

*Supplementary Information for “Polarizable Embedding for
Simulating Redox Potentials of Biomolecules”*

Ruslan N. Tazhigulov,^a Pradeep Kumar Gurunathan,^b Yongbin

Kim,^b Lyudmila V. Slipchenko,^b and Ksenia B. Bravaya^{*a}

^aDepartment of Chemistry, Boston University, Boston, Massachusetts 02215, USA

^bDepartment of Chemistry, Purdue University, West Lafayette, Indiana 47907, USA

Contents

S1. Energetics of lumiflavin in vacuo	S2
S2. Basis set correction to vertical energy gaps	S3
S3. $\lambda^{\text{LRA},f}$ dependence on the size of solvation shell	S4
S4. Dependence of polarization contribution to VEGs on the size of solvation shell	S5
S5. Ensemble-averaged vertical energy gaps and differential solvation free energy	S6
S6. $\Delta\Delta G_{\text{solv}}^f$ of phenolate water clusters	S7
S7. Redox potential of FAD in Cry1At under non-standard conditions	S8
References	S9

S1. ENERGETICS OF LUMIFLAVIN IN VACUO

Geometries of oxidized and unprotonated semireduced forms of flavin were obtained at DFT level with long-range corrected hybrid ω B97X-D functional [1]. Vertical ionization energies and electron affinities, as well as adiabatic ionization energies obtained using different one-electron bases (both for energy gaps evaluations and for geometry optimization) are listed in Table S1. All electronic structure calculations were performed in Q-Chem [2].

TABLE S1: Vertical electron affinity (VEA), vertical ionization energy (VIE), and adiabatic ionization energy (AIE) calculated for multiple geometries of lumiflavin obtained with different basis sets. All energies are given in eV and computed with the same density functional ω B97X-D.

		Geometry optimization		
		6-31G(d)	6-311++G(d,p)	aug-cc-pVTZ
	6-31G(d)	1.24	—	—
VEA	6-311++G(d,p)	—	1.63	—
	aug-cc-pVTZ	1.61	1.58	1.56
	6-31G(d)	1.78	—	—
VIE	6-311++G(d,p)	—	2.18	—
	aug-cc-pVTZ	2.14	2.12	2.11
	6-31G(d)	1.51	—	—
AIE	6-311++G(d,p)	—	1.91	—
	aug-cc-pVTZ	1.84	1.84	1.84

As expected, the basis set size (6-31G(d) vs. 6-311++G(d,p) vs. aug-cc-pVTZ) can impact the gas-phase energy gaps significantly (up to ≈ 0.3 - 0.4 eV, Table S1). However, as follows from Table S1 different basis sets used for geometry optimization almost do not affect vertical and adiabatic energy gaps: energy gaps computed with ω B97X-D/aug-cc-pVTZ for equilibrium geometries obtained with different one-electron basis differ by less than 0.05 eV with VEA being the most sensitive. Considering this small basis set effect on the geometry, the local geometry optimization of lumiflavin in the fixed MM environment for multiple MM MD configurations performed at ω B97X-D/6-31G(d) level (see ‘‘Computational Details’’ section in the main text) is not expected to notably affect the accuracy of the energetic parameters. The basis set correction

to the aug-cc-pVTZ basis was further added to the QM/NP-BioEFP and QM/BioEFP vertical energy gaps (VEGs), evaluated originally at the same ω B97X-D/6-31G(d) level as well (see Sec. S2).

Since the geometry lumiflavin was locally optimized, its degrees of freedom did not contribute to the computed energy gaps. To estimate these contributions, vibrational and thermochemical analysis was performed for oxidized and unprotonated semireduced forms of lumiflavin in the gas phase. The results are listed in Table S2.

TABLE S2: Differences in zero-point vibrational energies (Δ ZPVE) and thermodynamical quantities ($\Delta_r H$, $\Delta_r S$, $\Delta_r G^{\text{thermo}}$, $T = 298.15$ K) computed for different forms of lumiflavin, oxidized and semireduced. Quantities are given in eV (Δ ZPVE, $\Delta_r H$, $\Delta_r G^{\text{thermo}}$) and in meV / K ($\Delta_r S$).

