15N SOLID-STATE NMR DETECTION OF FLAVIN PERTURBATION BY H-BONDING IN MODELS AND ENZYME ACTIVE SITES

Dongtao Cui

University of Kentucky, dongtao.cui@gmail.com

Recommended Citation
Cui, Dongtao, "15N SOLID-STATE NMR DETECTION OF FLAVIN PERTURBATION BY H-BONDING IN MODELS AND ENZYME ACTIVE SITES" (2010). University of Kentucky Doctoral Dissertations. 48.
http://uknowledge.uky.edu/gradschool_diss/48

This Dissertation is brought to you for free and open access by the Graduate School at UKnowledge. It has been accepted for inclusion in University of Kentucky Doctoral Dissertations by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
ABSTRACT OF DISSERTATION

Dongtao Cui

The Graduate School
University of Kentucky
2010
15N SOLID-STATE NMR DETECTION OF FLAVIN PERTURBATION BY H-BONDING IN MODELS AND ENZYME ACTIVE SITES

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Arts and Sciences at the University of Kentucky

By
Dongtao Cui
Lexington, Kentucky

Director: Dr. Anne-Frances Miller, Associate Professor of Chemistry
Lexington, Kentucky
2010
Copyright © Dongtao Cui 2010
Massey and Hemmerrich proposed that the different reactivities displayed by different flavoenzymes could be achieved as a result of dominance of different flavin ring resonance structures in different binding sites. Thus, the FMN cofactor would engage in different reactions when it had different electronic structures. To test this proposal and understand how different protein sites could produce different flavin electronic structures, we are developing solid-state NMR as a means of characterizing the electronic state of the flavin ring, via the $^{15}$N chemical shift tensors of the ring N atoms. These provide information on the frontier orbitals. We propose that the $^{15}$N chemical shift tensors of flavins engaged in different hydrogen bonds will differ from one another. Tetraphenylacetyl riboflavin (TPARF) is soluble in benzene to over 250 mM, so, this flavin alone and in complexes with binding partners provides a system for studying the effects of formation of specific hydrogen bonds. For N5, the redox-active N atom, one of the chemical shift principle values (CSPVs) changed 10 ppm upon formation of a hydrogen bonded complex, and the results could be replicated computationally. Thus our DFT-derived frontier orbitals are validated by spectroscopy and can be used to understand reactivity. Indeed, our calculations indicate that the electron density in the diazabutadiene system diminishes upon H-bond complex formation, consistent with the observed 100 mV increase in reduction midpoint potential. Thus, the current studies of TPARF and its complexes provide a useful baseline for further SSNMR studies aimed at understanding flavin reactivity in enzymes.

KEYWORDS: Solid-state NMR, DFT, electronic structure, TPARF, flavoprotein.

Dongtao Cui
08/26/2010
15N SOLID-STATE NMR DETECTION OF FLAVIN PERTURBATION
BY H-BONDING IN MODELS AND ENZYME ACTIVE SITES

By
Dongtao Cui

Dr. Anne-Frances Miller
Director of Dissertation

Dr. John E. Anthony
Director of Graduate Studies

August 26, 2010
RULES FOR THE USE OF DISSERTATIONS

Unpublished dissertations submitted for the Doctor’s degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with permission of the author, and with the usual scholarly acknowledgements.

Extensive copying or publication of the dissertation in whole or in part also requires the consent of the Dean of the Graduate School of the University of Kentucky.

A library that borrows this dissertation for use by its patrons is expected to secure the signature of each user.

<table>
<thead>
<tr>
<th>Name</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
15N SOLID-STATE NMR DETECTION OF FLAVIN PERTURBATION BY H-BONDING IN MODELS AND ENZYME ACTIVE SITES

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Arts and Sciences at the University of Kentucky

By
Dongtao Cui
Lexington, Kentucky

Director: Dr. Anne-Frances Miller, Associate Professor of Chemistry
Lexington, Kentucky
2010

Copyright © Dongtao Cui 2010
This dissertation is dedicated to
my husband, Mr. Minao Tang
my daughter, Ms. Yining Tang
and my parents
Mr. Jiye Cui
Mrs. Zhaoqin He
ACKNOWLEDGEMENTS

Though only my name appears on the cover of this dissertation, a great many people have contributed to its production. I could never have reached the heights or explored the depths without your help, support, guidance and efforts.

First, my deepest gratitude is to my advisor, Dr. Anne-Frances Miller. I have been amazingly fortunate to have an advisor who gave me the freedom to explore on my own and at the same time the guidance to recover when my steps faltered. Her trust, understanding, encouragement, patience and support helped me overcome many crisis situations and finish this dissertation.

I am grateful to them members of my Ph.D committee, Dr. Bert Lynn, Dr. Stephen Testa, Dr. Boyd Haley, Dr. Trevor Creamer, and Dr. Wolfgang Korsch as the outside examiner. I appreciate your knowledge, time and patience throughout my graduate studies.

I am indebted to Dr. Ronald Koder for his help on making samples, Mr. John Layton and Mr. Art Sabasta for their assistance on NMR. I would like to thank all the members of the Miller lab, both past and present, Dr. Peng Zhang, Xiaonan Mei, Ting Wang, Jennifer Watt, and Mallory Muller. It has been such a pleasure to work with all of you. Additionally I wish to thank all of my friends at UK, especially Chunxia Zhao and Pei Gao. I will always cherish the times and memories we shared together.

Most importantly, none of this would have been possible without the love and support of my family. My husband, Minao Tang and my daughter Yining Tang, to whom this dissertation is dedicated to, have been a constant source of
love, concern, support and strength all these years. I would also like to thank my mother Zhaoqin He, father Jiye Cui. You provided me a lot of support and encouragement while at the University of Kentucky.

Thanks again to all of the above people!
# TABLE OF CONTENTS

Acknowledgments........................................................................................................... iii

List of Schemes................................................................................................................ viii

List of Tables.................................................................................................................... ix

List of Figures.................................................................................................................. x

Chapter One: Flavins and solid state NMR introduction
  1.1 Flavins and flavoproteins......................................................................................... 1
  1.2 Solid state NMR and chemical shift principle values on flavin system................... 3
  1.3 Interpreting shielding principal values via calculations..................................... 4

Chapter Two: SSNMR Studies of a Flavin Model - TPARF
  2.1 Introduction.............................................................................................................. 10
  2.2 Experimental Section............................................................................................. 15
    2.2.1 Synthesis.......................................................................................................... 15
    2.2.2 Optical spectra................................................................................................ 15
    2.2.3 NMR Conditions............................................................................................. 15
  2.3 Results ..................................................................................................................... 17
    2.3.1 Formation of a complex between TPARF and DBAP................................. 17
    2.3.2 Observation of TPARF’s H-bonding with water....................................... 21
  2.4 Discussion............................................................................................................... 22

Chapter Three: DFT Calculation of Flavin Electronics
  3.1 Introduction.............................................................................................................. 41
    3.1.1 Motivation........................................................................................................ 41
    3.1.2 Density Functional Theory (DFT).................................................................. 43
    3.1.3 Basis Sets....................................................................................................... 43
    3.1.4 Self-consistent reaction field (SCRF) method............................................ 45
    3.1.5 NMR Chemical Shielding.......................................................................... 46
    3.1.6 Natural Population Analysis (NPA).............................................................. 47
    3.1.7 Basis set superposition error (BSSE)........................................................ 48
    3.1.8 Molecular Orbitals......................................................................................... 48
    3.1.9 Molecular Dynamics (MD).......................................................................... 49
  3.2 Experimental Section............................................................................................. 50
  3.3 Results..................................................................................................................... 50
    3.3.1 Geometry Optimization............................................................................... 50
    3.3.2 NMR Chemical Shielding........................................................................... 52
    3.3.3 Molecular Dynamics (MD).......................................................................... 52
    3.3.4 Flavin dimmers............................................................................................. 53
    3.3.5 Solvent Effects.............................................................................................. 54
    3.3.6 Computations of TPARF’s complex with DBAP..................................... 55
4.3.6 Solution NMR of $^{15}$N FD.................................................. 99
4.3.7 Solid state NMR of $^{15}$N FD (1d CP-MAS)..............................101
4.3.8 His-tagged TEV-OYE1 cloning into pET32a vector........... 101
4.3.9 Overexpression and purification of
   Trx-His-TEV-OYE1.................................................................102
4.3.10 His-TEV-OYE1 cloning into pET3b vector..................... 102
4.3.11 Overexpression and purification of His-TEV-OYE1..... 102
4.3.12 Preparation, activity assay and titration
   with p-chlorophenol of native OYE1,
   apo-FMN OYE1 and FMN-reconstituted OYE1............ 103
4.4 Discussion..................................................................................104

Chapter Five: Conclusions and Future Perspectives
5.1 Conclusions..................................................................................137
5.2 Future Perspectives........................................................................139
   5.2.1 SSNMR comparison of different protein sites
       and calculation.................................................................139
   5.2.2 SSNMR detection of charge transfer complex
       formation in OYE1..........................................................140

Appendices
   Appendix A: Setting up and Running Gaussian Jobs..............142
   Appendix B: Example .com File................................................144

References.......................................................................................147

Vita....................................................................................................157
LIST OF SCHEMES

Scheme 1.1: The structures of flavins. For Riboflavin R = H, for FMN R = \( \text{PO}_3^{2-} \) and for FAD R = \((\text{PO}_3^-)-\text{adenosine} \)…………………6

Scheme 1.2: The oxidation states of flavins…………………………………………………7

Scheme 2.1: General scheme of a flavin-dependent dehydrogenation reaction………………………………………………………………………………24

Scheme 2.2: Structure of flavodoxin active site (A) (37), and TPARF (black) in complex with DBAP (gray) (B). R = tetraphenylacetylribityl for TPARF………………………………………25

Scheme 2.3: Reduction of FMN co-factor by one-electron. The reduction potential is indicated for flavodoxin (vs. N.H.E). Only one of several resonances structures is shown for the semiquinone………………………………………26
LIST OF TABLES

Table 1.1: Possible mechanism of typical flavin mediated reactions. 
Mechanisms are courtesy of Dr. Anne-Frances Miller ............8

Table 2.1: Experimental $^{15}$N Chemical Shift Principal Values 
(CSPVs) of N5 of $[^{15}$N-N5] solid TPARF, 
N5 of $[^{15}$N-N5] TPARF in DriSolv$^\circledR$ benzene, 
N5 of $[^{15}$N-N5] TPARF in DriSolv$^\circledR$ benzene with DBAP 
and N5 of $[^{15}$N-N5] TPARF in DriSolv$^\circledR$ benzene with H$_2$O, 
at different MAS spinning speeds........................................28

Table 2.2: Experimental CSPVs of N5 of $[^{15}$N-N5] TPARF 
and effects of solvent, H-bonding partners ....................29

Table 3.1: Cartesian coordinates of LF geometry optimized 
with B3PW91 and 6-311++G(2d,2p).................................61

Table 3.2: Cartesian coordinates of the complex of LF with DBAP, 
geometry optimized with B3PW91 and 6-311++G(2d,2p)........63

Table 3.3: Calculated Chemical Shift Principal Values (CSPVs) of 
N5 of LF solid, N5 of LF solid in DriSolv$^\circledR$ benzene, 
N5 of LF solid in DriSolv$^\circledR$ benzene with DBAP and 
N5 of LF solid in DriSolv$^\circledR$ benzene with H$_2$O.................64

Table 3.4: Counterpoise calculation results for complex of 
LF and DBAP, LF and one H$_2$O....................................65

Table 4.1: Isotropic chemical shift of $^{15}$N-FMN in various 
different environments in ppm vs. liquid NH$_3$ (11)............106

Table 4.2: List of Plasmids.............................................107

Table 4.3: Comparison of the FMN content of native OYE1, 
apo-FMN OYE1 and FMN-reconstituted OYE1....................108

Table 4.4: comparison of enzyme activity assay of native OYE1, 
and FMN-reconstituted OYE1 by measuring 
the rate of NADPH oxidation at 340 nm............................109
Figure 1.1: Binding features of the isoalloxazine ring in different flavoproteins. Images are courtesy of Dr. Anne-Frances Miller.........................................................9

Figure 2.1: Optical spectrum of TPARF in DriSolv® benzene showing the effects of complexation with DBAP. Also see Cerda et al (36)................................................................. 30

Figure 2.2: Optical spectrum of TPARF in DriSolv® benzene showing the effects of dimerization with 1 mm cuvette (A) and the absorbance at 445 nm vs. concentration, from the experiment shown above (B).............31

Figure 2.3: Solution NMR of [15N-N5] TPARF in DriSolv® benzene at different concentration. Data were collected at 60 MHz at 25 °C employing a 45 °C excitation pulse followed by 0.25 s acquisition and then a 4 s delay for relaxation. Chemical shift was referenced to 77 ppm for external 15N-urea in DMSO at 25 °C......................................................... 32

Figure 2.4: Solution NMR spectra of [15N-N5] TPARF in DriSolv® benzene, alone, with added DBAP or water (5 stoichiometric equivalents). Data were collected at 60 MHz at 25 °C employing a 45 °C excitation pulse followed by 0.25 s acquisition and then a 4 s delay for relaxation. Chemical shift was referenced to 77 ppm for external 15N-urea in DMSO at 25 °C......................................................... 33

Figure 2.5: 15N CP-MAS spectrum of [15N-N5] TPARF solid at room temperature with 20 ms acquisitions and 5 s between scans.................................................................34

Figure 2.6: Comparison of the 15N CP-MAS spectra of [15N-N5] TPARF solid (A), with that in DriSolv® benzene (B). 8 ms CP was used at -60 °C, with 20 ms acquisitions and 5 s between scans..............................35

Figure 2.7: 15N CP-MAS spectrum of [15N-N5] TPARF in DriSolv® benzene with DBAP. 8 ms CP was used at -60 °C, with 20 ms acquisitions and 5 s between scans.................................................................36

Figure 2.8: Comparison of the 15N CP-MAS spectra of dry [15N-N5] TPARF, [15N-N5] TPARF in DriSolv® benzene and [15N-N5] TPARF in DriSolv® benzene
with DBAP, all at 3000 Hz MAS speed. (Central peaks in the red box are enlarged for comparison of the $\delta_{\text{iso}}$.) Other conditions as in Figure 2.6 and 2.7...

**Figure 2.9:** $^{15}$N CP-MAS spectrum of $[^{15}\text{N-N5}]$ TPARF in DriSolv® benzene with H$_2$O. 8 ms CP was used at -60 °C, with 20 ms acquisitions and 5 s between scans...

**Figure 2.10:** Comparison of the $^{15}$N CP-MAS spectra of $[^{15}\text{N-N5}]$ TPARF in benzene and $[^{15}\text{N-N5}]$ TPARF in DriSolv® benzene with 5 and then 15 stoichiometric equivalents of water, at 3000 Hz MAS speed. (Side bands in the blue box are enlarged for comparison of their intensities: before (red) and after (blue) addition of water......

**Figure 2.11:** Effects of different H-bonding partner on CSPVs of TPARF...

**Figure 3.1:** DFT (B3PW91 and B3LYP) energies versus Basis Set Size. (A) LF (B) Complex of LF and DBAP...

**Figure 3.2:** Isotropic chemical shifts of N5 versus basis size and choice of functional used for each of GIAO and geometry optimization calculations...

**Figure 3.3:** Energy vs. time trajectory of MD calculation Beginning with the MM+ geometry-optimized structure. Entire 2 ns trajectory, insert: 1 ps of the trajectory surrounding the global energy minimum...

**Figure 3.4:** Comparison of the global energy-minimum structure and the structure representing the peak of the energy oscillations that persisted after equilibration (red circles in Figure 3.3)...

**Figure 3.5:** Optimized structure of LF dimers...

**Figure 3.6:** Distribution of electron density changes. (A) LumiFlavin; (B) complex of lumiflavin and DBAP. The most positive atom is C2 and its charge changed from 0.79501 to 0.80681; the most negative atom is N3 and its charge changed from -0.64057 to -0.66457...

**Figure 3.7:** Distribution of NPA electron density in LF (left) and LF complexed with DBAP (right). The total charge of LF changed from 0.01 to 0.04 upon complexation with DBAP. Structures are oriented
as in Scheme 1, red indicates excess electron density, green indicates electron deficiency…

Figure 3.8: Lengths of H-bonds measured from the LF•DBAP structure optimized using B3PW91 with 6-311++G(2d,2p).

Figure 3.9: HOMO (left) and LUMO (right) of LF (above), and LF in complex with DBAP (below).

Figure 3.10: Four possible optimized structures of LF with one H$_2$O molecule.

Figure 3.11: Six possible optimized structures of LF with two H$_2$O molecules.

Figure 3.12: Three possible optimized structures of LF with three H$_2$O molecules.

Figure 3.13: Three possible optimized structures of LF with four H$_2$O molecules.

Figure 3.14: One of the possible optimized structures of LF with six H$_2$O molecules (left) and with eight H$_2$O molecules (right).

Figure 4.1: The ribbon diagram of flavodoxin. It is generated by interpolating a smooth curve through the polypeptide backbone. α-helices are shown as coiled ribbons, β-strands as arrows, and lines for random coils. The direction of the polypeptide chain is indicated by a color ramp along the length of the ribbon. The cofactor, FMN, is shown in stick model. Carbon, oxygen, nitrogen, and phosphorus atoms are colored in gray, red, blue, and orange, respectively. Image is courtesy of Dr. Anne-Frances Miller. (1J8Q.pdb).

Figure 4.2: The ribbon diagram of old yellow enzyme. It is generated by interpolating a smooth curve through the polypeptide backbone. α-helices are shown as coiled ribbons, β-strands as arrows, and lines for random coils. The direction of the polypeptide chain is indicated by a color ramp along the length of the ribbon. The cofactor, FMN, is shown in stick model. Carbon, oxygen, nitrogen, and phosphorus atoms are colored in green, red, blue, and orange, respectively.
Figure 4.3: Plasmid map of Trx-6xHis-TEV-FD-pET32a (base number addresses are provided in parentheses)...

Figure 4.4: PCR program for amplifying the TEV-FD insert...

