UNIVERSITY OF CALIFORNIA

Santa Barbara

Designer Flavoproteins for the Elucidation of the Fundamental Mechanism of Potential Inversion

A thesis submitted in partial satisfaction of the requirements for the degree Master of Science in Chemistry

by

Mary Lerner

Committee in charge:

Professor Brandon Greene, Chair

Professor Lior Sepunaru

Professor Kevin Plaxco

Professor Yang Hai

September 2022

The thesis of Mary Lerner is approved.

_____

Kevin Plaxco

_____

Yang Hai

_____

Lior Sepunaru

_____

Brandon Greene, Committee Chair

September 2022

ABSTRACT

Designer Flavoproteins for the Elucidation of

the Fundamental Mechanism of Potential Inversion

by

Mary Lerner

Flavins are arguably one of the most versatile cofactors in biology, resulting from their ability to facilitate both $1e^-$ and $2e^-$ transfer reactions, as well as their widely-varying reduction potentials that can be tuned by the local protein environment. Flavins have three redox states: oxidized (OX), $1e^-$-reduced semiquinone (SQ), and $2e^-$-reduced hydroquinone (HQ). Their reactivity in redox processes is determined by $E'_{OX/SQ}$ and $E'_{SQ/HQ}$, which is affected by electrostatic and H-bonding interactions in the flavin binding site. Free in solution, flavins exhibit "inverted potentials", where the $1e^-$ reduced state is more unstable than the fully-reduced state. The flavin binding site of flavodoxins, a class of electron transfer flavoproteins, alters the $1e^-$ reduction potentials such that they become "normally" ordered, stabilizing $E'_{OX/SQ}$ relative to $E'_{SQ/HQ}$. In other cases, the protein environment can increase potential inversion. A dramatic example can be found in flavin-based electron bifurcation (FBEB), which utilizes significant potential separation (ΔE) to conserve energy from exergonic redox reactions by concomitantly driving endergonic ET. The properties of the flavin binding site that induce this inversion are unknown, largely due to the limited mutagenesis space in authentic EB-ases and lack of physical methods for measuring highly-inverted $1e^-$ reduction potentials. Since potential separation is fundamental to the

mechanism of energy conservation in FBEB, understanding its cause is both interesting from a basic science perspective, and valuable to the field of electrocatalysis. To move towards this goal, it is essential to develop minimalist models for potential inversion, since the study of this phenomenon in native EB-ases is nontrivial. We have demonstrated that iLOV can be used as a model flavoprotein to evidence the effects of near-flavin mutations on redox behavior and potential inversion, substituting Lys to illustrate the influence of an ionizable and H-bonding functional group on semiquinone stability. We have designed expression and reconstitution methods to recombinantly produce iLOV WT, Q104K, and Q104A in high holoprotein yield. The $2e^-$ reduction potential of iLOV WT and Q104K was measured by protein film voltammetry, and equilibration studies with a redox dye were conducted to confirm that protein-film electrochemistry corresponded to bulk properties. The semiquinone protonation state was characterized by EPR and UV-vis. Using spectral data to infer concentrations of the three flavin redox states at electrochemical equilibrium, ΔE was calculated for iLOV WT, iLOV Q104K and iLOV Q104A. The results suggest that iLOV WT has the most inverted $1e^-$ potentials, followed by Q104A, with Q104K having the least potential separation. The substitution of Ala at the N5 position appeared to stabilize the SQ, although this is likely due to exposure to ambient light. The addition of Lys at N5 appears to influence the thermodynamic properties of the flavin such that $E'_{OX/SQ}$, and seemingly $E'_{SQ/HQ}$, is more positive. This can be rationalized by its relative acidity and closer pK_a matching between the flavin and Lys, as compared to Gln in iLOV WT. Our discoveries on the crucial elements of potential inversion in flavin redox chemistry may shed light on the

mechanisms of potential tuning in natural systems, particularly extreme potential inversion of EB-ases.

Table of Contents

Table of Figures

Chapter 1. Background and Introduction

1.1 General Reactivity of Flavins in Biology

Flavins, a family of organic compounds containing the tricyclic heterocycle 7,8-dimethylisoalloxazine, are one of the most chemically diverse prosthetic groups in enzymatic catalysis. They are arguably one of the most versatile cofactors in biology, resulting from the wide variety of redox, ionic, and electronic states that the flavin can adopt during its reaction cycle. Acting as the cofactor in flavoenzymes, they can facilitate both $1e^-$ and $2e^-$ transfer in oxidoreductases, react with oxygen to perform hydroxylation reactions, induce signaling responses to visible light, enable bioluminescence, perform halogenation reactions, and photochemically repair DNA damage.[1] Flavins have three redox states: oxidized (OX), $1e^-$-reduced semiquinone (SQ), and $2e^-$-reduced hydroquinone (HQ). The semiquinone has both a neutral (NSQ) and anionic (ASQ) form, depending on its protonation state (**Figure 1.1**). The stability of the SQ radical allows flavins to act as an intermediary between obligate $2e^-$ donors (*e.g.* NAD(P)H) and obligate $1e^-$ acceptors (*e.g.* hemes, FeS clusters), a unique property that is not found in other biological electron carriers. This explains why flavins are utilized in a variety of metabolic processes (*e.g.* aerobic respiration, photosynthesis, denitrification, sulfur respiration) (**Figure 1.1**).[1]

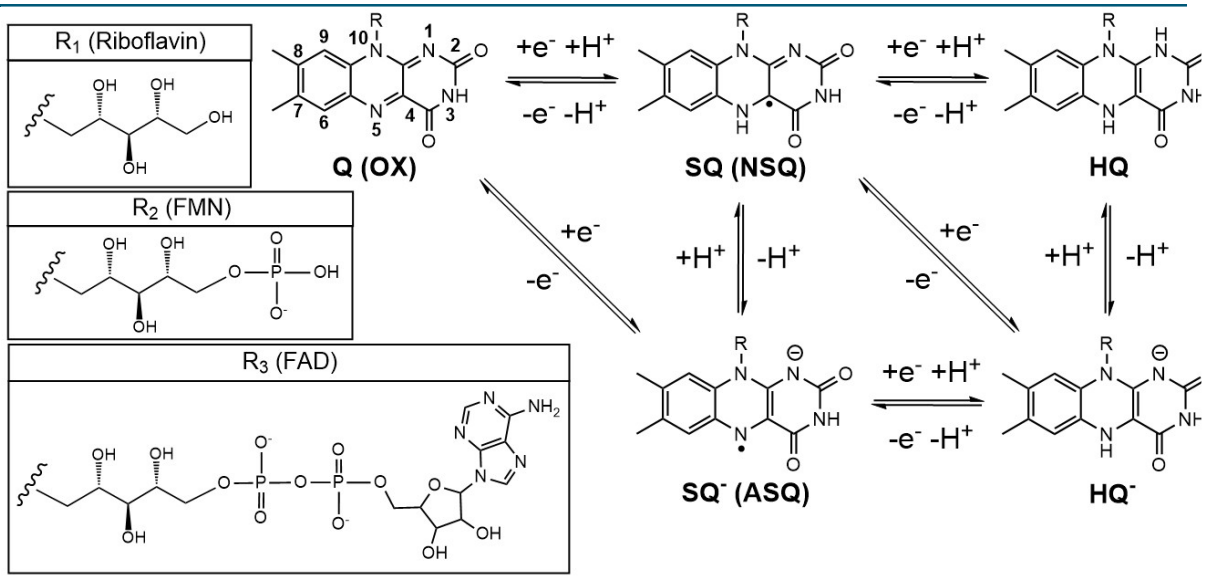

Figure 1.1. *Shows the basic structure of flavins, which all contain the 7,8-dimethyl-10-alkylisoalloxazine moiety. The protein binding affinity is increased by the addition of a phosphoryl or ADP R-group. Generally, flavoproteins preferentially bind either FMN or FAD, although some are known to be promiscuous and bind either form. Riboflavin is rarely used as a prosthetic group.*

Different flavoproteins can conduct electron transfer (ET) reactions at distinct potentials if they have unique protein environments that affect flavin energetics. This allows flavoproteins to catalyze a wide breadth of redox reactions, as compared to free flavins. Upon reduction from the OX to HQ state, the flavin molecular geometry changes from planar to a butterfly or bent shape.[2,3] Since the flavin remains non-covalently bound to the protein throughout the entirety of its redox cycle, changes in the flavin geometry affect its stability via enthalpic contributions from the binding site H-bonding network, as well as electrostatic interactions from nearby residues. These fluctuations affect flavin electronic structure and alter its reduction potentials, explaining why flavoproteins can have variable potentials.[2] Thus, it is no surprise that flavoenzymes catalyze a diverse array of redox reactions in biology, since their reactivity can be tuned by the flavin's protein environment.

Flavodoxins (Flds) are protein electron carriers that beautifully illustrate the versatility of flavins in biological redox chemistry. They serve a similar purpose to ferredoxins (to which they owe their name) but use flavin mononucleotide (FMN) in place of an FeS cluster to transfer low-potential electrons between redox enzymes. Using a protein-bound flavin for inter-protein ET is advantageous over a free flavin because its reduction potentials can be augmented to suit its function, and its tertiary structure can be recognized by specific binding partners. Compared to Flds, the ΔE between the $1e^-$ reduction potentials of free FMN is relatively small, with $E'_{OX/SQ}$ = -314 mV and $E'_{SQ/HQ}$ = -124 mV.[4] Flds have highly modular reduction potentials, with $E'_{OX/SQ}$ from -100 to -250 mV *vs.* SHE, and $E'_{SQ/HQ}$ from -380 to -520 mV *vs.* SHE.[5,6] It is suspected that potential variation is achieved by stabilization of the SQ state by the donation of a hydrogen bond from the protonated N5 of the NSQ to a backbone carbonyl[7,8,9] and destabilization of the HQ^- state by charge repulsion of flavin electron density by anionic residues and hydrophobic/π-stacking effects by aromatic residues sandwiching the flavin's isoalloxazine ring.[6,10]

The potentials of Flds and free FMN also differ by their order. FMN exhibits "inverted potentials", where the $1e^–$ reduced state is more unstable than the fully-reduced state (**Figure 1.2**). Potential separation (ΔE) can be quantified as the difference between the first ($E'_{OX/SQ}$) and second ($E'_{SQ/HQ}$) reduction potentials, as shown by (1).

$$\Delta E' = E'_{SQ/HQ} - E'_{OX/SQ} \qquad (1)$$

Flavins are described as having inverted potentials when ΔE > 0, which is common for organic molecules that have unstable carbon radicals. This is the case for free

FMN. When ΔE < 0, the potentials are said to be normally ordered, as observed in Flds. Not only is ΔE greater in Flds *vs.* free FMN, but the potential order is flipped, which demonstrates the significant influence of binding site residues on flavin redox behavior. Flds typically cycle between the HQ and SQ states to deliver electrons from low-potential metabolic substrates or Photosystem I to partner proteins, including hydrogenase, nitrogenase, and others.

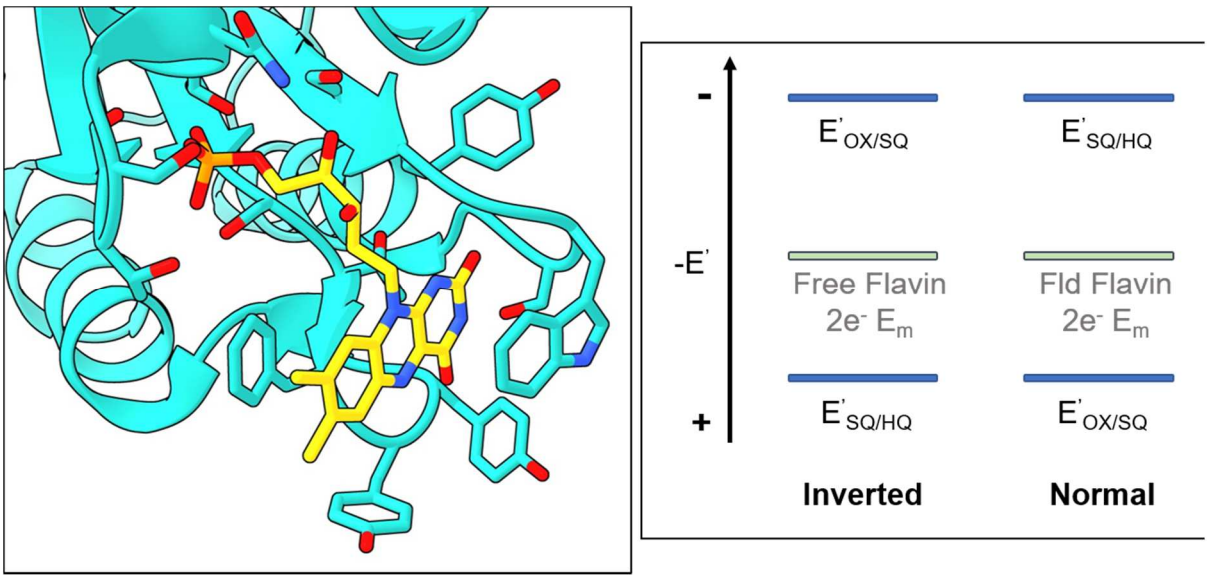

Figure 1.2 *Left: Depicts the structure of flavodoxin from M. acetivorans (PDB ID: 5WID), with aromatic residues that are crucial to flavin E'$_{SQ/HQ}$ destabilization. Right: Inverted potential scheme vs. normally-ordered flavin cofactor to illustrate differences between Flds and EB-ases.*

1.2 EB Exploits Extreme Potential Inversion for Redox Catalysis

In contrast to Flds, some flavoprotein environments can induce highly inverted potentials. An extreme example of protein-induced potential inversion can be found in flavin-based electron bifurcation (FBEB), which utilizes a large ΔE to conserve energy

from exergonic redox reactions by concomitantly driving endergonic ET. The net reaction is the reduction of two unique $1e^-$ acceptors from a common $2e^-$ donor. Since each electron travels down an independent path of ET cofactors before reaching its respective acceptor, the mechanism is termed “electron bifurcation” (EB) from the Latin root bifurcus, meaning "two-pronged” (**Figure 1.3**). These paths are referred to as the exergonic and endergonic ET branches, named according to their ultimate electron acceptor. Electron bifurcating enzymes (EB-ases) all use either a flavin or quinone cofactor, and those using the latter display less extreme potential inversion than in FBEB. Significant potential inversion has been observed in all enzymes that catalyze FBEB.

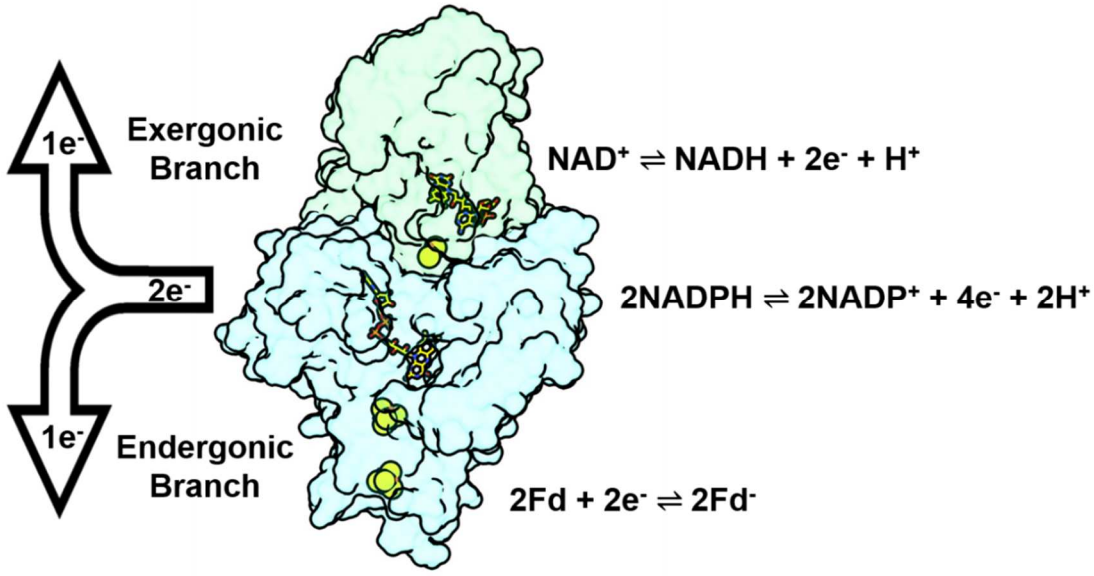

Figure 1.3 *Shows the reversible bifurcation scheme of NADPH-dependent ferredoxin:NADH oxidoreductase (NfnI)[11], which catalyzes reduction of NAD^+ and Fd by NADPH in the bifurcating direction. NfnI is regarded as the simplest EB-ase, with three iron sulfur clusters (shown as space-filling model in orange (Fe) and yellow (S)) and two FAD cofactors.*

FBEB was first observed in 2008, when Herrmann *et. al*. discovered a soluble electron-transfer flavoprotein in the strict anaerobe *C. kluyveri* that catalyzed the seemingly impossible generation of H_2 (E' = -420 mV) by NADH (E' = -320 mV), coupled to the exergonic reduction of an acyl-CoA-derivative.[12] The authors hypothesized that the electron-transfer flavoprotein in question had the ability to reduce crotonyl-CoA as well as ferredoxin, the latter acting as the electron donor for an independent hydrogenase to reduce protons to H_2. Successive work has shown that this was, indeed, the activity of this enzyme. Since then, numerous enzymes have been documented that catalyze the endergonic reduction of a protein-electron carrier (ferredoxin/flavodoxin) only in the presence of its specific exergonic acceptor, comprising twelve distinct types of enzymes (categorized by their substrates and ET partners).[13] Electron bifurcation is utilized in commonplace biological redox reactions involving various exergonic acceptors, such as pyruvate, coenzymes, and disulfides (**Figure 1.4**). To date, twelve distinct types of enzymes have been biochemically characterized that conduct FBEB.[13] These enzymes catalyze a wide breadth of metabolic reactions with conserved redox energy, including hydrogen production, generation of CH4 from fatty acid metabolism, and nitrogen fixation.[14,15,16]

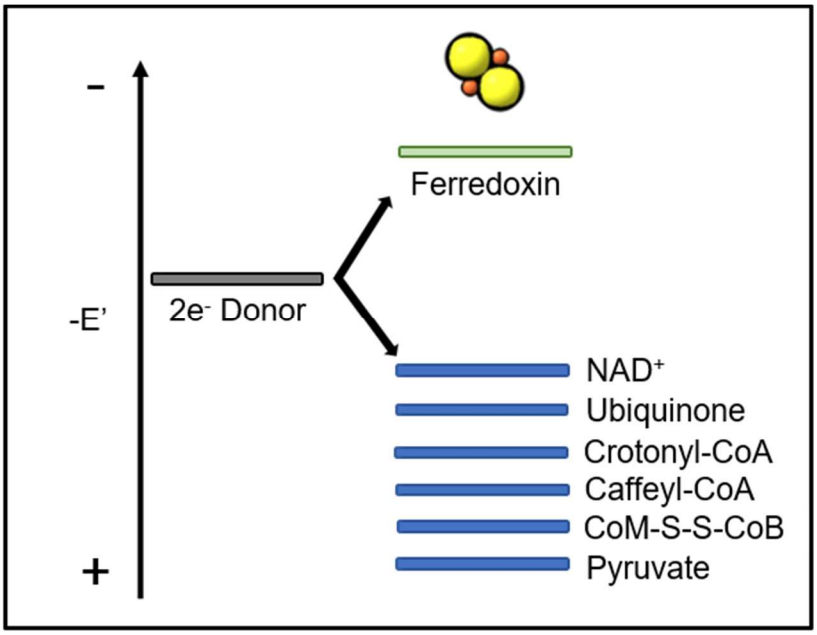

Figure 1.4 *Depicts the versatility of FBEB in the reduction of various biologically-relevant exergonic acceptors (E' of acceptors not to scale). Ferredoxin is depicted as the endergonic acceptor, although in some cases, flavodoxin is used instead.*

Significant separation between the flavin's reduction potentials explains how EB-ases can accept electrons from NAD(P)H or H_2 yet reduce lower-potential acceptors in an apparent violation of thermodynamics. Flavins' ability to facilitate both $1e^-$ and $2e^-$ transfer allows for its reduction by $2e^-$ at an intermediate energy, followed by sequential $1e^-$ transfer from the flavin at distinct energies. A large ΔE essentially raises the reduction potential of exergonic $1e^-$ transfer, relative to the flavin $2e^-$ E', in exchange for lowering the reduction potential of the endergonic ET step (**Figure 1.5**). Many demanding biochemical reactions require potent reductants, and FBEB allows for their generation using ubiquitous electron donors like NAD(P)H without expending ATP or dissipating ion gradients. A greater difference between $E'_{OX/SQ}$ and $E'_{SQ/HQ}$ allows more energy to be conserved. This evolutionary pressure has resulted in

significant potential inversion in EB-ases, with one estimate of ΔE = 1.3 V for the NADPH-dependent ferredoxin:NADH oxidoreductase (NfnI).[17] However, unlike in Flds, the properties of the flavin binding site that induce this inversion are unknown. Since potential separation is fundamental to the mechanism of energy conservation in FBEB, understanding its cause is both interesting from a basic science perspective, and valuable to the field of electrocatalysis.

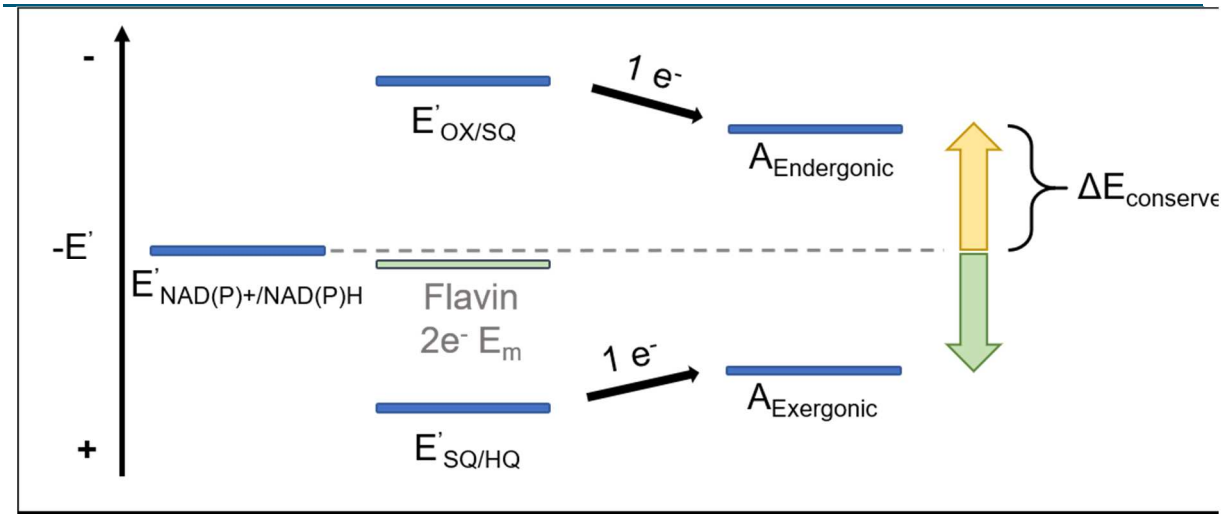

Figure 1.5 *Shows a thermodynamic bifurcation scheme for a nicotinamide-dependent EB-ase. The energy difference between NAD(P)H and the exergonic acceptor (green arrow) would normally be lost as heat, but via FBEB, it is conserved by the reduction of the endergonic acceptor (yellow arrow).*

1.3 Proposed Mechanisms of EB

The mechanism by which electrons are partitioned between the exergonic and endergonic ET branches in FBEB is highly debated in the field and may differ among phylogenetically-distinct EB-ases due to convergent evolution. Two distinct mechanisms have been proposed, the first being conformational, and the second dependent only on energetic poise.