	Δ ZPVE	$\Delta_r H$	$\Delta_r S$	$\Delta_r G^{\text{thermo}}$	Δ ZPVE + $\Delta_r G^{\text{thermo}}$
ω B97X-D/6-311++G(d,p)	0.073	0.070	-0.011	0.073	0.145
ω B97X-D/6-31G(d)	0.069	0.067	-0.002	0.068	0.137

Previously, experimental $\text{AIE} + \Delta\text{ZPVE} = 1.86 \pm 0.1$ eV was obtained by Fourier transform ion cyclotron resonance mass spectrometry in the work by Ridge et al [3]. Given our computed estimate of Δ ZPVE to be around 0.07 eV, our $\text{AIE} + \Delta\text{ZPVE}$ (AIE , computed at ω B97X-D/aug-cc-pVTZ level; Δ ZPVE, computed at ω B97X-D/6-311++G(d,p) level) = 1.91 eV is in good agreement with the experimental value and another computational value, reported in the same work by Ridge et al. [3]: $\text{AIE} + \Delta\text{ZPVE}$ (M06-L/6-31+G(d,p)) = 1.92 eV.

S2. BASIS SET CORRECTION TO VERTICAL ENERGY GAPS

The basis set corrections to the vertical energy gaps (VEG) were computed as an ensemble-averaged difference in VEGs, calculated with two different basis sets, but with the same method (density functional, ω B97X-D): aug-cc-pVTZ and 6-31G(d) (Eq. S1), in a similar way as in our previous work [4]. The configurations were taken from MM MD, followed by local optimization of lumiflavin at the fixed MM environment. After that, the environment was removed, effectively putting lumiflavin under *in vacuo* conditions.

$$\langle \Delta \text{VEG} \rangle^{\text{BSC}} = \langle \text{VEG} \rangle^{\omega\text{B97X-D/aug-cc-pVTZ}} - \langle \text{VEG} \rangle^{\omega\text{B97X-D/6-31G(d)}} \quad (\text{S1})$$

TABLE S3: VEGs and basis set corrections ($\langle\Delta\text{VEG}\rangle^{\text{BSC}}$), calculated as shown in Eq. S1. The corrections were added to the free energy and redox potential (Eq. 6, 7, 9, main text). All quantities are given in eV.

	$\omega\text{B97X-D}/6\text{-}31\text{G(d)}$	$\omega\text{B97X-D}/\text{aug-cc-pVTZ}$	$\langle\Delta\text{VEG}\rangle^{\text{BSC}}$
$\langle\text{VEA}\rangle$	1.20 ± 0.01	1.58 ± 0.01	0.38 ± 0.01
$\langle\text{VIE}\rangle$	1.77 ± 0.01	2.14 ± 0.01	0.37 ± 0.01

S3. $\lambda^{\text{LRA},f}$ DEPENDENCE ON THE SIZE OF SOLVATION SHELL

The linear response reorganization energy ($\lambda^{\text{LRA},f}$) was computed as discussed in the main text (Eq. 4). The resulting values as a function of the radius of solvation shell (R) around the protein are plotted in Fig. S1. The graphs corresponding to QM/NP-BioEFP and QM/BioEFP have very similar trends, and the values are slowly converging at large values of R .

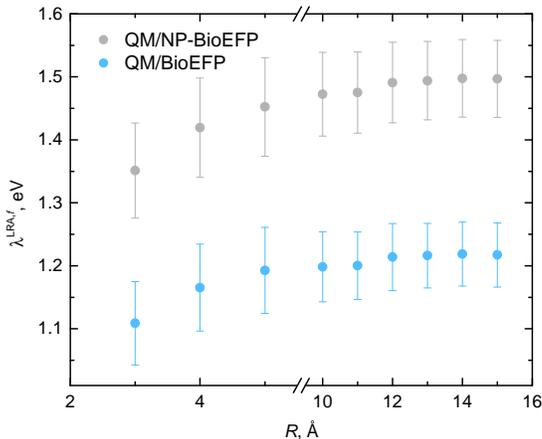


FIG. S1: $\lambda^{\text{LRA},f}$ dependence on the number of water molecules defined by the distance from protein atoms R .