Figure 4.5: SDS-PAGE gel (stained with Coomassie blue) showing overexpression of FD-His in MM9. Lane 1, prestained SDS-PAGE standards, low range (Bio-Rad); lane 2-7, cells before addition of IPTG (0 hours, 1 hour, 2 hours, 3 hours and 4 hours after inoculation, respectively); lane 6-7, cells after addition of IPTG (1 hour and 4 hours after addition of IPTG, respectively); lane 8, kaleidoscope polypeptide standards (Bio-Rad)...

Figure 4.6: The Ni-NTA agarose column shows the dark purple semiquinone produced when reduced FD released from the anaerobic interior of E coli becomes oxidized by air)...

Figure 4.7: SDS-PAGE gel (stained with Coomassie blue) showing purity of FD-6xHis. Lane 1, kaleidoscope polypeptide standards (Bio-Rad); lane 2-4, purified FD growth in LB at different concentration, 9x, 3x, 1x; lane 5-7, purified FD growth in MM9 at different concentration, 9x, 3x, 1x...

Figure 4.8: Plasmid map of FD-His-pET23d (base number addresses are provided in parentheses)...

Figure 4.9: 1% agarose gel analysis of the products of PCR (amplification of FD insert). Lane 1, DNA ladder (pBR322 DNA-BstN I Digest); lane 2, FD insert (452bp)...

Figure 4.10: 1% agarose gel analysis of pET23d vector and FD insert after digestion with Ncol and XhoI. Lane 1, DNA ladder (Hind III digest); lane 2, pET23d vector after Ncol and XhoI digestion; lane 3, DNA ladder (pBR322 DNA-BstN I Digest); lane 4, FD insert after Ncol and XhoI digestion...

Figure 4.11: Comparison of the optical spectra of $^{14}$N FD, $^{14}$N apo-FD and $^{15}$N FMN-reconstituted $^{14}$N FD.
The native FD sample was considerably more concentrated than the reconstituted FD.)

Figure 4.12: Flow chart for preparing $^{15}$N FMN reconstituted $^{14}$N FD for SSNMR.

Figure 4.13: Proton wet1d NMR spectrum of FD.

Figure 4.14: $^1$H-$^{15}$N HSQC spectrum of uniformly $^{15}$N-labeled FD at 20 °C.

Figure 4.15: $^{15}$N 1d NMR spectrum of uniformly $^{15}$N-labeled FD.

Figure 4.16: SSNMR 1d CP-MAS spectrum for $^{15}$N FD at 4000 Hz spinning speed. Blue line indicates the potential signal of N5 on flavin ring and the red line on protein backbone.

Figure 4.17: Plasmid map of Trx-His-TEV-OYE1-pET32a (base number addresses are provided in parentheses).

Figure 4.18: PCR program for amplifying TEV-OYE1 insert (His-TEV-OYE1 insert).

Figure 4.19: 1% agarose gel analysis of the products of PCR (amplification of TEV-OYE1 insert). Lane 1, DNA ladder (pBR322 DNA-BstN I Digest); lane 2, TEV-OYE1 insert (1204 bp).

Figure 4.20: 1% agarose gel analysis of the His-tagged TEV-OYE1-pET32a DNA plasmid. Lane 1, DNA ladder (Hind III digest); lane 2, specific DNA standard (1100 bp); lane 3, TEV-OYE1-pET32a DNA plasmid after Ncol and Xhol digestion; lane 4, TEV-OYE1-pET32a DNA plasmid before Ncol and Xhol digestion.

Figure 4.21: SDS-PAGE electrophoresis (stained with Coomassie blue) showing overexpression and protein fractions at different stages of purification of Trx-His-TEV-OYE1. Lane 1, prestained SDS-PAGE standards, low range; lane 2, cells before add IPTG; lane 3, cells after add IPTG; lane 4, supernatant after French Press; lane 5, final flow through after bind to Ni-NTA agarose column; lane 6, final wash-through after washing off impurities; lane 7 to lane 9, elution fractions of Trx-His-TEV-OYE1 (5 mL fractions, 10 μL samples loaded of each).
Figure 4.22: Plasmid map of His-TEV-OYE1-pET3b (base number addresses are provided in parentheses) .......................... 131

Figure 4.23: 1% agarose gel analysis of the His-TEV-OYE1-pET3b DNA plasmid. Lane 1, His-TEV-OYE1-pET3b DNA plasmid after NdeI and BamHI digestion; lane 2, DNA ladder (Hind III digest); lane 3, His-TEV-OYE1-pET3b DNA plasmid, before NdeI and BamHI digestion ................................................. 132

Figure 4.24: Comparison of the optical spectra of native OYE1, apo-FMN OYE1 and FMN-reconstituted OYE1 .......................... 133

Figure 4.25: Enzyme activity assay of native OYE1, apo-FMN OYE1 and FMN-reconstituted OYE1 by measuring the rate of NADPH oxidation at 340 nm ................................................. 134

Figure 4.26: Spectral changes of titration of (A) native OYE1 and (B) FMN-reconstituted OYE1 with p-chlorophenol. The enzyme concentration was 2.6 x 10^{-2} mM in 0.1 M potassium phosphate, pH=6.45. The concentration of p-chlorophenol was 1 x 10^{-2} mM, 2 x 10^{-2} mM, 3 x 10^{-2} mM, 4 x 10^{-2} mM, 6 x 10^{-2} mM, 60 x 10^{-2} mM respectively. Binding curves of titration of (C) native OYE1 and (D) FMN-reconstituted OYE1 with p-chlorophenol using the absorbance at 645 nm at 25 °C ................................................................. 136
LIST OF FILES

DongtaoCuiETD.pdf (3.42 MB)
Chapter One: Flavins and solid state NMR introduction

1.1 Flavins and flavoproteins

Flavins (7,8-dimethyl-10-alkylisoalloxazines) are essential to the activity of a number of important oxidoreductases. Indeed, it is often the flavin itself that mediates the chemistry. Flavins are capable of undergoing oxidation-reduction reactions, and can mediate either one-electron or two-electron chemistry (1).

The flavin cofactors most commonly found in enzymes are FMN (flavin mononucleotide) and FAD (flavin adenine dinucleotide). These are generated from riboflavin. Riboflavin (vitamin B2) is an easily-absorbed micronutrient with a key role in maintaining health in animals. It was first isolated from whey in 1879 by Blyth (2), and the structure was determined by Kuhn and coworkers in 1934 (3).

The structures of riboflavin, FMN and FAD are shown in Scheme 1.1. The redox-active isoalloxazine ring is the core structure of the various flavins. It is a highly-conjugated system containing several heteroatoms. The flavins have a characteristic bright yellow color and take their name from the Latin flavus for ‘yellow’. Flavin can exist in any of three different oxidation states. Fully-oxidized flavin is converted to a semiquinone upon acceptance of one electron and a proton, as shown in Scheme 1.2. When the blue neutral radical looses a proton at higher pH values, it becomes a red radical anion. The semiquinone radical is particularly stable, owing to extensive delocalization of the unpaired electron across the π-electron system of the isoalloxazine. A second one-electron reduction converts the semiquionone to the completely reduced dihydroflavin (Scheme 1.2).

Flavoenzymes, containing FAD or FMN as a prosthetic group, catalyze a wide range of biochemical reactions, including dehydrogenation, monooxygenation, disulfide reduction, signaling, DNA repair, magnetic sensing and more, as shown in Table 1.1 (1, 4). The stability of the flavin semiquinone state allows flavoproteins to function as effective one-electron carriers in respiratory electron transfer.
The use of such chemically versatile cofactors provides the metabolic economy that relatively few different cofactors are needed for life, but requires in return that individual enzymes be able to restrict their bound cofactor's reactivity to a fraction of its innate repertoire. Massey and Hemmerich proposed that different flavoenzymes could achieve their distinct reactivities by favoring different resonance structures of their bound flavins (5). Our modern rephrasing of this idea is to propose that different reactivities will be associated with different modulations of the natures and relative energies of the flavin frontier orbitals. In most cases this must be accomplished via non-covalent interactions between the cofactor and the protein that vary from enzyme to enzyme. Interactions commonly observed to modulate flavin reactivity include steric distortion of the ring system, placement of local charges, stacking of aromatic side chains, modulation of the pKs of redox-coupled protons, polarity, and, importantly, hydrogen-bonding (H-bonding) (4, 6), the attractive interaction of hydrogen atom with an electronegative atom.

For example, Flavodoxin from Desulfovibrio vulgaris (FD, 1J8Q.pdb) and Old Yellow Enzyme (OYE, 1OYA.pdb) have different flavin binding sites (Figure 1.1). FD's function as electron transfer agents in a variety of microbial metabolic processes, and the FMN of FD forms H-bonds at O2 with Cys102 and Asp95, at N3(H) with Tyr100 and at O4 with Asp62 (Figure 1.1 (A)). While for OYE, which has NADPH oxidase activity and unusual phenol binding activity, the bound FMN has H-bonds at O2 and N3(H) with Gln114 and at O4 with Gly72 and Tyr37, and N5 H-bonds with Tyr37 (Figure 1.1 (B)). Indeed, flavins in different active site environments are found to have different reactions. As a step towards understanding how proteins modulate flavin reactivity, we are developing methodology for observing and understanding how H-bonding alters the flavin's electronic structure.

Flavin electronics have been studied by several methods. Absorption spectroscopy and circular dichroism features correlate with the flavin binding environment (7). Stark spectroscopy has revealed the extent of charge transfer between the flavin and nearby chromophores (8). Vibrational spectroscopy of
various sorts has been used to report on nearby electron-donating groups, steric
distortion upon binding to the protein, and H-bonding with the protein (9, 10). The
ultimate spatial resolution is provided by NMR’s detection of signals from
individual atoms. Solution state NMR has been used to infer bond order, the
existence and strength of H-bonding, and more (11, 12). Indeed, the isotropic
chemical shifts ($\delta_{iso}$s) measured in solution for individual flavin ring C and N
atoms vary by over more than 10 and 40 ppm respectively, among flavins in
different environments (13).

1.2 Solid state NMR and chemical shift principle values of flavin systems

Besides the information that has been provided by all the above methods,
we are developing tools for revealing the valence electron distribution in flavins
and herein advance the possibility of testing for a correlation between patterns of
electron density distribution and reactivity. Thus, we seek to understand
variations in flavin reactivity at the fundamental and unifying level of flavin
electronics. In order to be able to probe flavins’ electron density distributions over
the ring system, solid state NMR measurements were used to observe chemical
shift principal values (CSPVs) of one of the redox-active diazabutadiene
nitrogens, N5. SSNMR has the advantage over solution NMR that it can reveal
the presence of anisotropic (directionally dependent) interactions. It is a useful
method for structural analyses of a wide range of inorganic, organic and
biologically relevant systems which do not crystallize easily and which are
unsuitable for solution state NMR (14). By SSNMR all three CSPVs ($\delta_{11}$, $\delta_{22}$ and
$\delta_{33}$), can be measured for each atom. Hence SSNMR can detect chemical shift
changes that affect different principal values in opposite ways, and thus cancel
out of the isotropic average.

Moreover the flavin N atoms N1 and N5 of the flavin’s redox-active
diazabutadiene core both present uniquely favorable cases for characterization
by SSNMR (Scheme 1.1). The diamagnetic term contains an integral that
describes the electron density around the nucleus. The paramagnetic term arises
from the perturbation of the ground state wave function due to the coupling
between electronic orbital momentum and the external magnetic field. NMR chemical shifts reflect diamagnetic shielding due to ground-state electron density, as well as paramagnetic shielding due to electron orbital angular momentum. The paramagnetic term directly reflects the nature of the valence orbitals (15). The paramagnetic term also reflects the energy separations between the HOMO and accessible excited states, especially the LUMO. The HOMO and LUMO tend to be the very orbitals that underlie reactivity. For conjugated systems, the energy gap between the HOMO and LUMO is relatively small, so the participating nuclei are strongly deshielded. N1 and N5 have a non-bonded lone pair of electrons, and the energy gap between this orbital and the $\pi^*$ LUMO is smaller than the $\pi-\pi^*$ energy gap, explaining why extreme deshielding should be observed (16). The different CSPVs of such sites have been shown to each report primarily on a different pairs of frontier orbitals (17), Grant and Miller confirm that the largest CSPV, $\delta_{11}$, reflects those orbitals, and is very sensitive to interactions that modulate the energy of the n orbital (16, 17). Because H-bonding can stabilize the lone pair, this CSPV is particularly responsive to such interactions (18). Indeed, the $\delta_{11}$ of pyridine shifts upfield by 36 ppm when pyridine is diluted in a protic solvent (19). Furthermore, because the HOMO and LUMO are also the orbitals that play the largest role in reactivity, the CSPVs are also expected be related to reactivity, and its modulation by proteins. Thus SSNMR not only yields three times more descriptors of each atom than solution NMR, but the CSPVs may be related to the very orbitals that are expected to dominate the chemistry.

SSNMR of $^{15}$N has been established as a method of detecting perturbations in peptides (20). It has also proved very sensitive to the natures of H-bonds to imidazole (21). It promises to be equally applicable to the organic cofactors that are central to the reactivity of many enzymes. A better understanding of the electronic structures of these cofactors would greatly improve our understanding of catalytic mechanisms, and our ability to manipulate or build upon them. Moreover, understanding basis for reactivity is the basis for engineering.
1.3 Interpreting shielding principal values via calculations

In order to understand SSNMR results in terms of electron density distributions, we combine our SSNMR measurement with density functional theory (DFT) computations. DFT is successful and sufficiently economical to allow us to calculate electron density distributions and CSPVs for flavins and models of a flavin H-bonding with a partner (22-24). The SSNMR experimental CSPVs provide extremely critical tests of the calculations and the calculations permit interpretation of the CSPVs in terms of flavin electronics (16, 25). Thus, comparison of the calculated electron densities for isolated and H-bonded flavins will reveal how H-bonding interactions perturb the distribution of valence electron density (24, 26). Understanding the effects of H-bonds on valence electrons represents an essential first step towards rational chemical understanding of why different H-bonds can produce different flavin reactivity in different protein active sites (27, 28).

Thus, we propose that SSNMR and DFT calculations represent an exceptionally powerful combination ideally suited to revealing the electronic bases for variations and control of flavin reactivity, at the level of individual atoms and orbitals.
Scheme 1.1: The structures of common biological flavins. For Riboflavin $R = H$, for FMN $R = PO_3^{2-}$ and for FAD $R = (PO_3^-)$-adenosine.
Scheme 1.2: The oxidation states of flavins.
<table>
<thead>
<tr>
<th>Type of reaction</th>
<th>Possible reaction mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Oxidases</strong></td>
<td></td>
</tr>
<tr>
<td><img src="image" alt="Oxidases reaction mechanism" /></td>
<td></td>
</tr>
<tr>
<td><strong>Monooxygenases</strong></td>
<td></td>
</tr>
<tr>
<td><img src="image" alt="Monooxygenases reaction mechanism" /></td>
<td></td>
</tr>
<tr>
<td><strong>Electron transferases, dehydrogenases</strong></td>
<td></td>
</tr>
<tr>
<td><img src="image" alt="Electron transferases reaction mechanism" /></td>
<td></td>
</tr>
<tr>
<td><strong>Reductases</strong></td>
<td></td>
</tr>
<tr>
<td><img src="image" alt="Reductases reaction mechanism" /></td>
<td></td>
</tr>
</tbody>
</table>

**Table 1.1**: Possible mechanism of typical flavin mediated reactions. Mechanisms are courtesy of Dr. Anne-Frances Miller.
Figure 1.1: Binding features of the isalloxazine ring in different flavoproteins. Images are courtesy of Dr. Anne-Frances Miller.
Chapter Two: SSNMR Studies of a Flavin Model - TPARF

SSNMR was used to study a flavin model system - tetraphenylacetyl riboflavin (TPARF) and the complex of TPARF and its H-bonding partner dibenzylamidopyridine (DBAP). DBAP mimics the active site of flavodoxin, and thus enables us to assess the ability of H-bonding to perturb flavin electronics.

2.1 Introduction

Flavins have been recognized as being capable of both one- and two-electron transfer processes, playing a pivotal role in coupling the two-electron oxidation of most organic substrates to the one-electron transfers of the respiratory chain. Flavins can also function as electrophiles and nucleophiles. Flavoproteins have the unique ability to catalyze a wide range of biochemical reactions, including but not limited to oxidation-reduction, mono-oxidation, and dehydrogenation (1, 4). There is little relationship between folding topology and function. Topologically similar flavoproteins can catalyze different reactions, whereas proteins performing similar functions can have different folding topology (4). The chemical versatility of flavoproteins is controlled by specific interactions with the proteins with which they are bound (1). The protein milieu played a crucial role in fine-tuning the chemical properties of the flavin (4).

In flavoproteins, enzyme-cofactor interactions including hydrogen bonding and π-stacking are responsible for tuning the redox-properties of enzyme-bound flavin (5, 29, 30). The differing catalytic capabilities of the flavin in different enzymes result from modulation of the highly conjugated flavin isoalloxazine electronic structure by the protein environment (5, 29). A protein-supplied positive charge at N1-C2=O2 is functionally relevant because it can stabilize the

 Reproduced with permission from the Journal of the American Chemical Society, and Biochemistry, submitted for publication. Unpublished work copyright 2010 American Chemical Society.
anionic form of the reduced flavin and increase the cofactor’s reduction midpoint potential (Scheme 2.1) (30). Thus, the N1-C2=O2 locus serves to favour flavin binding and to regulate the redox properties of the cofactor. The N5 locus is another crucial site for flavin-binding and tuning (4). In most flavoproteins, N5 is within H-bond distance of a H-bond donor, typically a backbone or side chain nitrogen atom with attached hydrogen. Interaction involving N5 will affect catalysis.