The first EB-ase discovered was ETF-Bcd, in the electron-transfer flavoprotein (ETF) class of bifurcating enzymes, which utilize a conformational gating mechanism to achieve efficient ET. Structural evidence of Etf-Bcd in *C. difficile, A. fermentans,* and *T. maritima* suggests that ferredoxin (Fd) and NADH-binding events are correlated with allosteric changes that alter the inter-cofactor distances of the electron transfer pathway.[18] X-ray crystal structures have been obtained for ETF-Bcd in two stable states that show an increase in distance by 20 Å between the two FAD cofactors.[18,19] A third structure has been solved by cryo-EM that is likely in an intermediate position between the two previously-determined states (**Figure 1.6**). This lowers the probability of the endergonic electron being transferred toward the exergonic acceptor by increasing the physical distance between cofactors, effectively lowering the electronic coupling parameter of Marcus ET.[20] A newly-discovered bifurcating hydrogenase, HydABCSL, has also been imaged by cryo-EM in multiple conformational states.[21] Conformational gating is also employed by coenzyme Q:cytochrome c oxidoreductase (complex III) by the movement of a Reiske-FeS cluster, which prevents short-circuiting by blocking the exergonic ET branch during part of the Q-cycle.

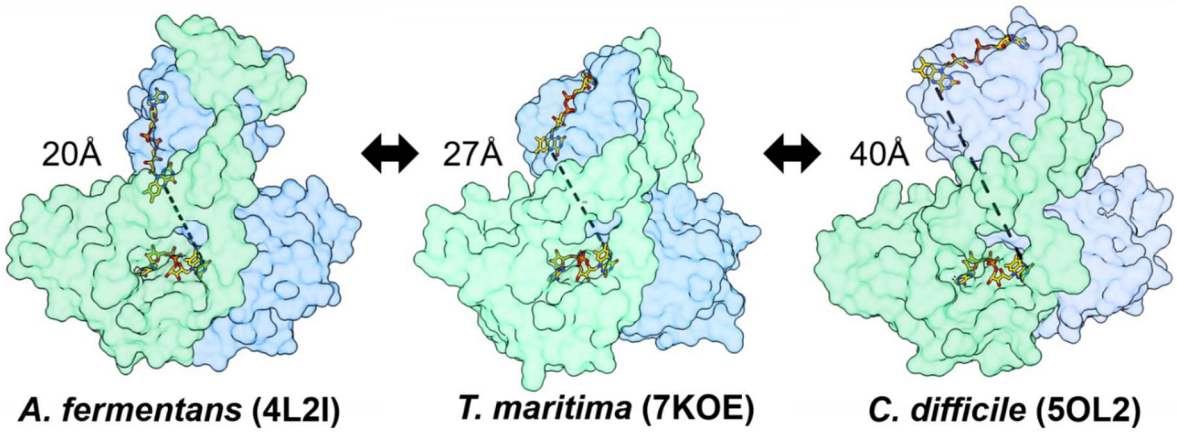

Figure 1.6 *Etfs exhibit two major conformational states; one where the ET-flavin is proximal to the BF-flavin (left), and another where the ET-flavin "swings" towards the Bcd subunit (right; Bcd not depicted). A recent cryo-EM image from T. maritima depicts an intermediate conformational state.*

Other EB-ases, such as NfnI, lack sufficient evidence to support a dynamic ET mechanism, and it is postulated that some enzymes can conduct EB without structural changes, driven purely by thermodynamics. A theoretical model has been proposed that is consistent with efficient bifurcation, called the “Z-scheme”, which relies on Boltzmann-predicted occupancy of electrons on near-flavin cofactors to minimize short circuits.[22] The term “Z-scheme” describes an energy landscape that promotes ET by successive exergonic transfer steps that spatially separate high-energy electrons from electron holes, preventing their recombination. By effectively sweeping the low-potential electrons away from the flavin cofactor down a thermodynamically-favorable transfer pathway, the probability of short circuits is reduced. This creates the upper slant of the “Z” landscape, which represents the endergonic ET branch of electron bifurcation (**Figure 1.7**). The exergonic branch is sloped in the opposite manner, which protects against short circuits by an electron occupancy blockade effect. The lowest point of the “Z” acts as a thermodynamic trap that lowers the

probability of an electron hole at the entrance point to the exergonic branch. Some EB-ases have energy landscapes that follow the “Z-scheme”, including ETFs and NfnI.[11,23] Electron-transfer cofactors in complex III, a quinone-based EB-ase, also have reduction potentials that follow this trend.[22] Independent from EB, this energy landscape is also observed in photosynthesis, where rapid endergonic ET is necessary to compete with the photochemical excitation and relaxation rates in Photosystem II. Thus, the “Z-scheme” describes known trends in biological catalysis that can be extended to electron bifurcation. Although the “Z-scheme” appears to be sufficient for efficient ET, it is debated whether energetic poise alone is necessary to conduct EB. While complex III has been used as a model for understanding EB efficiency through the “Z-scheme”, this has yet to be extended to FBEB, which displays more significant potential inversion. The construction of a simple and evaluable artificial EB system may be necessary to understand the roles of potential inversion, conformational gating, and thermodynamic poise.

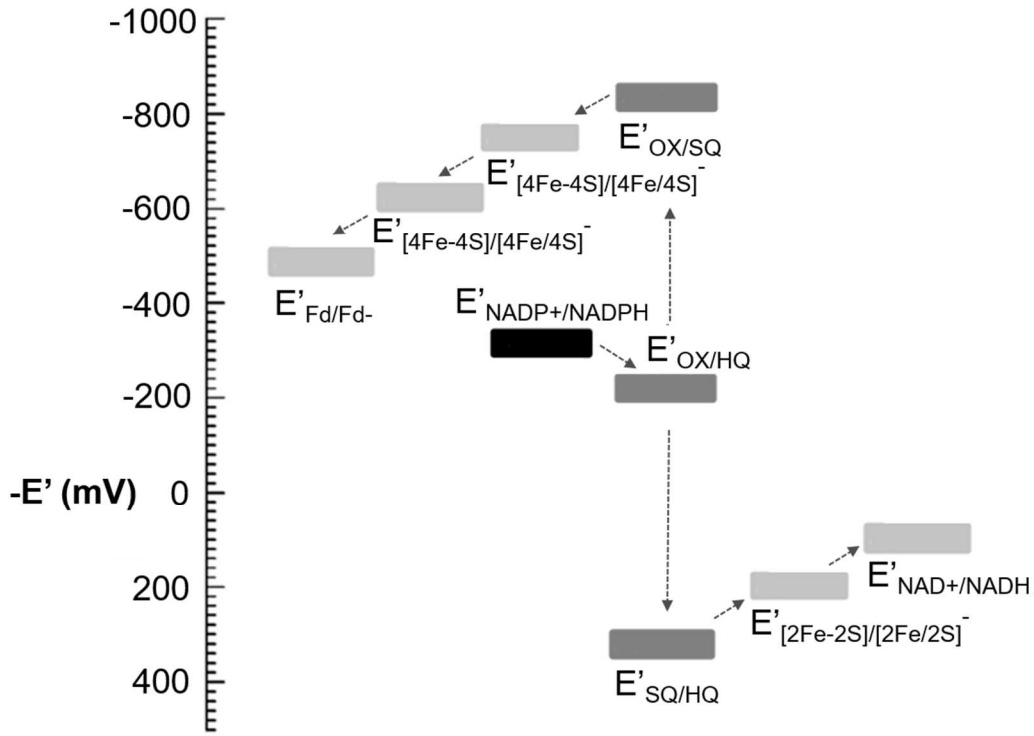

Figure 1.7 *Depicts the Z-scheme of electron transfer in EB-ase NfnI.*[11]

1.4 Project Goals

From Flds to EB-ases, the protein environment can tune the $1e^-$ potentials ~800 mV each (total of 1.6 V), yet how this is achieved is unknown due to lack of structural characterization, loss of function by mutagenesis in native enzymes, and the difficulty of potentiometric titration in multi-cofactor systems.[24] The extreme potential separation observed in EB-ases could be exploited in artificial systems via coupled electrochemical reductions. Coupling of endergonic and exergonic redox reactions would enable more efficient catalysis by lowering the overall energy demand for the net reaction, as compared to catalyzing them individually. As observed in EB,

reduction of both acceptors could be achieved using electrons at an intermediate potential between the two half-reactions. Artificial EB would maximize the energy efficiency of multi-electron reactions by eliminating the need for an applied overpotential, which is usually necessary to drive endergonic redox processes via electrochemistry. Therefore, understanding the cause for potential inversion in flavoproteins, namely EB-ases, is of fundamental interest. ***To move towards this goal, it is essential to develop minimalist models for potential inversion, since the study of this phenomenon in native EB-ases is nontrivial.*** Quinone-based EB-ases, such as Complex III, are large membrane proteins with lipo-soluble electron carriers, which makes them difficult to express, reconstitute, and study. While enzymes that catalyze FBEB are usually cytosolic with soluble electron carriers, they are similarly complex, composed of many subunits and cofactors that are essential for function. Therefore, a basic flavoprotein platform would serve as a useful tool for evaluating the cause of potential inversion. Since the proton-coupled electron transfer (PCET) chemistry of flavins involves proton transfer to both N1 and N5 atoms, and the proton transfer coordinate exhibits significant control on the redox properties of PCET reactions, site-directed mutagenesis of the protein binding site near these positions is a logical starting place.

We initially have turned our attention to the conserved structural elements of the bifurcating flavin binding pocket in authentic EB-ases. Analysis of EB-ase sequence and structural similarity has brought light some shared attributes. Six types of EBases have been structurally characterized by X-ray crystallography or electron microscopy. These structures include the ETF-Bcd complex from *C. difficile*[18], *A. fermentans*[19],

NfnI from *P. furiosus*[11], Caffeyl-CoA Reductase complex from *A. woodii*[25], Ni-Fe Hydrogenase complex from *M. thermolithotrophicus*[26], Ni-Fe Hydrogenase complex HydABCSL from *A. mobile*[21], and N_2-fixing FixABCX from *T. maritima.*[27] Alignment of the suspected "bifurcating" flavin in these structures allows us to visualize a positively-charged residue near the N5 of the flavin isoalloxazine ring (**Figure 1.8**).

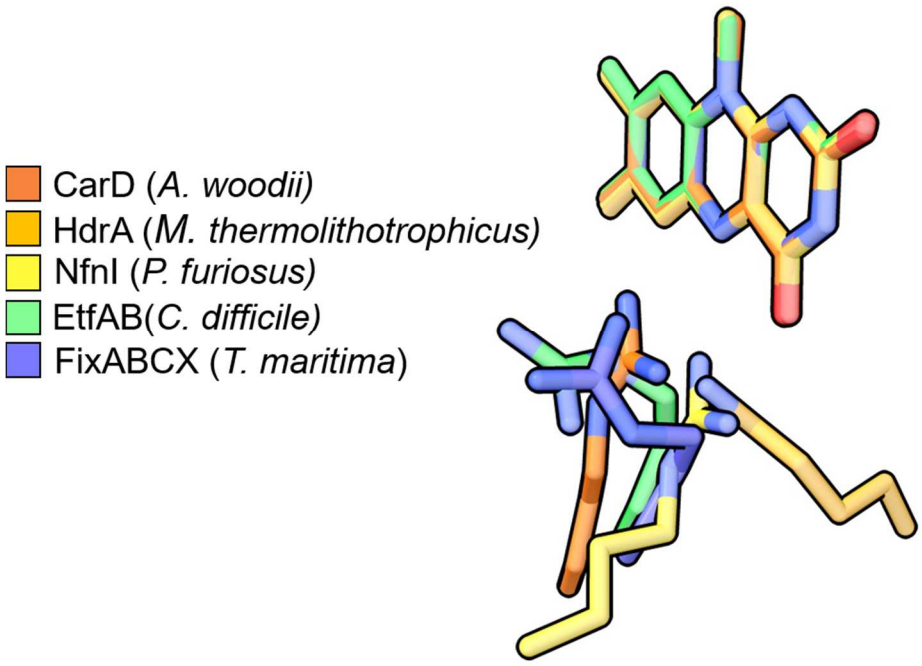

Figure 1.8 *Shows the structural alignment of residues near the flavin N5 in native EB-ases, highlighting a conserved Arg or Lys.*

Bioinformatic analysis of HdrA analogs also indicates that Lys/Arg is conserved at this position. Since potential inversion has been observed in all known EB-ases, this residue is suspected to contribute to this effect, as evidenced by its sequence and structural homology. It has been postulated that this conserved Arg or Lys H-bonds to the N5 and O4 of the flavin isoalloxazine ring, somehow lowering the redox potential

of E'OX/SQ by an unknown mechanism and increasing the difference between $E'_{OX/SQ}$ and $E'_{SQ/HQ}$.[17,28] Site-directed mutagenesis of ETF-Bcd to replace the arginine residue in both flavin binding sites has been performed, with deleterious effects on flavin binding and conformational motion.[24] ***The use of a small model system that lacks conformational dynamics may evidence the effects of such mutations on flavin redox behavior.*** The complexity of native bifurcating systems inspired the implementation of a basic flavoprotein scaffold to elucidate the underpinnings of FBEB.

I have identified a flavoprotein that will serve as a platform to explore the fundamental mechanism of potential inversion, the "**i**mproved **L**ight **O**xygen and **V**oltage"-sensing domain (iLOV), a truncated fluorescent mutant of the PHOT2 protein from *A. thaliana*.[29] iLOV is easily expressed in high yield (50+ mg/g wet cell weight) and has a high tolerance to heat, which enabled the development of a facile thermal cofactor reconstitution method (**3.3**). Most importantly, iLOV is small (13.4 kDa) and can withstand flavin binding site mutations without substantial loss of binding affinity or structural integrity (**Figure 1.9**). This allows us to probe the effects of mutagenesis on flavin redox behavior to better understand how the protein environment contributes to potential inversion.

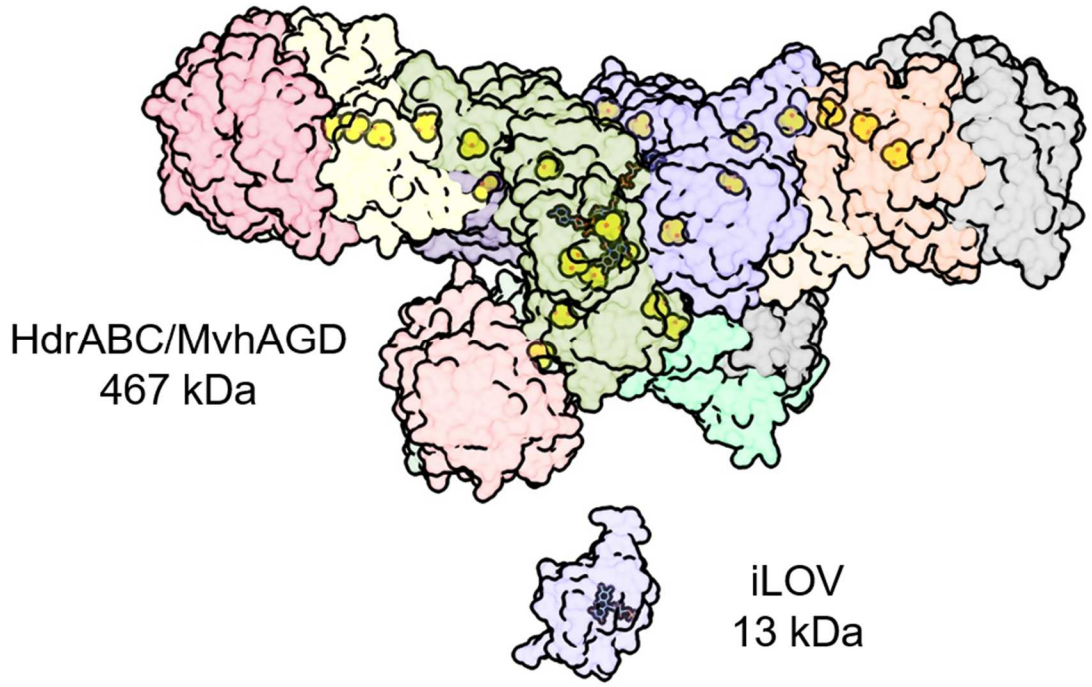

Figure 1.9 *Shows the immensity and cofactor complexity of native EB-ase HdrABC/MvhAGD relative to iLOV (to-scale). Hdr has 11 iron sulfur clusters (shown as space-filling model in orange (Fe) and yellow (S)) and two FAD cofactors. iLOV, by comparison, is small and simple with only one FMN cofactor.*

Using our iLOV model, we will investigate the influence of binding site residues with various functional groups on flavin thermodynamics by site-directed mutagenesis at the N5 position, inspired by structural conservation in EB-ases (**Figure 1.8**). We have designed an iLOV mutant, Q104K, that has Lys near the N5 position to study its effect on the flavin redox landscape and potential inversion. Contrastingly, we've also substituted Ala in the same position to infer how a lack of an ionizable or H-bonding functional group in this position would affect flavin redox behavior. Since each redox form has a unique UV-vis signature, we can determine the redox cycle for iLOV mutants and evaluate differences in PCET/ET pathways that may be induced by ionizable residues (**Figure 1.1**). Our discoveries on the crucial elements of potential

inversion in flavin redox chemistry may shed light on the mechanisms of potential tuning in natural systems, particularly extreme potential inversion of EB-ases.

References

[1] Mansoorabadi, S.O., Thibodeaux, C.J., Liu, H.W. The Diverse Roles of Flavin Coenzymes Nature's Most Versatile Thespians. *J. Org. Chem.,* **2007,** *72* (17), 6329-6342

[2] Walsh, C. Flavin Coenzymes: at the Crossroads of Biological Redox Chemistry. *Acc. Chem. Res.*, **1980**, *13* (5), 148-155

[3] Zheng, Y.J., Ornstein, R.L. A Theoretical Study of the Structures of Flavin in Different Oxidation and Protonation States. *J. Am. Chem. Soc.*, **1996,** *118* (39), 9402-9408

[4] Anderson, R.F. Energetics of the One-Electron Reduction Steps of Riboflavin, FMN and FAD to Their Fully Reduced Forms*. Biochim. Biophys. Acta. Bioenerg.,* **1983,** *722* (1), 158-162

[5] Deistung, J., Thorneley, R.N. Electron Transfer to Nitrogenase. Characterization of Flavodoxin from Azotobacter chroococcum and Comparison of its Redox Potentials with Those of Flavodoxins from Azotobacter vinelandii and Klebsiella pneumoniae (nifF-Gene Product). *Biochem. J.*, **1986,** *239* (1), 69-75

[6] Alagaratnam, S., Van Pouderoyen, G., Pijning, T., Dijkstra, B.W., Cavazzini, D., Rossi, G.L., Van Dongen, W.M., Van Mierlo, C.P., Van Berkel, W.J., Canters, G.W. A Crystallographic Study of Cys69Ala Flavodoxin II from Azotobacter vinelandii: Structural Determinants of Redox Potential. *Protein Sci.*, **2005,** *14* (9), 2284-2295

[7] O'Farrell, P.A., Walsh, M.A., McCarthy, A.A., Higgins, T.M., Voordouw, G., Mayhew, S.G. Modulation of the Redox Potentials of FMN in Desulfovibrio vulgaris Flavodoxin: Thermodynamic Properties and Crystal Structures of Glycine-61 Mutants. *Biochemistry*, **1998,** *37* (23), 8405-8416

[8] Ludwig, M.L., Pattridge, K.A., Metzger, A.L., Dixon, M.M., Eren, M., Feng, Y., Swenson, R.P. Control of Oxidation−Reduction Potentials in Flavodoxin from Clostridium beijerinckii: The Role of Conformation Changes. *Biochemistry*, **1997,** *36* (6), 1259-1280

[9] Chang, F.C., Swenson, R.P. The Midpoint Potentials for the Oxidized−Semiquinone Couple for Gly57 Mutants of the Clostridium beijerinckii Flavodoxin Correlate with Changes in the Hydrogen-Bonding Interaction with the

Proton on N(5) of the Reduced Flavin Mononucleotide Cofactor as Measured by NMR Chemical Shift Temperature Dependencies. *Biochemistry*, **1999,** *38* (22), 7168-7176

[10] Sancho, J. Flavodoxins: Sequence, Folding, Binding, Function and Beyond. *Cell Mol. Life Sci.*, **2006,** *63* (7), 855-864

[11] Lubner, C.E., Jennings, D.P., Mulder, D.W., Schut, G.J., Zadvornyy, O.A., Hoben, J.P., Tokmina-Lukaszewska, M., Berry, L., Nguyen, D.M., Lipscomb, G.L., Bothner, B., Jones, A.K., Miller, A.F., King, P.W., Adams, M.W.W., Peters, J.W. Mechanistic Insights into Energy Conservation by Flavin-Based Electron Bifurcation. *Nat. Chem. Biol.*, **2017,** *13*, 655-659

[12] Li, F., Hinderberger, J., Seedorf, H., Zhang, J., Buckel, W., Thauer, K. Coupled Ferredoxin Coenzyme A (CoA) and Crotonyl-CoA Reduction with NADH Catalyzed by the Butyryl-CoA Dehydrogenase/Etf Complex from Clostridium kluyveri. *J. Bacteriol.*, **2008**, *190* (3), 843-850

[13] Poudel, S., Dunham, E.C., Lindsay, M.R., Amenabar, M.J., Fones, E.M., Colman, D.R., Boyd, E.S. Origin and Evolution of Flavin-Based Electron Bifurcating Enzymes. *Front. Microbiol.*, **2018**, *9*, 1762

[14] Schut G.J., Adams M.W.W. The Iron-Hydrogenase of Thermotoga maritima Utilizes Ferredoxin and NADH Synergistically: a New Perspective on Anaerobic Hydrogen
Production. *J. Bacteriol.,* **2009**, *191*, 4451-4457

[15] Agne, M., Estelmann, S., Seelmann, C., Kung, J., Wilkens, D., Koch, H., van der Does, C., Albers, S., Ballmoos, C., Simon, J., Boll, M. The Missing Enzymatic Link in Syntrophic Methane Formation from Fatty Acids. *Proc. Natl. Acad. Sci. U.S.A.,* **2021**, *118* (40), e2111682118

[16] Ledbetter R., Garcia-Costas A.M., Lubner C.E., Mulder D.W., Tokmina-Lukaszewska M., Artz J.H., Patterson A., Magnuson T.S., Jay Z.J., Duan H.D., Miller J. The Electron Bifurcating FixABCX Protein Complex from Azotobacter vinelandii: Generation of Low-Potential Reducing Equivalents for Nitrogenase Catalysis. *Biochemistry*, **2017**, *56* (32), 4177-4190

[17] Demmer J.K., Huang H., Wang S.N., Demmer U., Thauer R.K., Ermler U. Insights into Flavin-Based Electron Bifurcation via the NADH-Dependent Reduced Ferredoxin:NADP Oxidoreductase Structure. *J. Biol. Chem.*, **2015**, *290*, 21985-21995

[18] Demmer, J.K., Chowdhury, N.P., Selmer, T., Ermler, U., Buckel, W. The Semiquinone Swing in the Bifurcating Electron Transferring Flavoprotein/Butyryl-CoA Dehydrogenase Complex from Clostridium difficile. *Nat. Commun.,* **2017,** *8* (1), 1-10

[19] Mowafy, A.M., Chowdhury, N.P., Demmer, J., Upadhyay, V., Kolzer, S., Jayamani, E., Kahnt, J., Demmer, U., Ermler, U., Buckel, W. Electron Transferring Flavoprotein of Acidaminococcus fermentans: Towards a Mechanism of Flavin-Based Electron Bifurcation. *J. Biol. Chem.,* **2014,** *289,* 5145-5157

[20] Marcus R.A. Chemical and Electrochemical Electron-Transfer Theory. *Annu. Rev. Phys. Chem.,* **1964**, *15* (1), 155-196

[21] Feng, X., Schut, G., Haja, D., Adams, M.W.W., Li, H. Structure and Electron Transfer Pathways of an Electron-Bifurcating NiFe-Hydrogenase. *Sci. Adv.,* **2022**, *8* (8), eabm7546

[22] Yuly, J., Zhang, P., Lubner, C., Peters, J., Beratan, D. Universal Free-Energy Landscape Produces Efficient and Reversible Electron Bifurcation. *Proc. Natl. Acad. Sci. U.S.A*., **2020,** *117* (35), 21045-21051

[23] Sucharitakul, J., Buckel, W., Chaiyen, P. Rapid Kinetics Reveal Surprising Flavin Chemistry in Bifurcating Electron Transfer Flavoprotein from Acidaminococcus fermentans. *J. Biol. Chem.*, **2021,** *296,* 100124

[24] Mohamed-Raseek, N., Miller, A-F. Contrasting Roles for Two Conserved Arginines: Stabilizing Flavin Semiquinone or Quaternary Structure, in Bifurcating Electron Transfer Flavoproteins. *J. Biol. Chem.*, **2022**, *298* (4), 101733

[25] Demmer, J.K., Bertsch, J., Oppinger, C., Wohlers, H., Kayastha, K., Demmer, U., Ermler, U., Muller, V. Molecular Basis of the Flavin-Based Electron-Bifurcating Caffeyl-CoA Reductase Reaction. *FEBS Lett.,* **2018**, *592*, 332-342

[26] Wagner, T., Koch, J., Ermler, U., Shima, S. Methanogenic Heterodisulfide Reductase (HdrABC-MvhAGD) Uses Two Noncubane [4Fe-4S] Clusters for Reduction. *Science*, **2017,** *357*, 699-703

[27] Feng, X., Schut, G., Lipscomb, G., Li, H., Adams, M.W.W. Cryoelectron Microscopy Structure and Mechanism of the Membrane-Associated Electron-Bifurcating Flavoprotein Fix/EtfABCX. *Proc. Natl. Acad. Sci. U.S.A.,* **2021,** *118* (2), e2016978118

[28] J.L. Yuly, P. Zhang, X. Ru, K. Terai, N. Singh, D.N. Beratan**.** Efficient and Reversible Electron Bifurcation with Either Normal or Inverted Potentials at the Bifurcating Cofactor. *Chem,* **2021,** *7* (7), 1870-1886

[29] Christie, J.M., Hitomi, K., Arvai, A.S., Hartfield, K.A., Mettlen, M., Pratt, A.J., Tainer, J.A., Getzoff, E.D. Structural Tuning of the Fluorescent Protein iLOV for Improved Photostability. *J. Biol. Chem.,* **2012**, *287* (26), 22295-22304

2.1 Bioinformatics Identifies Conserved Lysine in Homologs of Native EB-ase HdrABC/MvhAGD

It is clear from the flavin redox cycle that chemistry at the N1 and N5 positions must dictate the reduction potentials, since their molecular orbitals undergo electron density changes upon reduction that influence binding site electrostatics, and they are the only ionizable sites (**Figure 2.1**). We selected the N5 site as our focus to determine the effect of the protein environment on $E'_{OX/SQ}$ and propensity to adopt the ASQ form *vs.* the protonated NSQ. As described in **1.4,** a Lys or Arg residue near the flavin N5 is conserved in all known structures and suspected to play a role in flavoprotein potential inversion by destabilization of the SQ.[17,28] In an effort to evidence the conservation and necessity of the basic residue near N5, we examined the metagenomic sequence homology of the EB-ase HdrABC/MvhAGD. We found that Lys is highly conserved at sequence positions aligned with Lys^{409} in HdrA, the residue closest to the flavin N5. This finding inspired mutagenic substitution of an ionizable, positively-charged residue at the N5 position in our model iLOV flavoprotein to deduce its implications on semiquinone protonation state and stability.