S4. DEPENDENCE OF POLARIZATION CONTRIBUTION TO VEGS ON THE SIZE OF SOLVATION SHELL

The ion (point charge)-induced dipole interaction energy is proportional to R^{-4} , where R is the distance between the ion and the induced dipole. The total number of particles dN in a spherical shell of radius R and shell thickness dR is $4\pi R^2 dR$. Therefore, the total interaction energy between the ion (charge) and all induced dipoles within a sphere of a given radius R will be proportional to R^{-1} . Thus, we expect R^{-1} dependence for polarization contribution to vertical energy gaps, defined as the difference in VEGs computed with BioEFP (fully polarizable) and NP-BioEFP (fully non-polarizable) approach (Eq. S2). Note, that fully oxidized flavin has zero net charge, while the semireduced flavin has -1 net charge.

$$\langle \Delta \text{VEG} \rangle^f = \langle \text{VEG} \rangle^{\text{QM/BioEFP}} - \langle \text{VEG} \rangle^{\text{QM/NP-BioEFP}} \quad (\text{S2})$$

The resulting values as a function of the radius of solvation shell (R) around the protein are plotted in Fig. S2.

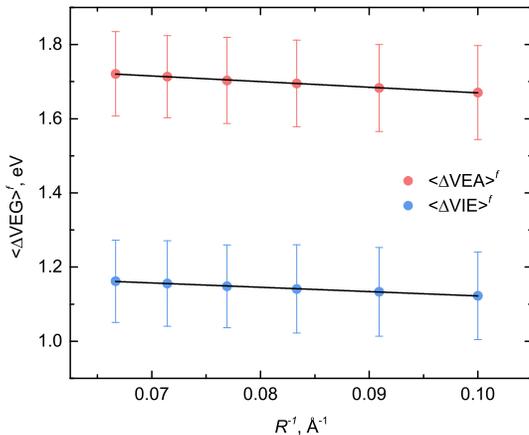


FIG. S2: $\langle \Delta \text{VEG} \rangle^f$ dependence on the inverse radius of solvation shell R .

Importantly, as the contribution of the environmental polarization is found to be substantial and, therefore, an account of environment polarization is critical for accurate estimates of the vertical energy gaps, QM/BioEFP calculations are only approximately 2-2.5x more expensive than QM/NP-BioEFP ($R = 10\text{--}15 \text{ \AA}$, 1 core, Intel Xeon Gold 6132 processor).

S5. ENSEMBLE-AVERAGED VERTICAL ENERGY GAPS AND DIFFERENTIAL SOLVATION FREE ENERGY

Below we summarize the data used for calculation of oxidation free energies and redox potentials.

TABLE S4: Computed $\langle \text{VEA} \rangle^f$, $\langle \text{VIE} \rangle^f$, and $\Delta\Delta G_{\text{solv}}^f$ as a function of water shell size around the protein defined by R . Bare $\omega\text{B97X-D}/6\text{-}31\text{G(d)}/\text{NP-BioEFP}$ and $\omega\text{B97X-D}/6\text{-}31\text{G(d)}/\text{BioEFP}$ values are reported, $\Delta\Delta G_{\text{solv}}^f$ is computed as discussed in the main text. All energies are given in eV.

R , Å	$\langle \text{VEA} \rangle^{\text{QM/NP-BioEFP}}$	$\langle \text{VEA} \rangle^{\text{QM/BioEFP}}$	$\langle \text{VIE} \rangle^{\text{QM/NP-BioEFP}}$	$\langle \text{VIE} \rangle^{\text{QM/BioEFP}}$	$\Delta\Delta G_{\text{solv}}^f$
3	0.32 ± 0.11	1.87 ± 0.10	3.02 ± 0.10	4.08 ± 0.09	0.32
4	0.38 ± 0.11	1.95 ± 0.10	3.22 ± 0.12	4.28 ± 0.10	0.29
5	0.38 ± 0.10	1.97 ± 0.10	3.29 ± 0.12	4.35 ± 0.10	0.27
10	0.80 ± 0.10	2.47 ± 0.08	3.74 ± 0.09	4.87 ± 0.07	0.21
11	0.85 ± 0.09	2.54 ± 0.08	3.80 ± 0.09	4.94 ± 0.08	0.20
12	0.88 ± 0.09	2.58 ± 0.08	3.87 ± 0.09	5.01 ± 0.07	0.19
13	0.94 ± 0.09	2.64 ± 0.08	3.92 ± 0.09	5.07 ± 0.07	0.19
14	0.97 ± 0.08	2.68 ± 0.07	3.97 ± 0.09	5.12 ± 0.07	0.18
15	1.02 ± 0.09	2.74 ± 0.07	4.01 ± 0.09	5.18 ± 0.07	0.17

The standard errors (Δ) have been computed for 95% confidence interval as follows: $\Delta = \frac{t(95\%,n-1)\sigma}{\sqrt{n}}$, where t is Student's t test, σ is a standard deviation, n is a sample size, and $n = 50$ as the ensemble-averaged quantities were computed for 50 configurations.