The electron density distribution within bound flavins, and hence enzyme-cofactor interactions in the flavoenzymes, has been studied on the molecular level using NMR spectroscopy of the apoenzymes reconstituted with isotopically labeled flavins (31). The nitrogen atoms in flavins can be characterized as either pyridine-type nitrogens, N1 and N5 in the oxidized molecule, or pyrrole-type nitrogens, N10 and N3 in oxidized flavin and all four nitrogen atoms in two-electron reduced flavin. Information about H-bonding between the flavin and the amino acid residues of the protein could be obtained from the flavin 15N chemical shifts and 15N-1H coupling constants. For example, the 15N chemical shifts of pyridine-like nitrogen atom are very sensitive to H-bonding interactions, which lead to upfield shifts. In flavodoxins, 15N chemical shifts due to N5 of FMN reveal whether N5 is H-bonded or not, and whether H-bonding is weak or strong. N5 is not H-bonded in Megasphaera elsdenii (32) and Clostridium MP flavodoxins (33), weakly H-bonded in Azotobacter vinelandii (33) and Desulfovibrio vulgaris flavodoxins (34) and strongly H-bonded in Anabaena 7120 flavodoxin (35). Although the resonances of all four nitrogen atoms in protein-bound flavin can be observed without any interference from 15N natural abundance nuclei, it is not possible to attribute the chemical shift changes to specific interactions, including hydrogen bonding, π-stacking and protein environment (dipolar interactions).

To determine the effects of specific hydrogen bonds on NMR chemical shifts, and hence electron distributions, Rotello’s group developed a family of receptors designed to bind flavin derivatives using specific H-bonding interactions (36). A synthetic host molecule providing a model for specific flavoenzyme-cofactor interactions, H-bonding to O2, N3(H) and O4, was used to
observe the effect of H-bonding on flavin NMR. In a study of the complex of 2-methylpropyl flavin and host diethylaminopyridine, their $^{13}$C NMR data revealed the influence of H-bond interactions at O2, N3(H) and O4 on the π-electron density distribution in the flavin isoalloxazine ring system (36).

Solution NMR studies of flavins have traditionally used tetraacetylriboflavin (TARF) in CHCl$_3$ as a model system. It serves as a reference where H-bonding interaction with the solvent is diminished and self-association of flavin is minimized (31). Therefore, oxidized TARF was also used in the first SSNMR studies of a flavin in powder form. These studies revealed that SSNMR is a good tool for directly observing reactive positions of the flavin ring and thereby obtaining information on their frontier orbitals. These studies also demonstrated that very different CSPVs clearly distinguish the N5 and N1 sites, consistent with N5 and N1’s distinct electronic properties and reactivities (17, 22). That work provided a valuable reference point for understanding flavin reactivity in enzyme active sites, but in order to extend the studies to H-bonded complexes it is necessary to form these in at high concentration in solution. Although the flavin must be soluble to high concentration, the solvent itself should not engage in any H-bonding with the flavin rings. However the limited solubility of TARF precludes its use for such studies. Therefore, a new analog, tetraphenylacetyl riboflavin (TPARF) was designed and synthesized, wherein the acetyl groups of TARF are replaced with benzyl groups that dissolve readily in benzene (37). TPARF is soluble in benzene to over 250 mM, and electrochemical analysis of both TPARF alone and complexes with the benzene-soluble H-bonding partner, dibenzylamidopyridine (DBAP) (Scheme 2.2(B)), demonstrated that this model system behaves similarly to TARF model systems (37, 38).

Use of benzene as the solvent instead of the commonly-used CHCl$_3$ provides the crucial advantage that benzene does not H-bond to the flavin, whereas CHCl$_3$ does (37). We wanted to use a solvent not compete H-bonding for the analysis of the effects of hydrogen bonding, in order to be able to interpret spectral changes and energies as absolute effects of the H-bonds formed, rather
than differential effects upon exchange of weaker H-bonds for stronger ones. This could be illustrated by the following thermodynamic comparison.

For solvent not compete H-bonding, such as benzene:

\[\text{Flavin} + \text{H-bonding partner} \rightarrow \text{Flavin} \cdot \text{H-bonding partner}\]

\[\Delta G = \Delta G_{\text{H-bonding with H-bonding partner}}\]

For solvent which can form H-bonding with flavin, such as CHCl₃:

\[\text{Flavin} \cdot \text{Solvent} + \text{H-bonding partner} \rightarrow \text{Flavin} + \text{Solvent} + \text{H-bonding partner}\]

\[\rightarrow \text{Flavin} \cdot \text{H-bonding partner} + \text{Solvent}\]

\[\Delta G = \Delta G_{\text{H-bonding with H-bonding partner}} - \Delta G_{\text{H-bonding with solvent}}\]

For example, ranges of solvent effects on the nitrogen shielding of pyridine were reported by the Webb group (19). The range is significant. Pyridine’s nitrogen chemical shift is 38 ppm different when cyclohexane is the solvent vs when CF₃CH₂OH is the solvent. Thus formation of H-bonds between flavin and solvent strongly affects the nitrogen shielding. The change in nitrogen shielding upon a change in solvent polarity is also significant but less than that due to hydrogen bonding (19).

The use of the benzene-soluble H-bonding partner DBAP was inspired by earlier work by Yano and Rotello on diethylamidopyridine (DAP) (38, 39). The benzyl analog DBAP, binds to TPARF with a \(K_d\) of 420 \(\mu\)M and raises the flavin’s Oxidized/Semiquinone reduction potential by 100 mV (37, 38). The extreme solubility of this analogue in nonpolar, noninteractive solvent allowed us to perform SSNMR. Furthermore, the TPARF and DBAP model mimics the active site interactions of flavodoxin (Scheme 2.2(A)) (38). The observed changes of chemical shifts upon H-bond formation reflect electronic density distribution changes, which should be related to flavin reactivity. Thus, the TPARF in benzene and the complex with its H-bonding partner DBAP provide us with a stable benzene soluble model system wherein we can study how H-bond formation modulates the reactivity of a flavin.

High-resolution \(^{15}\text{N}\) SSNMR spectroscopy with magic angle spinning (MAS) has been applied to study H-bonding in proteins since early 1980’s (40). It
permits the measurement of isotropic chemical shifts in samples exhibiting anisotropic spin interactions, making it attractive for study of heterogeneous systems. From the study of $^{15}$N NMR of $^{15}$N$^{\pi}$-labeled histidyl $\alpha$-lytic protease in lyophilized powders with pH values ranging from 4.9 to 9.3, the $^{15}$N isotropic shifts demonstrate that N$^{\pi}$ of His-57 participates in a strong H-bond, as the H-bond donor, in powders prepared at both high pH and low pH ($^{40, 41}$). Thus, the isotropic shift is a good indicator of the protonation state, and can provide information concerning H-bonding.

Anisotropic chemical shifts have the potential to reveal more detail about the H-bonding interactions. The chemical shift anisotropy study of the H-bonding of carboxyl groups revealed that three tensor elements reveal completely different facets of the interaction ($^{42, 43}$). $^{15}$N chemical shift anisotropy for a variety of crystalline histidine and histidine-containing peptides at natural abundance were reported by the McDermott group ($^{44}$). These studies provide databases of chemical shift tensor values that are useful for studies of imidazole groups in enzymes and permit interpretation of SSNMR parameters in terms of protonation state and H-bonding. SSNMR chemical shifts have been a most insightful tool for studying H-bonds in enzyme active sites.

Given our goal of understanding how different protein sites can modulate flavin reactivity, we begin by focusing on the N5 site of the flavin ring (Scheme 2.2(B)), as this is one of the two redox-active nitrogens and the site of hydride acceptance. For example, Flavodoxin (FD) from Desulfovibrio vulgaris, a 16 kDa protein, stabilizes the neutral (blue) flavin radical at reduction potentials between -260 mV and -320 mV vs. N.H.E. ($^{45}$), uptake of an electron is accompanied by acquisition of a proton by the N5 site (Scheme 2.3) ($^{46}$). Thus, N5 participates directly in much of flavin chemistry and constitutes a reactive site of the flavin ring. H-bonds donated to N1, O2 and O4 have been proposed to draw electron density out of the ring and make N5 a better oxidant ($^{36}$).

Ramsey’s equation predicts that NMR chemical shifts will reflect shielding due to ground-state electron density, as well as deshielding due to electron orbital angular momentum ($^{15, 17}$). The chemical shift tensor of N5 is over 720
This is consistent with Ramsey’s equation, that understand the large deshielding of the $\delta_{11}$ CSPV as a reflection of the relatively small energy between the N5 non-bonded lone pair and the flavin’s extensively conjugated $\pi^*$ orbitals (15, 16). Thus any interactions that modulate the energy of the $\pi^*$ orbitals or the non-bonded lone pair should be reflected in $\delta_{11}$. In order to focus on N5, we used TPARF incorporating $^{15}$N at that position, and now report the CSPVs of N5 and their response to solvation in benzene, formation of a well-defined H-bonded complex with DBAP, and engagement by weak transient H-bonds with a H-bonding solvent, water.

2.2 Experimental Section

2.2.1 Synthesis

$[^{15}\text{N-N5}]$ TPARF and DBAP were generously provided by Dr. Ronald L. Koder, Department of Physics, The City College of New York, who synthesized them according to published methods (37). All the benzene used was DriSolv® benzene, purchased from EMD and maintained under argon.

2.2.2 Optical spectra

All optical spectra were obtained using a HP-8452B diode array spectrophotometer (Agilent). Solid TPARF and DBAP were each dissolved in DriSolv® benzene. The optical spectra collected in Figure 2.1 were using a 1 cm path length cuvette, which was the same as in literature. A 1 mm cuvette made optical spectral collection possible for higher concentration samples are shown in Figure 2.2.

2.2.3 NMR Conditions

All solution state $^{15}$N 1d NMR spectra were recorded on a Varian Inova™ 600 MHz spectrometer. Data were collected at 60 MHz at 25 °C employing a 45 ° excitation pulse followed by 0.25 s acquisition and then a 4 s delay for relaxation. Chemical shift was referenced to 77 ppm for external $^{15}$N-urea in DMSO at 25 °C.
The sweep width was typically 50,000 Hz, and at least 102,400 scans were accumulated.

All SSNMR data were collected at a nominal temperature of -60 °C. This was chosen on the basis of $^{15}$N signal-to-noise obtained, and the requirement that benzene (freezing point = 5.5 °C) be frozen solid. The temperature value was not corrected for frictional heating due to spinning etc. However, the choice of temperature was found not to affect the CSPVs obtained for TARF, in previous work (17). The signal-to-noise of spectra improved as the temperature was lowered from 0 °C to -30 °C to -60 °C. No further improvement was obtained upon cooling to -90 °C.

Solid TPARF (0.124 g, 0.146 mmol) was packed in a 5 mm zirconia rotor. 80 μL 200 mM TPARF in DriSolv® benzene, 80 μL 200 mM TPARF in DriSolv® benzene with equal molar DBAP, or 80 μL 200 mM TPARF in DriSolv® benzene with 1.6 μL H₂O were each contained in delrin vials custom machined to fit snugly into the 5 mm zirconia rotors. 1.6 μL of water provides five stoichiometric equivalents of water per flavin. A sample containing three times as much water was also characterized and found to produce the same CSPVs as the 1.6 μL-containing sample. Rotors were frozen vertically in dry ice / acetone, loaded into the probe at 0°C and then cooled to -60°C while spinning at 5000 Hz.

$^{15}$N MAS spectra were collected at 40 MHz for $^{15}$N with signal enhancement via ramped Cross Polarization (CP) from $^1$H in 5 mm Zirconia rotors in an HX Chemagnetics-type probe using a Varian Inova™ spectrometer. $^{15}$N spectra were collected with interscan delays of 5 sec, CP contact times of 8 ms at a CP field of 50 kHz for $^1$H and $^1$H TPPM2 decoupling at 50 kHz during 20 ms acquisitions (47). Spinning speeds of 3000, 4000 and 5000 Hz were used for the $^{15}$N 1d CP-MAS spectra used to determine $^{15}$N CSPVs (Table 2.1). All $^{15}$N chemical shifts are quoted relative to liquid ammonia, and spectra were referenced indirectly based on the $^{15}$NO₃⁻ signal of NH₄NO₃ at 376 ppm (48). The intensity distribution among the spinning sidebands yielded the three CSPVs of the signal via Herzfeld-Berger analysis of each spectrum (49, 50). Standard deviations reported reflect variations in the CSPVs and $\delta_{iso}$ values obtained from
different MAS speeds for a specific sample. Thus they incorporate experimental as well as fitting errors. All chemical shifts are also subject to systematic offset error related to chemical shift referencing in combination with the effect of temperature. We addressed this by comparing the chemical shift of $^{15}$N-urea in DMSO at 25 °C obtained via our NH$_4$NO$_3$ reference with the literature value of 77 ppm, and by comparing the chemical shifts obtained for TARF obtained at room temperature vs. at -60 °C based on the same reference. Experimental uncertainties associated with the CSPVs were estimated at 1 ppm based on the repeatability of the values obtained from determinations made of the same sample on different occasions at the same temperature (the corresponding uncertainty in the $\delta_{iso}$ was 0.3 ppm).

2.3 Results

2.3.1 Formation of a complex between TPARF and DBAP

In order to exploit the distribution of different electronic effects into different CSPV, and to detect changes in individual CSPVs without losses due to cancellation in their average, we wished to use SSNMR. However $^{15}$N is a relatively insensitive NMR nucleus, so even with the use of CP, SSNMR of $^{15}$N requires very high concentration of TPARF, in addition to freezing of the benzene. Benzene did not freeze to a glass in our hands, even when various mixtures of benzene and toluene were tried. Thus, upon freezing, we do not retain a truly homogenous medium. Moreover we cannot be sure that the TPARF does not precipitate out of solution in the course of freezing. Finally, at the very high concentrations necessary for $^{15}$N SSNMR, in combination with the non-interacting solvent, we must consider the possibility that TPARF molecules associate with one another.

We obtained an estimate for the dimerization constant of TPARF from the dependence of the optical spectrum on concentration. Although previous optical spectroscopic work had concluded that TPARF does not dimerize significantly in solution, those studies employed considerably lower concentrations of TPARF (Figure 2.1). A 1 mm path length cuvette makes the optical spectroscopic study
possible for more concentrated samples. In the concentration range of TPARF up to 0.6 mM, although there is an almost linear relationship between the absorbance at 446 nm and concentration, the non-zero intercept means that we do not have ideal behavior (Figure 2.2).

The possibility of TPARF dimerization was further addressed via a study of the $\delta_{\text{iso}}$ vs. concentration of TPARF by solution NMR at room temperature (Figure 2.3 and Figure 2.4). The N5 $^{15}\text{N} \delta_{\text{iso}}$ of $[^{15}\text{N}-\text{N5}]$ TPARF in DriSolv® benzene was 346.8 ppm when extrapolated to zero concentration. In contrast, the N5 $^{15}\text{N} \delta_{\text{iso}}$ of $[^{15}\text{N}-\text{N5}]$ TPARF in DriSolv® benzene with DBAP was increased if the concentration was higher. The chemical shift was 349.6 ppm when extrapolated to zero concentration. So, the extrapolated effect of H-bonding to DBAP was 2.9 ppm. However, the relationship between N5 $^{15}\text{N} \delta_{\text{iso}}$ of $[^{15}\text{N}-\text{N5}]$ TPARF in DriSolv® benzene with DBAP and concentration was not perfect linear. There appears to be curvature when the TPARF concentration is near 2 mM. Because each sample had added excess DBAP, I consider that when the concentration of TPARF was smaller than 2 mM, TPARF•DBAP complex is the dominant form, which showing the trend of sharply increased chemical shift. However, when the concentration of TPARF was bigger than 2 mM, TPARF could form TPARF dimer showing less sharply increased chemical shift changes.

Thus it is quite likely that TPARF exists substantially as dissolved dimers in our very concentrated solutions in benzene. Some possible molecular configurations of such associations were compared by computational methods, which are described in detail in the next chapter. However we note here that the computed $\delta_{\text{iso}}$ of monomeric lumiflavin differs from that of lumiflavin dimers by a change that is in the same direction and similar in magnitude to that observed in solution. Thus, considering that the $\delta_{\text{iso}}$ is concentration dependent and the difference agrees with the computational prediction for dimerization, we infer that our SSNMR data reflect TPARF dimers rather that isolated TPARF molecules. Consequently, the changes observed upon addition of DBAP could be interpreted as due to displacement of TPARF by DBAP. Crystallographic experiments are planned to refine our understanding of TPARF-TPARF.
interactions, however we note that several different configurations are possible and may co-exist in solution.

The chemical shifts of TPARF powder at room temperature and -60 °C assess the effect of temperature independent of the phase transition (Figure 2.5). At room temperature, $^{15}$N one-dimensional cross polarized magic angle spinning NMR spectra (1 d CP-MAS spectra) showed that the single signal spans 700.53 ppm (Table 2.1). The frequency of the central peaks indicates the isotropic chemical shift, which is 342.08 ppm. For each different spinning speed, the side bands spanned over the spectral width and parted by the corresponding spinning speed. The three CSPVs are 669.9 ± 3.8 ppm, 386.9 ± 2.1 ppm, and -30.6 ± 3.3 ppm based on determination of the CSPVs at three different MAS speeds including replicate spectra at individual speeds. At -60 °C, TPARF’s N5 $^{15}$N $\delta_{iso}$ was 342.8 ppm, only 0.8 ppm bigger than that of solid TPARF at room temperature. Similarly, all three CSPVs were within error of the values obtained for the same sample at room temperature. Thus, low temperature, -60°C, could increase the $^{15}$N signal-to-noise without altering signal shape or intensity distributions.

We also assessed non-glassing of benzene and the innocence of benzene as a solvent at low temperature, by comparing the $^{15}$N spectra of $[^{15}$N-N5] TPARF solid and $[^{15}$N-N5] TPARF dissolved in DriSolv® benzene and frozen. For the dry TPARF, $^{15}$N 1d CP-MAS spectra showed that the signal spans 709.0 ppm with a $\delta_{iso}$ of 342.8 ppm (Table 2.1). The three CSPVs are 674.7 ± 6.1 ppm, 388.1 ± 3.8 ppm, and -34.3 ± 5.9 ppm based on determination of the CSPVs at three different MAS speeds including replicate spectra at individual speeds. In benzene, TPARF N5 $^{15}$N $\delta_{iso}$ was 342.0 ppm, only 0.8 ppm smaller than that of solid TPARF (Figure 2.6). Similarly, all three CSPVs were within error of the values obtained for dry TPARF. Thus, dissolving TPARF in benzene does not significantly change its CSPVs and we confirm that benzene acts as a solvent not compete H-bonding. Any TPARF•TPARF associations present in solution are also apparently present in the powder, consistent with the fact that our powders were dried out of ethyl acetate at very high concentrations of TPARF.
Spectroscopic analyses of the interactions between the oxidized TPARF and DBAP in DriSolv® benzene were first performed in a 1 cm path length cuvette in a concentration of 10⁻² mM as in literature. TPARF in DriSolv® benzene had a maximum absorbance at 445 nm (Figure 2.1). Addition of DBAP to TPARF in DriSolv® benzene resulted in a 20% increase in the extinction coefficient of the flavin in concert with a red shift the main absorption band from 445 nm to 447 nm, in agreement with published work (Figure 2.1) (37).