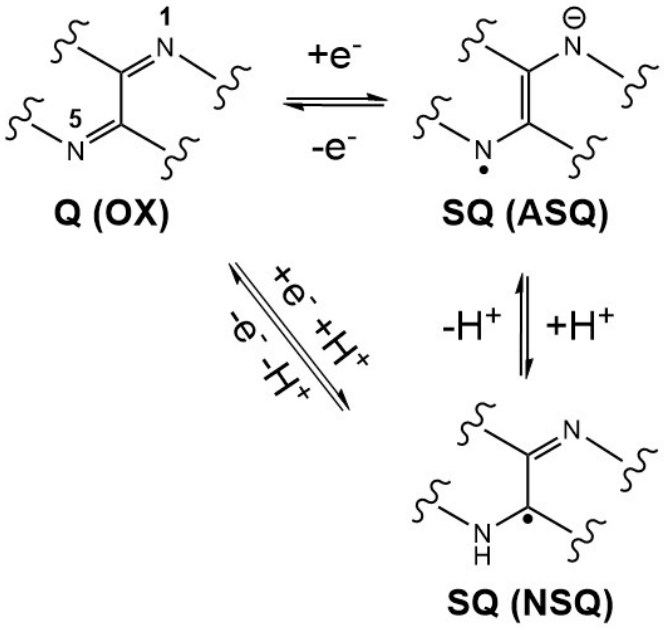

Figure 2.1 *A simplified depiction of the isoalloxazine ring indicates the importance of the N1 and N5 nitrogen atoms during the flavin redox cycle. Influencing semiquinone formation by altering the protein environment at N5 is the only logical choice, since this is the site of NSQ protonation and the ASQ radical.*

We used bioinformatics to document the prevalence of this binding site motif among structurally uncharacterized homologs of the structurally-characterized EB-ase HdrABC/MvAGD (M. thermolithotrophicus).[26] Using the methods described in **3.1**, a Sequence Similarity Network (SSN) was generated for the 10,000 closest homologs (% identity > 70%) of the bifurcating subunit, HdrA. The SSN was viewed with Cytoscape and the homologs (blue nodes) were sorted into clusters with a grey connecting edge if they shared greater than 70% sequence homology (**Figure 2.2**). No unifying families of suspected EB-ases were detected, as shown by the large number of groups with few members. It was clear that sequence similarity alone was insufficient to correlate sequence and function, which is why the gene neighborhood of each homolog was explored.

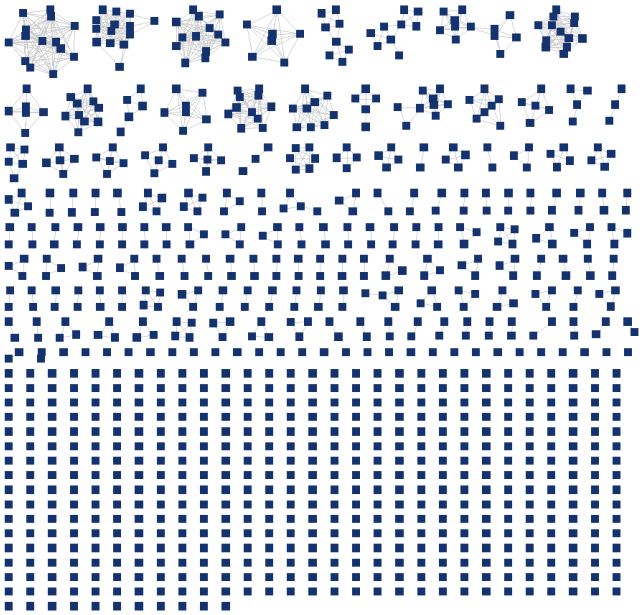

Figure 2.2 *The SSN of 10,000 HdrA homologs, clustered into families related by sequence similarity. Blue nodes represent each HdrA homolog, grey edges indicate a relationship between nodes of >90% sequence identity. No unifying families of suspected EB-ases were detected, indicated by co-expression of HdrBC in all clusters with greater than 5 members, rather than one distinct group.*

Using the EFI Genome Browser tool and a Genome Neighborhood Network (GNN) generated from this SSN (**3.1**), the gene neighborhoods within the largest SSN clusters (with > 5 members) were analyzed for the co-expression of HdrBC analogs. A gene neighborhood shows annotated predictions for gene products expressed nearby in the genome, which is useful because the genes of multi-subunit proteins are often co-transcribed and found next to each other. It was expected that HdrA homologs that exist as a HdrABC-analogous trimer would share distinct sequence similarity, and likely also bifurcating activity. HdrB is categorized in the Interpro database, which classifies protein families, as “PF02754”, a cysteine-rich domain usually found at the C-terminus of Hdr-like proteins. HdrC belongs to the Interpro family “PF13183”, a 4Fe-4S dicluster domain superfamily. HdrA homologs of the SSN

were sorted by co-expression of genes belonging to these interpro families within ± 5 genes of the HdrA-like gene in each organism's sequenced metagenome. At least one node from each cluster co-expressed homologs of HdrBC, but surprisingly, these nodes did not all belong to the same group. We expected that bifurcating homologs of HdrA would have genes similar to HdrBC up- or down-stream in their gene neighborhood, and that these would share high sequence similarity. Since there was no cluster to which all the suspected EB-ase homologs belonged, every HdrABC homolog was extracted from the dataset for manual analysis.

HdrA has a Lys residue in its FMN binding site that is suspected to affect the redox cycle and energies of the flavin. We determined the prevalence of Lys at this position by aligning suspected bifurcating homologs of HdrABC with the query sequence, HdrA. Using structured query language (SQL) commands for database management (**3.1**), a subset of the GNN dataset was created that contained all HdrA homologs (5337 in total) with HdrBC-like genes present in their gene neighborhood. These sequences were aligned using Clustal Omega's multiple sequence alignment (MSA) tool alongside the original HdrA query sequence (EMBL-EBI).

This significant residue (Lys^{409}) was found in the original HdrA query sequence of the MSA, and the residues aligned with that position were analyzed (**3.1**). Homologs with no aligned residue at this sequence position were removed (24% of the sequences). The remaining sequences were used to generate a new MSA (including the original HdrA sequence) that excluded homologs without a residue at this position. The MSAs were used to generate sequence logos to show the relative abundance of residues at this position (**Figure 2.3)**. The results indicate that Lys is highly conserved

at this position, with His and Arg as the second and third most abundant, respectively. This suggests that a positively-charged basic residue it is important for protein structure or function. For this reason, we chose to further explore the effect of Lys, His and Arg on redox properties of general flavoproteins to understand how it may contribute to potential inversion. 5% of HdrA homologs were found to have some other residue in this position, including Ala, Gly, Asp, Tyr, and Phe. Since the function of these proteins is unknown, we cannot speculate on the effect of these residues on potential inversion, but it may indicate that there are other residues crucial to flavin redox behavior besides that at N5. The lack of residue aligned with Lys^{409} in 24% of HdrABC homologs may indicate that residues of their flavin binding sites may not be ordered sequentially, since the gene sequence does not account for protein tertiary structure.

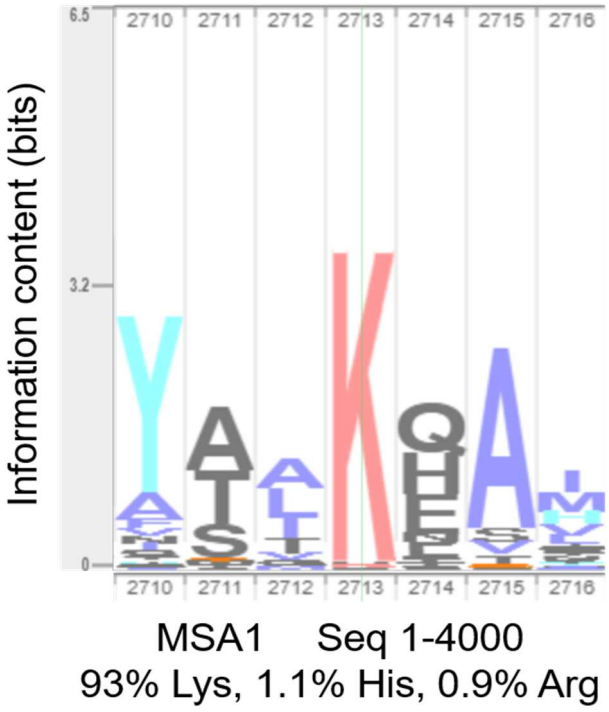

Figure 2.3 *iLOV as a Flavoprotein Platform for Investigating Potential Inversion Sequence logo depicting the relative abundance of residues aligned with Lys[409], representing 98.5% of the dataset.*

2.2 Mimicking an Authentic EB-ase using iLOV: Mutagenesis Strategy

Since iLOV is small (13.4 kDa) and has few aromatic residues, initial attempts to measure protein concentration via Bradford Assay resulted in under-quantitation when comparing to a calibration curve made with bovine serum albumin. To resolve this, iLOV was expressed as a fusion construct with an N-terminal small ubiquitin-like modifier (SUMO) tag. The addition of SUMO (11.1 kDa) increased the number of aromatic residues by 1 Tyr and 4 Phe, and nearly doubled the MW of the construct. Since Bradford Reagent binds the amide backbone and reacts with pi-stacking residues Trp, Tyr and Phe, the fusion protein increased the sensitivity of the Bradford

assay. Moreover, SUMO tags can increase solubility and may assist in the expression of iLOV mutants prone to aggregation. We designed our construct to have an N-terminal 6-His tag on SUMO, rather than iLOV. Since SUMO has an associated protease that recognizes tertiary structure and specifically cleaves at the C-terminal end of SUMO, we were able to isolate iLOV without a His-tag after proteolysis. This ensures that the tag will not interfere with redox behavior.

Using iLOV as a platform amenable to augmentation, we aim to determine the implications of residues near the flavin N5 on redox behavior, which will enable us to elucidate the role of the secondary coordination sphere of the flavin on potential inversion, a hallmark of all known EB-ases. Even conservative mutagenesis of native EB-ases can result in poor flavin binding affinity and changes in protein conformation.[24] A benefit of our system is that iLOV is small (13.4 kDa) with no known conformational motion, and it is relatively resilient to mutations close to the flavin. This allows us to observe the effect of mutations that could otherwise be deleterious on native EB-ase function.

The iLOV protein has a Gln (Q104) within 3.5 Å of the FMN N5.[33] We selected this residue for systematic SDM of iLOV to mimic the protein environment of HdrA by the replacement of Q104 with Lys near the flavin N5. As a comparison, we substituted Ala in the same position to interrogate how a non-H-bonding sidechain would affect flavin redox behavior. We performed several experiments to determine how these mutations influenced the extent of potential separation between the flavin's $1e^-$ reduction potentials. This thesis will focus on the differences in semiquinone stability resulting from these mutations, Q104K and Q104A.

2.3 Preparation of iLOV Mutants

Plasmids were designed and generated by Gibson Assembly as detailed in the methods section (**3.2**). The purification of SUMO-iLOV WT, Q104K, and Q104A followed the procedures described in **3.3**, with each step shown in the representative SDS-PAGE gel in **Figure 2.4.**

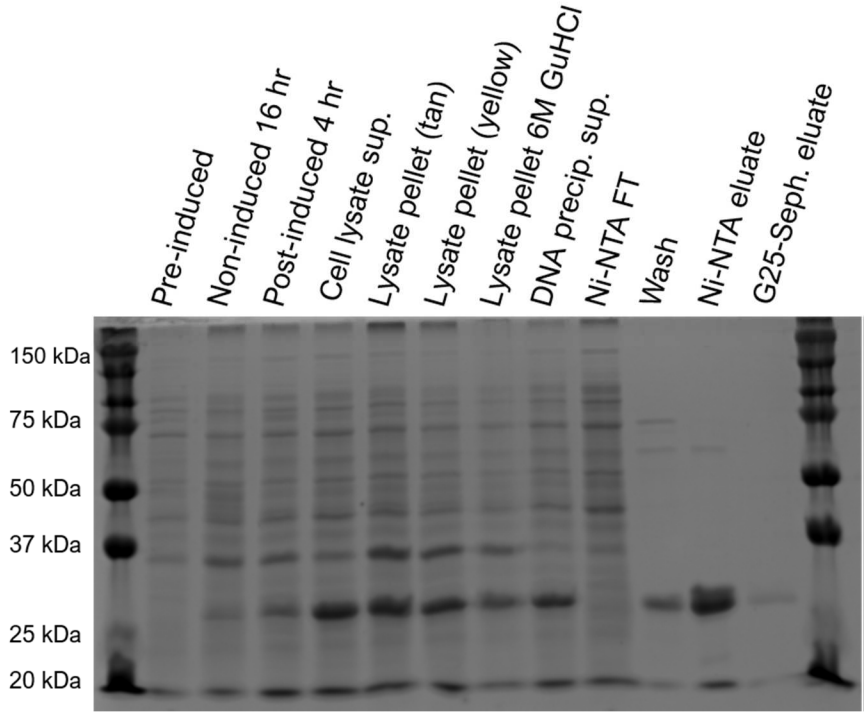

Figure 2.4 *Shows an SDS-PAGE gel of samples taken at every stage of protein purification. SUMO-iLOV is seen at ~26 kDa. The tan and yellow samples from the cell debris pellet were unable to be separated, as shown by the presence of SUMO-iLOV in the tan layer, which should only be cell membranes and other debris. The yellow layer is either unlysed cells and/or inclusion bodies.*

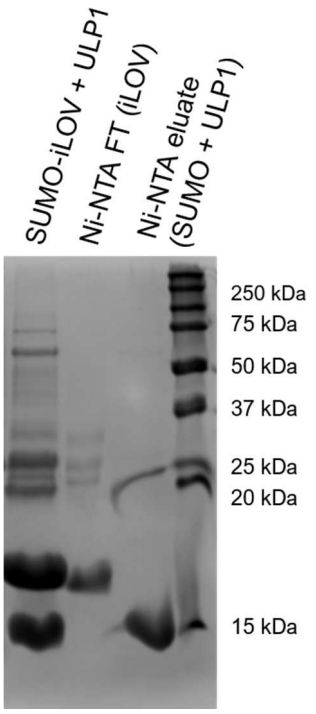

Figure 2.5 *Shows an SDS-PAGE gel of the fusion protein/protease mixture, the isolated iLOV, and the ULP1/SUMO eluted with 250 mM imidazole. SUMO-iLOV appears at ~26 kDa, iLOV at ~17 kDa, SUMO at ~15 kDa, and ULP1 at ~25 kDa. Other bands indicate the presence of contaminating proteins, which were removed in future purifications by increasing the wash volume from 20 CVs to 30 CVs.*

SUMO-iLOV was reconstituted with FMN using a thermal protocol based on protein folding principles, detailed in **3.3**. We are able to increase flavin loading of all mutants from ≤ 20% as-purified to at least 60% (**Table 2.1**). SUMO tags were removed from each protein after reconstitution and quantitation were complete (**3.3**), as shown by SDS-PAGE (**Figure 2.5**). Influence of the mutations were observed as a blue-shift of 8 nm relative to WT of the oxidized absorbance maxima (**Figure 2.6**). Reduced $iLOV_{HQ}$ was generated by chemical reduction with dithionite, using methyl viologen as an electron transfer mediator (**3.3**). Some experiments were conducted with the SUMO tag, such as protein film voltammetry.

Table 2.1. *Measurements of iLOV reconstitution after flavin loading, averaged from triplicate*

Protein Name	Highest Flavin Loading
SUMO-iLOV WT	99 ± 1%
SUMO-iLOV Q104K	69 ± 2%
SUMO-iLOV Q104A	63 ± 2%

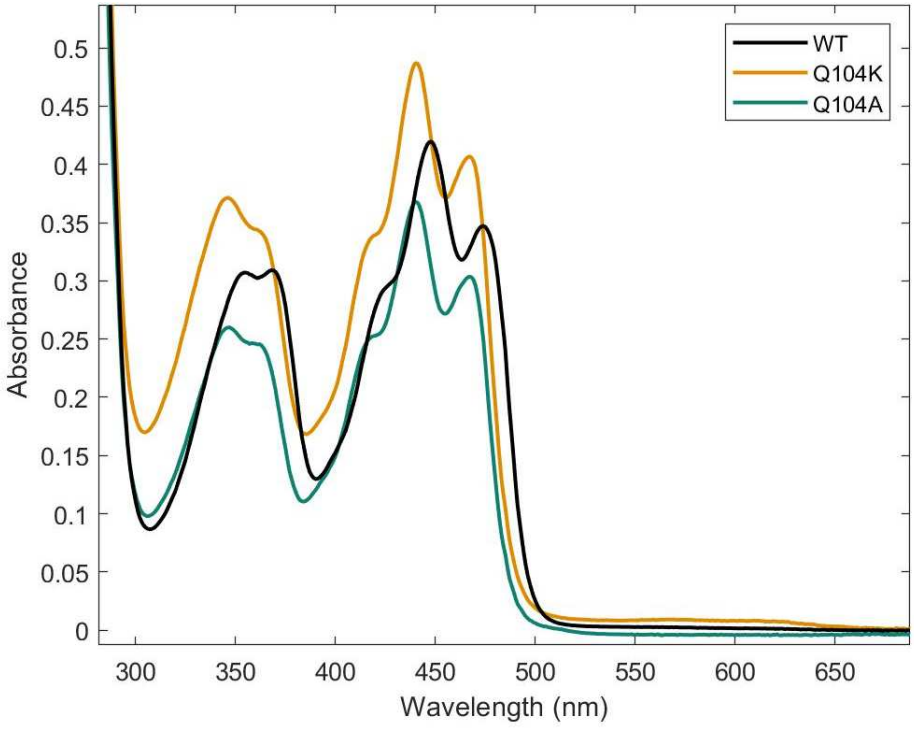

Figure 2.6 *Depicts oxidized spectra of iLOV WT and mutants at the 104 position. Influence of the mutations were observed as a blue-shift of 8 nm relative to WT.*

2.4 Determination of E' of iLOV by Protein Film Voltammetry

If a flavoprotein has inverted potentials, only the $2e^-$ midpoint potential $E'_{OX/HQ}$ can be directly accessed by electrochemical techniques. This is explained by the fact that

Gibbs free energy is a state function, and the energy associated with $2e^-$ reduction can be calculated as the sum of both $1e^-$ reduction steps (2):

$$\Delta G^0_{OX/HQ} = \Delta G^0_{OX/SQ} + \Delta G^0_{SQ/HQ} \quad (2)$$

By the relation $\Delta G^0 = nFE^0$, this corresponds to:

$$E^0_{OX/HQ} = \frac{\left(E^0_{OX/SQ} + E^0_{SQ/HQ}\right)}{2F} \quad (3)$$

Therefore, the $2e^-$ midpoint potential is observed at the average of the two $1e^-$ reduction potentials (3), although the degree of inversion cannot be determined from the $2e^-$ measurement alone. Seeing as only the $2e^-$ reduction potential of PHOT2 from *A. sativa* has been measured (E' = -308 mV), this indicates potential inversion.[30] Since iLOV is a derivative of PHOT2, it is suspected to also have inverted potentials. To determine this experimentally, protein-film cyclic voltammetry (CV) and differential pulse voltammetry (DPV) was conducted on SUMO-iLOV WT and Q104K (**Figure 2.7, Figure 2.8)**.