Vertical energies were also computed with the environment being treated as point charges (PC), using CHARMM36 FF and TIP3P water model charges. The results, comparing NP-BioEFP (static multipoles) with PC model, are provided below.

TABLE S5: Computed $\langle \text{VEA} \rangle^f$ and $\langle \text{VIE} \rangle^f$ for the water shell size around the protein $R = 10$ Å. Bare $\omega\text{B97X-D}/6\text{-}31\text{G(d)}/\text{NP-BioEFP}$ and $\omega\text{B97X-D}/6\text{-}31\text{G(d)}/\text{PC}$ values (in eV) are reported.

R , Å	$\langle \text{VEA} \rangle^{\text{QM/NP-BioEFP}}$	$\langle \text{VEA} \rangle^{\text{QM/PC}}$	$\langle \text{VIE} \rangle^{\text{QM/NP-BioEFP}}$	$\langle \text{VIE} \rangle^{\text{QM/PC}}$
10	0.80 ± 0.10	0.79 ± 0.09	3.74 ± 0.09	3.88 ± 0.09

S6. $\Delta\Delta G_{\text{solv}}^f$ OF PHENOLATE WATER CLUSTERS

To support our approach in computation of the difference in the solvation energies ($\Delta\Delta G_{\text{solv}}^f$) of non-spherical model systems corresponding to two different electronic states (charge distributions), we calculated the difference in solvation free energies of phenoxyl radical and phenolate spherical water clusters, solving Poisson-Boltzmann equation (PBEQ) numerically with APBS [5]. We used geometry configurations discussed in our previous work [4]. The snapshots were obtained from MM MD simulations, and then the spherical water clusters were created, defined by the radius R from the center of mass of phenolate molecule. If the O-atom of water molecule was within a sphere of radius R , the water molecule would be included in the model system. Similarly as it was discussed in the main text and performed for protein-solvent model systems, the environmental charges (water molecule charges) were zeroed-out to prevent artifacts in solvation energy calculations, caused by non-polarizable MM environment. The solute molecule was represented by point charges taken from the modified forcefield [4]. In such approach, the solvent water molecules only determine the shape of the cavity, but do not interact with the polarizable continuum. The results, averaged over 100 configurations, were compared with analytical Born solvation energy for spherical point charge/cluster (Eq. S3, in atomic units), and are provided in Table S6.

$$\Delta\Delta G_{\text{solv}}^{f,\text{Born}} = \frac{1}{2R} \left(1 - \frac{1}{\epsilon}\right) \quad (\text{S3})$$

where ϵ is the static dielectric constant of solvent.

TABLE S6: Difference in solvation free energies (eV) between oxidized and reduced forms of phenolate computed by solving Poisson-Boltzmann equations (PBEQ) numerically ($\Delta\Delta G_{\text{solv}}^{f,\text{PBEQ}}$), and analytically, according to Eq. S3 ($\Delta\Delta G_{\text{solv}}^{f,\text{Born}}$).

$R, \text{ \AA}$	$\Delta\Delta G_{\text{solv}}^{f,\text{PBEQ}}$	$\Delta\Delta G_{\text{solv}}^{f,\text{Born}}$
10	0.78	0.71
20	0.38	0.36

With the size increasing, the discrepancy between PBEQ and Born $\Delta\Delta G_{\text{solv}}^f$ is decreasing owing to non-spherical shape of the small clusters: the deviation between computed and analytical solvation free energies decreases from 0.07 eV to 0.02 eV once the radius of the solvation

shell increases from 10 Å to 20 Å. Therefore, for the large protein-solvent clusters, we expect the error introduced from PBEQ solvation to be insignificant.