Solution NMR spectra were collected to assay for formation of a TPARF•DBAP complex and compare the N5 ¹⁵N δiso of [¹⁵N-N5] TPARF alone in DriSolv® benzene vs TPARF complexed with DBAP. The N5 ¹⁵N δiso of [¹⁵N-N5] TPARF alone in DriSolv® benzene was 346.1 ppm whereas upon addition of DBAP, the δiso increased by 3.7 ppm (Figure 2.4). This is a significant change but small compared to the changes observed by SSNMR. To study TPARF’s complex with DBAP we dissolved TPARF ± DBAP in benzene and froze the solutions to achieve the solid state. Upon complexation with DBAP, the TPARF’s N5 δiso increased 6.7 ± 0.8 ppm to 348.7 ppm (Figure 2.7 and 2.8). Of the three CSPVs, δ₁₁ was the most responsive to the H-bonding, with a change of 10.4 ± 5.0 ppm (Tables 2.1 and 2.2). δ₁₁’s strong deshielding and responsiveness to electronic perturbation associated with noncovalent interactions is expected based on theory, prior experiments and calculations (16, 19, 22). However our experimental results extend these findings to the more complex, and biologically important flavin system. Furthermore, previous studies addressed direct interactions with the observed atom (36), whereas DBAP does not form an H-bond with N5 and must exert its effect via perturbation of molecular orbitals (37). Similarly, the most stable TPARF•TPARF dimers do not involve interactions at the N5 site (see Chapter 3 for detail, Figure 3.5), so the chemical shift change is also unlikely to reflect losses of interactions at N5. Regardless of underlying molecular details, any effects observed can only be attributed to changes in intermolecular hydrogen bonds, yet the changes in δiso and δ₁₁ are readily detected, and significant compared with experimental error. Thus, we anticipate
that it should be possible to detect the effects on N5 of H-bonding between proteins and flavins.

2.3.2 Observation of TPARF’s H-bonding with water

The effect of H-bonding directly to the N5 site could also be detected, although we do not have the benefit of a synthetic binding partner designed to H-bond to N5. Instead, we exploited the fact that the flavin N5 can H-bond with H2O. Because our calculations indicated that an individual flavin molecule can bind at least four H2O molecules before H2O molecules begin to preferentially self-associate (Chapter 3), we can expect that individual H2O molecules may fluctuate between binding in different locations and being free, and that an ensemble of flavin molecules with a few H2O will include a statistical collection of all the different possible H-bonding options for each H2O molecule. Therefore, a five-fold stoichiometric excess of H2O was added. This results in a concentration of water in excess of the solubility of water in benzene at room temperature, 0.0355 M, and at 5 °C where benzene freezes (51). Thus, H2O molecules will be available in excess to interact with benzene. Moreover we also confirmed that the concentration used was saturating with respect to the effects observed by SSNMR.

In solution NMR spectra, the presence of 5 equivalents of H2O produced an $\delta_{iso}$ of 345.5 ± 0.1 ppm (Figure 2.4), barely changed from TPARF in benzene alone at 346.1 ± 0.1 ppm. However the CSPVs measured by SSNMR revealed that the added H2O perturbs the flavin electronic structure, as $\delta_{11}$ changed by -6.1 ± 4.0 ppm (Figure 2.9 and Table 2.2) (note changes in intensity distribution among spinning side bands in Figure 2.10). The SSNMR spectra of this sample were experimentally indistinguishable from those obtained upon addition of 15 stoichiometric equivalents of water, indicating that H2O in excess of 5 equivalents produces negligible additional effects (Figure 2.10). Thus, although chemical shift changes are invisible to solution NMR, SSNMR detects a significant effect on $\delta_{11}$. Moreover this effect is clearly distinct from the effect of H-bonding with DBAP and H2O (Table 2.2 and Figure 2.11). H-bonding with DBAP increases all three
CSPVs, especially $\delta_{11}$. However, H-bonding with $\text{H}_2\text{O}$ has different effect on different CSPVs. There is a relatively big negative effect on $\delta_{11}$, a relatively big positive effect on $\delta_{33}$ and a relatively small positive effect on $\delta_{22}$. Thus, the average of CSPVs cancels each other shows a relatively small positive change which is very hard to tell by solution state NMR.

2.4 Discussion

Although NMR spectroscopy is inherently insensitive, $15\text{N}$ NMR is particularly poor. The sensitivity is improved dramatically and the spectra are rendered selective for flavin by use of $^{15}\text{N}$ labeled flavins. Dr. Koder synthesized [$^{15}\text{N-N5}$] TPARF and its H-bonding partner DBAP for us. Because TPARF could soluble in benzene up to 250 mM, we can produce well-defined complexes in solution at high concentration.

SSNMR is used to detect the model of protein site for the first time and it is possible to get the differences modulated by different H-bonding partner. It is exciting fact that different signatures are associated with different interactions. The TPARF•DBAP system in DriSolv® Benzene is an ideal system to perform this study, because the solvent will not compete H-bonding at all. The irony that use of such a genuinely solvent has made flavin•flavin interactions essentially inevitable. TPARF could form dimers while the concentration is higher than 0.08 mM. However flavin•flavin interactions have the possibility of being better defined. Since we can detect the changes with adding H-bonding partner to this system, it is might be true that the CSPVs changes that we detected is the difference between TPARF dimer and TPARF•DBAP. This further indicate that SSNMR is sensitive to detect H-bonding, because we have shown that the $^{15}\text{N}$ CSPVs not only detect H-bonding, but display qualitatively different responses to different sorts of H-bonding. Thus, different patterns of change may be produced by different non-covalent interactions.

So, SSNMR is a good method to detect the H-bonding interactions in flavin system and it is capable to detect the atom specific and even orbital specific interactions, because CSPVs reflect specific combination of HOMO and
LUMO. It provides a way to detect the electronic structure changes of flavin system while H-bonding with the partner. Thus, it enables us to use SSNMR to understand why different flavoprotein containing the same cofactor can do very different chemistry.
Scheme 2.1: General scheme of a flavin-dependent dehydrogenation reaction.
Scheme 2.2: Structure of flavodoxin active site (A) (38), and TPARF (black) in complex with DBAP (gray) (B). R = tetraphenylacetylribityl for TPARF.
Scheme 2.3: Reduction of FMN co-factor by one-electron. The reduction potential is indicated for flavodoxin (vs. normal hydrogen electrode, N.H.E.). Only one of several resonances structures is shown for the semiquinone.
<table>
<thead>
<tr>
<th>TPARF Solid</th>
<th>MAS Speed (Hz)</th>
<th>δ₁₁ (ppm)</th>
<th>δ₂₂ (ppm)</th>
<th>δ₃₃ (ppm)</th>
<th>δᵢₒ = (δ₁₁ + δ₂₂ + δ₃₃) / 3 (ppm)</th>
<th>Span = δ₁₁ - δ₃₃ (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>at RT</td>
<td>3000</td>
<td>672</td>
<td>384</td>
<td>-29</td>
<td>342.2</td>
<td>701</td>
</tr>
<tr>
<td></td>
<td>4000</td>
<td>674</td>
<td>387</td>
<td>-35</td>
<td>342.2</td>
<td>709</td>
</tr>
<tr>
<td></td>
<td>5000</td>
<td>667</td>
<td>388</td>
<td>-29</td>
<td>341.9</td>
<td>696</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>671</td>
<td>387</td>
<td>-31</td>
<td>342.1</td>
<td>702</td>
</tr>
<tr>
<td></td>
<td>Average⁺</td>
<td>670</td>
<td>387</td>
<td>-31</td>
<td>342.1</td>
<td>701</td>
</tr>
<tr>
<td></td>
<td>STDEV</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>0.2</td>
<td>7</td>
</tr>
<tr>
<td>at -60 °C</td>
<td>3000</td>
<td>676</td>
<td>384</td>
<td>-31</td>
<td>342.8</td>
<td>707</td>
</tr>
<tr>
<td></td>
<td>4000</td>
<td>670</td>
<td>392</td>
<td>-33</td>
<td>342.8</td>
<td>702</td>
</tr>
<tr>
<td></td>
<td>5000</td>
<td>682</td>
<td>389</td>
<td>-42</td>
<td>342.8</td>
<td>724</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>676</td>
<td>388</td>
<td>-35</td>
<td>342.8</td>
<td>711</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>675</td>
<td>388</td>
<td>-34</td>
<td>342.8</td>
<td>709</td>
</tr>
<tr>
<td></td>
<td>STDEV</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>0.3⁺</td>
<td>11</td>
</tr>
<tr>
<td>in DriSolv® Benzene</td>
<td>3000</td>
<td>674</td>
<td>388</td>
<td>-36</td>
<td>342.2</td>
<td>709</td>
</tr>
<tr>
<td></td>
<td>4000</td>
<td>676</td>
<td>391</td>
<td>-40</td>
<td>342.2</td>
<td>716</td>
</tr>
<tr>
<td></td>
<td>5000</td>
<td>680</td>
<td>391</td>
<td>-47</td>
<td>341.6</td>
<td>727</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>677</td>
<td>390</td>
<td>-41</td>
<td>342.0</td>
<td>717</td>
</tr>
<tr>
<td></td>
<td>Average⁺</td>
<td>676</td>
<td>390</td>
<td>-40</td>
<td>342.0</td>
<td>716</td>
</tr>
<tr>
<td></td>
<td>STDEV</td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>0.4</td>
<td>8</td>
</tr>
<tr>
<td>with DBAP</td>
<td>3000</td>
<td>689</td>
<td>391</td>
<td>-33</td>
<td>348.9</td>
<td>721</td>
</tr>
<tr>
<td></td>
<td>4000</td>
<td>682</td>
<td>400</td>
<td>-33</td>
<td>349.5</td>
<td>715</td>
</tr>
<tr>
<td></td>
<td>5000</td>
<td>689</td>
<td>397</td>
<td>-41</td>
<td>348.0</td>
<td>730</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>686</td>
<td>396</td>
<td>-36</td>
<td>348.8</td>
<td>722</td>
</tr>
<tr>
<td></td>
<td>Average⁺</td>
<td>687</td>
<td>395</td>
<td>-36</td>
<td>348.7</td>
<td>723</td>
</tr>
<tr>
<td></td>
<td>STDEV</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>0.8</td>
<td>7</td>
</tr>
<tr>
<td>in DriSolv® Benzene with 1:5 H₂O</td>
<td>3000</td>
<td>670</td>
<td>387</td>
<td>-29</td>
<td>342.8</td>
<td>699</td>
</tr>
<tr>
<td></td>
<td>4000</td>
<td>673</td>
<td>392</td>
<td>-38</td>
<td>342.5</td>
<td>711</td>
</tr>
<tr>
<td></td>
<td>5000</td>
<td>668</td>
<td>393</td>
<td>-34</td>
<td>342.2</td>
<td>702</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>670</td>
<td>391</td>
<td>-34</td>
<td>342.5</td>
<td>704</td>
</tr>
<tr>
<td></td>
<td>Average⁺</td>
<td>670</td>
<td>391</td>
<td>-34</td>
<td>342.5</td>
<td>704</td>
</tr>
<tr>
<td></td>
<td>STDEV</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>0.3</td>
<td>6</td>
</tr>
</tbody>
</table>
Table 2.1: Experimental $^{15}$N Chemical Shift Principal Values (CSPVs) of N5 of $[^{15}$N-N5] solid TPARF, N5 of $[^{15}$N-N5] TPARF in DriSolv® benzene, N5 of $[^{15}$N-N5] TPARF in DriSolv® benzene with DBAP and N5 of $[^{15}$N-N5] TPARF in DriSolv® benzene with H$_2$O, at different MAS spinning speeds.

a. Chemical shift relative to liquid ammonia. Average is the arithmetic mean and average’ is weighted according to the chi squared values associated with the Hertzfeld-Berger fits and standard deviations are provided too. Reproducibility was 2, 3, 3, 0.6 ppm in replicate spectra. Standard deviations are provided for the averages.

b. Precision limited only by instrumental error, estimated at 0.3 ppm.
<table>
<thead>
<tr>
<th></th>
<th>$\delta_{11}$ (ppm)</th>
<th>$\delta_{22}$ (ppm)</th>
<th>$\delta_{33}$ (ppm)</th>
<th>$\delta_{\text{iso}}$ (ppm)</th>
<th>Span (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPARF Solid</td>
<td>670</td>
<td>387</td>
<td>-31</td>
<td>342.1</td>
<td>701</td>
</tr>
<tr>
<td>Effect of temperature</td>
<td>5</td>
<td>1.2</td>
<td>-4</td>
<td>0.8</td>
<td>9</td>
</tr>
<tr>
<td>Effect of dissolving</td>
<td>2</td>
<td>1.7</td>
<td>-6</td>
<td>-0.8</td>
<td>7</td>
</tr>
<tr>
<td>in DriSolv® benzene</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effect of addition of DBAP</td>
<td>10</td>
<td>5</td>
<td>4</td>
<td>6.7</td>
<td>6</td>
</tr>
<tr>
<td>Effect of addition of H$_2$O</td>
<td>-6</td>
<td>1</td>
<td>6</td>
<td>0.5</td>
<td>-12</td>
</tr>
</tbody>
</table>

**Table 2.2:** Experimental CSPVs of N5 of $[^{15}$N-N5] TPARF and effects of solvent, H-bonding partners.
Figure 2.1: Optical spectrum of TPARF in DriSolv® benzene showing the effects of complexation with DBAP. Also see Cerda et al (37).
Figure 2.2: Optical spectrum of TPARF in DriSolv® benzene showing the effects of dimerization with 1 mm cuvette (A) and the absorbance at 445 nm vs. concentration, from the experiment shown above (B).
Figure 2.3: Solution NMR of \([^{15}\text{N}-\text{N5}]\) TPARF in DriSolv\textsuperscript{®} benzene at different concentration. Data were collected at 60 MHz at 25 °C employing a 45 °C excitation pulse followed by 0.25 s acquisition and then a 4 s delay for relaxation. Chemical shift was referenced to 77 ppm for external \(^{15}\text{N}\)-urea in DMSO at 25 °C.
Figure 2.4: Solution NMR spectra of $[^{15}\text{N-N5}]$ TPARF in DriSolv® benzene, alone, with added DBAP or water (5 stoichiometric equivalents). Data were collected at 60 MHz at 25 °C employing a 45 °C excitation pulse followed by 0.25 s acquisition and then a 4 s delay for relaxation. Chemical shift was referenced to 77 ppm for external $^{15}\text{N}$-urea in DMSO at 25 °C.
Figure 2.5: $^{15}$N CP-MAS spectrum of $^{[15}$N-N5] TPARF solid at room temperature with 20 ms acquisitions and 5 s between scans.
Figure 2.6: Comparison of the $^{15}$N CP-MAS spectra of $[^{15}$N-N5] TPARF solid (A), with that in DriSolv® benzene (B). 8 ms CP was used at -60 °C, with 20 ms acquisitions and 5 s between scans.
Figure 2.7: $^{15}$N CP-MAS spectrum of $[^{15}\text{N-N5}]$ TPARF in DriSolv® benzene with DBAP. 8 ms CP was used at -60 °C, with 20 ms acquisitions and 5 s between scans.
Figure 2.8: Comparison of the $^{15}$N CP-MAS spectra of dry $[^{15}$N-N5$]$ TPARF, $[^{15}$N-N5$]$ TPARF in DriSolv® benzene and $[^{15}$N-N5$]$ TPARF in DriSolv® benzene with DBAP, all at 3000 Hz MAS speed. (Central peaks in the red box are enlarged for comparison of the $\delta_{iso}$S.) Other conditions as in Figure 2.6 and 2.7.
Figure 2.9: $^{15}$N CP-MAS spectrum of $[^{15}$N-N5] TPARF in DriSolv® benzene with H$_2$O. 8 ms CP was used at -60 °C, with 20 ms acquisitions and 5 s between scans.
Figure 2.10: Comparison of the $^{15}$N CP-MAS spectra of $[^{15}\text{N-N5}]$ TPARF in benzene and $[^{15}\text{N-N5}]$ TPARF in DriSolv® benzene with 5 and then 15 stoichiometric equivalents of water, at 3000 Hz MAS speed. (Side bands in the blue box are enlarged for comparison of their intensities: before (red) and after (blue) addition of water.)
Figure 2.11: Effects of different H-bonding partner on CSPVs of TPARF.
Chapter Three: DFT Calculation of Flavin Electronics

Density Functional Theory (DFT) was used to calculate the electronic structure of flavin. The calculated NMR chemical shift changes produced by H-bonding to DBAP reproduced the results obtained by SSNMR. Thus, SSNMR can provide an experimental read-out or probe of the flavin electronic structure, and DFT calculations enable us to interpret the SSNMR data in terms of electron density distribution, molecular orbital natures and energy spacing.

3.1 Introduction
3.1.1 Motivation

As chemists, we expect that the wide diversity of reactions mediated by flavins in enzymes rests upon a diversity of flavin electronic structures. Our long-term goal is to test this assumption by comparing the electronic structures of flavins bound in the active sites of enzymes that execute different reactions. Quantum mechanical calculations can provide detailed models for the electronic structures including the natures and relative energies of all the molecular orbitals. However flavins are relatively large molecules and their frontier electrons are highly correlated (22). Thus, while computations have been performed, and published results succeed in correlating calculated electron affinities (EA) or ionization potentials (IP) with those observed (52, 53), these calculations were restricted to simple flavins differing only with respect to small covalent modifications. In contrast, most enzymes modulate the activity of their bound flavins via non-covalent interactions. These are more subtle, and more demanding with respect to their computational treatment as larger basis sets are needed and the optimal geometry may be less well-defined. Rotello has applied

Reproduced with permission from the Journal of the American Chemical Society, and Biochemistry, submitted for publication. Unpublished work copyright 2010 American Chemical Society.
a simple DFT treatment to the H-bonded complex between 2-methylpropyl flavin and diethylaminopyridine, and found that the calculated redistribution of electron density concurs with the changes in EPR hyperfine structure observed for the semiquinone formed upon one-electron oxidation of the naphthalimide (54).