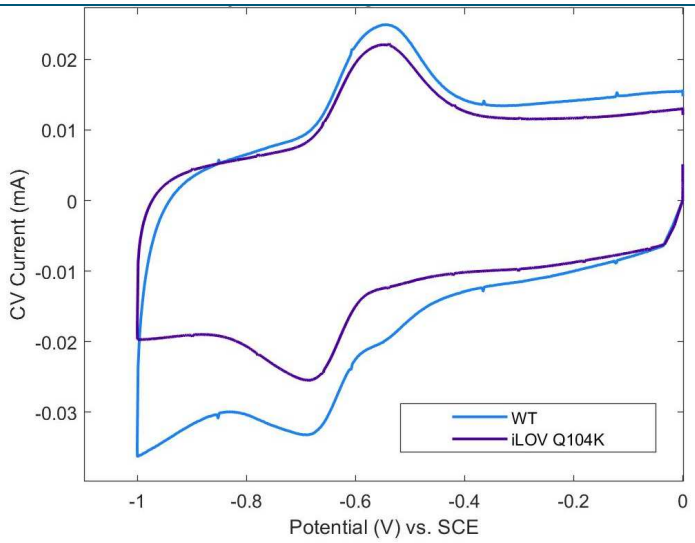

Figure 2.7 *Cyclic voltammograms of SUMO-iLOV WT and Q104K show similar E' at pH 8 (100 mM Tris HCl, 200 mM NaCl). CVs were conducted with a pyrolytic graphite edge working electrode, Pt counter electrode, and standard calomel reference electrode (sat. KCl) at a scan rate of 1 V/s).*

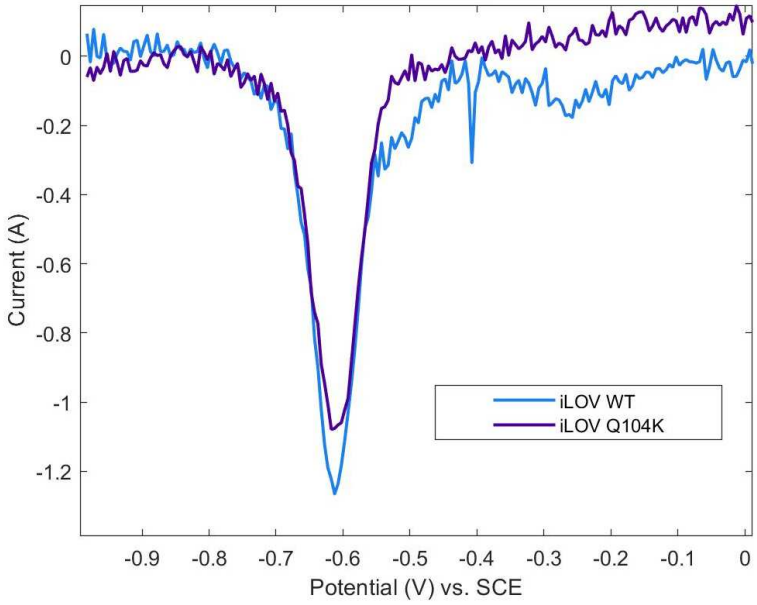

Figure 2.8 *Differential pulse voltammogram of SUMO-iLOV WT and Q104K at pH 8 show nearly identical E' (100 mM Tris HCl, 200 mM NaCl), conducted with a step height of 5 mV and a pulse height of 25 mV, with a pulse time of 10 ms using a pyrolytic graphite edge working electrode, Pt counter electrode, and standard calomel reference electrode (sat. KCl).*

The resulting voltammograms show that the reduction potentials of the WT vs. Q104K mutant are E'_{WT} = -416 ± 3 mV *vs.* SHE and E'_{Q104K} = -412 ± 7 mV *vs.* SHE, respectively, at pH 8. Error was determined from the standard deviation between triplicate measurements. The midpoint potentials determined from the cyclic voltammograms matched those obtained from DPV and indicate that the midpoint potentials are not significantly different between iLOV WT and Q104K. Since iLOV Q104K is expected to have an altered $E'_{OX/SQ}$, but an unaffected $E'_{SQ/HQ}$ (due to changes in the flavin binding site at N5, but not N1), it is surprising that $E'_{OX/HQ}$ appears unchanged, since it is the average of the two $1e^-$ potentials.

SUMO-iLOV WT showed only one signal in the DPV and CV measurements, suggesting that this is a $2e^-$ potential. While this suggests that the WT and Q104K have inverted potentials, it could also be that either the potentials are normal and so close that they appear as overlapping peaks, or that the second reduction potential is less than -1.0 V. The number of electrons transferred was calculated using the Randles-Sevcik equation (4) from the peak current obtained from the CVs taken with a 1 V/s scan rate (**Table 2.2**). A hard spheres estimation of surface coverage was performed by dividing the electrode surface area by the cross-sectional area of iLOV (1590 $Å^2$), estimated by πr^2 from the maximum distance across the protein (**3.14**).

$$i_p = \frac{n^2F^2}{4RT}\nu A\Gamma^* \qquad (4)$$

***Table 2.2**. Results from Randles-Sevcik analysis suggest iLOV undergoes a 4e- transfer from the electrode during CV*

Sample	*n* (mol e^- transferred)
SUMO-iLOV WT cathodic	4.02
SUMO-iLOV WT anodic	4.35
SUMO-iLOV Q104K cathodic	3.45
SUMO-iLOV Q104K anodic	4.23

Although $n \approx 4$ in these calculations, these values are likely an over-estimate due to erroneous approximation of the electrode surface coverage. The radius of iLOV was chosen to account for the largest cross-sectional area, since proteins are more likely to adsorb when the surface area contact is high. However, if a smaller radius is selected (ex. 12 Å, the shortest radius of iLOV), this would increase the surface coverage and lower *n* closer to 2. The hard spheres coverage estimation assumes that the electrode surface is a perfect plane; thus, higher experimental surface coverage may be observed if there are imperfections due to polishing and wear or electrode microstructure. This would also lower the value of *n* closer to the expected value of $2e^-$/protein.

2.5 Determination of E' of iLOV by Equilibration with Methyl Viologen

To confirm that the E' determined by protein film voltammetry is the same as the bulk property, reduced $iLOV_{HQ}$ was mixed electro-stoichiometrically (1:2) with methyl viologen, a $1e^-$ acceptor. Since the reduced form of methyl viologen has a characteristic absorbance at 600 nm, the concentration of the reduced form (MV^+) can be quantified spectrophotometrically and the apparent E' ($E'_{OX/HQ}$ of iLOV) can be found using the Nernst Equation (5).

$$E' = E^{\circ}_{MV+} + \frac{2.3RT}{nF}\log\left(\frac{[MV^+]}{[MV^{2+}]}\right) \qquad (5)$$

The experiment was performed under low-light conditions, with the only excitation source being the UV-Vis light source, to minimize side photochemistry (**3.4**). Formation of reduced methyl viologen cation, MV^+, was at its maximum after 3.8 hours (**Figure 2.9, Figure 2.10)**.

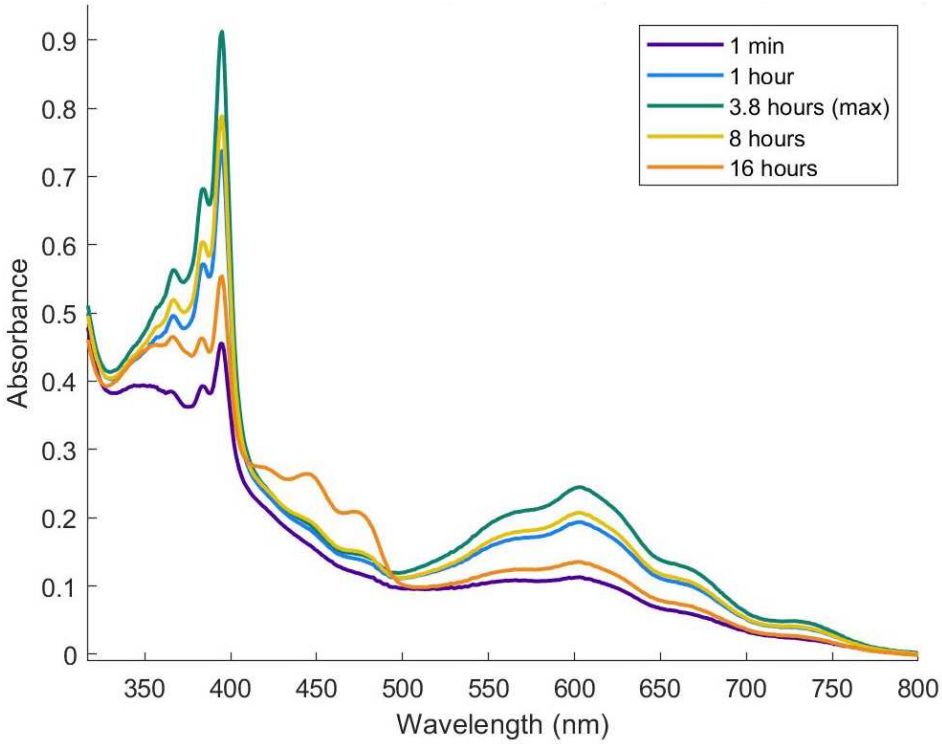

Figure 2.9 *Shows the formation and decay of MV^{+} during equilibration with $iLOV_{HQ}$ WT.*

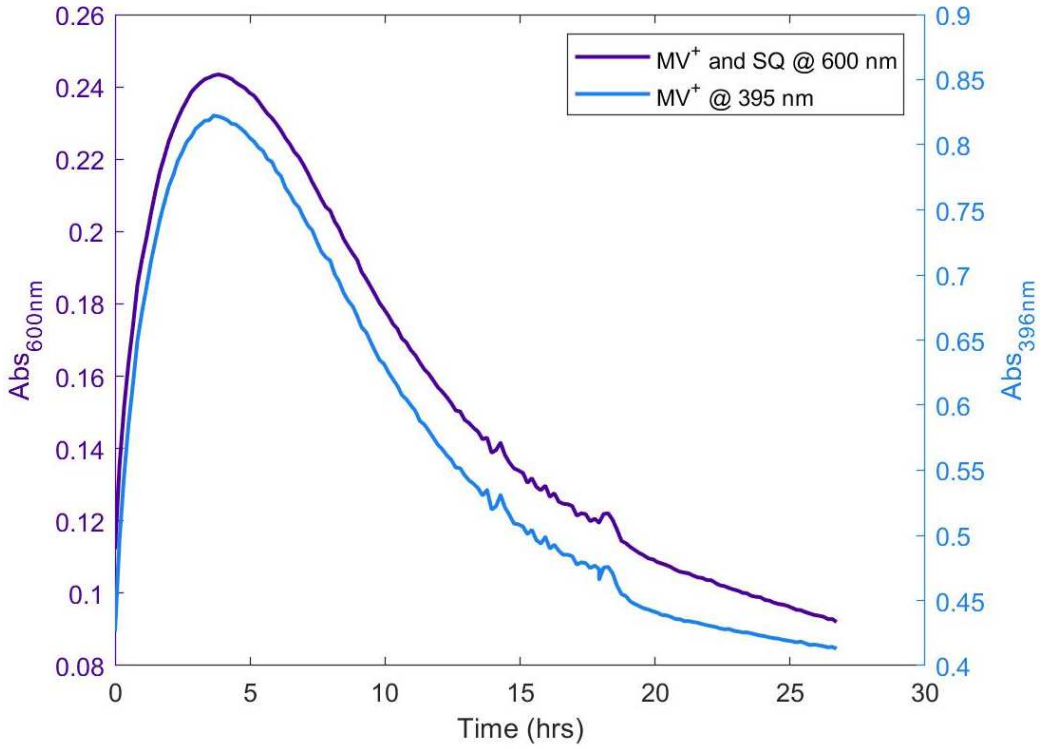

Figure 2.10 *Shows the kinetic scheme of MV^+ reduction by $iLOV_{HQ}$ WT. The cause of MV^+ decay is unknown, but suspected to be slow oxidation by O_2 leaking into the cell.*

Using an experimentally-determined ε_{600} value of 11.2 mM^{-1} cm^{-1}, the absorbance of MV^+ at 600nm was used to determine the concentration of the reduced viologen species (assuming that the contribution from $iLOV_{SQ}$ at this wavelength was negligible). Using the Nernst equation, the E' of iLOV was calculated as –407 mV *vs*. SHE, in agreement with the potential measured by protein film voltammetry (**2.4**). However, since oxidation of MV^+ prevented the measurement of the true equilibrium concentration, we cannot say that this is an equilibrium measurement.

2.6 Characterization of Semiquinone State by EPR Spectroscopy

While NfnI and ETFs preferentially exhibit proton-coupled electron transfer (PCET) from the HQ state to yield an anionic semiquinone (ASQ), it is not known how the protein environment influences this propensity.[11,18] We sought to compare the protonation state of the iLOV WT and iLOV Q104K SQ. To determine the identity of the iLOV WT semiquinone, UV-vis spectroscopy was used to characterize an equimolar mixture of $iLOV_{OX}$ and $iLOV_{HQ}$, which forms $iLOV_{SQ}$ by comproportionation. The resulting spectra showed a characteristic absorbance feature between 520 nm – 650 nm, indicative of the NSQ. This was also true of the Q104K mutant, which shows the same UV-vis feature (**Figure 2.11**). ASQ absorbance features at ~380 nm were not observed.

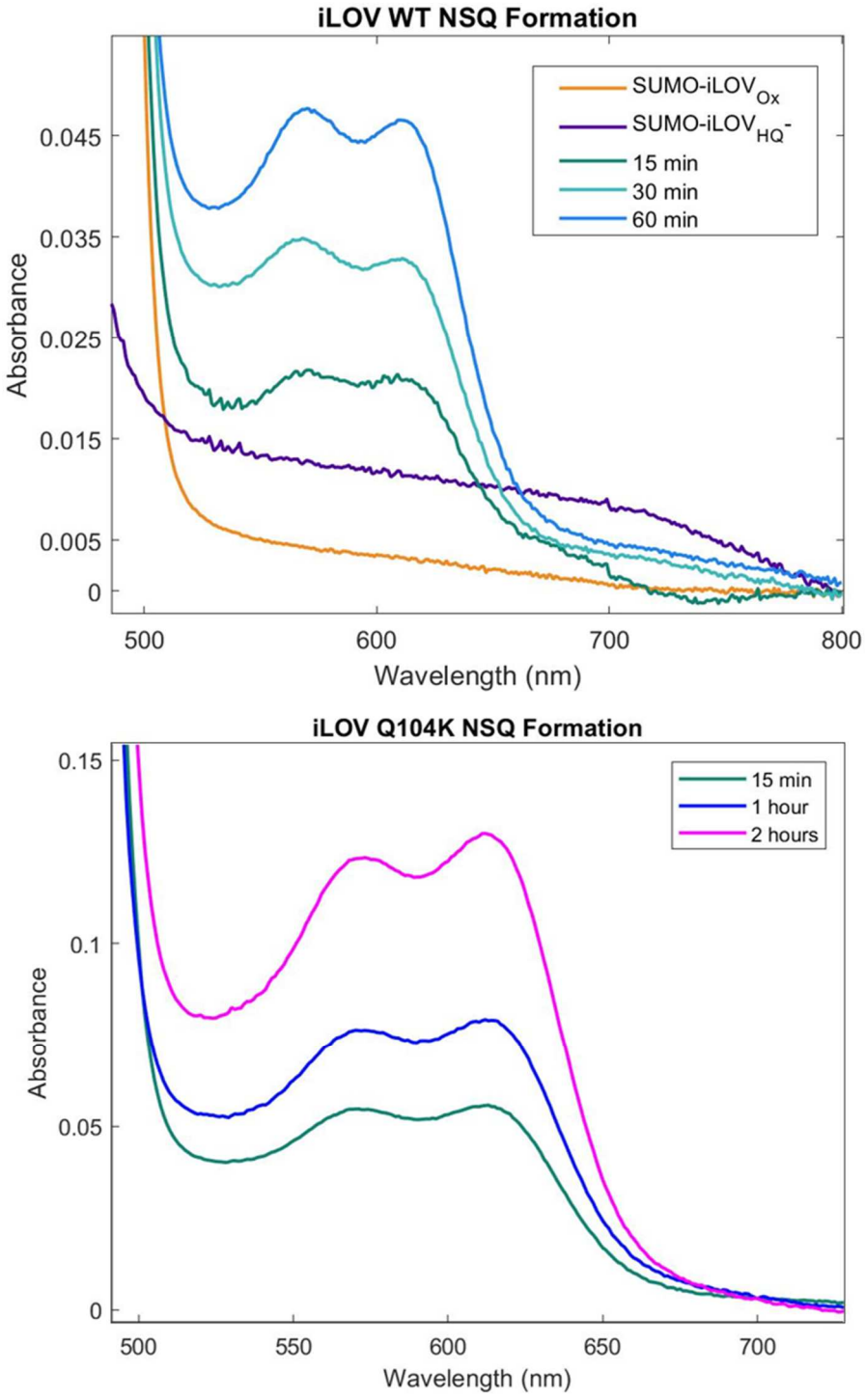

Figure 2.11 *Shows the formation of NSQ over time in iLOV WT (left) and Q104K (right), in the presence of ambient light.*

To further characterize the NSQ, an X-band EPR spectrum was obtained for iLOV$_{SQ}$ WT (**Figure 2.12**). The resulting signal indicates that an unpaired electron was present in the sample, corresponding to the semiquinone radical species. The protonation states of protein-bound flavin radicals can be differentiated by the peak-to-peak line widths of their EPR signals in the first derivative. ASQ radicals have shorter peak separation (1.2-1.5 mT) than NSQ radicals (1.8-2.0 mT) due to hyperfine coupling from the N5 proton[31]. Because H-bonding can occur between the flavin and surrounding residues in the binding site, further EPR signal width broadening can occur. We measured the peak-to-peak distance as 20 mT, suggesting a NSQ, and indicating that significant H-bonding occurs between the flavin and iLOV protein environment.

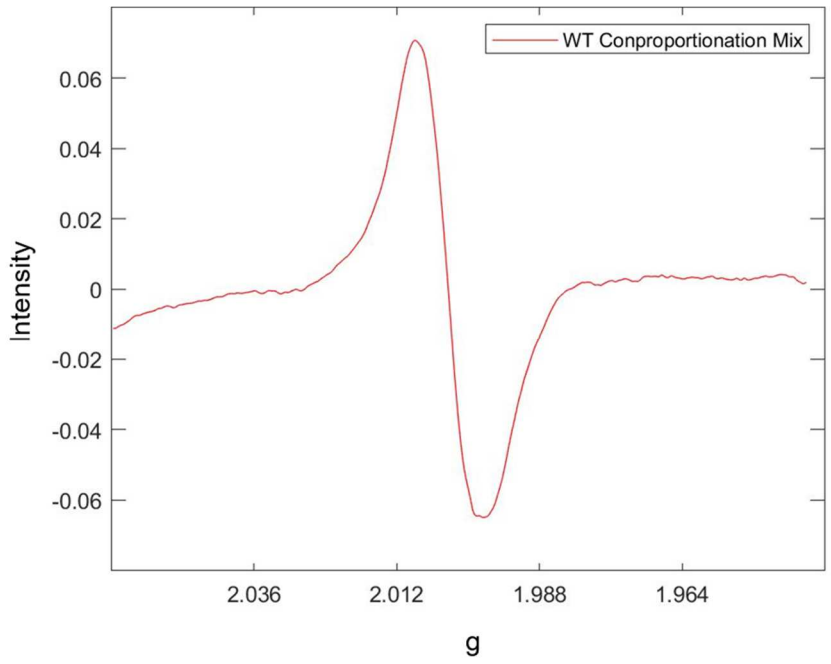

Figure 2.12 *Shows the X-band EPR spectrum obtained from a mixture of WT iLOV$_{OX}$ and iLOV$_{HQ}$ after 3 hours.*

2.7 Determination of ΔE for iLOV 1e- Potentials

While the $1e^-$ midpoint potentials of inverted flavins cannot be accessed via electrochemistry, they can be inferred from the equilibrium constant. Due to comproportionation, the $[SQ]_{eq}$ is a measure of the splitting between the two 1e- potentials. $[SQ]_{eq}$ was generated by the comproportionation of equal parts $iLOV_{HQ}$ and $iLOV_{OX}$ to form $iLOV_{SQ}$ (6). The stability of the semiquinone is inherently related to its reduction potential, which is reflected by its equilibrium concentration (7).

$$iLOV_{HQ} + iLOV_{OX} \rightleftharpoons 2\mathrm{iLO}V_{SQ} \quad (6)$$

$$K_{eq} = \frac{[SQ]^2}{[HQ][OX]} \quad (7)$$

K_{eq} is related to ΔE by (8)[28] :

$$K_{eq} = \frac{[SQ]^2}{[HQ][OX]} = e^{(\frac{-nF\Delta E}{RT})} \quad (8)$$

Rearranging in terms of ΔE (9).

$$\Delta E = -\frac{RT}{nF}\ln(K_{eq}) \quad (9)$$

This relation allows us to calculate ΔE from K_{eq}, which can be determined experimentally from a system where iLOV is at electronic equilibrium (reduced by $1e^-$

eq.). $[SQ]_{eq}$ can be determined spectrophotometrically by its characteristic absorbance in the visible spectrum. Since the initial concentrations of $iLOV_{HQ}$ and $iLOV_{OX}$ are known, their equilibrium concentrations can be determined (**3.4**). This allows for the spectrophotometric determination of K_{eq}, and therefore ΔE.

The comproportionation of iLOV is affected by the presence of white light, therefore it is crucial that these experiments are performed in the dark with short scan rates (~600 nm/min) and/or restriction to the red end (800 nm – 500 nm) of the visible spectrum. Exposure to light causes an increase of semiquinone concentration, likely produced by photochemical side reactions. This is observed by the loss of $iLOV_{HQ}$, and an increase in $iLOV_{SQ}$ and $iLOV_{OX}$ (**Figure 2.13, Figure 2.14**).