S7. REDOX POTENTIAL OF FAD IN Cry1At UNDER NON-STANDARD CONDITIONS

In the experiment, both unprotonated (Flv_{SR}^-) and protonated (FlvH_{SR}) semireduced forms of flavin coexist at pH 7.4 [6]. Yet, the computed quantity is a reduction potential of $\text{Flv}_{\text{OX}} \xrightarrow{e^-} \text{Flv}_{\text{SR}}^-$ half-reaction. From the available experimental data, we were able to evaluate the possible discrepancy in the redox potential between computational and experimental quantities, occurring from the coexistence of two semireduced forms at given pH.

Note that pK_a (N5 of Flv_{SR}^-) of semireduced flavin mononucleotide (FMN) in water is 8.3 [7]. However, pK_a (N5 of Flv_{SR}^-) is slightly shifted in the protein environment, based on the variation of yield of protonated semireduced form (FlvH_{SR}) reported in multiple experimental works [6, 8, 9], and it is likely located in the range between 7.4 and 8.0, given the reported peak yield of $\text{FlvH}_{\text{SR}}^-$ around 37 % at pH 7.4 [6, 9] and around 18 % at pH 8.0 [8]. For T (experimental) = 283.15 K and neglecting the term occurring from the ratio $\frac{[\text{Flv}_{\text{OX}}]}{[\text{FlvH}_{\text{SR}}]} \approx 1$, the Nernst equation for the redox potential of $\text{Flv}_{\text{OX}} \xrightarrow{\text{H}^+, e^-} \text{FlvH}_{\text{SR}}$ half-reaction is as follows:

$$\begin{aligned} E(\text{Flv}_{\text{OX}}/\text{FlvH}_{\text{SR}}) &= E^\circ(\text{Flv}_{\text{OX}}/\text{FlvH}_{\text{SR}}) + \frac{RT}{F} \ln \frac{[\text{Flv}_{\text{OX}}][\text{H}^+]}{[\text{FlvH}_{\text{SR}}]} \\ &= E^\circ(\text{Flv}_{\text{OX}}/\text{Flv}_{\text{SR}}^-) - \frac{\Delta_r G^\circ(\text{Flv}_{\text{SR}}^- + \text{H}^+ \rightleftharpoons \text{FlvH}_{\text{SR}})}{F} + \frac{RT}{F} \ln \frac{[\text{Flv}_{\text{OX}}][\text{H}^+]}{[\text{FlvH}_{\text{SR}}]} \\ &\approx E^\circ(\text{Flv}_{\text{OX}}/\text{Flv}_{\text{SR}}^-) - \frac{RT}{F} \ln K_a + \frac{RT}{F} \ln [\text{H}^+] \\ &\approx E^\circ(\text{Flv}_{\text{OX}}/\text{Flv}_{\text{SR}}^-) + 0.0562 \cdot (\text{pK}_a - \text{pH}) \end{aligned}$$

Based on the assumptions above, the difference between experimental ($E(\text{Flv}_{\text{OX}}/\text{FlvH}_{\text{SR}})$) and computed quantities ($E^\circ(\text{Flv}_{\text{OX}}/\text{Flv}_{\text{SR}}^-)$), appearing from the difference in pK_a (N5 of Flv_{SR}^-) and experimental pH, should be within 0.02–0.05 V (assuming pK_a (N5 of Flv_{SR}^-) = 7.8–8.3), relatively small with respect to the typical error bars of DFT and hybrid QM/MM schemes.