Finally, most DFT calculations rest on the premise that the system will assume its optimal geometry. However in enzymes, portions of the active site (catalytic residues) may be poised in locally non-optimal configurations by interactions with neighboring residues. If the neighboring residues are not included in the calculations, the catalytic residues may relax to optimal geometries that are not catalytic, and moreover non-native (55). Thus, DFT calculations of enzyme active sites confront the difficulty that the size of the system selected for quantum mechanical treatment can have important consequences for the results of the calculations. However we have the privileged situation that we can probe the flavin electronics \textit{in-situ}, experimentally, via SSNMR. Thus, we can perform DFT calculations of flavins in non-covalent complexes, and test the validity of our calculations based on their abilities to replicate the spectroscopic results. This ability to test the faithfulness of the calculations allows that we can undertake more complicated systems by DFT, with confidence that we will be able to distinguish relevant results from results that are computationally valid, but not genuinely representative of the experimental system.

DFT calculations have only been applied to one flavin active site so far, that of photolyase (56), and SSNMR has not been applied to any flavin non-covalent system before. Thus, it is important that we demonstrate that this pair of methods can support one another and address the important biochemical questions that motivate them. We have already shown that SSNMR can perceive effects of H-bonding even at a remote site (Chapter 2). Here we show that DFT can reproduce the experimental results, to within error. Indeed, DFT reproduces the effect of H-bonding to DBAP, and also distinguishes that complex from H-bonding with water. Thus, in the following we describe our experiments to identify computational protocols that provide a realistic description of the flavin and
capture the effects of H-bonding, as detected by SSNMR. The DFT conditions we identify succeed in replicating chemical shift tensor principal value changes produced by H-bonding to TPARF. This is the first time this has been done for a flavin. It paves the way for using SSNMR to identify successful computations of flavins, and thus for using computation to understand flavin reactivity.

3.1.2 Density Functional Theory (DFT)

DFT is a quantum mechanical theory used in physics and chemistry to investigate the electronic structure of molecules and it is a leading method for electronic structure calculations in chemistry and solid-state physics now (57, 58). In many cases the results of DFT calculations for solid-state systems agreed quite satisfactorily with experimental data, especially for the description of ground state properties of metals, semiconductors, and insulators (59-61). The success of DFT not only encompasses standard bulk materials but also complex materials such as proteins and carbon nanotubes (62, 63).

With this theory, the properties of a many-electron system can be determined by using functionals, i.e. functions of another function, which in this case is the spatially-dependent electron density. Hence the name density functional theory comes from the use of functionals of the electron density. The main idea of DFT is to describe an interacting system of fermions via its density and not via its many-body wave function. For N electrons in a solid, which obey the Pauli principle and repel each other via the Coulomb potential, this means that the basic variable of the system depends only on three variables, the spatial coordinates x, y, and z, rather than 3*N degrees of freedom.

3.1.3 Basis Sets

A basis set is a set of functions used to describe the molecular orbitals, which are expanded as linear combinations of atomic orbitals with the weights or coefficients to be determined. The smallest of these are called minimal basis sets, and they are typically composed of the minimum number of basis functions
required to represent all of the electrons on each atom. The largest of these can contain literally hundreds of basis functions on each atom.

A minimum basis set is a single basis function used for each orbital on the free atom. It describes only the most basic aspects of the orbitals using one basis function for each atomic orbital angular momentum component. For example, each atom in the second period of the periodic system (Li - Ne) would have a basis set of five functions (two s functions and three p functions). In the minimum basis set STO (Slater Type Orbital)-NG, N represents the number of Gaussian type orbitals that approximate the Gaussian type orbital (64, 65). Most methods require that a basis set be specified. When using the Gaussian™ package, if no basis set keyword is included in the route section, then the STO-3G basis will be used.

The extended basis sets are basis sets that describe the orbitals in greater detail, and often include multiple functions for each angular momentum component in the electronic configuration description. One way the basis-set can be made larger is to increase the number of basis functions per atom. Split valence basis sets have two or more sizes of basis functions for each valence orbital. For example, the ‘double-zeta’ split valence basis sets use linear combinations of two sizes of functions for each atomic orbital. Split valence basis sets allow orbitals to change size but not shape. For the example of 3-21G, ‘3’ is the number of Gaussian functions summed to describe the inner shell orbital, ‘2’ is the number of Gaussian functions that comprise the first STO of the double zeta, and ‘1’ is the number of Gaussian functions summed in the second STO. Additional numbers, e.g. the right most ‘1’ in 6-311G indicate that it is a triple split valence basis-set and there is a third STO with 1 gaussian. Split valence basis sets are the most commonly used basis-sets.

The most common addition to split valence basis sets is addition of polarization functions (66), denoted by an asterisk, *. Two asterisks, **, indicate that polarization functions are also added to light atoms (hydrogen and helium). Also, the addition of the letters in parenthesis e.g. (d,p) indicate that ‘polarization’ functions have been added. Polarized basis sets add molecular orbitals beyond
what is required for the ground state description of that atom e.g. polarized basis sets add d functions to carbon and p functions to hydrogen. These are auxiliary functions with one additional node. This is an important result when considering accurate representations of bonding between atoms, because the very presence of the bonded atom makes the energetic environment of the electrons spherically asymmetric.

Another common addition to basis sets is the addition of diffuse functions (67), denoted in Pople-type sets by a plus sign, +, and in Dunning-type sets by ‘aug’ (from ‘augmented’). Two plus signs indicate that diffuse functions are also added to light atoms (hydrogen and helium). Diffuse functions are basically larger sized versions of the s and p type functions to allow the orbitals to occupy a larger region of space. These are very shallow Gaussian basis functions, which more accurately represent the ‘tail’ portion of the atomic orbitals, which are distant from the atomic nuclei. These additional basis functions can be important when considering anions and other large, ‘soft’ molecular systems.

To identify the proper basis set for calculation, some of the more often used basis sets were assessed, including the most widely used minimal basis set, STO-3G, the Pople sets, 3-21G, 6-31G, and 6-311G, with the extra functions described previously. All of them are widely used for quantitative results, particularly for organic molecules (68). We sought to identify a basis set that is capable of providing a realistic description, yet not unnecessarily large.

3.1.4 Self-consistent reaction field (SCRF) method

The self-consistent reaction field (SCRF) method is a way of accounting for solvation effects in ab initio calculations. The calculation is performed in the presence of a solvent by placing the solute in a cavity within the solvent reaction field (69). There are quite a number of variations on this method. One point of difference is the shape of the solvent cavity. Various models use spherical cavities, spheres for each atom, or an isosurface of electron density. The second difference is the description of the solute, which could employ a dipole, multipole expansion, or numerical integration over the charge density.
There are many technical details involved in SCRF calculations. The most popular of the SCRF methods is the polarized continuum method (PCM) developed by Tomasi and coworkers (70), which is the default SCRF method. This technique uses a numerical integration over the solute charge density. There are several variations, each of which uses a nonspherical cavity. The generally good results and ability to describe the arbitrary solute make this a widely-used method. However, it is sensitive to the choice of a basis set. Some software implementations of this method may fail for more complex molecules.

3.1.5 NMR Chemical Shielding

Nuclear magnetic resonance (NMR) spectroscopy is a valuable technique for obtaining chemical information. This is because the resonant frequencies of the different signals are very sensitive to changes in the molecular structure and chemical functionality. This same sensitivity makes NMR a difficult case for molecular modeling.

The most important perturbation of the NMR frequency for $^1$H applications of NMR is the 'shielding' effect of the surrounding electrons, which is a local property of each nucleus, and acts on the external magnetic field. In general, this electronic shielding reduces the magnetic field at the nucleus (which is what determines the NMR frequency). As a result the Zeeman energy gap is reduced, and the frequency required to achieve resonance is also reduced. This shift in the NMR frequency due to the ground state electron density distribution in the corresponding molecular orbitals explains why NMR is able to probe the chemical structure of molecules.

Under sufficiently fast magic angle spinning (MAS), or in solution-state NMR, the directionally dependent character of the chemical shielding is removed, leaving only the isotropic chemical shift (71).

NMR chemical shifts can be computed using ab initio methods, which actually compute the shielding tensor. Once the shielding tensors have been computed, the chemical shifts can be determined by subtracting the isotropic
shielding values for the molecule of interest from the standard values (eg. that of TMS for $^1$H or $^{13}$C).

It is extremely important that the shielding tensors be computed for equilibrium geometries with the same method and basis that were used to complete the geometry optimization. It is also important that sufficiently large basis sets are used. The 6-31G(d) basis set should be considered the absolute minimum for reliable results (72). Some studies have used locally-dense basis sets, which have a larger basis on the atom of interest and a smaller basis on the other atoms. In general, these results in only minimal improvement since the spectra are due to interaction between atoms, rather than the electron density around one atom (73).

Computing shielding tensors is difficult because of gauge problems (dependence on the coordinate system’s origin). A number of techniques for correcting this are in use (74, 75). One of the most popular techniques is called GIAO (gauge-independent atomic orbitals). The GIAO technique removes dependence on the coordinate system when computing NMR chemical shifts or optical activity, and it is based on perturbation theory (75-79). This provides a means for computing shielding tensors from DFT wave functions. Density functional theory calculations have shown promise in recent studies where they have been used to calculate chemical shielding tensors for imidazole and carboxylate side chains engaged in H-bonds (80, 81), as well as isotropic shieldings for entire proteins (82, 83), to name just a few biochemical applications. Usually, it is necessary to employ a moderately large basis set with polarization and diffuse functions along with DFT functionals.

3.1.6 Natural Population Analysis (NPA)

Natural bond order analysis (NBO) is the name of a whole set of analysis techniques (84-88). One of these is the natural population analysis (NPA), which analyzes of the electron density distribution in a molecular system based on orthonormal natural atomic orbitals (89). It is for obtaining occupancies, the number of electrons assigned to each atom and orbital, as well as atomic
charges. Rather than using the molecular orbitals directly, NPA uses the natural orbitals. The natural orbitals are those for which the first-order density matrix is diagonal; each will contain some non-integer number of electrons between 0 and 2 \((90, 91)\). This results in a population analysis scheme that is less basis set dependent than the Mulliken scheme \((92, 93)\). However, basis set effects are still readily apparent. NPA is also a popular technique because it is a convenient way of classifying the different orbitals according to familiar types such as s, p-\(\sigma\) and p-\(\pi\).

### 3.1.7 Basis set superposition error (BSSE)

Basis set superposition error (BSSE) lowers the energy of a complex of two molecules with respect to the sum of the individual molecule energies, simply as a result of each molecule's electrons having access to more basis functions in the complex than they do in a description lacking the basis functions of the other molecule. Thus, the complex benefits from extra stabilization, simply due to the greater number of basis functions. Upon subtraction of the energies of the two individual molecules, this produces complexation energies that are too large. In the limit of exact basis sets, there would be no superposition error. The error is also small for minimal basis sets, which do not have functions diffuse enough to describe an adjacent atom. The largest errors occur when using moderate-size basis sets \((94)\).

We corrected for BSSE using a procedure called a counterpoise correction \((95, 96)\). In this procedure, the complex of molecules is first computed. The individual molecule calculations are then performed using all the basis functions from the complex. For this purpose, many \textit{ab initio} software programs contain a mechanism for defining basis functions that are centered at a location which is not on one of the nuclei. The interaction energy is expressed as the energy for the complex minus the individual molecule energies computed in this way.

### 3.1.8 Molecular Orbitals
Chemists use a mathematical function that describes the wave-like behavior of an electron in a molecule. Molecular orbitals are usually constructed by combining atomic orbitals or hybrid orbitals from each atom of the molecule, or other molecular orbitals from groups of atoms. They can be quantitatively calculated using the Hartree-Fock or the Self-Consistent Field method. The highest occupied molecular orbital and lowest unoccupied molecular orbital are often referred to as the HOMO and LUMO, respectively. The difference of the energies of the HOMO and LUMO, termed the band gap, can sometimes serve as a measure of the excitability of the molecule: the smaller the band gap, the more easily it will be excited.

3.1.9 Molecular Dynamics (MD)

Molecular dynamics (MD) is a simulation of the time dependent behavior of a molecular system, including internal motions such as vibrational motion or overall displacement such as Brownian motion to obtain information about accessible conformations and their energies, and so on.

To perform MD, the MM+ force field was designed to describe intermolecular forces and vibrations away from equilibrium (97). At each time point, the energy of the system is calculated, as is the energy gradient with respect to the system's spatial coordinates, internal and overall. The energy gradients yield forces. The internal coordinates are modified in accordance with the calculated actions of these forces over a small interval of time. The choice of time step is important as a time step that is too large will cause atoms to move too far along a given trajectory, thus poorly simulating the motion. A time step that is too small will make it necessary to run more iterations, thus taking longer to run the simulation. One general rule of thumb is that the time step should be one order of magnitude less than the timescale of the shortest motion (vibrational period or time between collisions). This gives a time step on the order of tens of femtoseconds for simulating a liquid of rigid molecules, and tenths of a femtosecond for simulating vibrating molecules.
3.2 Experimental Section

All DFT calculations were performed using the Gaussian 03 package (98), implemented on the University of Kentucky’s N- and X-class HP supercomputer. Lumiflavin (LF) was used to model the flavin ring system. LF simplifies the computations because it has a methyl group at the N10 position instead of the functionalized ribityl side chain of TPARF. However, it retains the characteristic flavin reactivity (99). Coordinates of optimized models are provided at the end of this chapter (Table 3.1 and 3.2).

Chemical shieldings were calculated using the GIAO method implemented in Gaussian 03 (98). Calculated $^{15}$N shieldings were converted to calculated chemical shifts by subtracting the former from the shielding of liquid ammonia of 244.6 ppm (100).

NPA charges were used to calculate the distribution of electron density in optimized geometries (89). GFPrint was used as an additional keyword to generate the current basis set and density fitting basis set in tabular form in output file. The energies of H-bonds between LF and DBAP were calculated from the energy of the complex minus the energies of the participants, upon counterpoise correction for the basis set superposition error (96, 101).

3.3 Results

I am able to duplicate our SSNMR results with DFT calculations, thus obtaining descriptions of electronic consequences of H-bond formation. The substantial responses we observe in the CSPVs, and their agreement with calculations confirm that $^{15}$N chemical shift principal values are sensitive and informative tools for perceiving and distinguishing non-covalent interactions, such as those used by proteins to modulate flavin reactivity.

3.3.1 Geometry Optimization

The geometry of a molecule determines many of its physical and chemical properties. This is why it is very important that we use the correct geometry of a molecule when running computations.
For all the geometry optimizations in this work, I specified a beginning geometry retain coordinates drawn from prior work by J. D. Walsh (22), either in the form of internal coordinates (a Z-matrix), Cartesian coordinates, or in mixed format. I also specified a basis set. The Gaussian package then computes the energy and the gradient at that point, determines if it has reached a stopping point (convergence), and if not modifies the geometry based on the size of the gradient in such a direction as to lower the energy of the system. New integrals are calculated, new self-consistent field calculations are done, and a new energy and gradient are calculated. These steps are repeated until the calculation converges i.e. finds a stationary point at a minimum energy. The coordinates of each atom are then written to a coordinate file which constitutes the energy-optimized geometry.

Initially, geometry optimizations of oxidized LF, DBAP and the complex between LF and DBAP were done with each of B3PW91 (102) and B3LYP (103) and a range of different basis sets. Plots of the energy of LF or the LF·DBAP complex versus the number of basis functions used in the optimization reveal that larger basis sets allow attainment of lower energies, to a point, called basis set saturation (Figure 3.1). The value of this lowest attainable energy also depends on the choice of functional. Thus we identified functionals best able to minimize the system energy (and thus approach the global minimum which characterizes the real system).

Because NMR chemical shifts represent very small energies, the accuracy of their calculation depends on details of the procedures used. Therefore, we began by evaluating different density functionals and levels of theory, based on their convergence to low energy in geometry optimizations and abilities to replicate experiment. Figure 3.1 shows the energies obtained upon geometry optimization of lumiflavin (LF) and the complex with DBAP. The data reveal that basis sets smaller than 6-311G(d,p) produce elevated energies and therefore constitute poor approximations to the expected energy minimum. Hence only basis sets of 6-311G(d,p) (414 primitives for LF) and larger were evaluated for performance in NMR calculations.
For each of the more successful functionals, basis sets deemed to have reached the point of diminishing returns with respect to energy were additionally tested for their ability to reproduce experimental CSPVs. Thus, we conclude that the 6-311+G(d,p) represents a minimum acceptable basis set for geometry optimization. However calculation of chemical shifts may be more sensitive and require a better description of the system still.

3.3.2 NMR Chemical Shielding

For calculations of NMR chemical shifts, the GIAO method was chosen based on its success in comparable systems (22). Both B3PW91 and B3LYP with STO-3G, 3-21G, 6-31G, 6-311G, 6-31G(d), 6-31G(d,p), 6-311G(d), 6-311G(d,p), 6-311+G(d,p), 6-31G(2d,2p), 6-311++G(d,p), 6-311G(2d,2p), 6-311+G(2d,2p) and 6-311++G(2d,2p) basis sets were used (Figure 3.2). For each basis set, each the two functionals was used for geometry optimization. Then, each resulting structure was used as the basis for NMR shielding calculation with each of the two functional, yielding a total of four shielding values for each basis. Use of the B3PW91 functional with the 6-311++G(2d,2p) basis set for both geometry optimization and GIAO chemical shifts calculations best reproduced the experimental $\delta_{iso}$ of TPARF’s N5 in benzene, of 342.01 ppm. Therefore, reported calculations employed B3PW91 and 6-311++G(2d,2p).