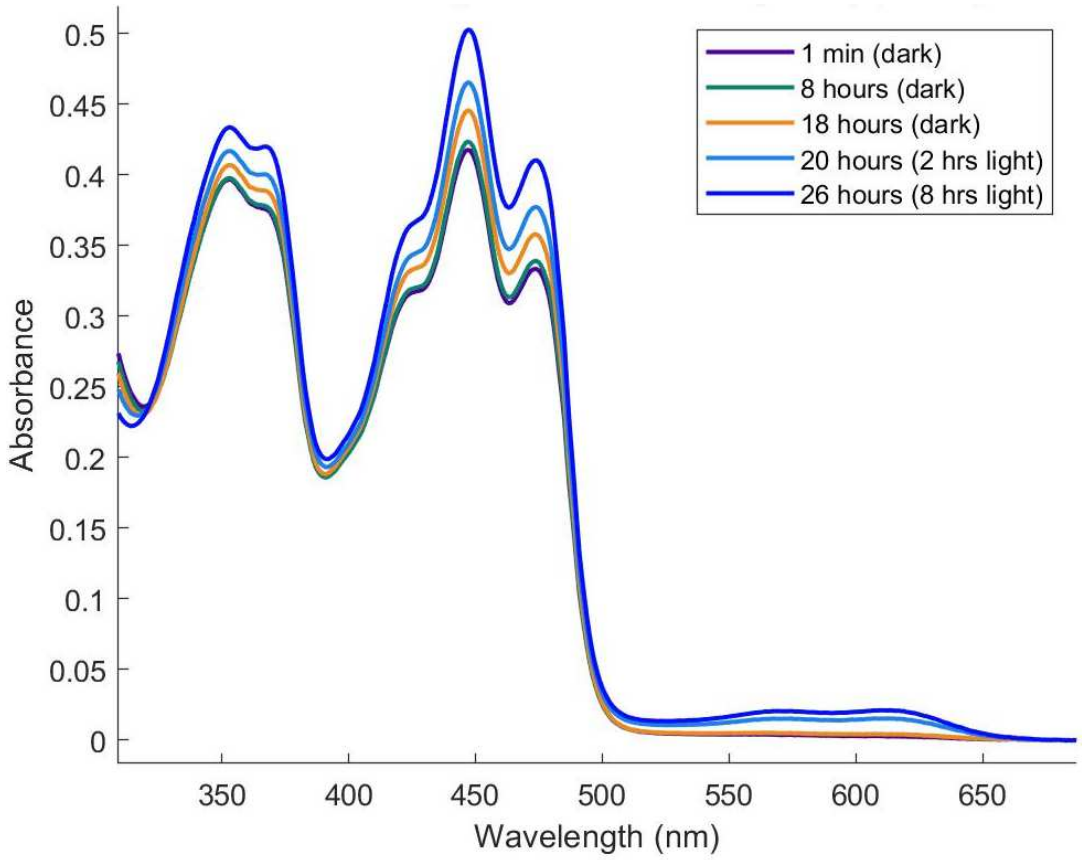

Figure 2.13 *Shows the production of neutral semiquinone (520 nm - 650 nm) when the sample is exposed to ambient white light.*

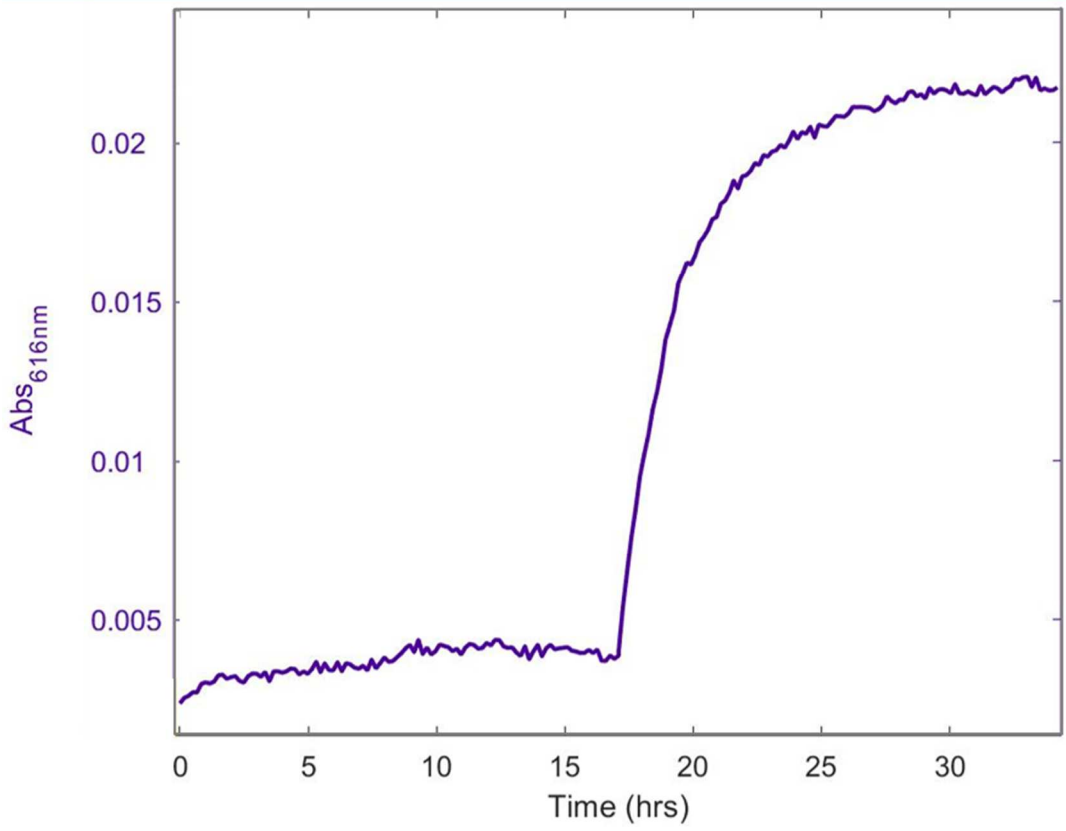

Figure 2.14 *Tracks the change in absorbance of the semiquinone over time. The sample was protected from ambient light until 18 hours, when an exponential growth of SQ is observed.*

From the iLOV WT comproportionation data, the semiquinone appears to reach its equilibrium concentration after 2 hours (**Figure 2.15, Figure 2.16**). The absorbance at 616 nm was used to quantify $[SQ]_{eq}$, using the molar extinction coefficient of 4.25 mM^{-1} cm^{-1} from *D. vulgaris* flavodoxin NSQ.[32] The equilibrium concentrations of $iLOV_{HQ}$ and $iLOV_{OX}$ were determined by the methods described in **3.5.**

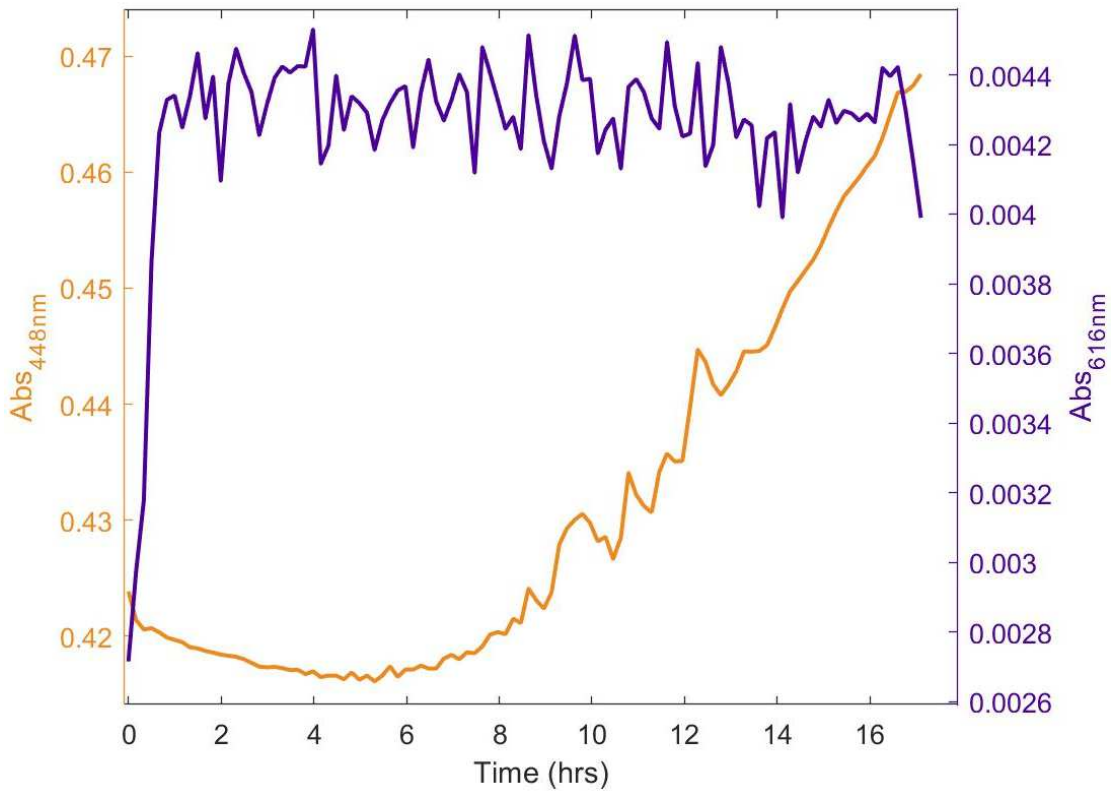

Figure 2.15 *Shows the change in absorbance of the neutral semiquinone (purple) vs. the oxidized form in the absence of light. Differences in kinetic trends likely indicate that side chemistry is occurring to produce* $iLOV_{ox}$

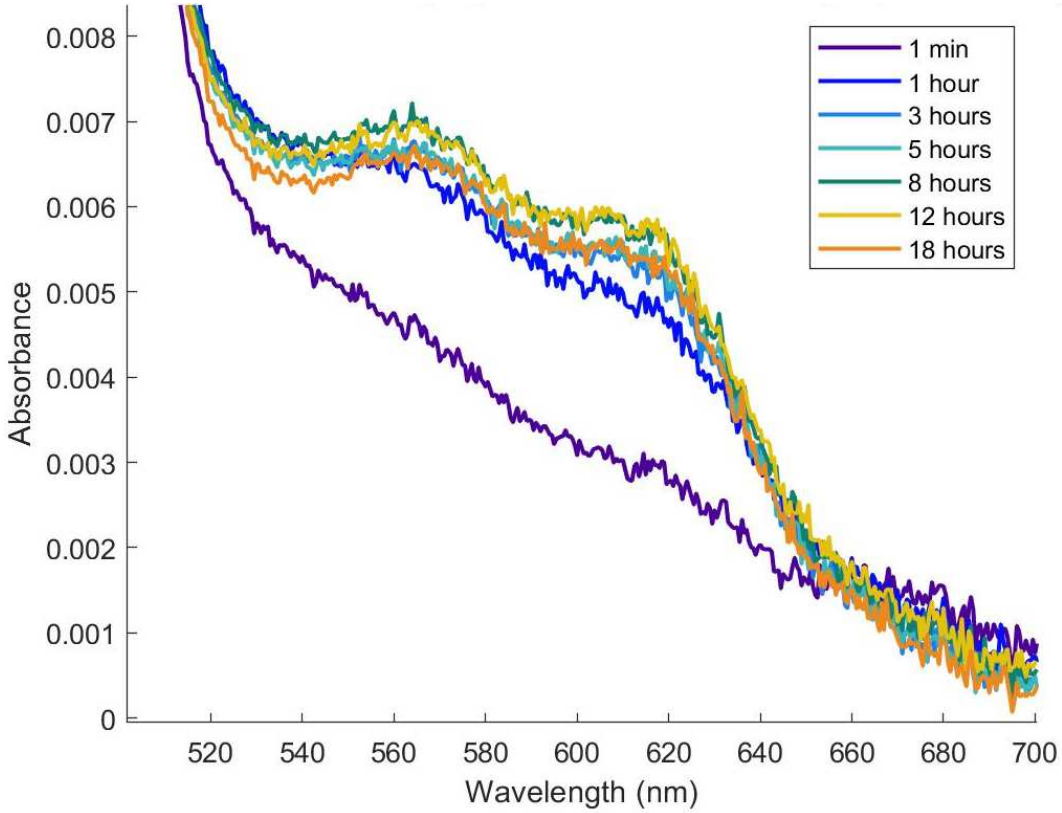

Figure 2.16 *Shows the formation of $iLOV_{SQ}$ with no ambient light, and scans restricted to 800-500 nm. Sloping is likely due to the absorbance of the oxidized form between 400-500 nm.*

The WT data is the only case where the sample was kept completely in the dark, and therefore represents the most accurate thermodynamic $[NSQ]_{eq}$. The iLOV Q104K and iLOV Q104A spectra were taken after approximately 5 minutes of light exposure, which likely did not increase the observed semiquinone concentration by a significant amount (**Figure 2.17, Figure 2.18**).

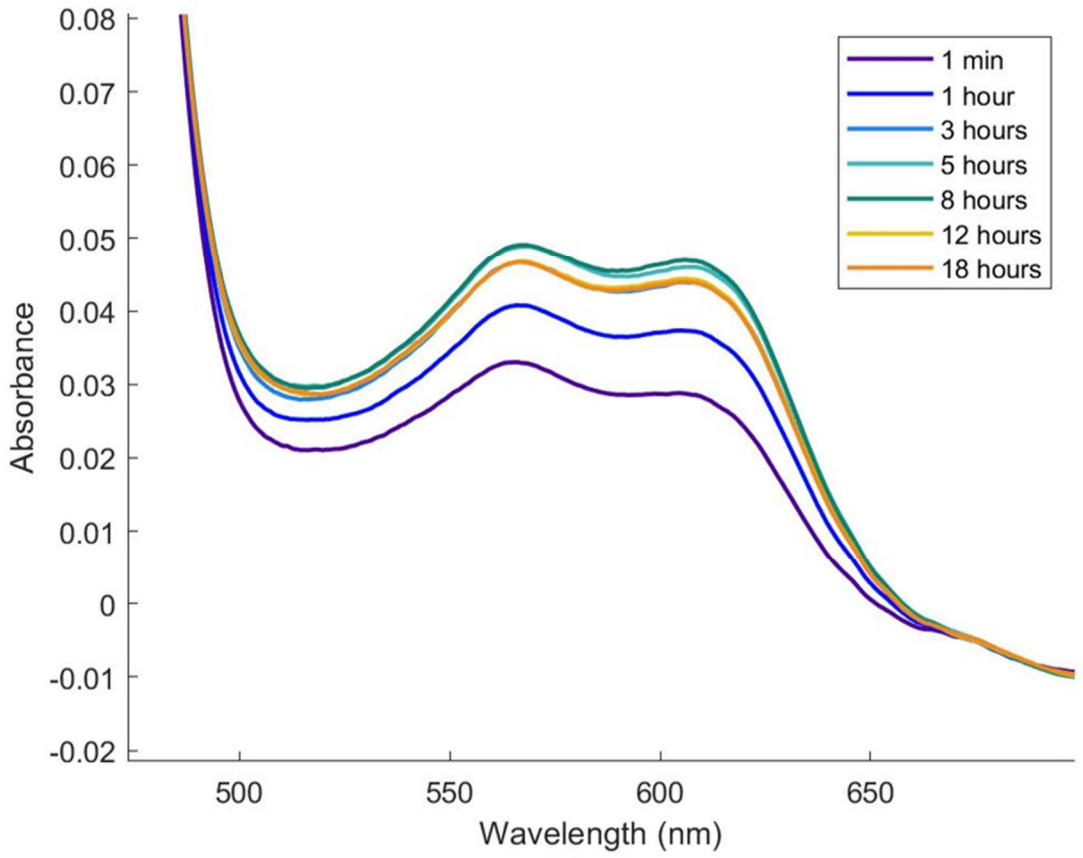

Figure 2.17 *Shows the semiquinone absorbance feature from the comproportionation of iLOV Q104K, which is markedly larger than the WT or Q104A mutant.*

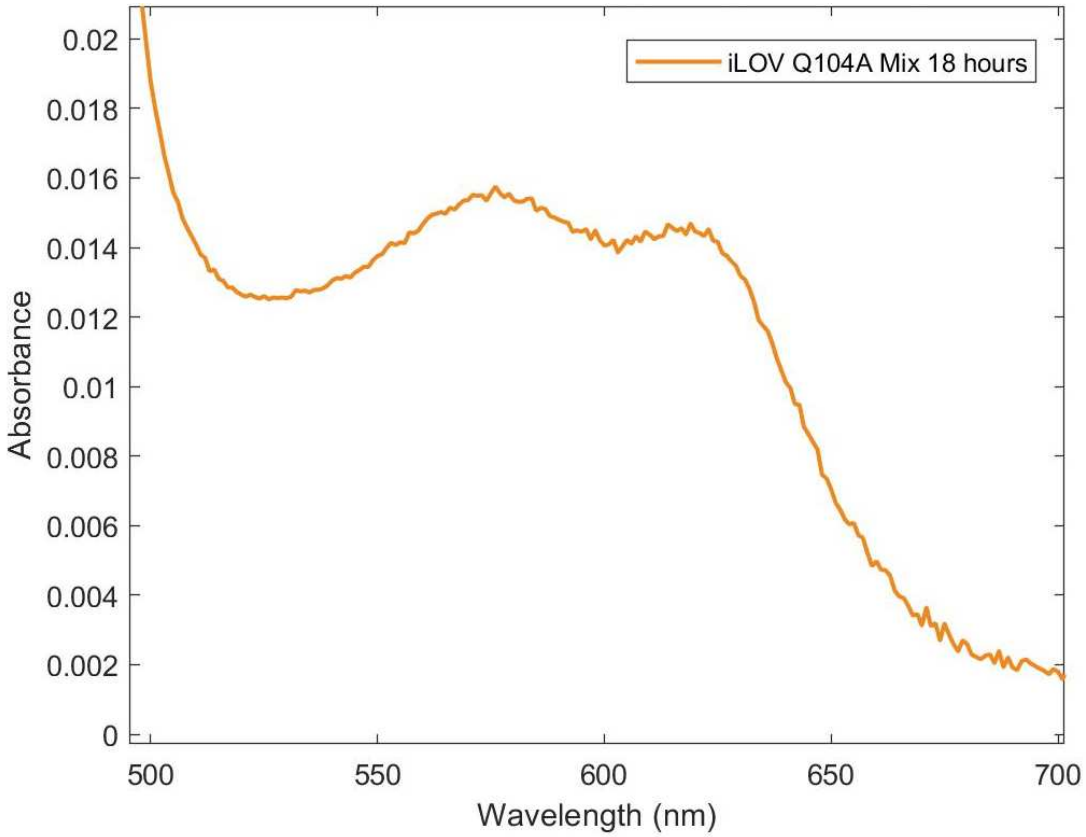

Figure 2.18 *Shows the endpoint spectrum of the iLOV Q104A comproportionation after 18 hours in the glovebox, with minimal ambient light.*

Table 2.3*. Compares potential inversion of iLOV WT and mutants.*

Sample	SQ $A_{616\ nm}$	[SQ] (μM)	ΔE (mV)
iLOV WT	0.006	1.4	-146
iLOV Q104K	0.045	10.6	-32
iLOV Q104A	0.015	3.5	-97

Using the spectral data to infer equilibrium concentrations of the reactants and products, ΔE was calculated for iLOV WT, iLOV Q104K and iLOV Q104A. The results of the comproportionation experiments suggest that iLOV WT has the most inverted $1e^-$ potentials, since it has the smallest equilibrium population of semiquinone (**Table 2.3**). Since Ala is non-ionizable and cannot donate H-bonds to the flavin, it is expected to destabilize the NSQ. The substitution of Ala at the N5 position appeared to stabilize the NSQ, as evidenced by a greater NSQ equilibrium concentration, although this is likely due to exposure to ambient light immediately prior to obtaining the spectrum from the comproportionation experiment. The addition of Lys at N5 appears to influence the thermodynamic properties of the flavin such that $E'_{OX/SQ}$ is more positive. Since the protein film voltammetry measurement of $E'_{OX/HQ}$ did not differ between iLOV WT and Q104K, we can deduce that $E'_{SQ/HQ}$ must also be more positive.

Among LOV proteins, the Gln proximal to the flavin N5 is highly conserved.[33] It has been shown to participate in conformational changes of the LOV photocycle by changing its role from an H-bond donor to acceptor by flipping of the amide sidechain. In the oxidized “dark” state, Gln donates an H-bond from the amide to the flavin O4, and perhaps also to N5. After photo-induced reduction to the NSQ state by an adjacent Cys, Gln flips to become a H-bond acceptor of the NSQ N5 H-bond to the amide carbonyl oxygen.[34] While iLOV lacks blue-light sensing ability due to its C45A mutation, it is likely that Q104 in iLOV WT plays a similar role in stabilization of the NSQ.

Replacement of Gln with Lys may stabilize the NSQ by a similar mechanism, assuming Lys can act as the proton donor for NSQ reduction by PCET. Deprotonated

Lys yields a lone pair which could form a H-bond with N5 of the NSQ, and the side chain structural flexibility may allow for repositioning to form this H-bond. Considering that the pK_a of Lys is highly tunable, and that protonated amines are energetically unfavorable within the low-dielectric environment of a protein pocket, it is suspected that K104 exhibits a lower pK_a (<10) than when free in solution. Compared to Q104 (pK_a > 15), which is not acidic, Lys better matches the pK_a of the iLOV FMN ASQ/NSQ transition, which is estimated to be below 8 (based on UV-vis evidence of NSQ formation at pH 8 and below). The unique ability of Lys to act as the proton donor would likely favor PCET and yield higher equilibrium concentrations of NSQ.

If Lys indeed has a pK_a close to that of the flavin, this would also stabilize the H-bond between Lys and N5 due to their similar proton affinities. Since the ΔpK_a between the flavin and Lys is likely smaller than the ΔpK_a between the flavin and Gln, it is suspected that H-bond stabilization by Lys would more significantly lower the energy of the NSQ, and therefore raise $E'_{OX/SQ}$. In addition to H-bond strength, the replacement of Gln by Lys would lower the energetic barrier to flavin protonation by acting as both the H-bond donor and the proton donor. The H-bond would effectively transfer ownership from Lys to the flavin, which would allow Lys to switch roles from H-bond donor to acceptor without the energetic cost of breaking the bond (**Figure 2.19**). Contrastingly, Gln can act as an H-bond donor but not as an acid, which means that its H-bond would need to be broken before protonation of the flavin by another source could occur. Therefore, it is likely that the enthalpic cost of this Gln H-bond breaking and re-forming during SQ protonation in the WT destabilizes the NSQ relative to the Q104K mutant. The flipping of the Gln sidechain to re-orient the carbonyl to act as a H-bond acceptor

to the NSQ would also have a higher associated entropic cost from solvent reorganization.

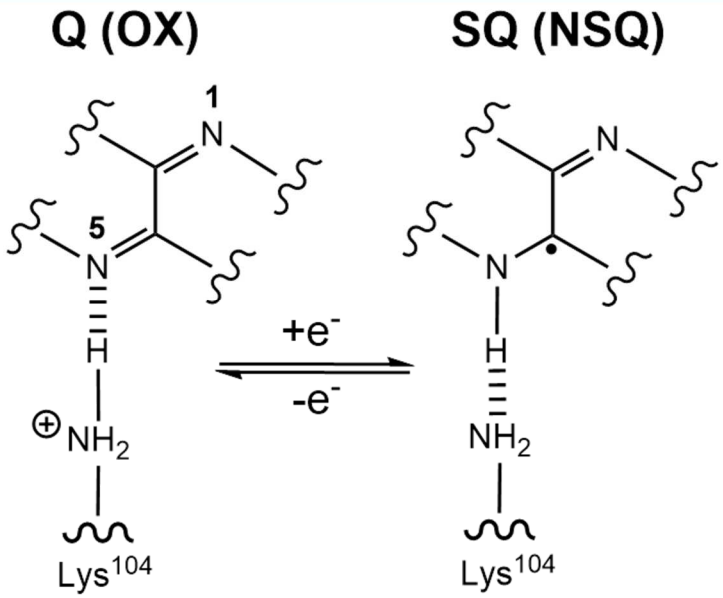

Figure 2.19 *Shows a possible mechanism for acidic H-bond transfer by K104 to form the protonated NSQ*

Considering that the 2e$^-$ E'$_{OX/HQ}$ of iLOV Q104K did not differ significantly from the WT (**2.4**), and that the 2e$^-$ midpoint potential is the thermodynamic average of the two 1e$^-$ couples, I must be true that E'$_{OX/SQ}$ stabilization must be accompanied by E'$_{SQ/HQ}$ destabilization of equal magnitude. Since E'$_{SQ/HQ}$ is predicted to be affected by alteration of the flavin N1 environment, and no mutations occurred at this site, it is not clear how this may occur. One possibility is that our protein film voltammetry data does not reflect E'$_{OX/HQ}$ as a bulk property, if iLOV surface-bound chemistry is distinct. Structural characterization of iLOV Q104K would allow for a deeper understanding of how the Lys interacts with the flavin to stabilize the NSQ, and possibly cause changes

in binding configuration to destabilize the HQ. Future studies of mutagenesis near the N1 position will also elucidate changes in $E'_{OX/HQ}$ as a result of $E'_{SQ/HQ}$ destabilization.

2.8 Spectroelectrochemical Measurement of iLOV

It is also possible to reduce $iLOV_{OX}$ halfway, with $1e^-$ eq., using an electrode to supply current at a potential equal to the E' of iLOV. This would allow for the procurement of a clean UV-vis spectrum of the SQ and HQ forms of iLOV without contributions from chemical reductants, such as dithionite. Although UV-vis spectra of comproportionated mixes did not indicate ASQ absorption features at ~380 nm, this region is obscured by the sloping absorbance of the HQ. Spectroelectrochemistry would enable UV-vis characterization of the semiquinone to confirm that only the NSQ form is present. Simultaneous electrochemistry and spectrophotometry (SpecEchem) of iLOV alone would theoretically allow for the measurement of $[SQ]_{eq}$ at electrochemical equilibrium without spectral contributions from chemical reductants or dyes.

After measuring E' by cyclic voltammetry, $iLOV_{OX}$ was subjected to chronoamperometry at the same E' in a spectrophotometric cell. A platinum flag electrode was used to allow light to pass through to the detector, which also maximizes the surface area in contact with the protein. Initial attempts to reduce iLOV with a bare electrode showed slow and resistant reduction, even when an overpotential was applied (**Figure 2.20**).