-
- [1] Chai, J.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion interactions. *Phys. Chem. Chem. Phys.*, **2008**, *10*, 6615–6620.
- [2] Y. Shao, Z. Gan, E. Epifanovsky, A.T.B. Gilbert, M. Wormit, J. Kussmann, A.W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P.R. Horn, L.D. Jacobson, I. Kaliman, R.Z. Khaliullin, T. Kus, A. Landau, J. Liu, E.I. Proynov, Y.M. Rhee, R.M. Richard, M.A. Rohrdanz, R.P. Steele, E.J. Sundstrom, H.L. Woodcock III, P.M. Zimmerman, D. Zuev, B. Albrecht, E. Alguires, B. Austin, G.J.O. Beran, Y.A. Bernard, E. Berquist, K. Brandhorst, K.B. Bravaya, S.T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S.H. Chien, K.D. Closser, D.L. Crittenden, M. Diedenhofen, R.A. DiStasio Jr., H. Do, A.D. Dutoi, R.G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M.W.D. Hanson-Heine, P.H.P. Harbach, A.W. Hauser, E.G. Hohenstein, Z.C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyayev, J. Kim, J. Kim, R.A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C.M. Krauter, K.U. Laog, A. Laurent, K.V. Lawler, S.V. Levchenko, C.Y. Lin, F. Liu, E. Livshits, R.C. Lochan, A. Luenser, P. Manohar, S.F. Manzer, S.-P. Mao, N. Mardirossian, A.V. Marenich, S.A. Maurer, N.J. Mayhall, C.M. Oana, R. Olivares-Amaya, D.P. O'Neill, J.A. Parkhill, T.M. Perrine, R. Peverati, P.A. Pieniazek, A. Prociuk, D.R. Rehn, E. Rosta, N.J. Russ, N. Sergueev, S.M. Sharada, S. Sharma, D.W. Small, A. Sodt, T. Stein, D. Stuck, Y.-C. Su, A.J.W. Thom, T. Tsuchimochi, L. Vogt, O. Vydrov, T. Wang, M.A. Watson, J. Wenzel, A. White, C.F. Williams, V. Vanovschi, S. Yeganeh, S.R. Yost, Z.-Q. You, I.Y. Zhang, X. Zhang, Y. Zhou, B.R. Brooks, G.K.L. Chan, D.M. Chipman, C.J. Cramer, W.A. Goddard III, M.S. Gordon, W.J. Hehre, A. Klamt, H.F. Schaefer III, M.W. Schmidt, C.D. Sherrill, D.G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A.T. Bell, N.A. Besley, J.-D. Chai, A. Dreuw, B.D. Dunietz, T.R. Furlani, S.R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D.S. Lambrecht, W.Z. Liang, C. Ochsenfeld, V.A. Rassolov, L.V. Slipchenko, J.E. Subotnik, T. Van Voorhis, J.M. Herbert, A.I. Krylov, P.M.W. Gill, and M. Head-Gordon. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.*, **2015**, *113*, 184–215.
- [3] Zhang, T.; Pappas, K.; Ochran, R.; Ridge, D. P. Stability of flavin semiquinones in the gas phase: The electron affinity, proton affinity, and hydrogen atom affinity of lumiflavin. *J. Phys. Chem. A*, **2013**, *117*, 11136–11141.
- [4] Tazhigulov, R. N.; Bravaya, K. B. Free energies of redox half-reactions from first-principles calcu-

- lations. *J. Phys. Chem. Lett.*, **2016**, *7*, 2490–2495.
- [5] Jurrus, E.; Engel, D.; Star, K.; Monson, K.; Brandi, J.; Felberg, L. E.; Brookes, D. H.; Wilson, L.; Chen, J.; Liles, K.; Chun, M.; Li, P.; Gohara, D. W.; Dolinsky, T.; Konecny, R.; Koes, D. R.; Nielsen, J. E.; Head-Gordon, T.; Geng, W.; Krasny, R.; Wei, G. W.; Holst, M. J.; McCammon, J. A.; Baker, N. A. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.*, **2018**, *27*, 112–128.
- [6] Müller, P.; Bouly, J. P.; Hitomi, K.; Balland, V.; Getzoff, E. D.; Ritz, T.; Brettel, K. ATP binding turns plant cryptochrome into an efficient natural photoswitch. *Sci. Rep.*, **2014**, *4*, 5175.
- [7] Müller, F. *The flavin redox-system and its biological function*. Springer, Berlin, Heidelberg, 1983.
- [8] Lin, C.; Robertson, D.; Ahmad, M.; Raibekas, A.; Jorns, M.; Dutton, P.; Cashmore, A. Association of flavin adenine dinucleotide with the Arabidopsis blue light receptor CRY1. *Science*, **1995**, *269*, 968–970.
- [9] Balland, V.; Byrdin, M.; Eker, A.; Ahmad, M.; Brettel, K. What makes the difference between a cryptochrome and DNA photolyase? A spectroelectrochemical comparison of the flavin redox transitions. *J. Am. Chem. Soc.*, **2008**, *131*, 426–427.