The coordinates of LF and LF·DBAP geometry-optimized with this basis set and functional are included at the end of this chapter.

3.3.3 Molecular Dynamics (MD)

To assess the uncertainty associated with our computational results, MD of LF was executed in HyperChem 8.0. I took the MM+ geometry-optimized structure of lumiflavin (LF) and allowed it to equilibrate for 2 ns at a temperature of 213 K (-60°C) with the MM+ force field. A 2 fs step size was used (Figure 3.3).

An expansion of the 1 ps including the global energy minimum shows that oscillations nonetheless persist with a period of 0.95 ps and an amplitude of 64 cal/mol even after 2 ns equilibration. Thus we estimate that our calculated
energies should be considered to be associated with uncertainties on the order of 130 J/mol. The structure with the global minimum energy was extracted, along with that of the following peak energy, and eight structures from evenly-spaced time points in between. The global minimum structure and the following peak-energy structure are overlaid in Figure 3.4. The principle difference is the rotational angle of the methyl groups at positions 7 and 8.

The ten frames described above from 1 ps of equilibrated MD were used as input for GIAO chemical shielding calculations in Gaussian (without geometry re-optimization). The standard deviations of the obtained CSPVs serve as our estimates of the uncertainties associated with NMR chemical shift calculations. The standard deviations were 4.9, 1.4, 0.68 and 1.8 ppm for $\delta_{11}$, $\delta_{22}$ and $\delta_{33}$, and $\delta_{\text{iso}}$, respectively. These values therefore provide estimates of the uncertainty associated with our chemical shift calculations. Similarly, we estimated the uncertainty associated with NPA charges based on deviations between those calculated on different structures from the equilibrated MD.

The calculated $\delta_{\text{iso}}$ for N5 of LF was calculated to be 363 ± 2 ppm, 21 ppm larger than the experimental value. The CSPVs were 723 ± 5 ppm, 399 ± 1 ppm and -33 ± 1 ppm, respectively. Of these, the $\delta_{11}$, in particular, is much larger than the experimental value of 676.32 ppm. Such a discrepancy was also observed in calculations employing a range of different basis sets and functionals (22). The prior studies considered only LF in vacuo, whereas solvents have been shown experimentally to decrease $\delta_{\text{iso}}$ by up to 20 ppm, via effects that are expected affect predominantly $\delta_{11}$ (19). In the current work, we intentionally chose a solvent, benzene, which should interact relatively weakly with the flavin ring system, in an effort to minimize solvent effects. Indeed, Witanowski reports that $\delta_{\text{iso}}$ of pyridine need only be corrected by 3.4 ppm for benzene, vs. use of cyclohexane, corrected by 7.6 ppm, vs. use of chloroform, the solvent traditionally used for NMR studies of flavins (104).

3.3.4 Flavin dimers
The samples for both optical study and solution state NMR study are much less concentrated than the one for SSNMR study. Concentrated TPARF might form dimers in benzene. Thus, we also addressed dimerization between flavin moieties, computationally. Because our solution NMR studies indicated that at the very high concentrations of flavins required for this work, the population of flavin dimers is significant, we generated the four most plausible H-bonding coplanar flavin dimers and subjected them to geometry optimization. Three achieved low energy minima (Figure 3.5). Counterpoise-corrected calculations of the binding energies indicated that one dimer would predominate in a Boltzmann-weighted population, but that the other two would nonetheless contribute 3% or more of the population. Counterpoise corrected dimer energies were done without SCRF solvent as this is not implemented in Gaussian 03. The Boltzmann-weighted SCRF-corrected isotropic chemical shift of LF dimer differed from the SCRF-corrected monomer LF by having a 1 ppm lower $\delta_{\text{iso}}$. This agrees with the experimental findings, supporting the utility of the calculations. The calculated effects of LF dimerization on the CSPV are a change of -2 ppm to $\delta_{11}$ (and changes of -1 and 0 ppm to $\delta_{22}$ and $\delta_{33}$, respectively). Because these corrections are all small relative to error, and are in agreement with our experimental data, we conclude that our Boltzmann-weighted dimer is a good model for the state of TPARF in our high-concentration solutions. Thus, the chemical shifts calculated based on the Boltzmann population of dimers are used as the reference state for the comparisons that follow, and we note that despite our earnest efforts to devise a system in which we could study formation of H-bonds de-novo, the inevitability of favourable interactions between solution components that could not be diluted has placed us in the situation of assessing the effects of forming new H-bonded complexes upon disruption of flavin dimers.

3.3.5 Solvent Effects

Most of the modeling methods model gas-phase molecular behavior, in which it is reasonable to assume that there is no interaction with other molecules.
However, most laboratory chemistry is done in solution where the interaction between the species of interest and the solvent is not negligible.

To evaluate the possible effects of benzene as a solvent, optimization with NMR GIAO calculation were done by Gaussian with and without SCRF. More recently, it has become possible to implement a bulk dielectric correction to account for solvent in GIAO calculations performed by Gaussian. When this was done, $\delta_{\text{iso}}$ decreased by 2 ppm and $\delta_{11}$ decreased by 7 ppm (in addition to changes of +2 and +1 ppm on $\delta_{22}$ and $\delta_{33}$). These effects are qualitatively similar to the -12 and -36 ppm correction employed previously to account for solvent, but roughly 5 times smaller, consistent with benzene's being a solvent not compete H-bonding, which is better than those used in previous work (19).

For flavin monomer, there is no significant difference as upon adding SCRF. Also, flavin dimers give very similar NMR GIAO results as flavin dimer with consideration of SCRF.

3.3.6 Computations of TPARF's complex with DBAP

The observed changes in TPARF's N5's CSPV upon complexation with DBAP could be replicated computationally. The geometries of oxidized lumiflavin (LF), DBAP, the complex of LF and DBAP was optimized using density functional theory (DFT) and the B3PW91 functional with the 6-311++G(2d,2p) basis set.

We were expecting the same or at least similar results from calculation as we got from experiments. However, the calculated results are different as the experimental results got by SSNMR. These differences are likely due to imperfect accounting for solvent dielectric, temperature effects, or chemical shift referencing. Such issues should not affect comparisons between calculated chemical shifts. Indeed the changes in chemical shifts produced upon H-bond formation agreed very well with changes in experimental values. Upon H-bonding with DBAP, the calculated $\delta_{\text{iso}}$ of N5 changed by 5.6 ppm. The calculated CSPVs changed 14, 4 and -1 ppm, respectively. Besides agreeing within error with our SSNMR experiments, these calculations confirm that $\delta_{11}$ is considerably more responsive to H-bonding with DBAP than the other CSPVs, or $\delta_{\text{iso}}$. Although the
experimental change in $\delta_{11}$ is a little smaller than the calculated one, we do not attribute this to incomplete replacement of the flavin dimer by TPARF•DBAP complexes, as the $\delta_{22}$ and $\delta_{33}$ experimental changes are either larger or different from the calculated changes. Nonetheless, the calculations reproduce the relative sizes of the CSPVs from most positive ($\delta_{11}$) to least ($\delta_{33}$) and produces a very similar $\delta_{iso}$.

Our calculated change in $\delta_{iso}$ due to H-bonding with DBAP deviates from the experimental value by 1.1 ppm, which at an NMR frequency of 40 MHz corresponds to an energy of 44 Hz. In comparison, the disparity between observed and calculated changes in vibrational frequencies was on the order of 3 cm$^{-1} = 9 \times 10^{10}$ Hz for neutral lumazine (105). Thus, our calculations do very well in reproducing experimental changes. Since our purpose is to understand changes in flavin reactivity brought about by interactions with proteins, it is the changes in CSPVs produced by H-bonding that are important.

The qualitative as well as quantitative agreement between our calculations and experiments indicates that the calculations can provide a realistic description of the flavin. Moreover, because none of DBAP’s H-bonds, nor those of the flavin dimer involve N5, the changes we observe in N5’s CSPV must reflect changes in the flavin electronics. Therefore, the calculations succeed in capturing the changes in flavin electronics produced by complexation with DBAP, and we can use the calculated electronic structures to understand the effect of DBAP binding on reactivity. The natural population analysis (NPA) distributions of electron density in LF and LF•DBAP indicate that the most positive atom is C2 consistent with its double bond to O and two bonds to N, and the most negative atoms are N3, N1 and O2 consistent with their electronegative natures (Figure 3.6). Complexation with DBAP did not significantly change this very strong polarization of bonds within the uracil ring. However the flavin ring propagates the effects of H-bonding to more distant sites. Comparing the effects of complexation on the diazabutadiene system (N5, C4a, C10a, N1) with the effects on sites directly involved in H-bonding (C2, N3, C4) and sites at the far end of the flavin (C7, C7a, C8, C8a), we find changes in excess charge of $0.10 \pm 0.04$ in the diazabutadiene
system whereas the changes at sites directly involved in H-bonding, and at the far end of the flavin are both comparable to the uncertainty. For N5 plus C4a, the two sites associated with most flavin reactivity, the change in net charge upon binding to DBAP is $0.12 \pm 0.04$ (Figure 3.7). These results are consistent with the measured 100 mV increase in TPARF’s reduction midpoint potential upon complexation with DBAP.

The counterpoise-corrected energy of complex formation is -$27$ kJ/mol (Table 3.4). Considering that this does not take into account entropic contributions or solvation effects and had to be calculated without SCRF, this value is consistent with the standard-state Gibbs free energy of -$19.2$ kJ/mol calculated from the 420 μM dissociation constant for TPARF and DBAP in benzene (37). These values imply H-bond energies on the order of $10$ kJ/mol each, consistent with the good bond geometries and distances of 2.9 - 3.0 Å between participating N or O atoms (Figure 3.8). Thus, calculated properties of the complex are internally consistent and consistent with experimental observations. Nonetheless, the natures of the flavin-based HOMO and LUMO do not change substantially upon complexation of DBAP, nor does the energy separating them (Figure 3.9). Indeed, the five HOMOs and the five LUMOs are qualitatively the same for LF•DBAP as for LF, although HOMO-3 and HOMO-4 exchange positions in the energy-order upon complexation with either DBAP or water (below). Thus, we do not expect a large change in the nature of the reactivity of the oxidized state, consistent with experiment. These comparisons also indicate that SSNMR is able to detect even very subtle perturbations that are not obvious at the level of the frontier MOs.

### 3.3.7 Computation of TPARF’s complex with water

DBAP forms a chelate-like interaction with TPARF that favors formation of a unique complex. In contrast, H-bonds with H$_2$O are much less site specific, as geometry optimization identified at least 4 possible optimized configurations of LF with one H$_2$O bound (Figure 3.10). Structures were also optimized with up to 8 H$_2$O molecules employing many different starting positions for the water
molecules (Figures 3.11 to 3.14). The general lack of agreement between experiment and CSPV calculated from any one configuration, and the fact that calculated $\delta_{11}$ values from individual water-LF complexes were scattered widely above and below the experimental value indicates that the sample of TPARF and $H_2O$ is not well described by any one configuration of H-bonded $H_2O$.

The lack of improvement as the number of water molecules in the calculation was increased could be explained the fact that when more waters were present they tended to interact among themselves rather than interact with all the functionalities of the flavin, thus we observed a tendency to partition, rather than solvate the LF (Figure 3.13 and 3.14). The calculations of individual one-water configurations demonstrated that when $H_2O$ H-bonded directly with N5, both $\delta_{11}$ and $\delta_{iso}$ decreased a lot (-50 ppm and -16 ppm, respectively, Figure 3.10) whereas when $H_2O$ formed H-bonds with atoms other than N5 $\delta_{11}$ and $\delta_{iso}$ increased (by average increases of 4 ppm and 3 ppm respectively). The same trends were observed with larger numbers of waters too. The fact that the experimental changes of -5 and 0.5 were intermediate between these two cases supports an approach that combines the effects of multiple different H-bonding configurations, at once. A Boltzmann-weighted average of CSPVs produced by the different single-water configurations improved the agreement with experiment (Table 3.3). Thus, although more sophisticated models can readily be constructed, our data do not provide motivation or justification for doing so, with the current experimental uncertainties.

Counterpoise-corrected energies of binding between LF and one water ranged from -25 to -30 kJ/mol. Water has greater freedom than DBAP to adopt the position that optimizes individual H-bonds. The competition between various sites on LF that have comparable affinities for water conspires with cancellation of contributions from the different CSPVs to produce a small net effect on the $\delta_{iso}$. However, more responsive to water, the experimental effect on $\delta_{11}$ is opposite in sign to the effect produced by DBAP binding. Thus, our necessarily simple treatment of competing statistical associations succeeds in reproducing the qualitatively distinct effect of water, and adjustment of the number of waters
considered within a reasonable range obtains quantitative agreement within error, too.

Even for this more complicated case, our experiments demonstrate that SSNMR provides added sensitivity to non-covalent interactions that will greatly advance our ability to detect and evaluate interactions in proteins that tune flavin reactivity.

3.4 Discussion

The DFT calculations reveal that the small effects of possible flavin dimerization also extend to the CSPVs. The changes in CSPVs calculated to occur for formation of each of the three most plausible dimers were all less than our error, as was their Boltzmann-weighted average.

In accordance with our solution results and the fact that our solutions do not form glasses upon freezing, we consider the flavin dimer population as the state present prior to addition of DBAP or H₂O. However, given the small effects dimerization has on CSPVs, we have acceptable agreement with experiment regardless of whether we consider that DBAP binds to TPARF monomers or that TPARF dimers dissociate and form TPARF•DBAP complexes.

\[ \text{TPARF monomer} + \text{DBAP} \rightarrow \text{TPARF} \cdot \text{DBAP} \]

or

\[ \text{TPARF dimer} \rightarrow \text{TPARF monomer} + \text{DBAP} \rightarrow \text{TPARF} \cdot \text{DBAP} \]

NMR CSPVs advance the possibility of understanding flavin electronic structure and reactivity via theoretical methods, as the CSPVs can be calculated using the same packages as are also used to optimize geometries and calculate electronic properties. Comparison of the calculated CSPVs with the experimental values then provides a means of validating the calculated structures and electronic descriptions. We are able to duplicate our observed large changes in \(^{15}\text{N}\) CSPVs with DFT calculations, indicating that the latter provide realistic descriptions of electronic consequences of H-bond formation for our flavin. Spectroscopically-validated calculations not only permit detailed insight into the protein induced variations in flavin activity, they also can also greatly increase the value of the NMR results. Calculations allow us to interpret CSPV changes in
terms of changes in electron density distribution. For the current case of $\delta_{11}$ of N5, a large decrease in conjunction with a small increase in $\delta_{33}$ can now be taken as evidence for acceptance of a H-bond, at least (other interactions may in future be found to produce similar effects).

H-bonds donated to N1, O2 and O4 have been reported to draw electron density out of the flavin and make it a better oxidant;\cite{44} our spectroscopically-validated calculations find that the redox-active diazabutadiene becomes less net-negative, consistent with the observed increase in reduction midpoint potential, even though electron density is little changed at the actual sites of H-bonding. In addition, our spectroscopically-validated calculations yield a binding energy for DBAP and LF that is very close to the experimental free energy of binding. Thus, our calculations are consistent with two different types of activity, indicating that they will be useful for understanding reactivity, and possibly even predicting it.

In the current two cases, electronic effects were subtle, likely because the complexes were both favorable and relatively weak. Nevertheless, significant changes in CSPVs indicate that there were repercussions for the flavin electronics, as we could detect the effects of H-bonds formed at remote positions. CSPVs will aid us in identifying correct description of interactions in effect in active sites. For FMN and FAD bound in proteins, where binding can be very tight and exploit the ribose and phosphate for affinity, the flavin-binding site can bring more energy to bear in distorting the flavin and its electronic structure, and larger solution NMR effects are observed than those described here. Thus, larger effects are expected in SSNMR as well, and those we document here on models underestimate what is possible, and thereby provide a particularly stringent proof of principle that SSNMR will be able to perceive perturbations applied by proteins.
<table>
<thead>
<tr>
<th>Center Number</th>
<th>Atomic Number</th>
<th>Atomic Type</th>
<th>Coordinates (Angstrom)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>0</td>
<td>0.000000</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>0</td>
<td>0.000000</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>0</td>
<td>1.196957</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>0</td>
<td>2.419404</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>0</td>
<td>2.402161</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>0</td>
<td>1.225468</td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>0</td>
<td>3.618931</td>
</tr>
<tr>
<td>8</td>
<td>6</td>
<td>0</td>
<td>2.424939</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>0</td>
<td>3.604188</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0</td>
<td>-0.947855</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>0</td>
<td>3.362829</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>0</td>
<td>5.655059</td>
</tr>
<tr>
<td>13</td>
<td>7</td>
<td>0</td>
<td>2.417313</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>0</td>
<td>4.798255</td>
</tr>
<tr>
<td>15</td>
<td>7</td>
<td>0</td>
<td>3.616169</td>
</tr>
<tr>
<td>16</td>
<td>7</td>
<td>0</td>
<td>1.231572</td>
</tr>
<tr>
<td>17</td>
<td>6</td>
<td>0</td>
<td>4.929463</td>
</tr>
<tr>
<td>18</td>
<td>8</td>
<td>0</td>
<td>3.673676</td>
</tr>
<tr>
<td>19</td>
<td>8</td>
<td>0</td>
<td>6.007250</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
<td>0</td>
<td>-1.303024</td>
</tr>
<tr>
<td>21</td>
<td>1</td>
<td>0</td>
<td>-1.387756</td>
</tr>
<tr>
<td>22</td>
<td>1</td>
<td>0</td>
<td>-1.387756</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
<td>0</td>
<td>-2.149997</td>
</tr>
<tr>
<td>24</td>
<td>6</td>
<td>0</td>
<td>1.237200</td>
</tr>
<tr>
<td>25</td>
<td>1</td>
<td>0</td>
<td>0.727073</td>
</tr>
<tr>
<td>26</td>
<td>1</td>
<td>0</td>
<td>0.727073</td>
</tr>
<tr>
<td>27</td>
<td>1</td>
<td>0</td>
<td>2.257370</td>
</tr>
<tr>
<td>28</td>
<td>6</td>
<td>0</td>
<td>-0.017331</td>
</tr>
<tr>
<td>29</td>
<td>1</td>
<td>0</td>
<td>-0.596100</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>0</td>
<td>0.232020</td>
</tr>
<tr>
<td>31</td>
<td>1</td>
<td>0</td>
<td>-0.596100</td>
</tr>
</tbody>
</table>