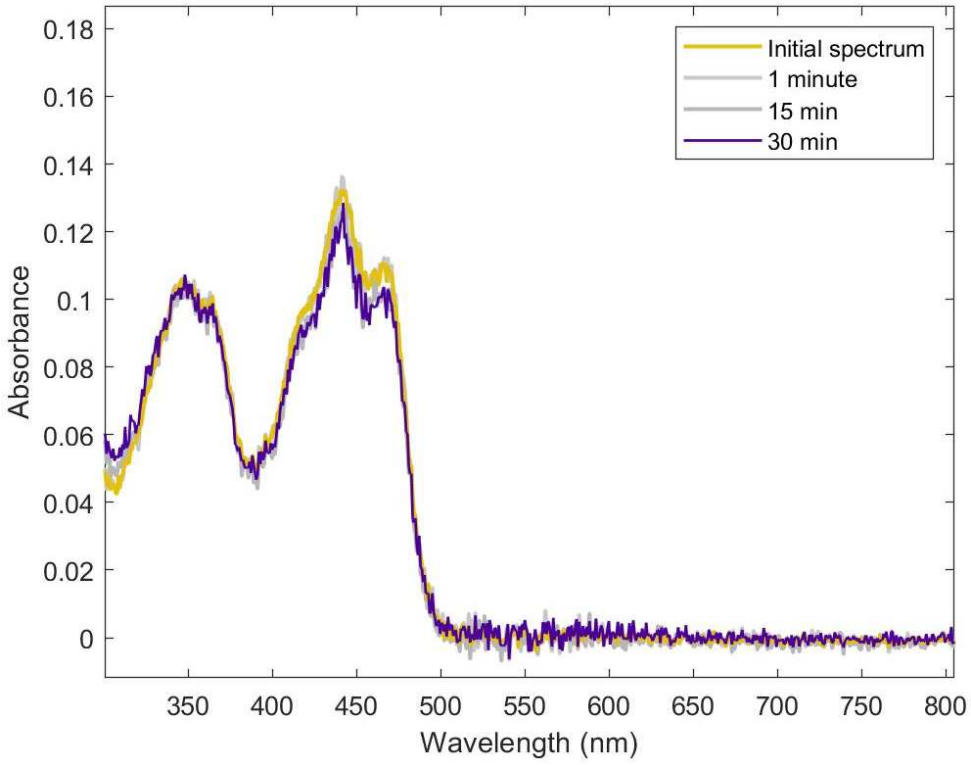

Figure 2.20 *Reduction of iLOV without electron transfer mediators was not as rapid as expected, despite an applied overpotential of 137 mV.*

Upon addition of equimolar benzyl viologen to act as a small molecule electron transfer mediator, the electrochemical reduction of iLOV was observed. However, the spectral interference of benzyl viologen is significant and obscures the neutral semiquinone feature between 520 - 650 nm. To accurately measure $[SQ]_{eq}$ spectrophotometrically, the methyl viologen contribution must be subtracted from the convoluted spectra. This can be done by measuring the absorbance spectrum of benzyl viologen alone at the same applied potential. Surprisingly, when the E' of iLOV Q104K (-597 mV *vs*. Ag/AgCl (3 M NaCl) at pH 8, measured by Pourbaix analysis (data not shown)) was applied to the specEchem cell containing equimolar $iLOV_{OX}$ and benzyl viologen, more than half of iLOV was reduced (**Figure 2.21**).

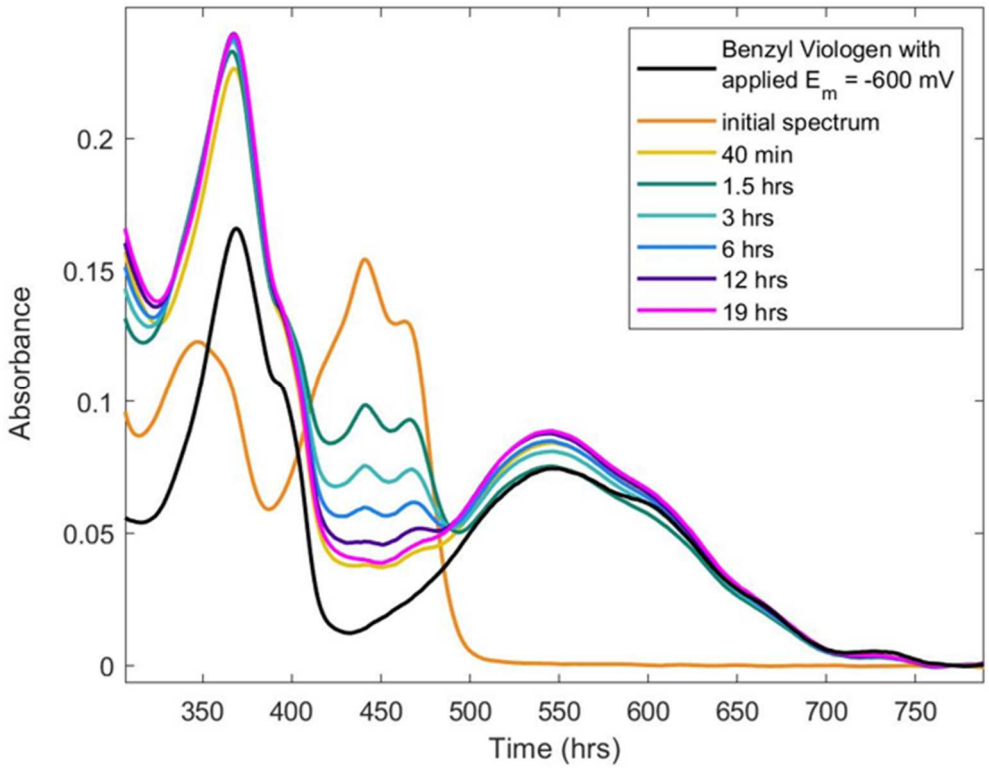

Figure 2.21 *Shows the spectroelectrochemical reduction at E' of iLOV Q104K with electron transfer mediator, benzyl viologen. Black trace was overlaid to show expected* $A_{600\ nm}$ *of benzyl viologen at the applied E'.*

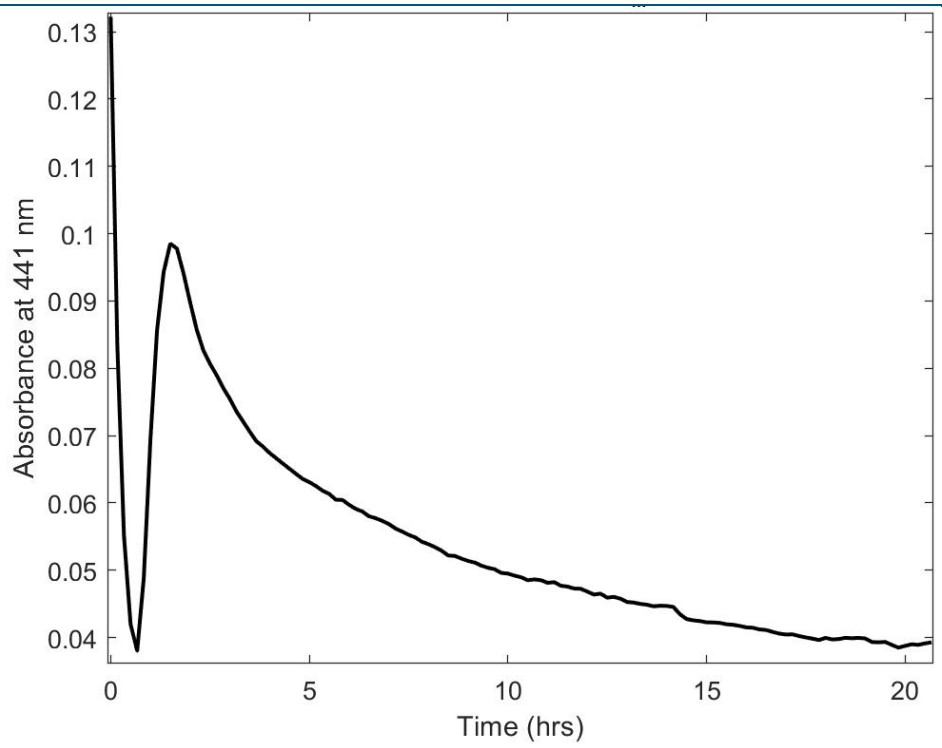

Figure 2.22 *Shows the reduction kinetics of iLOV Q104K with an unusual rapid reduction phase, followed by re-oxidation and a logarithmic reduction phase.*

Further, unusual kinetics of reduction were observed when tracking the absorbance of the oxidized flavin (448 nm) over time. While the noise can be attributed to unequal loss of light across the visible spectrum through the fiber-optic couples of our spectrometer light path, this does not explain the rapid reduction and re-oxidation observed upon initiation of chronoamperometry, nor the final concentration of reduced iLOV (**Figure 2.22**). While it was suspected that the E' of iLOV may have been measured as more positive than reality, this is unlikely because the benzyl viologen equilibration experiment results are in agreement with the cyclic voltammetry values (**2.4, 2.5**), although the benzyl viologen equilibration did not reach true equilibrium. Further investigation into these issues will address this unexpected response.

References

[30] Arents, J., Perez, M., Hendriks, J., Hellingwerf, K. On the Midpoint Potential of the FAD Chromophore in a BLUF-Domain Containing Photoreceptor Protein. *FEBS Lett.*, **2011**, *585* (1), 167-172

[31] Schleicher, E., Weber, S. Radicals in Flavoproteins. *Top. Curr. Chem.*, **2012**, *321*, 41-65

[32] Miller, A.-F., Duan, H.D., Varner, T.A., Mohamed Raseek, N. Reduction Midpoint Potentials of Bifurcating Electron Transfer Flavoproteins. In *Methods in Enzymology;* New Approaches for Flavin Catalysis*,* Vol. 620; Palfey, B., Ed.; Elsevier, **2019**; 365-398

[33] Dietler, J., Gelfert, R., Kaiser, J., Borin, V., Renzl, C., Pilsl, S., Ranzani, A.T., García de Fuentes, A., Gleichmann, T., Diensthuber, R.P., Weyand, M. Signal Transduction in Light-Oxygen-Voltage Receptors Lacking the Active-Site Glutamine. *Nat. Commun.*, **2022,** *13* (1), 1-16

[34] Nozaki, D., Iwata, T., Ishikawa, T., Todo, T., Tokutomi, S., Kandori, H. Role of Gln1029 in the Photoactivation Processes of the LOV2 Domain in Adiantum Phytochrome3. *Biochemistry*, **2004,** *43* (26), 8373-8379

Chapter 3. Methods

3.1 Bioinformatics: Identification of Suspected-Ebases Homologous to HdrA

Using the Enzyme Function Initiative Enzyme Similarity Tool (EFI-EST, Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign), a SSN (Sequence Similarity Network) was generated for homologs of subunit HdrA from native EB-ase HdrABC/MvAGD (PDB ID: 5ODC, chain A), which is the site of electron bifurcation for this complex. The SSN was generated from the results of a BLAST-p search of the Uniprot database, using the sequence of HdrA as the search query (E-value = 10^{-5}, alignment score threshold = 35% seq ID (default), max number of sequences = 10,000). From this SSN, a GNN (Genome Neighborhood Network) was

created using the Enzyme Function Initiative Genome Neighborhood Tool (EFI-GNT, Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign) using a neighborhood size of 10 genes up- and down-stream from each HdrA homolog, and a minimal co-occurrence percentage lower limit of 20%. The GNN information was viewed and analyzed using the program DB Browser (Sqlite). To determine which HdrA homologs have gene neighbors that are homologous to HdrB/C, the following SQL command was executed twice on the GNN database:

SELECT attributes.accession FROM attributes JOIN neighbors ON attributes.sort_key = neighbors.gene_key WHERE neighbors.family LIKE '%PFAM%'

Where "PFAM" is the Pfam family ID corresponding to either HdrB (PF02754) or HdrC (PF13183). The two lists of hits were compared, and accession IDs co-occurring in both lists were identified as HdrA homologs co-expressing HdrBC. There were 5337 homologs that met this condition. The accession IDs were used to retrieve their associated sequences using the Uniprot Retrieve/ID mapping tool (https://www.uniprot.org/id-mapping). These sequences were aligned using Clustal Omega's MSA tool (EMBL-EBI), along with the original HdrA sequence. Since the upper limit is 4000 sequences, two alignments were made using two halves of the data set (sorting alphabetically).

The MSA was opened in Excel, and the sequence segments containing the residue aligned with HdrA Lys^{409} were identified by the "find" function, searching for "5ODC_1|Chain" and manually searching for the HdrA sequence "NVCCMYAMKNSQLIK" using "find next". Homologous sequence segments were isolated in a separate column using "=RIGHT(A3, 8)" where A3 is the cell containing

the homolog name and sequence segment, and 8 is the position of the aligned residue counting left from the end of the entry. Another column was created to list only the residues aligning with HdrA Lys^{409} using "=LEFT(B3, 1)" where B3 is the column containing the right-most nine characters of each segment, and 1 is the position of the character corresponding to the homologous aligned residue to be extracted to column C. This column was then analyzed to create lists of accession ID's for homologs that have a K, R, H, or – in the position aligned with HdrA Lys^{409} by using the MATCH function "=IF(MATCH(D1, C3, 0), LEFT(A3, (FIND(" ", A3, 1)-1)), "#N/A")" where D1 is a reference cell containing only "K", "R", "H", or "-", C3 is the cell containing the residue aligned with HdrA Lys^{409}, and A3 is the cell containing the accession ID and sequence segment aligned with this position. Here, the "LEFT" function takes just the accession ID as the entry in column D, E, F or G (**Figure 3.1**).

This method generated four lists in columns D-G that contained accession IDs of homologs that have a K, R, H, or – in the position aligned with HdrA Lys^{409}. The IDs of homologs with no residue aligned at this position were removed from the analysis (total of 1275 hits). This was done by creating a column of accession IDs using "=LEFT(A3, (FIND(" ", A3, 1)-1))" to pull all characters left of the space from column A, and the corresponding aligned residue was obtained from column C by "=C3". The accession IDs were refined by trimming the excess characters on the right with "=LEFT(I3, FIND("|", I3, 10)-1)", and the excess characters on the left with "=RIGHT(J3, 10)" to yield a column containing the trimmed IDs. These two columns (columns K and L: Trim Left, Residue Aligned with Lys^{409}) were copied and pasted

using special paste (values (V)) to obtain the character strings instead of the functions (**Figure 3.2**).

These two columns were sorted together by the residue-containing column as Z-A. This placed all accession IDs with an alpha-numeric character aligned with the HdrA Lys^{409} position at the top of the list. Of the remaining homologs, 3782 contained Lys at this position, 43 contained Arg, 36 contained His, and 200 contained some other residue. The 4061 IDs with a residue at this position were copied from this list and their sequences obtained using the Uniprot Retrieve/ID mapping tool. One ID (A0A7J2NI99) could not be mapped. These sequences were used to generate a new MSA (including the original HdrA sequence) that excluded homologs without a residue at this position. Since the maximum number of sequences for Clustal Omega is 4,000, the sequences were grouped into the first 4,000 (+ HdrA) and the last 4,000, saved as "HdrA_HdrBC_residueN5_FASTA_for_MSA1" and "HdrA_HdrBC_residueN5_FASTA_for_MSA2" in the Greene Lab Google Drive. The MSAs were used to generate sequence logos using Skylign (Skylign.org; Jody Clements, Travis Wheeler & Robert Finn) to show the relative abundance of residues at this position (**2.1**).

Sequence of HdrA (PDB ID: 5ODC, chain A):

MEEPRIGVYVCHCGVNIGGTVDCPDVTEFAKTLKNVVVARDYKYMCADPGQEMIK
KDIKEHNLNRVVVAACSPRLHEPTFRRCVAEAGLNPFLFEFANIREHCSWVHMHEK
EKATEKAKDLVRMAVAKARLLEPLEFIKVGVTQRALVIGGGVAGIQTALDLGDMGFE
TILVEKTPSVGGRMAQLDKTFPTNDCSICILAPKMVDVAKHPNVKLYAYSEVVDVQ
GYVGNFKVKMKKARYIDETKCTGCGQCSEVCPIDVPNEFDMGIGMRKAIYKPFPQ

AVPAKYTIDKEHCIECGLCAKVCGPNAIDFDQEPEIIEAEVGTIICAIGYDAFDPTVRE

EYGYGVYDNVVTALELERMINASGPTGGKVIRLSDGQKPKRIAFIQCVGSRDAKVG

NKYCSNVCCMYAMKNSQLIKEKSPDTEIDIYYMDIRAFSKGYEEFYERSAKQYGIKF

MRGRPSQVIEDPETGNLVVRAEDTLLGEILEKEYDLVVLSVGMVPTKSADEVQKILG

ISRTPDQFFMEAHPKLRPVDTATDGVYLAGACQGPKDIPASVAQGSAAASRAAIPL

AKGEVEVEPIIASVDAEICGGCGVCVKQCPYGAPRLVEKDGKVVAEVISALCKGCG

TCPAGCPSGALEQDHFKTIQLFKQIEGMFRDTA

A	B	C	D	E	F	G
			K	R	H	-
Sequence Segment	=RIGHT(A3, 8)	=LEFT(B3, 1)	Has K @ Lys409	Has R @ Lys409	Has H @ Lys409	Has nothing @ Lys409
tr|M1E6N1|M1E6N1_9FIRM M----ILCVGSRDRKV----------DNNICS-RVCCMYSLKQA-QLIM	KQA-QLIM	K	tr|M1E6N1|M1E6N1_9FIRM	#N/A	#N/A	#N/A
tr|A0A7C4JZ68|A0A7C4JZ68_9BACT Y----ILCTGSRDSSLE----------NVMSCSGECSNNPICS-QICCMYSIKQA-QLLM	KQA-QLLM	K	tr|A0A7C4JZ68|A0A7C4JZ68_9BACT	#N/A	#N/A	#N/A
tr|A0A524JZD4|A0A524JZD4_9BACT Y----VLCTGSRDRSV----------NNPICS-RVCCMYSLKQA-QLLL	KQA-QLLL	K	tr|A0A524JZD4|A0A524JZD4_9BACT	#N/A	#N/A	#N/A
tr|A0A662D0J6|A0A662D0J6_9BACT Y----ILCTGSRDRTV----------DNPICS-QVCCMYSIKQA-QLLM	KQA-QLLM	K	tr|A0A662D0J6|A0A662D0J6_9BACT	#N/A	#N/A	#N/A
tr|A0A7C0VER5|A0A7C0VER5_9BACT Y----ILCTGSRDHTV----------GNPLCS-QVCCMYSIKQA-QLLM	KQA-QLLM	K	tr|A0A7C0VER5|A0A7C0VER5_9BACT	#N/A	#N/A	#N/A
tr|A0A5S4ZNR5|A0A5S4ZNR5_9FIRM Y----VLCTGSRDCTV----------DNPLCS-RVCCMYSVKQN-QLIM	KQN-QLIM	K	tr|A0A5S4ZNR5|A0A5S4ZNR5_9FIRM	#N/A	#N/A	#N/A

Figure 3.1 *Accession IDs for homologs with no residue aligned with Lys^{409} were identified using Excel functions*

I	J	K	L	M	N	O
Accession ID	Trim right	Trim Left	Residue Aligned with Lys409		(Sorted) Assession ID	(Sorted) Residue Aligned with Lys409
tr|A0A1J0GHC1|A0A1J0GHC1_9CLOT	tr|A0A1J0GHC1	A0A1J0GHC1	K		A0A352D7J4	Y
tr|A0A1T4KDL1|A0A1T4KDL1_9FIRM	tr|A0A1T4KDL1	A0A1T4KDL1	K		A0A1F9SLZ7	Y
tr|A0A4R2KD27|A0A4R2KD27_9CLOT	tr|A0A4R2KD27	A0A4R2KD27	K		A0A349PB84	Y
tr|A0A369XYV7|A0A369XYV7_9FUSO	tr|A0A369XYV7	A0A369XYV7	K		A0A2H5V3N5	Y
tr|A0A414PR43|A0A414PR43_FUSMR	tr|A0A414PR43	A0A414PR43	K		A0A0S6UGI8	W

Figure 3.2 *Accession IDs and their corresponding residue aligned with Lys^{409} were sorted Z-A and the ID's with no residue at this position were omitted from the subsequent analysis.*

3.2 Construction of Vector pET-SUMO-iLOV

Since iLOV is a small protein (13 kDa) with only one tryptophan and three tyrosine residues, it is under-quantified by Bradford Assay when referencing a calibration curve made with Bovine Serum Albumin (66.4 kDa, 2 Trp, 20 Tyr). To more accurately quantify yield and cofactor reconstitution efficacy, iLOV was expressed as a SUMO-fusion construct (25.35 kDa, 2 Trp, 1 Tyr) and isolated just prior to spectrophotometric experiments, using SUMO-specific proteolysis with Ubiquitin-Like Protease I (ULP1). This allowed for more soluble protein during purification/cofactor loading, more

accurate protein quantitation, and the removal of the affinity tag from iLOV before conducting electron transfer experiments.

Plasmid Design of Fusion Construct pET-SUMO-iLOV

The iLOV gene was obtained as a gift from the Mukherjee lab, in UCSB's Chemical Engineering Department. It was originally cloned into a pQE80L vector backbone (**Figure 3.3**). The iLOV gene was cloned into the pET SUMO vector (pHYRSF53, www.addgene.org/64696) via Gibson Assembly, using primers ordered from Integrated DNA Technologies (Coralville, IA). The iLOV Gibson segment was amplified by PCR using "iLOV Gibson forward primer" at the N-terminus of iLOV, and "iLOV Gibson reverse primer" at the C-terminus of iLOV. The reverse iLOV primer inserted a XhoI restriction site at the C-terminus of the iLOV gene after the amber codon, with the XhoI restriction site ending 27 base pairs from the start of the T7 terminator sequence on the pET-SUMO backbone. The pET-SUMO Gibson segment was amplified by PCR using "SUMO Gibson forward primer" ending 27 base pairs from the start of the T7 terminator sequence, and "SUMO Gibson reverse primer" at the C-terminus of SUMO. The forward primer inserted an XhoI restriction site at the C-terminus of the iLOV gene after the amber codon.

Primer Sequences:

Capitalized bases are complementary to the iLOV gene, and lowercase corresponds to the pET-SUMO vector. XhoI restriction site was inserted at the C-terminus of iLOV (bolded).

iLOV Gibson forward primer:

5’-agattggtggatcccatatgATGATTGAAAAAAACTTTGTGATTACCGACCCG-3’

iLOV Gibson reverse primer:

5’-gcggtggcagcagccta**ctcgag**TTATACGTGGTCAGAACCATCCAG-3’

SUMO Gibson forward primer:

5’-ATGGTTCTGACCACGTATAA**ctcgag**taggctgctgccaccgc-3’

SUMO Gibson reverse primer:

5’-ACAAAGTTTTTTTCAATCATcatatgggatccaccaatctgttctct-3’

Figure 3.3 *SUMO-iLOV vector, visualized in Snapgene*

Amplification of Gibson Fragments for the Assembly of pET-SUMO-iLOV

The iLOV gene and pET-SUMO vector were amplified in two separate 50 µL mixtures, each containing ~25 ng of the appropriate template DNA, 1X Phusion HF buffer (NEB), 200 µM dNTPs (NEB), 0.5 µL of Phusion polymerase (NEB), and 10 µM of the corresponding fwd/rev primers (gene-specific) in autoclaved milli-Q water. The template DNA was added last, and a 10 uL aliquot was kept separate prior to its addition to serve as a negative control. The 40 µL remaining of each mix was split into thirds, and each 13 µL fraction was amplified at a unique annealing temperature. The thermocycler program ran according to **Table 3.1**.

Table 3.1. *Thermocycler steps for PCR amplification of SUMO-iLOV WT Gibson fragments*

Step 1	30 sec at 98 C°
Step 2	15 sec at 98 C°
Step 3	1 min at **Tm** (58 C°, 62 C°, or 67 C°)
Step 4	5 min at 72 C°
Step 5	10 min at 72 C°
Step 6	5 min at 4 C°

*After step 5, return to step 2. Loop for 30 cycles before step 6.

Amplification products were run on a 0.8% agarose gel at 90 V for 20 min using Apex DNA loading dye (Apex) and Generuler™ 1 kB DNA ladder (Thermo Scientific). The iLOV Gibson fragment corresponded to the expected MW of 379 bp and had no non-

specific amplification, so it was regarded as pure after digesting the remaining template DNA with DpnI (NEB) in 1X CutSmart buffer (NEB) for 12 hours at 37 C°. The pET-SUMO Gibson fragment, corresponding to the heaviest amplicon band (expected MW of 3,847 bp), was isolated via gel purification (**Figure 3.4**).

