}}|$. ^c The mutant, Y94A, for which the sign of $\Delta\Delta E^{\circ}_{\text{ox/sq}}{}^{\text{EXP}}$ and $\Delta\Delta E^{\circ}_{\text{ox/sq}}{}^{\text{CAL}}$ differ.

significantly more negative than the experimental value of -212 mV,¹⁸ we attribute this error to the classical Hamiltonian used in our study. These results can be compared with recent application of the semiempirical QM/MM SCF-DFTB/CHARMM22 level of theory, also within a TI formalism, to calculate redox potentials for the equivalent first reductive process of FAD in cholesterol oxidase (CHOX) and medium chain acyl-CoA dehydrogenase (MCAD); the calculated values were 285 mV and 265 mV more negative than experimental data, respectively.³⁷ Thus, even with relatively high levels of theory, the absolute value calculated deviates significantly from that observed experimentally. The sign and magnitude of the change in first reduction potential ($\Delta\Delta E^{\circ}_{\text{ox/sq}}$) calculated in our study is in good agreement with experimental observations for 5 cases (Table 1 and Figure 2). This level of agreement appears noteworthy in light of recent reports.³⁶ For example, a combined semiempirical SCC-DFTB/MM FEP study of FAD in cholesterol oxidase (CHOX) computed the effect of a single protein mutation on the first reduction potential to be 113 mV too negative compared to the experimental value.³⁸ Furthermore, it was recently concluded that only qualitative estimates of the change in redox potential imparted by amino acid substitutions could be calculated in a recent classical TI study of Azurin mutants.³⁹ Though the experimental error in reduction potential measurement was ±5 mV, it is nevertheless very encouraging that the average deviation of calculated values from experimental values was 22 mV.

W57 Mutations. Our calculations of $\Delta\Delta E^{\circ}_{\text{ox/sq}}$ for W57Y, W57F and W57A enzyme faithfully reproduce the rank order of the experimental values,¹⁸ W57 < W57A < W57F < W57Y. Our atomistic studies therefore add further support to the proposed basis for this effect,¹⁸ that the naturally occurring tryptophan in this position renders the protein better able to stabilize oxidized FMN relative to the semiquinone species than any of the mutations studied, including those involving a change to another aromatic residue.

Y94 Mutations. The *si-face* tyrosine residue, Y94, is highly conserved in flavodoxins. The experimentally observed effect¹⁸ of increasing the $\Delta E^{\circ}_{\text{ox/sq}}$ value by introducing the mutations Y94F and Y94W was reproduced by the calculations. Perhaps the most surprising outcome was the inability of this approach to predict the correct sign (and magnitude) of the shift in reduction potential imparted by Y94A. It is noteworthy that the homologous *Desulfovibrio vulgaris* mutation Y98A imparted an experimental $\Delta\Delta E^{\circ}_{\text{ox/sq}}$ value of -38 mV,⁴⁰ which is in excellent agreement with the calculated *Anabaena* Y94A $\Delta\Delta E^{\circ}_{\text{ox/sq}}$ value of -32 mV. It is conceivable that one or more water molecules (not modeled in our study) occupy the space created by deletion of the side chain of Y94 *in vivo*. This has been speculated previously for the homologous Y98A *Des-*

ulfovibrio vulgaris flavodoxin mutation,⁴⁰ although no crystal structure for either *Anabaena* or *Desulfovibrio vulgaris* tyrosine to alanine *si-face* mutation is available to confirm this.

Because the free energy is calculated from an integrand at finite intervals, inaccuracies arise when the simulation is not run for long enough to obtain a converged value of the averaged integrand. Convergence of the free energy derivative ($\langle\delta U/\delta\lambda\rangle$) was monitored for all 7 simulations. This is illustrated for wild-type *Anabaena* flavodoxin in Figure 3B; note the linear relationships between λ and the Gibbs free energy derivative. Exploratory calculations with longer trajectories did not improve the results - this effect has been observed previously in MDFES using AMBER 9.⁴¹

Concluding Remarks

The changes in first reduction potential relative to the wild-type protein, $\Delta\Delta E^{\circ}_{\text{ox/sq}}$, imparted by 6 flavodoxin active site mutations were computed using MDFES. The sign of the computed $\Delta\Delta E^{\circ}_{\text{ox/sq}}$ values were consistent with experimental values for 5 out of 6 of the *in silico* flavodoxin mutants. Minimum, maximum and mean absolute deviation of experimental $\Delta\Delta E^{\circ}_{\text{ox/sq}}$ values from calculated $\Delta\Delta E^{\circ}_{\text{ox/sq}}$ values for the neutral reductive half-reaction were 10, 39 and 22 mV (<0.5 kcal mol⁻¹). Modeling the reductive process proved robust for the modeled aromatic → aromatic mutations. These findings suggest that single topology classical MDFES of modest duration (1.5 ns) could be employed with confidence to predict mutations imparting desired thermodynamic properties in eT flavoproteins of biotechnological interest where experimental reduction potentials and structural data pertaining to the wild-type protein are available. The AMBER ff03¹⁶ and GAFF¹⁷ force fields were shown to be useful for MDFES involving eT cofactors for which no published molecular mechanics parameters are available, thereby facilitating future modeling of nonstandard residues using TI. Although computed $\Delta E^{\circ}_{\text{ox/sq}}$ values deviated from experimental values by ~400 mV, it emerged that $\Delta\Delta E^{\circ}_{\text{ox/sq}}$ values can be meaningful in the context of flavoprotein engineering. The results improve on contemporary reports of $\Delta\Delta E^{\circ}_{\text{ox/sq}}$ calculations imparted by flavoprotein mutations in which computed values typically deviate from experimental values by > 100 mV, even within hybrid semiempirical QM/MM MDFES formalisms.⁴² Qualitative estimates of $\Delta\Delta E^{\circ}_{\text{ox/sq}}$ in other contemporary reports³⁹ were improved upon with the finding that the mean unsigned deviation of calculated from experimental $\Delta\Delta E^{\circ}_{\text{ox/sq}}$ values for 5 flavodoxin mutations were 22 mV (~0.5 kcal mol⁻¹). Classical MDFES were used to calculate with chemical accuracy the change in first reduction potential for 5 flavodoxin mutations. Application of the method to future interdisciplinary flavoprotein engineering strategies will facilitate prescreening of mutations that are likely to impart desired thermodynamic properties in flavoproteins of biotechnological interest.

Acknowledgment. This work was funded by the UK Biotechnology and Biological Sciences Research Council (BBSRC). We are grateful to Nigel Scrutton, Sam Hay, Sam de Visser and Guy Grant for useful discussions.

References and Notes

- (1) Marcus, R. A.; Sutin, N. *Biochim. Biophys. Acta* **1985**, *811*, 265–322.
- (2) Massey, V. *Biochem. Soc. Trans.* **2000**, *28*, 283–296.
- (3) Reinert, K. C.; Dunbar, R. L.; Gao, W.; Chen, G.; Ebner, T. J. *J. Neurophysiol.* **2004**, *92*, 199–211.