Table 3.1: Cartesian coordinates of LF geometry optimized with B3PW91 and 6-311++G(2d,2p).
<table>
<thead>
<tr>
<th>Center Number</th>
<th>Atomic Number</th>
<th>Atomic Type</th>
<th>Coordinates (Angstrom)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
<td>0</td>
<td>7.457504</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>0</td>
<td>7.635380</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>0</td>
<td>6.641772</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>0</td>
<td>5.459562</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>0</td>
<td>5.293179</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>0</td>
<td>6.282799</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0</td>
<td>8.228706</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0</td>
<td>8.548126</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>0</td>
<td>6.790675</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0</td>
<td>4.384681</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>0</td>
<td>6.134398</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>0</td>
<td>4.423431</td>
</tr>
<tr>
<td>13</td>
<td>6</td>
<td>0</td>
<td>3.768329</td>
</tr>
<tr>
<td>14</td>
<td>6</td>
<td>0</td>
<td>4.062491</td>
</tr>
<tr>
<td>15</td>
<td>6</td>
<td>0</td>
<td>2.565453</td>
</tr>
<tr>
<td>16</td>
<td>6</td>
<td>0</td>
<td>3.172818</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>0</td>
<td>4.974146</td>
</tr>
<tr>
<td>18</td>
<td>6</td>
<td>0</td>
<td>1.990647</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>0</td>
<td>3.358400</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
<td>0</td>
<td>3.265325</td>
</tr>
<tr>
<td>21</td>
<td>6</td>
<td>0</td>
<td>2.552895</td>
</tr>
<tr>
<td>22</td>
<td>6</td>
<td>0</td>
<td>4.645232</td>
</tr>
<tr>
<td>23</td>
<td>6</td>
<td>0</td>
<td>3.206478</td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>0</td>
<td>1.482328</td>
</tr>
<tr>
<td>25</td>
<td>6</td>
<td>0</td>
<td>5.300944</td>
</tr>
<tr>
<td>26</td>
<td>1</td>
<td>0</td>
<td>5.215777</td>
</tr>
<tr>
<td>27</td>
<td>6</td>
<td>0</td>
<td>4.581906</td>
</tr>
<tr>
<td>28</td>
<td>1</td>
<td>0</td>
<td>2.635487</td>
</tr>
<tr>
<td>29</td>
<td>1</td>
<td>0</td>
<td>6.374883</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>0</td>
<td>5.089321</td>
</tr>
<tr>
<td>31</td>
<td>6</td>
<td>0</td>
<td>4.417379</td>
</tr>
<tr>
<td>32</td>
<td>1</td>
<td>0</td>
<td>4.509776</td>
</tr>
<tr>
<td>33</td>
<td>1</td>
<td>0</td>
<td>4.607263</td>
</tr>
<tr>
<td>34</td>
<td>6</td>
<td>0</td>
<td>2.976606</td>
</tr>
<tr>
<td>35</td>
<td>8</td>
<td>0</td>
<td>2.483114</td>
</tr>
<tr>
<td>36</td>
<td>7</td>
<td>0</td>
<td>2.195850</td>
</tr>
<tr>
<td>37</td>
<td>1</td>
<td>0</td>
<td>1.192839</td>
</tr>
<tr>
<td>38</td>
<td>7</td>
<td>0</td>
<td>1.681351</td>
</tr>
<tr>
<td>39</td>
<td>7</td>
<td>0</td>
<td>1.036442</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>40</td>
<td>1</td>
<td>0</td>
<td>0.051558</td>
</tr>
<tr>
<td>41</td>
<td>6</td>
<td>0</td>
<td>1.214748</td>
</tr>
<tr>
<td>42</td>
<td>8</td>
<td>0</td>
<td>0.249452</td>
</tr>
<tr>
<td>43</td>
<td>6</td>
<td>0</td>
<td>2.591083</td>
</tr>
<tr>
<td>44</td>
<td>1</td>
<td>0</td>
<td>3.255987</td>
</tr>
<tr>
<td>45</td>
<td>1</td>
<td>0</td>
<td>2.409188</td>
</tr>
<tr>
<td>46</td>
<td>6</td>
<td>0</td>
<td>-1.563434</td>
</tr>
<tr>
<td>47</td>
<td>6</td>
<td>0</td>
<td>-2.134607</td>
</tr>
<tr>
<td>48</td>
<td>6</td>
<td>0</td>
<td>-3.550036</td>
</tr>
<tr>
<td>49</td>
<td>6</td>
<td>0</td>
<td>-3.795349</td>
</tr>
<tr>
<td>50</td>
<td>6</td>
<td>0</td>
<td>-5.762540</td>
</tr>
<tr>
<td>51</td>
<td>6</td>
<td>0</td>
<td>-6.110672</td>
</tr>
<tr>
<td>52</td>
<td>6</td>
<td>0</td>
<td>-6.774266</td>
</tr>
<tr>
<td>53</td>
<td>6</td>
<td>0</td>
<td>-7.459174</td>
</tr>
<tr>
<td>54</td>
<td>6</td>
<td>0</td>
<td>-8.101027</td>
</tr>
<tr>
<td>55</td>
<td>1</td>
<td>0</td>
<td>-6.461994</td>
</tr>
<tr>
<td>56</td>
<td>6</td>
<td>0</td>
<td>-8.444526</td>
</tr>
<tr>
<td>57</td>
<td>1</td>
<td>0</td>
<td>-7.752555</td>
</tr>
<tr>
<td>58</td>
<td>7</td>
<td>0</td>
<td>-2.866028</td>
</tr>
<tr>
<td>59</td>
<td>7</td>
<td>0</td>
<td>-1.253711</td>
</tr>
<tr>
<td>60</td>
<td>7</td>
<td>0</td>
<td>-5.105428</td>
</tr>
<tr>
<td>61</td>
<td>7</td>
<td>0</td>
<td>-4.472327</td>
</tr>
<tr>
<td>62</td>
<td>8</td>
<td>0</td>
<td>-0.642665</td>
</tr>
<tr>
<td>63</td>
<td>8</td>
<td>0</td>
<td>-1.781845</td>
</tr>
<tr>
<td>64</td>
<td>1</td>
<td>0</td>
<td>-0.258741</td>
</tr>
<tr>
<td>65</td>
<td>6</td>
<td>0</td>
<td>-9.878633</td>
</tr>
<tr>
<td>66</td>
<td>1</td>
<td>0</td>
<td>-10.476186</td>
</tr>
<tr>
<td>67</td>
<td>1</td>
<td>0</td>
<td>-10.319481</td>
</tr>
<tr>
<td>68</td>
<td>1</td>
<td>0</td>
<td>-9.982143</td>
</tr>
<tr>
<td>69</td>
<td>6</td>
<td>0</td>
<td>-9.160396</td>
</tr>
<tr>
<td>70</td>
<td>1</td>
<td>0</td>
<td>-9.728790</td>
</tr>
<tr>
<td>71</td>
<td>1</td>
<td>0</td>
<td>-9.878398</td>
</tr>
<tr>
<td>72</td>
<td>1</td>
<td>0</td>
<td>-8.719133</td>
</tr>
<tr>
<td>73</td>
<td>1</td>
<td>0</td>
<td>-6.088920</td>
</tr>
<tr>
<td>74</td>
<td>6</td>
<td>0</td>
<td>-5.436208</td>
</tr>
<tr>
<td>75</td>
<td>1</td>
<td>0</td>
<td>-5.936562</td>
</tr>
<tr>
<td>76</td>
<td>1</td>
<td>0</td>
<td>-4.511518</td>
</tr>
</tbody>
</table>

**Table 3.2:** Cartesian coordinates of the complex of LF with DBAP, geometry optimized with B3PW91 and 6-311++G(2d,2p).
Table 3.3: Calculated Chemical Shift Principal Values (CSPVs) of N5 of LF solid, N5 of LF solid in DriSolv® benzene, N5 of LF solid in DriSolv® benzene with DBAP and N5 of LF solid in DriSolv® benzene with H2O.

<table>
<thead>
<tr>
<th></th>
<th>$\delta_{11}$ (ppm)</th>
<th>$\delta_{22}$ (ppm)</th>
<th>$\delta_{33}$ (ppm)</th>
<th>$\delta_{\text{iso}}=(\delta_{11}+\delta_{22}+\delta_{33})/3$ (ppm)</th>
<th>span=$\delta_{11}-\delta_{33}$ (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LF Solid</td>
<td>730.23</td>
<td>396.41</td>
<td>-31.70</td>
<td>364.98</td>
<td>761.93</td>
</tr>
<tr>
<td>LF Solid in DriSolv® Benzene</td>
<td>722.74</td>
<td>398.73</td>
<td>-32.93</td>
<td>362.85</td>
<td>755.67</td>
</tr>
<tr>
<td>LF dimmers in DriSolv® Benzene</td>
<td>722.55</td>
<td>398.03</td>
<td>-32.80</td>
<td>362.59</td>
<td>755.35</td>
</tr>
<tr>
<td>LF Solid in DriSolv® Benzene with DBAP</td>
<td>733.64</td>
<td>400.52</td>
<td>-33.34</td>
<td>366.94</td>
<td>766.98</td>
</tr>
<tr>
<td>LF Solid in DriSolv® Benzene with 1:1 H2O</td>
<td>712.63</td>
<td>397.93</td>
<td>-32.22</td>
<td>359.45</td>
<td>744.85</td>
</tr>
</tbody>
</table>
Table 3.4: Counterpoise calculation results for complex of LF and DBAP, LF and one H$_2$O.

<table>
<thead>
<tr>
<th>File name</th>
<th>Method / Basis Set</th>
<th>Energy of LumiFlavin and DBAP</th>
<th>Corrected energy of complex</th>
<th>H-bond energy (Hartree)</th>
<th>H-bond energy (kJ/mol)</th>
<th>BSSE energy (Hartree)</th>
<th>BSSE energy (kJ/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LD-CP009</td>
<td>RB3LYP/6-311+G(d,p)</td>
<td>-1999.0703</td>
<td>-1999.0823</td>
<td>0.01198</td>
<td>31.453</td>
<td>0.001706</td>
<td>4.4784</td>
</tr>
<tr>
<td>LD-CP012</td>
<td>RB3LYP/6-311G(2d,2p)</td>
<td>-1999.1063</td>
<td>-1999.118</td>
<td>0.01168</td>
<td>30.6608</td>
<td>0.005842</td>
<td>15.338</td>
</tr>
<tr>
<td>LD-CP013</td>
<td>RB3LYP/6-311+G(2d,2p)</td>
<td>-1999.1378</td>
<td>-1999.1493</td>
<td>0.01148</td>
<td>30.1442</td>
<td>0.000993</td>
<td>2.606</td>
</tr>
<tr>
<td>LD-CP014</td>
<td>RB3LYP/6-311++G(2d,2p)</td>
<td>-1999.138</td>
<td>-1999.1495</td>
<td>0.01151</td>
<td>30.2196</td>
<td>0.000983</td>
<td>2.5817</td>
</tr>
<tr>
<td>LD-CP023</td>
<td>RB3PW91/6-311+G(d,p)</td>
<td>-1998.2777</td>
<td>-1998.2884</td>
<td>0.01071</td>
<td>28.1132</td>
<td>0.002072</td>
<td>5.4397</td>
</tr>
<tr>
<td>LD-CP026</td>
<td>RB3PW91/6-311G(2d,2p)</td>
<td>-1998.3182</td>
<td>-1998.3288</td>
<td>0.01062</td>
<td>27.8851</td>
<td>0.005156</td>
<td>13.537</td>
</tr>
<tr>
<td>LD-CP027</td>
<td>RB3PW91/6-311+G(2d,2p)</td>
<td>-1998.3445</td>
<td>-1998.3548</td>
<td>0.01033</td>
<td>27.1132</td>
<td>0.001129</td>
<td>2.9653</td>
</tr>
<tr>
<td>LD-CP028</td>
<td>RB3PW91/6-311++G(2d,2p)</td>
<td>-1998.3447</td>
<td>-1998.355</td>
<td>0.01035</td>
<td>27.1646</td>
<td>0.001137</td>
<td>2.9863</td>
</tr>
<tr>
<td>LW-CP112</td>
<td>RB3PW91/6-311++G(2d,2p)</td>
<td>-948.4646</td>
<td>-948.4741</td>
<td>0.009488</td>
<td>24.9102</td>
<td>0.0005659</td>
<td>1.4857</td>
</tr>
<tr>
<td>LW-CP114</td>
<td>RB3PW91/6-311++G(2d,2p)</td>
<td>-948.4646</td>
<td>-948.4748</td>
<td>0.0102</td>
<td>26.7795</td>
<td>0.0007911</td>
<td>2.0772</td>
</tr>
<tr>
<td>LW-CP115</td>
<td>RB3PW91/6-311++G(2d,2p)</td>
<td>-948.4646</td>
<td>-948.4745</td>
<td>0.009827</td>
<td>25.8005</td>
<td>0.0006658</td>
<td>1.7481</td>
</tr>
<tr>
<td>LW-CP123</td>
<td>RB3PW91/6-311++G(2d,2p)</td>
<td>-948.4646</td>
<td>-948.4762</td>
<td>0.01154</td>
<td>30.2987</td>
<td>0.000831</td>
<td>2.1815</td>
</tr>
</tbody>
</table>
Figure 3.1: DFT (B3PW91 and B3LYP) energies versus Basis Set Size. (A) LF (B) Complex of LF and DBAP.
LumiFlavin (experiment = 342.0 ppm)

Isotropic Chemical Shift of N5 vs. NH$_3$ (ppm)

Number of Basis Functions

LumiFlavin NMR GIAO Calculations

Geometry Opt + NMR GIAO
- B3LYP+B3LYP
- B3PW91+B3PW91
- B3LYP+B3PW91
- B3PW91+B3LYP

(A) Lumiflavin (experiment = 342.0 ppm)
Figure 3.2: Isotropic chemical shifts of N5 versus basis size and choice of functional used for each of GIAO and geometry optimization calculations.
**Figure 3.3:** Energy vs. time trajectory of MD calculation beginning with the MM+ geometry-optimized structure. Entire 2 ns trajectory, insert: 1 ps of the trajectory surrounding the global energy minimum.
Figure 3.4: Comparison of the global energy-minimum structure and the structure representing the peak of the energy oscillations that persisted after equilibration (red circles in Figure 3.3).
Figure 3.5: Optimized structures of LF dimers.
Figure 3.6: Distribution of electron density changes. (A)LumiFlavin; (B) complex of lumiflavin and DBAP. The most positive atom is C2 and its charge changed from 0.79501 to 0.80681; the most negative atom is N3 and its charge changed from -0.64057 to -0.66457.
Figure 3.7: Distribution of NPA electron density in LF (left) and LF complexed with DBAP (right). The total charge of LF changed from 0.01 to 0.04 upon complexation with DBAP. Structures are oriented as in Scheme 1, red indicates excess electron density, green indicates electron deficiency.
Figure 3.8: Lengths of H-bonds measured from the LF•DBAP structure optimized using B3PW91 with 6-311++G(2d,2p).
Figure 3.9: HOMO (left) and LUMO (right) of LF (above), and LF in complex with DBAP (below).
Figure 3.10: Four possible optimized structures of LF with one H$_2$O molecule.
Figure 3.11: Six possible optimized structures of LF with two H$_2$O molecules.
Figure 3.12: Three possible optimized structures of LF with three H₂O molecules.
Figure 3.13: Three possible optimized structures of LF with four H₂O molecules.
Figure 3.14: Possible optimized structure of LF with six H$_2$O molecules (left) and with eight H$_2$O molecules (right).
Chapter Four: A Foundation for SSNMR Comparison of Flavoproteins

SSNMR is a very promising method for studying flavoproteins, and has the potential to provide atom-specific and even orbital-specific insight. SSNMR is being applied to a flavin bound in a protein, for the first time. We are preparing to obtain $^{15}$N SSNMR of FMN in two well-studied flavoproteins, Flavodoxin (FD) and Old Yellow Enzyme (OYE), with the long-term goal of learning how interactions with protein sites change the flavin reactivity by modifying the frontier orbitals.

4.1 Introduction to representative flavoproteins

4.1.1 Introduction to Flavodoxin (FD)

Flavodoxins are small bacterial flavoproteins that function as low potential electron transfer proteins in many organisms. They have about 140-150 amino acids and contain a non-covalently bound FMN as their only redox center. The redox properties of the bound cofactor differ markedly from those of the unbound molecule, suggesting that the holoprotein strongly regulates the bound prosthetic group. For example, the redox potential of FD’s semiquinone / hydroquinone couple is among the lowest of any known flavoprotein (approximately -450 vs. -175 mV for free FMN, pH 7), whereas that of the oxidized / semiquinone couple is generally more positive than that for free FMN ($106, 107$). Also, the one-electron reduced semiquinone is unusually stable and is stabilized as the blue neutral flavin radical species rather than the red anionic radical.

The flavodoxin from sulfate-reducing bacteria *Desulfovibrio vulgaris* mediates electron transfer at low redox potential. It is made up of a five-stranded parallel $\beta$-sheet, surrounded by $\alpha$-helices on either side of the sheet (Figure 4.1). *D. vulgaris* apoflavodoxin binds FMN very tightly, with a $K_d$ of 0.24 nM at pH 7.0

Reproduced with permission from the Journal of the American Chemical Society, and Biochemistry, submitted for publication. Unpublished work copyright 2010 American Chemical Society.
and 25 °C (107). The FMN lies mostly below the surface of the protein molecule. The isoalloxazine ring appears to be planar and is buried between two segments of polypeptide chain and the side groups Trp60 and Tyr98.