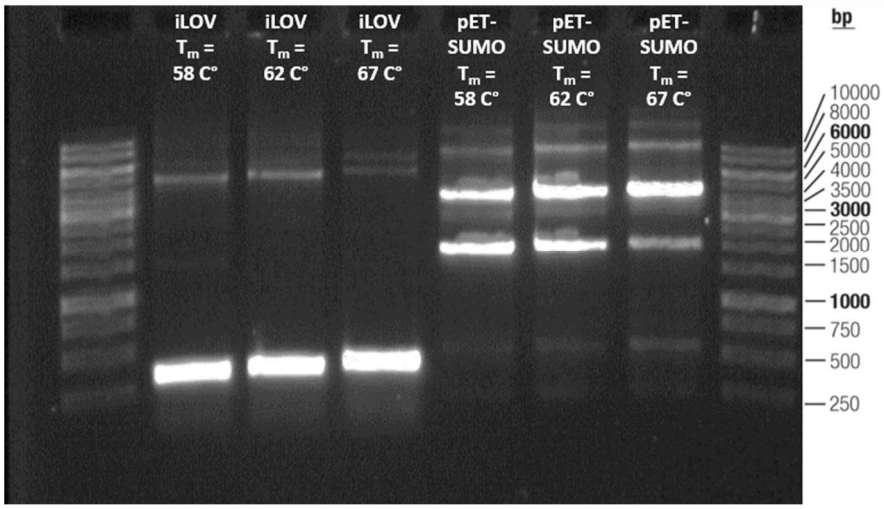

Figure 3.4 *Amplified iLOV and pET-SUMO fragments for Gibson Assembly, run on a 0.8% agarose gel.*

Gel Purification of pET-SUMO Gibson Fragment

The whole sample of the SUMO Gibson fragment PCR product was combined with 1X loading dye (Apex) and loaded into one large well in a 1% agarose gel for separation by electrophoresis. The DNA was run alongside the Generuler^{TM} 1 kB DNA ladder (Thermo Scientific), also stained with loading dye. The band corresponding to ~3.5 KB were cut out of the gel on a transilluminator and transferred to pre-tared eppendorf tubes.

The pet-SUMO Gibson fragment was obtained via gel purification using the Qiagen Gel Extraction Kit (Qiagen). The isolated gel bands containing the SUMO fragment were dissolved in 3 parts QG buffer (300 uL GQ buffer per 100 mg gel fragment) and heated at 60 C°for 10 min, mixing by inversion throughout. The dissolved gel and DNA sample was mixed with 1 part isopropanol and inverted to mix before it was applied to a Quiagen spin filter and centrifuged at 17,900 x g for 1 min at 20 C°. The DNA in the filters was washed with 500 uL of QG buffer and spun at the same conditions, followed by 750 uL of buffer PE. The PE buffer was left to sit for 5 min before centrifuging. The tube was spun twice at the aforementioned conditions to remove all traces of PE buffer. Then, 10 uL of hot (near-boiling) milliQ water was applied to the spin filter and allowed to stand for 5 min. The DNA was eluted by centrifuging the tube at the same conditions for 1 min into a clean eppendorf tube. The DNA was quantified by a Nanodrop® ND-1000 Spectrophotometer. The DNA fragments were stored at -20 C° for later use.

Gibson Assembly of Vector Construct pET-SUMO-iLOV

The pET-SUMO-iLOV vector was assembled using the NEBuilder Hi-Fi DNA Assembly Master Mix (NEB) using a 1:3 molar ratio of SUMO:iLOV Gibson fragments (0.03 pmol of the pET-SUMO fragment and 0.09 pmol of the iLOV fragment). DNA fragment concentration was determined using a Nanodrop® ND-1000 Spectrophotometer. Appropriate volumes of the DNA fragments were added to an PCR tube and diluted to 10 uL with autoclaved water. To that was added 10 uL of the Gibson HiFi DNA assembly master mix. The total volume of the assembly mix was 20 uL. The ligation mix was incubated at 50 C° for 30 min and then chilled on ice prior to

transformation. Ligation mixture was introduced into chemically-competent DH5-α *E. coli* (NEB) cells using our standard transformation protocol, using 2 µL of ligated DNA solution. Successful transformants were identified by T7 colony PCR. The plasmid pET-SUMO-iLOV was purified from DH5-α *E. coli* cultures grown from selected colonies and confirmed by Sanger DNA sequencing (UC Berkeley). This construct was used as the template for site-directed mutagenesis to generate mutants Q104K and Q104A.

Colony PCR Identifies Successful Gibson Assembly of pET-SUMO-iLOV

A Colony PCR mix was made up with 1X HF buffer, dNTP mix (200 uM final conc.), 1.5 uL of Phusion polymerase, T7 promoter/terminator primers (0.5 uM of each primer, final conc.), and autoclaved water up to 150 uL. The PCR mix was divided among 15 tubes in 10 uL aliquots. 12 colonies were chosen, along with a positive control (DNA with similar expected amplicon size as a successful Gibson transformant) and two negative controls (Original vector template and PCR mix with no DNA (bare toothpick) to check for contamination). 0.5 uL of pure plasmid was added to the positive control and original template PCR mix aliquots. Working near a flame, the colonies were touched with an autoclaved toothpick and dragged across a numbered LB-kanmycin agar plate (streaking just above the number), then the toothpick was swirled in the PCR mix aliquot corresponding to the colony #. The aliquots were vortexed, centrifuged briefly and placed in the thermocycler. The colony plate was incubated at 37 C° while the temperature cycled. The thermocycler protocol is outlined in **Table 3.2.**The PCR products were ran in a 0.8% agarose gel to observe the gene fragment copied by the T7 primers. Suspected successful transformants were identified by their

T7 amplicon MW, and by comparing to the migration distance of the positive control (**Figure 3.5**). Sequence of the pET-SUMO-iLOV construct confirmed by Sanger sequencing at UC Berkeley by sending purified plasmid from a miniprep of the numbered plate colony (**Figure 3.6**).

Table 3.2. *Thermocycler steps for cPCR amplification of T7 amplicons.*

Step 1	30 sec at 98 C°
Step 2	15 sec at 98 C°
Step 3	1 min at **Tm** (58 C°)
Step 4	5 min at 72 C°
Step 5	10 min at 72 C°
Step 6	5 min at 4 C°

After reaching step 4, steps 2-4 were repeated for 30 cycles

SUMO-iLOV WT Sanger-confirmed gene sequence:

Capitalized bases are complementary to the iLOV gene, and lowercase corresponds to the pET-SUMO vector.

5’-catcatcatcatcatcacagcactagtgactcagaagtcaatcaagaagctaagccagaggtcaagccagaa
gtcaagcctgagactcacatcaatttaaaggtgtccgatggatcttcagaaatcttctttaagatcaaaaagaccactc
ctttaagaaggctgatggaagcgttcgctaaaagacagggtaaggaaatggactccttaagattcttgtacgacggta
ttagaattcaagctgatcagacccctgaagatttggacatggaggataacgatattattgaggctcacagagaacag
attggtggatcccatatgATGATTGAAAAAAACTTTGTGATTACCGACCCGCGTCTGCCG
GATAACCCGATCATTTTCGCGTCTGATGGCTTCCTGGAACTGACTGAGTATAGC
CGTGAAGAAATCCTGGGCCGCAATGCTCGTTTTCTGCAGGGCCCGGAGACCG
ATCAAGCTACCGTGCAGAAGATTCGTGATGCAATTCGTGACCAGCGCGAAACG
ACTGTGCAGCTGATTAACTATACCAAGAGCGGTAAAAAATTCTGGAACCTGCTG
CACCTGCAGCCGGTCCGTGATCAGAAAGGCGAGCTGCAGTATTTCATCGGTGT
TCAGCTGGATGGTTCTGACCACGTATAA-3’

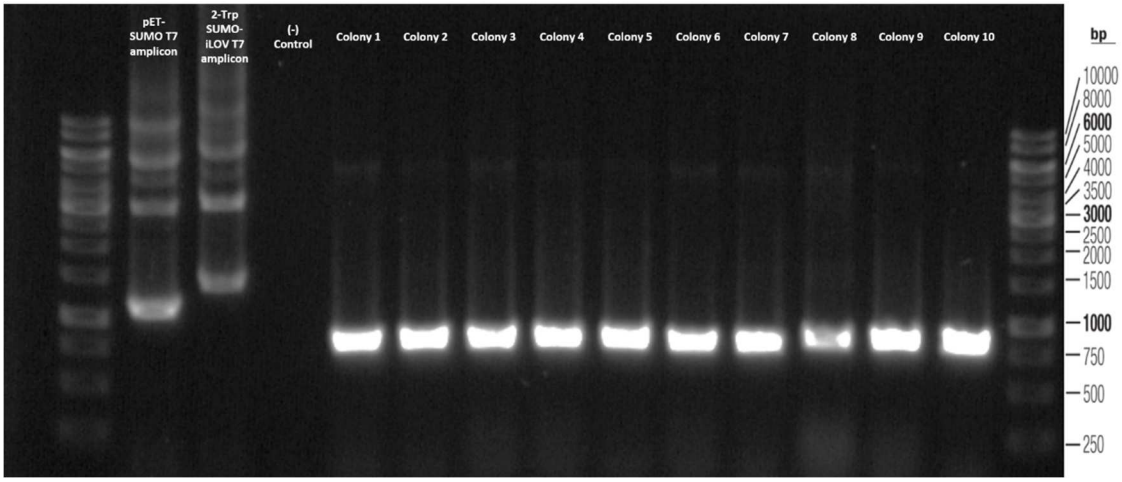

Figure 3.5 *cPCR of transformants from Gibson Assembly DNA have T7 amplicons close to expected MW (822 bp). pET-SUMO was used as a negative control (expected MW: 1237 bp), as well as previous construct 2-Trp-SUMO-iLOV (expected MW: 1403 bp)*

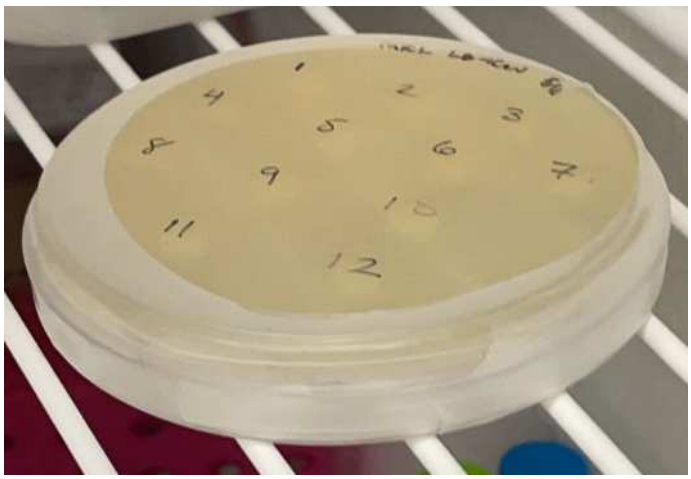

Figure 3.6 *Shows numbered colony plate from cPCR*

Site-Directed Mutagenesis of SUMO-iLOV to Generate Q104K and Q104A Plasmids

The SUMO-iLOV WT construct was used as the template for site-directed mutagenesis to generate mutants Q104K and Q104A. SDM was performed with

primers purchased from Integrated DNA Technologies using the same thermocycler protocol as the Gibson fragment amplification (**Table 3.1**).

Q104K SDM Forward Primer:

5'- gcagtatttcatcggtgttaagctggatggttctga-3'

Q104K SDM Reverse Primer:

5'- ccagcttaacaccgatgaaatactgc-3'

Q104A SDM Forward Primer:

5'- cggtgttgcgctggatggttctg-3'

Q104A SDM Reverse Primer:

5'-gaaccatccagcgcaacaccgatg-3'

3.3 Cell Culture, Protein Expression and Purification of SUMO-iLOV

SUMO-iLOV was recombinantly expressed in BL21-DE3 *E. coli* (NEB) using the *lac* operon and inducing protein expression with IPTG. Cells were grown in Miller LB media (Sigma) at 37 C° shaking at 190 RPM, and protein expression occurred under the same conditions. Proteins were purified using nickel affinity chromatography and SEC. After cofactor loading and quantitation the his-tagged SUMO was removed via ULP1, a SUMO-specific protease. iLOV and mutants of iLOV were used for all electron-transfer experiments, except protein film voltammetry, which used SUMO-iLOV.

Greene Lab Chemically-Competent Cell Protocol

Competent cells were made from cultures of cells propagated from stocks purchased from the manufacturer (DH5-α *E. coli*, NEB, #C2987H or BL21-DE3 *E. coli*, NEB, # C2527H). 100 mL cultures were grown in Miller LB medium (Sigma) to $O.D._{500\ nm}$ = 0.5, inoculated from a 1:100 dilution of overnight seed cultures grown in 5 mL of LB media with 1 uL of the manufacturer's competent cells. Culture was chilled on ice for 10 min after reaching the appropriate cell density. Cells were harvested by centrifugation at 6,000 x g for 10 min at 4°C, and the supernatant discarded. Cell pellet was re-suspended in TSS buffer (PEG 8000 (Sigma), $MgCl_2$ (Sigma), DMSO (Sigma), Miller LB (Sigma)), using 5% of the original culture volume. Re-suspended cells were frozen in liquid N_2 in 100-µL aliquots and stored at -80 C°.

Chemically-Competent Cell Transformation Protocol

Competent cells were thawed on ice for 20 min (just until the last ice crystal has just melted) before adding 1 uL of plasmid or 2 uL of PCR product, flicking to ensure homogeneity. The cells were kept on ice for exactly 30 min, then submerged in a heat bath at 42°C for 30 seconds. The cells were immediately transferred on ice to chill for 5 min, and then diluted to 1 mL with SOC media (use sterile technique). Cells were put in the warm room for an outgrowth period of 1 hour at 37°C, shaking at 200 rpm. The sample was centrifuged at 6000 x g for 10 min, the supernatant discarded, and the cell pellet re-suspended in ~100 uL of the remaining supernatant. The entire aliquot was plated using a triangular cell spreader and the plates were incubated overnight at 37°C.

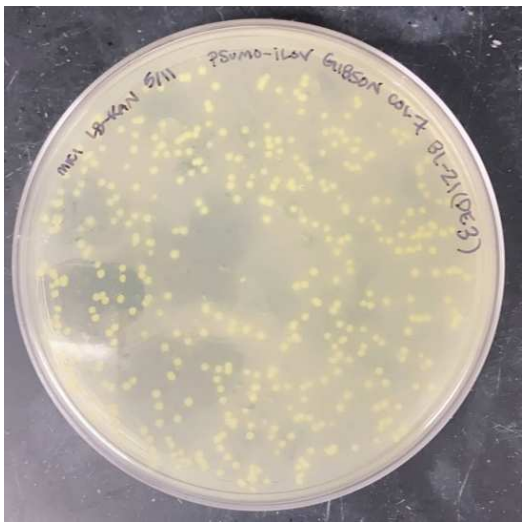

Figure 3.7 *A plate of BL21-DE3 E. coli transformed with the pET-SUMO-iLOV plasmid, purified from cPCR colony 7. Uninduced expression of SUMO-iLOV causes the yellow appearance of the colonies.*

SUMO-iLOV Protein Expression Protocol

SUMO-iLOV was expressed in 4-6 L batches in 6L Erlenmeyer flasks filled to no greater than 25% of the total capacity. Warm (37 C°), sterile Miller LB (Sigma) containing 50 ug/mL kanamycin was inoculated by 1:100 dilution from saturated seed cultures (Miller LB (Sigma), 50 ug/mL kanamycin (Sigma)) grown from a single colony of BL21-DE3 *E. coli* transformed with pET-SUMO-iLOV. Flasks were shaken at 37 C° at 190 RPM for 2.5 hours, until the optical density (O.D. 600 nm) of the media was 0.5. Then, protein expression was induced by the addition of 1 mM IPTG (final concentration). Cells expressed protein for at least 4 hours (no greater than 8) before harvesting by centrifugation at 3,000 x g for 10 min. Wet cell weight was recorded before flash-freezing in liquid nitrogen and storing at -80 C°.

SUMO-iLOV Protein Purification Protocol

E. coli cell pellet was thawed on ice and dissolved in cold Buffer A (6 mL/1g cells, 200 mM NaCl, 100 mM Tris-HCl, pH 8) with 1 mM PMSF (TCI) and 2mM DTT (Sigma) by adding 1/3rd of the buffer to the pellet at a time, shaking to dissolve. The cell solution was thoroughly homogenized in a glass cell homogenizer before mechanical lysis by French Press extraction. The cells were lysed in 3-4 passes (collecting lysate in a tube on ice between each pass), with the needle kicking to a maximum pressure of ~14,000 psi, and the resulting solution was centrifuged at 14,636 x g for 10 min at 4°C (max speed for the tabletop centrifuge). The debris pellet had a tan upper layer and a thin fluorescent yellow layer of intact cells/inclusion bodies underneath.

The supernatant was stirred in a beaker on ice while adding 20% v/v of 6% streptomycin sulfate (Sigma) solution dropwise. After stirring for 10 min, the solution was centrifuged at 20,000 x g for 10 min, and then again in a new centrifuge tube at 29,000 x g for 20 min in the Sorvall centrifuge to precipitate the DNA. 1 M, pH 8 imidazole (Sigma) was added dropwise until the final concentration was 30 mM. The solution was then vacuum filtered using 0.45 uM filter paper (Sigma). The supernatant was then added to a pre-equilibrated (with Buffer A) Ni-NTA column with ~20 mL of column resin. After collecting the flow-through, 30 column volumes of wash buffer were applied to the column by gravity flow. Then, 10 column volumes of elution buffer were added to the column and the eluent was collected in fractions. The fractions containing SUMO-iLOV were visibly fluorescent and pooled according to observation of yellow color. The protein was concentrated by salting out the protein with 39% ammonium sulfate (Sigma), adding dry salt slowly while stirring the protein on ice for

10 min and spinning at 10,000 x g at 4 C° for 10 min. The protein was resuspended in minimal Buffer A by pipetting and was desalted using a Hi-Trap desalting column (Cytvia). If necessary, protein was further concentrated using a 10 kDa 50 mL Amicon centrifugal filter. Glycerol was added up to 5%, and the protein was flash-frozen in liquid N2 and stored at -80 C°.

SUMO-iLOV Flavin Loading Protocol

Proteins were thawed and quantified via Bradford assay with Bradford Coomassie Plus (Bio-Rad) by comparing to a BSA calibration curve (BSA standards from Sigma). An appropriate volume of flavination buffer (200 mM NaCl (Sigma), 100 mM Tris-HCl (Sigma), 100 mM FMN(Thermo Fisher), pH 8) was added to the protein in a 15 mL falcon tube, so that the FMN concentration was 50X the protein concentration. The protein was heated at 60 C° in a water bath for 15 min, then cooled on ice for 5 min. The ideal temperature, 60 C°, was decided after review of CD spectra from 20-60 C°, which showed marked changes after 70 C° (**Figure 3.8, Figure 3.9, Figure 3.10**).

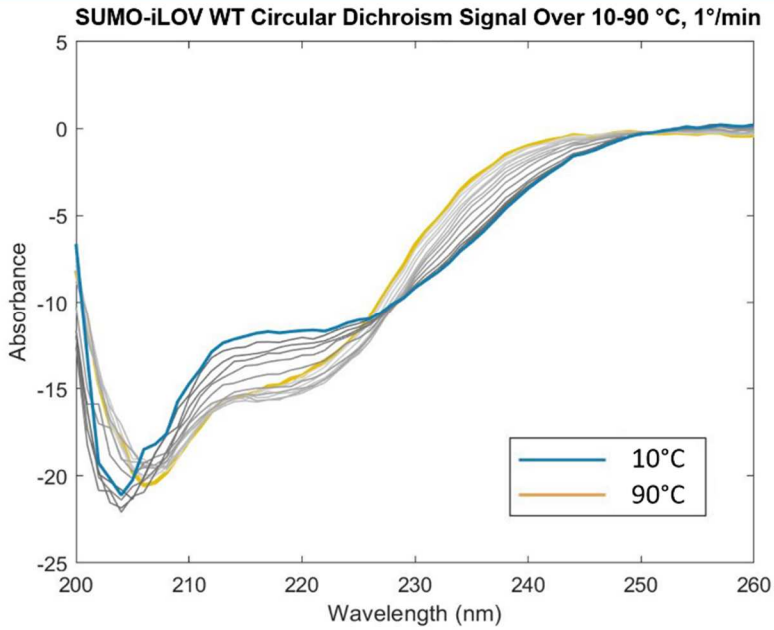

Figure 3.8 *The CD spectra of SUMO-iLOV WT with increasing temperature shows changes in protein structure after 60°C.*

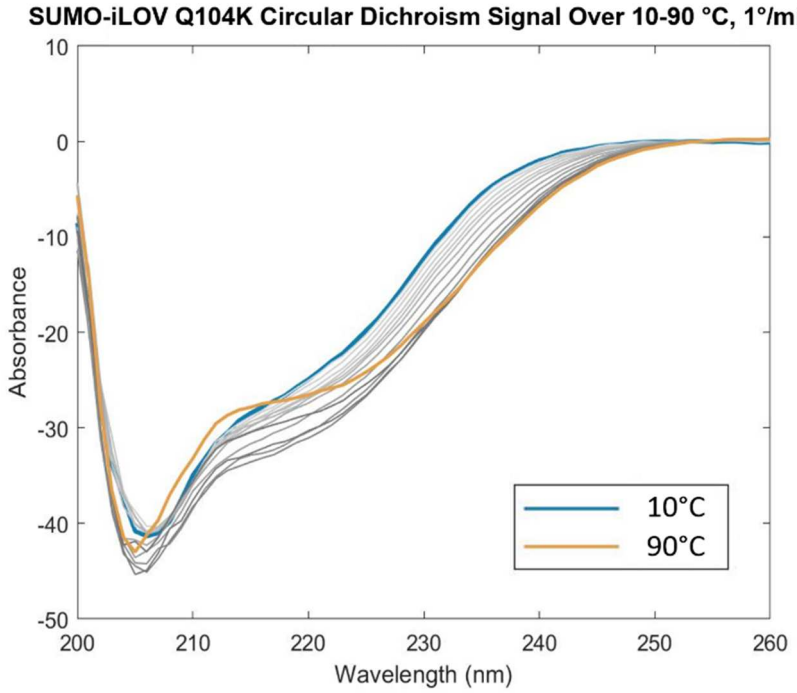

Figure 3.9 *The CD spectra of SUMO-iLOV Q104K with increasing temperature shows similar changes in protein structure after 60°C.*

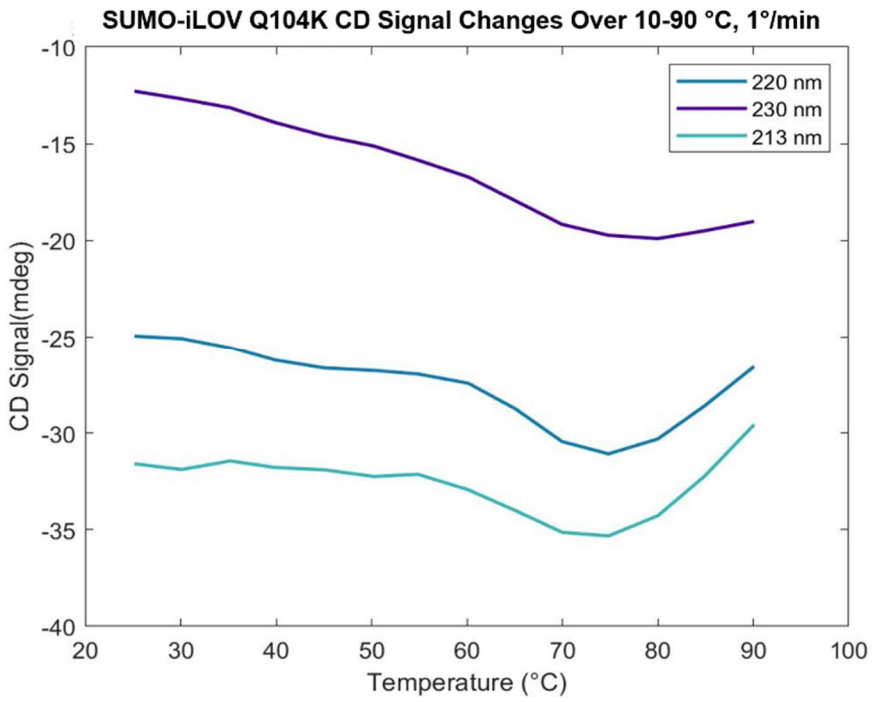

Figure 3.10 *Shows melt curves obtained from CD spectra of SUMO-iLOV Q104K, tracking changes in θ (mdeg) at three different wavelengths. Significant unfolding occurs beyond 60°C.*

Heated protein/FMN samples larger than 2 mL were pre-concentrated prior to SEC using an Amicon 10 kDa centrifugal filter (Sigma), specifically marked for this purpose. Samples were not concentrated beyond 4 mM SUMO-iLOV, because the protein precipitates above this concentration. The protein-FMN sample was loaded onto a size exclusion column (G-25 Superfine Sephadex, Sigma) to separate the holoprotein from free FMN. Protein was collected by eye in one fraction.

Flavin loading efficacy was analyzed by comparing the Bradford protein concentration to the FMN concentration, as determined spectroscopically by UV-vis (using the molar extinction coefficient for SUMO-iLOV, 14.8 mM^{-1} cm^{-1} for WT and 16.1 mM^{-1} cm^{-1} for Q104K, Q104A, and other blue-shifted mutants).[35] Bradford assay was conducted in triplicate from a 1 mL sample of diluted protein (using no less than

2.5 mL of protein for the serial dilution, to reduce error). This same sample was analyzed by UV-vis after removing a portion for the Bradford assay. Flavin loading was determined by dividing the FMN concentration (UV-vis) by the protein concentration (averaged from triplicate measurements). After flavin loading was assessed, protein was concentrated in a 10 kDa Amicon centrifugal filter to 1-3 mM. Proteins were frozen with 5% glycerol in liquid N_2 and stored at -80 C°.

Isolation of iLOV by ULP1 Digestion of SUMO-iLOV Protocol

Protein concentration was determined by Bradford assay with Bradford Coomassie Plus (Bio-Rad), in triplicate, comparing to a BSA calibration curve. The concentration was used to calculate the volume of 1 mg/mL ULP1 needed to reach a 1:500 ULP1:SUMO-iLOV molar ratio. The appropriate volume of ULP1 (recombinantly expressed and purified in our lab) was added to a 15-mL falcon tube containing the thawed SUMO-iLOV protein and inverted to mix. The reaction mixture was incubated at 30 C° for at least 3 hours (up to 5) in the dark, undisturbed.

During the incubation, ~20 mL Ni-NTA resin was prepared by washing the resin in several cycles. For each wash, an approximately equal volume of milliQ water was added to the resin, and the tube inverted until nearly homogenous. The tube was centrifuged at ~5,000 x g for 5 min, and the supernatant discarded. This rinsing process was repeated 4X. For the last rinse step, Buffer A (200 mM NaCl (Sigma), 100 mM Tris-HCl (Sigma), pH 8) was used instead of water. The supernatant was not poured off until the protein was ready to be added to keep the resin from drying out.

After the incubation time had elapsed, the protein reaction mixture was added to the falcon tube containing the pre-equilibrated Ni-NTA resin. This was attached to our lab

rotator/revolver and allowed to rotate for 30 min-1 hour to mix. During that time, a clean glass column was set up and pre-rinsed with buffer A. Approx. 5 mL of buffer A was left in the column after the rinse. The tube containing the nickel resin and protein mixture was then added to this closed column. The column was then hooked up to a pneumatic pump to feed in Buffer A (~100 mL). The flow-through was collected, containing the cleaved iLOV. The column was rinsed with buffer A until the eluent was no longer colored to the eye. Then, the column was cleaned with ~100 mL of elution buffer to remove the ULP1, SUMO, and uncleaved SUMO-iLOV. If the residual fusion protein concentration was significant, the eluted protein was stored for later purification.

Collected iLOV was concentrated in a 50 mL 10 kDa Amicon centrifugal filter to ~1 mM. The holoprotein concentration was determined from the absorbance at 448 nm (WT) or 440 nm (K, R, A mutants) from the UV-vis spectrum, using molar extinction coefficients of 14.8 mM-1 cm-1 and 16.1 mM-1 cm-1, respectively. Protein was diluted to 5% with glycerol, frozen in liquid N_2, and stored at -80 C°.

3.4 SUMO-iLOV Hydroquinone Reduction and Purification Protocol

Protein was thawed and centrifuged briefly at 10,000 x g to pull protein aggregates down to the bottom of the tube. Soluble protein was transferred to a 25-mL conical flask with an egg-shaped small stir bar. Remaining precipitated protein was re-

suspended in three cycles by using a pipette and 100 uL of Buffer A, centrifuging at 10,000 x g for 2 min, and combining the supernatant with the other soluble protein. Any remaining insoluble protein was discarded after three cycles. The holoprotein was quantified spectroscopically from the absorbance at 448 nm (WT) or 440 nm (K, R, A mutants) in the UV-vis spectrum, using molar extinction coefficients of 14.8 mM^{-1} cm^{-1} and 16.1 mM^{-1} cm^{-1}, respectively.[35] The soluble protein was de-gassed on the schlenk line using four purge cycles, allowing the solution to equilibrate with the N2 headspace for 5 minutes in-between. After vacuum was pulled on the flask, the protein was brought into the VAC glovebox and clamped above a magnetic stirrer.

Solid methyl viologen (~0.1 mmol) and sodium dithionite (~1 mmol) were weighed out in eppendorf tubes and brought into the glovebox alongside the protein. Solids were dissolved with 1 mL buffer A to make a ~100 mM and 1 M solution of methyl viologen and dithionite, respectively. The mass of the solids was used to determine their concentration, rather than their extinction coefficient. Equimolar methyl viologen was added to the protein, according to the holoprotein concentration estimate. 10X (10-fold molar excess of the holoprotein concentration) sodium dithionite was added drop-wise while stirring. The protein was left to reduce, stirring at medium speed, for 1 hour. Then, 5X dithionite was added drop-wise to ensure full reduction (15X in total). After 10 min of stirring, the protein was removed from the glovebox with a closed schlenk adapter. Vacuum was pulled on the flask before transferring to the Coy chamber. There, the protein/viologen/dithionite mixture was applied to a 120-mL SEC column (G-25 Superfine Sephedex resin, Sigma) pre-equilibrated with anaerobic Buffer A. The protein eluted in 15 min and was collected in three borosilicate test

tubes. The clear tube helped determine the beginning of the protein elution peak by eye, because the color of the hydroquinone is much lighter yellow. Protein-containing fractions were pooled in a 50-mL conical vial and capped with a closed schlenk adapter. The protein was removed from the Coy chamber and vacuum was applied to the flask before transferring it back to the VAC glovebox. The hydroquinone was quantified by UV-vis from a 1:10 diluted sample in a cuvette sealed with an anaerobic suba-seal septum. The spectrum of the reduced species was obtained first, and the molar absorptivity coefficient at 448 nm (or 440 nm for Q104K/A mutants) was compared to a WT reference spectrum to determine the extent of reduction. The cuvette was opened and allowed to air-oxidize by inverting ~10X before taking a spectrum of the oxidized protein. The holoprotein was quantified spectroscopically from the air-oxidized sample. The remaining protein in the glovebox was diluted up to 10% glycerol and stored in glass crimp vials. Oxidized and reduced iLOV samples were stored at -32 °C in 10% glycerol in an anaerobic freezer inside the glovebox. Care was taken to prevent unnecessary irradiation of samples by white light in the lab prior to comproportionation and equilibration experiments. This was done by working in low-light conditions with aluminum foil covering all tubes/cuvettes, using blackout cloth over the UV-vis during use.

For samples that were not fully reduced, the amount of oxidized protein remaining was calculated algebraically using the relationship below:

$$Abs_{Reduced\ iLOV\ at\ 448\ nm\ or\ 440\ nm} = \left(\varepsilon_{HQ} C_{tot} \ell\right) - \left(\varepsilon_{HQ}\ \chi\right) + \left(\varepsilon_{Ox}\ \chi\right)$$

where χ is the variable representing the concentration of remaining oxidized protein. ℓ = 1.0 cm, εHQ at 448 nm was determined from a reduced spectrum of iLOV WT as 0.76 mM-1 cm-1, and εOx was retrieved from the literature (14.8 mM^{-1} cm^{-1} for WT and 16.1 mM^{-1} cm^{-1} for Q104K/A).[35]

In the case of the Q104K mutant, the concentration of SQ is non-negligible, and the equation above needs to be adjusted to subtract the semiquinone concentration, C_{SQ}, from the total protein concentration, C_{tot}:

$$Abs_{Reduced\ iLOV\ at\ 448\ nm\ or\ 440\ nm} = \left(\varepsilon_{HQ}(C_{tot} - C_{SQ})\ell\right) - \left(\varepsilon_{HQ}\,\chi\right) + \left(\varepsilon_{Ox}\,\chi\right)$$

3.5 Protein Film Voltammetry of SUMO-iLOV

Protein-film Cyclic Voltammetry (CV) and Differential Pulse Voltammetry (DPV) were conducted on SUMO-iLOV WT and Q104K. To prepare the sample, 10 uL of a ~300 mM buffered protein solution (100 mM Tris-HCl) was dried on the surface of a pyrolytic graphite edge electrode (BASI EF-1368) by flowing N_2 over it in a closed container (falcon tube with a hole for N_2 entry, and a second hole for pressure release). Later, this protocol was improved by applying ~10 uL of a 50 uM SUMO-iLOV sample for 10 min, then removing the drop with a pipette and rinsing the electrode briefly with DI water to remove non-adsorbed protein. The electrolyte solution (Buffer A: 200 mM NaCl, 100 mM Tris, pH 8) was sparged with N_2 for at least 20 min prior to measurement to remove air from the solution. The counter electrode was platinum wire, and the reference was a standard calomel electrode (Sat. KCl). The DPV spectra were collected from 0.0 V to -0.9 V with a step height of 5 mV and a pulse height of

25 mV (pulse time = 10 ms). Cyclic voltammetry was conducted with the same electrode setup and electrolyte solution, scanning from 0.0 V to -1.0 V and back to 0.0 V at scan rates of 10 mV/s, 100 mV/s, 1 V/s, and 10 V/s. The midpoint potentials determined from the cyclic voltammograms matched those obtained from DPV.

To determine the number of electrons transferred, the Randles-Sevcik equation can be inferred from the peak current if the scan rate, electrode surface area, and surface coverage of the protein film are known **(Table 3.3).**[36] A hard spheres estimation of surface coverage was performed by dividing the electrode surface area by the cross-sectional area of iLOV, estimated by $A = \pi r^2$ from the radius of iLOV (measured as 22.5 Å, half the distance between Lys^{465} and Lys^{479} of PDB model 4EES in ChimeraX). The area of the electrode was calculated using $A = \pi r^2$, where r = 0.15 cm.

Table 3.3. *Calculation Table for n using Randles-Sevcik Equation*

iLOV Radius (Å)	n iLOV	mol iLOV (n/N_A)	Electrode SA (cm^2)	Surface Coverage (mol/cm^2)	Sample	i_p (A), v = 1 V/s	n (mol e^- transferred)
22.5	4.46E+11	7.41E-13	0.071	1.04E-11	SUMO-iLOV WT cathodic	1.12E-05	4.02
					SUMO-iLOV WT anodic	1.32E-05	4.35
					SUMO-iLOV Q104K cathodic	8.31E-06	3.45

	SUMO-iLOV Q104K anodic	1.25E-05	4.23

Hard Spheres Approximation of Surface Coverage:

Cross-sectional $SA_{iLOV} = \pi r^2$

r_{iLOV} = 22.5 Å (half the distance between Lys^{465} and Lys^{479} of PDB model 4EES)

$SA_{electrode} = \pi r^2$, r = 0.15 cm (BASI EF-1368)

$$n_{iLOV} = \frac{SA_{electrode}}{\pi(22.5\ \text{Å} * (10^{-8}\ \text{cm/ Å}))^2}$$

mol iLOV = *n* iLOV/N_A

Surface Coverage (Γ*) = mol iLOV / $SA_{electrode}$

3.6 Measuring of iLOV WT E' by Equilibration with Methyl Viologen

A ~10 μM sample of $iLOV_{HQ}$ in a sealed quartz cuvette was quantified by UV-vis before each experiment to account for variation in concentration as a result of freeze-thaw cycles. The UV-vis spectrum of the $iLOV_{HQ}$ sample was obtained first before opening up the cuvette to air for 3 minutes and mixing by inversion, which oxidized the protein. The $iLOV_{HQ}$ sample was quantified using the literature molar extinction coefficient.[35]

A 1-mL solution was made in a 1 cm path length stoppered quartz cuvette with a final concentration of 60 μM $iLOV_{HQ}$ WT, 120 μM methyl viologen, and buffer A (100 mM Tris HCl, 200 mM NaCl, pH 8). The pressure in the glovebox was lowered below 0 mm H_2O before sealing the cuvette with a greased glass stopper. The cuvette was removed from the glovebox without ambient light and placed in an Agilent Cary 60

spectrophotometer, covering with blackout cloth. The 'Align" program was used to determine the proper z-height of the cuvette (to ensure the light path did not travel above the meniscus) before initiating the 'Kinetics' program. Spectra were obtained from 800 nm – 200 nm at a scan rate of 600 nm/min every 10 minutes for 18 hours. At this point, the blackout cloth was removed and measurements were resumed in the presence of light.

3.7 EPR to Characterize iLOV NSQ

Anaerobic protein samples of $iLOV_{ox}$ and $iLOV_{HQ}$ were thawed in the glovebox and quantified by UV-vis prior to each experiment, since minor protein precipitation occurred between freeze-thaw cycles. Each protein sample was diluted to ~10 μM in a cuvette sealed with a suba-seal septum. $iLOV_{OX}$ was quantified by its absorbance at 448 or 440 nm using molar extinction coefficients obtained from the literature (14.8 mM^{-1} cm^{-1} for WT and 16.1 mM^{-1} cm^{-1} for Q104K/A).[35] The UV-vis spectrum of the $iLOV_{HQ}$ sample was obtained first before opening up the cuvette to air for 3 minutes and mixing by inversion, which oxidized the protein. The $iLOV_{HQ}$ sample was quantified using the literature molar extinction coefficient.[35]

Equimolar $iLOV_{HQ}$ and $iLOV_{OX}$ were mixed in the glovebox with lysis buffer (100 mM Tris HCl, 200 mM NaCl, pH 8) with 10% glycerol to a final concentration of 50 μM in an Eppendorf tube. After 3 hours, the mixture was removed with a long-tipped Pasteur pipet and transferred to a quartz EPR tube. The tube was quickly removed from the glovebox and immediately submerged into liquid nitrogen, slowly lowering the tube to minimize expansion of the water in the sample as it froze. The EPR tube was labeled and stored in liquid N_2 for 3 days, until measurements were taken. The EPR spectrum

was collected on a Bruker EMXplus EPR Spectrometer at 100 K as an average of 50 scans at 2 mW, with a modulation amplitude of 2 G.

3.8 Comproportionation of iLOVox and iLOVHQ to measure [SQ]eq

Anaerobic protein samples of $iLOV_{ox}$ and $iLOV_{HQ}$ were thawed in the glovebox and quantified by UV-vis prior to each experiment, since minor protein precipitation occurred between freeze-thaw cycles. Each protein sample was diluted to ~10 μM in a cuvette sealed with a suba-seal septum. $iLOV_{OX}$ was quantified by its absorbance at 448 or 440 nm using molar extinction coefficients obtained from the literature (14.8 mM^{-1} cm^{-1} for WT and 16.1 mM^{-1} cm^{-1} for Q104K/A).[35] The UV-vis spectrum of the $iLOV_{HQ}$ sample was obtained first before opening up the cuvette to air for 3 minutes and mixing by inversion, which oxidized the protein. The $iLOV_{HQ}$ sample was quantified using the literature molar extinction coefficient.[35]

Equimolar $iLOV_{HQ}$ and $iLOV_{OX}$ were mixed in the glovebox with lysis buffer (100 mM Tris HCl, 200 mM NaCl, pH 8) to a final concentration of 50 μM in a 1 mL stoppered cuvette. The pressure in the glovebox was lowered below 0 mm H_2O before sealing the cuvette with a greased glass stopper. The cuvette was removed from the glovebox without ambient light and placed in a Agilent Cary 60 spectrophotometer, covering with blackout cloth. The 'Align" program was used to determine the proper z-height of the cuvette (to ensure the light path did not travel above the meniscus) before initiating the 'Kinetics' program. Spectra were obtained from 800 nm – 200 nm at a scan rate of 600 nm/min every 10 minutes for 18 hours.

The iLOV WT sample was considered fully reduced and was used to determine $iLOV_{HQ}$ $\varepsilon_{448\ nm}$ and $\varepsilon_{440\ nm}$ (**Figure 3.11**). Since the mutant $iLOV_{HQ}$ samples were not

fully reduced, the concentration of $iLOV_{OX}$ and $iLOV_{SQ}$ was determined for these samples and taken into account when determining mixing stoichiometry (**Figure 3.12**). Since iLOV Q104A have strongly inverted reduction potentials, [$iLOV_{SQ}$] was considered insignificant for this calculation.

Determination of [$iLOV_{OX}$] in $iLOV_{HQ}$ Q104A:

$$A_{440nm} = \left(\varepsilon_{HQ,440\,nm} * C_{tot} * \ell\right) - \left(\varepsilon_{HQ,440\,nm} * X\right) + \left(\varepsilon_{OX,440\,nm} * X\right)$$

$\varepsilon_{HQ,440\,nm} = 0.758\,mM^{-1}\,cm^{-1}$ (**Figure 3.11**)

$\varepsilon_{OX,440\,nm} = 16.1\,mM^{-1}\,cm^{-1}$ [35]

$X = [iLOV_{OX}]$

Determination of [$iLOV_{OX}$] in $iLOV_{HQ}$ Q104K:

$$A_{440nm} = \left(\varepsilon_{HQ,440\,nm} * (C_{tot} - C_{SQ} * \ell\right) - \left(\varepsilon_{HQ,440\,nm} * X\right) + \left(\varepsilon_{OX,440\,nm} * X\right)$$

$\varepsilon_{HQ,440\,nm} = 0.758\,mM^{-1}\,cm^{-1}$ (**Figure 3.11**)

$\varepsilon_{OX,440\,nm} = 16.1\,mM^{-1}\,cm^{-1}$ [35]

$\varepsilon_{SQ,616\,nm} = 4.25\,mM^{-1}\,cm^{-1}$ [32]

$C_{SQ} = \frac{A_{616\,nm}}{\varepsilon_{SQ,616\,nm}}$

$X = [iLOV_{OX}]$

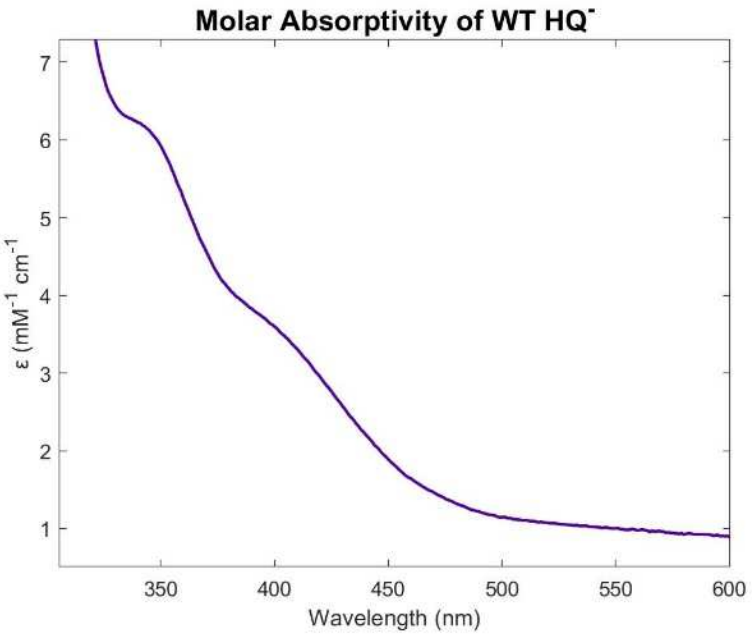

Figure 3.11 *Molar absorptivity of $iLOV_{HQ}$ WT was used to determine extinction coefficients for the calculation of $iLOV_{ox}$ remaining in partially-reduced mutant $iLOV_{HQ}$ samples.*

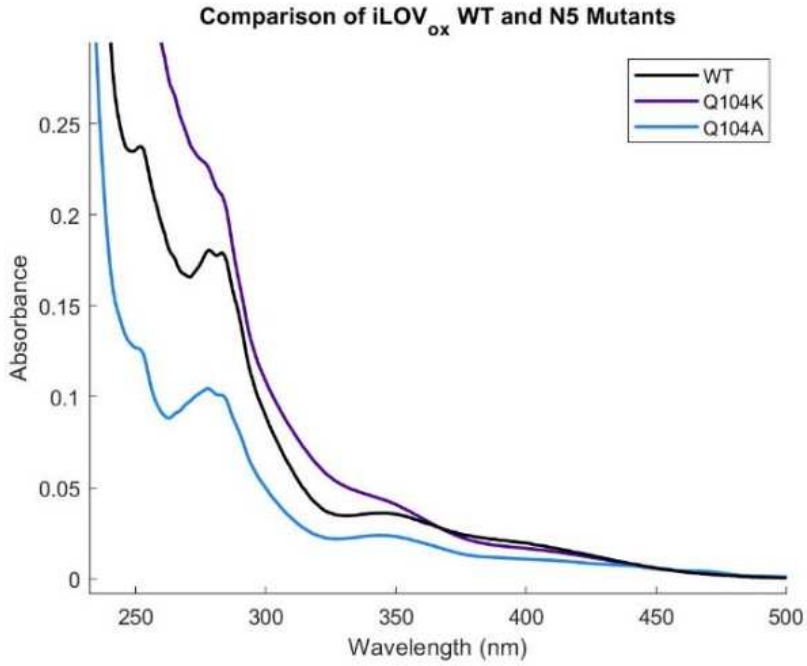

Figure 3.12 *Spectra of reduced iLOV WT and mutants. Incomplete reduction was assessed by absorbance at 440 nm.*

3.9 SpecEchem of iLOV

A concentrated (~ 2 mM) sample of $iLOV_{OX}$ was de-gassed on the Schlenk line and diluted to 500 µM with equimolar methyl viologen ($1e^-$ eq) in a spectroelectrochemical cell with a 2 mm path length. The platinum flag electrode (BASI EF-1355) was carefully lowered into the cell and checked to ensure proper alignment before placing into the cuvette holder. The cell was capped with its lid and the reference electrode (BASI MF-2052, Ag/AgCl, 3 M NaCl) and counter electrode (BASI EF-1356) were inserted into the solution. UV-vis data was collected using an Ocean Optics USB 2000 spectrometer and a DH-BAL-2000 light source. The electrodes were connected to a BioLogic SP-200 Potentiostat. Chronoamperometry was conducted at the E' of iLOV (-416 mV *vs* SHE at pH 8) for 8 hours while recording UV-vis spectra every 1 second.

References

[35] Davari, M.D., Kopka, B., Wingen, M., Bocola, M., Drepper, T., Jaeger, K.E., Schwaneberg, U., Krauss, U. Photophysics of the LOV-Based Fluorescent Protein Variant iLOV-Q489K Determined by Simulation and Experiment. *J Phys Chem*, **2016**, *120* (13) 3344-335

[36] Elgrishi, N., Rountree, K.J., McCarthy, B.D., Rountree, E.S., Eisenhart, T.T., Dempsey, J.L. A Practical Beginner's Guide to Cyclic Voltammetry. *J Chem Ed,* **2018**, *95*, 197-206