The three-dimensional structures of all three redox states of *D. vulgaris* flavodoxin have been determined by high-resolution X-ray crystallography (108), and NMR spectroscopy (109). The crystal structure is known to 2.0 Å resolution facilitating interpretation of the flavin binding site (110). In order to gain a better understanding how the redox properties of the flavin cofactor are regulated when bound to protein, the specific flavin-apoprotein interactions (Scheme 2.1) that exist in FD had been widely studied by site-specific mutagenesis (111, 112). FMN interacts with the protein through H-bonds as well as van der Waals and stacking interactions (108). The isoalloxazine ring is sandwiched between the aromatic rings of the side chains of Trp60 and Tyr98 and is approximately coplanar with Tyr98. Loop 100 (sequence 95-102) contains the residues that are in direct contact with the isoalloxazine ring and are known to influence the reactivity and redox properties of the bound flavin. Residues 95, 100 and 102 form backbone H-bonds to the atoms N3(H) and O2. Moreover, many different mutants have been characterized by X-ray crystallography (113, 114). Because FD is well behaved, readily prepared in large quantity, and well understood, FD has become the so-called “hydrogen atom of flavoproteins”: the ideal system on which to assess the feasibility of a new experiment.

As described in Chapter 2, SSNMR is a good tool for detecting site-specific H-bonding. The CSPVs are much more sensitive than the isotropic average and can be related to specific combinations of molecular orbitals. I used SSNMR as a tool to detect the active site flavin of FD with the goal of getting atom specific and even orbital specific information about how the cofactor electronics could be modulated by the host protein.

### 4.1.2 Introduction to Old Yellow Enzyme 1 (OYE1)

Old yellow enzyme (OYE, NADPH oxidoreductase, E.C. 1.6.99.1) from brewer’s yeast was the first flavoprotein to be discovered, having been isolated
by Warburg and Christian in 1933 (115). It is a dimer of approximately 49 kDa subunits with one noncovalently bound FMN per subunit. It is rapidly reduced by NADPH and can be reoxidized by oxygen. Although the physiological role of OYE has only recently begun to emerge (116, 117), the protein-flavin interaction has been extensively studied. Mutations affecting residues that bind the flavin have been made and their effects on reactivity and the flavin’s reduction midpoint potentials have been published (118, 119). Whereas free FMN readily undergoes one-electron redox chemistry, FMN bound in OYE undergoes predominantly two-electron reactions. Thus, the interactions with OYE modulate the FMN’s chemistry.

The crystal structures of oxidized and reduced OYE reveal an α/β-barrel topology. Half a dozen crystal structures have been published including structures of mutants and structures of OYE with bound substrate analogs (Figure 4.2) (119, 120). OYE’s complexes with phenolic compounds show characteristic charge transfer absorbance bands in the long wavelength region between 500 and 800 nm (121). The crystal structures show that these phenols stack on the si-face of the flavin. Studies of a series of phenols, and OYE reconstituted with a series of modified flavins, demonstrated that the spectral changes upon phenol binding are due to close physical association of the flavin and phenolate, that allows the two to interact electronically (120). Thus, OYE provides a well-characterized system with several well-understood mutated and complexed variants that affect flavin electronics, that we now plan to compare by SSNMR.

FD’s site contrasts with those of OYE in lacking positively charged residues near N1 and O2, and having a potentially negatively-charged one there instead. Thus, Asp95 is the closest Glu/Asp to the flavin, and mutagenesis of this side chain to a neutral has the biggest effect on the reduction potential of any carboxylate replacement (122), even though the side chain of Asp95 is not in direct contact with the flavin. Interactions between the flavin and the FD protein are sufficiently different to increase the protein’s $K_d$ for oxidized FMN by a factor of 5 (110). The flavin is essentially flat, like that of OYE, so comparison of the
chemical shift of OYE and FD could test our ability to perceive charge differences via CSPVs. Since the charge difference in question is adjacent to N1, we anticipate a large effect there. However, even N5 is expected to be quite responsive, due to propagation of effects in the conjugated orbitals of the oxidized state. In addition, the flavin of FD is stacked parallel to the Tyr98 ring. Indeed, FD’s flavin isotropic shifts are very different from those of OYE (Table 4.1) (13).

Studying the protein-cofactor interactions related to catalytic activity with the representative flavoenzyme OYE1 by SSNMR is a significant and challenging project. A comparison of free and OYE-bound FMN will test SSNMR’s ability to detect the effects of protein interactions on the FMN electronics and structures. A comparison of FMN in OYE alone and in OYE bound by phenolates will provide a control in that this is a case where optical signatures already demonstrate that the flavin electronics are affected; if SS NMR fails to detect a difference between these samples, we will have definitively demonstrated that it is not a good tool for such studies. This will be a disappointing result, but good science nonetheless.

4.2 Methods
4.2.1 Materials

Primers for PCR were ordered from Integrated-DNA Technologies. Pfu Ultra Polymerase, Pfu Ultra HF reaction buffer, XL1-Blue supercompetent cells were purchased from Stratagene. The gel Extraction Kit, QIAprep Spin miniprep kit and HiSpeed plasmid midi and maxi kits, Ni-NTA agarose gel were purchased from QIAGEN. dNTP mix, Ncol, Xhol, 100x BSA, NEBuffer 2, T4 ligase buffer and T4 DNA ligase were purchased from Bio-rad. The pET32a vector was the kind gift of Prof. David W. Rodgers (Department of Biochemistry, University of Kentucky). pET 23d and pET3b vector, BL21 (DE3) competent cells were purchased from Novagen and stored at -20 °C. Ampicillin, isopropyl β-thiogalactopyranoside (IPTG), imidazole were purchased from Sigma. 4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF) was purchased from Roche. CH₃COONa was purchased from J. T. Baker. KBr was purchased
from EMD. D$_2$O was purchased from Cambridge Isotope Laboratories. OYE1 DNA template was from Dr. Palfey (University of Michigan).

### 4.2.2 TEV-6xHis-TEV-FD cloning into pET32a vector

The first goal was to produce high quality protein samples. In order to facilitate purification of FD, the FD gene was engineered to include a 6xHis tag followed by DNA encoding a TEV (Tobacco Etch Virus) protease cleavage site (Glu-Asn-Leu-Tyr-Phe-Gln-(Gly/Ser)) at the N-terminus of the FD (123). The expressed tagged protein could be easily purified by Ni-NTA affinity chromatography. The 6xHis tag could then be removed by digestion with TEV protease after protein purification.

The above features were obtained by adding a TEV target sequence at the N terminus of FD by polymerase chain reaction (PCR). The insert encoding the TEV target sequence and FD was cloned into the pET32a vector using the NcoI and XhoI sites of the vector. Cloning into pET32a vector adds a thioredoxin domain as well as a 6xHis tag to the N terminus (Figure 4.3). All of these could be cut off using TEV protease after purification.

The forward primer Nterm\_NcoI\_TEV\_FD (5’ – gat acc cat ggc aga aaa cct gta ttt tca ggg cgc g – 3’) and reverse primer Cterm\_XhoI\_FD (5’ – tgg atc tcg agt caa atc gcg ccg cgc aca tc - 3’) were used to add a TEV cleavage site at the N terminus of FD by PCR using the regimen described in Figure 4.4. For a 50 μL PCR, 1 μL DNA template (50-100 ng), 1.5 μL forward and reverse primers (0.1 μg/μL), 1 μL dNTP (10 mM), 5.0 μL of 10X Pfu Ultra HF reaction buffer (Stratagene), 1 μL of Pfu Ultra Polymerase (Stratagene), and 39 μL sterile deionized water were mixed in a microamp reaction tube. Production of the correct sized fragment was confirmed by gel electrophoresis on 1% agarose. The fragment was purified from the agarose using the Gel Extraction Kit (Qiagen). Both the TEV-FD insert and pET32a vector were digested with NcoI and XhoI. Digestion products were purified by electrophoresis through 1% agarose, and recovered by using the Gel Extraction Kit (Qiagen).
After ligation of the insert and cut vector, using T4 DNA ligase, the resulting Trx-6xHis-TEV-FD-pET32a was used to transform XL1-Blue supercompetent *E. coli* cells (Stratagene). Cells were spread on an LB plate containing 50 μg/mL ampicillin to select those cells containing plasmid. Upon incubation at 37 °C overnight, a few single colonies were selected and each used to inoculate 3 mL of LB culture (10 g/L Bacto™ Peptone, 5 g/L Yeast Extract, 10 g/L NaCl. pH = 7.5) containing 50 μg/mL ampicillin. After shaking at 37 °C, 250 rpm for 12 hours, each culture was harvested by centrifugation and Trx-6xHis-TEV-FD-pET32a DNA plasmid was purified from each using the QIAprep Spin miniprep kit (Qiagen). The plasmid DNA from the different colonies was digested using Ncol and Xhol and resolved by agarose gel electrophoresis (1%). Plasmids containing the right size of insert (500 bp) were characterized further.

A strain with an apparently correct plasmid was grown in a 100 mL LB culture containing 50 μg/mL ampicillin. After incubation at 37 °C, shaking at 250 rpm for 14 hours, bacteria were harvested and plasmid was purified using HiSpeed plasmid midi or maxi kits (Qiagen). The Trx-6xHis-TEV-FD-pET32a plasmid DNA was sent to Davis Sequencing INC for nucleotide sequence determination.

4.2.3 Expression and purification Trx-6xHis-TEV-FD

Trx-6xHis-TEV-FD-pET32a plasmid was transformed into BL21 (DE3) competent cells. FD was overexpressed in *E. coli* and successfully purified by Ni-NTA agarose column chromatography.

**Overexpression Non-labeled Trx-6xHis-TEV-FD**

Trx-6xHis-TEV-FD-pET32a plasmid constructed as described above was transformed into BL21 (DE3) competent cells. Glycerol stock solutions of *E. coli* bearing the Trx-6xHis-TEV-FD-pET32a plasmid were made and stored at -80 °C.

Bacterial culture leading up to protein overexpression was initiated by spreading 30 μL of the glycerol stock solution on an LB plate containing 50 μg/mL ampicillin and incubating at 37 °C overnight. A single colony was picked
and used to inoculate a tube of 3 mL LB medium containing 50 μg/mL ampicillin. After incubation at 37 °C, shaking at 220 rpm for 12 hours, 100 μL were used to inoculate 100 mL of LB culture containing 50 μg/mL ampicillin. After incubation at 37 °C, shaking at 220 rpm for an additional 12 hours, the 100 mL LB culture was added to 2 L LB containing 50 μg/mL ampicillin and then shaken at 37 °C in a 6 L flask. When the optical density at 600 nm (OD₆₀₀) reached approximately 0.6 in about 4 hours, 0.5 M IPTG was added to a final concentration of 1 mM to induce overexpression of the plasmid-borne FD gene. The temperature was lowered to 32 °C and the culture was shaken at 220 rpm for an additional 4 hours whereupon the culture had reached an OD₆₀₀ of about 1.7. The cells were harvested by centrifugation at 4 °C, at 9500 rpm for 15 min. The cells were suspended in lysis buffer (50 mM NaH₂PO₄, pH = 7.30, containing 300 mM NaCl, 10 mM imidazole) and collected in a single pellet by centrifugation at 4 °C, at 9500 rpm for 15 min. Cells were stored at -20 °C until needed. The overexpression was tested by running SDS-PAGE for the samples collected before and 4 hours after adding IPTG (Figure 4.5).

**Overexpression**¹⁵N Uniformly-labeled Trx-6xHis-TEV-FD

100 mL of LB culture was grown in the same way as described for unlabeled wild type FD. The 100 mL LB culture was added to 2 L ¹⁵N labeled MM9 culture medium (12 g/L Na₂HPO₄, 6 g/L KH₂PO₄, 0.5 g/L NaCl, 1 g/L ¹⁵NH₄Cl, pH=7.50) with 1% glucose, 10 μM FeCl₂, 4 mM MgSO₄, 100 μM CaCl₂, 50 μg/mL ampicillin and 1mL of micronutrient solution (3 μM (NH₄)₆Mo₇O₂₄, 400 μM H₃BO₃, 30 μM CoCl₃, 80 μM MnCl₂ and 10 μM ZnSO₄; sterilized by Millipore syringe filter) per liter MM9. The culture was incubated at 37 °C, shaking at 220 rpm. When OD₆₀₀ reached about 0.5 in about 8 hours, 0.5 M IPTG was added to the final concentration of 1 mM to induce overexpression. The temperature was changed to 32 °C and the culture was incubated with shaking at 220 rpm for an additional 12 hours. The cells were harvested by centrifugation at 4 °C, 9500 rpm for 15 min. The cells were suspended in lysis buffer, and collected by centrifugation at 4 °C, 9500 rpm for 15 min. Cell pellets were stored at -20 °C.
**Purification Trx-6xHis-TEV-FD (\textsuperscript{15}N Trx-6xHis-TEV-FD)**

Frozen cells were suspended in lysis buffer (10 mL per gram cells) with 0.5 mg/mL 4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF) and 0.1 mmol/L FMN (for \textsuperscript{15}N Trx-6xHis-TEV-FD, no addition of FMN in this step). After homogenization by hand, cells were broken by passage through a French Press twice at 15k psi. The crude FD and the cell debris were separated by centrifugation at 4 °C, at 15000 x g for 45 min.

The supernatant was applied to a 20 mL Ni-NTA agarose column (Figure 4.6) which had been pre-equilibrated with 200 mL of lysis buffer. The flow-through was collected and reloaded three more times. The column was washed with wash buffer (50 mM NaH\textsubscript{2}PO\textsubscript{4}, pH = 7.30, containing 300 mM NaCl, 20 mM Imidazole) to remove non-specifically bound protein until OD\textsubscript{280} of the eluate was low and constant. 6xHis-tagged FD was eluted using high imidazole elution buffer (50 mM NaH\textsubscript{2}PO\textsubscript{4}, pH = 7.30, containing 300 mM NaCl, 250 mM imidazole). 5 mL fractions were collected and their OD\textsubscript{280} was monitored to identify fractions containing protein (presumably Trx-6xHis-TEV-FD). The eluted protein’s identity was determined on the basis of molecular weight by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 4.7) and yellow color. Fractions containing pure His-tagged FD were combined and dialyzed against dialysis buffer (50 mM NaH\textsubscript{2}PO\textsubscript{4}, pH = 7.30).

The eluate was concentrated using an Amicon Ultra-15 centrifugal filter device (10 k Nominal Molecular Weight Limit), to a final volume of less than 10 mL, by centrifuging at 4 °C, at 4000 x g for 15 to 45 min. The concentration was determined based on measured values of OD\textsubscript{272} and OD\textsubscript{445} by using a HP8452B diode array spectrophotometer (\(\varepsilon_{272} = 47,600 \text{ M}^{-1} \text{ cm}^{-1}\), \(\varepsilon_{445} = 10,200 \text{ M}^{-1} \text{ cm}^{-1}\) for FD) (124).

**Cleavage of Trx-6xHis-TEV-FD**

200 μL 2.0 mg/mL Trx-6xHis-TEV protease in 50 mM Tris, 2 mM BME, pH = 7.40, was added to Trx-6xHis-TEV-FD in 50 mM NaH\textsubscript{2}PO\textsubscript{4}, pH = 7.30. SDS-
PAGE was run for the sample mixture after rocking at 4°C or at 22°C or at 37°C overnight. Pre-equilibrated Ni-NTA agarose column with at least 200 mL 50 mM NaH₂PO₄, pH = 7.30. The sample mixture was applied to the column and the yellow elution fraction (cut FD) were collected.

4.2.4 FD-6xHis cloning into pET23d vector

In order to circumvent the need to cut off a Trx-6xHis tag with TEV, we sought a cloning vector that would not add Trx coding sequence to our construct. We chose the pET23d vector which lacks the Trx coding region and adds a C-terminal 6xHis-tag to the inset gene. DNA sequence encoding the gene for FD was amplified using the polymerase chain reaction (PCR). This insert was successfully cloned into pET23d vector using the Ncol and Xhol endonuclease sites (Figure 4.8).

pET23d vector was purchased from Novagen and stored at -20 °C. The forward primer Nterm_Ncol_FD (5’ – gat acc cat ggc gaa agc gct gat t - 3’) and reverse primer Cterm_FD_Xhol (5’ – tgg atc tcg aga atc tcg cgc cgc cgc cgc aca tc - 3’) were used to amplify the FD gene. After amplifying the insert gene by PCR (Figure 4.4), the product was identified by 1% agarose gel electrophoresis and purified using the Gel Extraction Kit (Qiagen) (Figure 4.9). Both FD insert and pET23d vector were digested with Ncol and Xhol. Digestion products were confirmed by 1% agarose gel electrophoresis (Figure 4.10) and purified using the Gel Extraction Kit (Qiagen).

After ligation of the FD insert and cut vector with T4 DNA ligase, the resulting plasmid FD-His-pET23d was transformed into XL1-Blue supercompetent cells (Stratagene). After spreading onto an LB-Ampicillin plate containing 50 μg/mL ampicillin and incubating at 37 °C overnight, different single colonies were selected and each used to inoculate a 3 mL LB cultures containing 50 μg/mL ampicillin. After shaking at 37 °C, 250 rpm for 12 hours, each culture was harvested by centrifugation and FD-His-pET23d DNA plasmids were purified using the QIAprep Spin miniprep kit (Qiagen) for each of the different colonies. The plasmid DNA from the different colonies was digested using Ncol and Xhol;
plasmids containing the right size of insert were identified by 1% agarose gel electrophoresis.

A strain with an apparently correct plasmid was grown in a 100 mL LB culture containing 50 μg/mL ampicillin. After shaking incubation at 37 °C, 250 rpm for 14 hours, bacteria were harvested and the FD-His-pET23d DNA plasmid was purified using HiSpeed plasmid midi or maxi kits (Qiagen). The FD-His-pET23d plasmid DNA was sent to Davis Sequencing INC for nucleotide sequence determination.

4.2.5 Overexpression and purification of FD-His

FD-His-pET32a plasmid was transformed into BL21 (DE3) competent cells. FD was overexpressed in E. coli and successfully purified by Ni-NTA agarose column chromatography. Overexpression and purification was similar as Trx-6xHis-TEV-FD. See 4.2.3 for details.

4.2.6 Preparation of $^{15}$N FMN reconstituted FD

FMN is generally tightly bound to FD. The dissociation constant ($K_d$) at pH 7.0 and 25 °C is 0.24 nM (107). It is therefore impractical to remove FMN by simple dialysis. However, it is useful to remove the $^{14}$N FMN from FD and reconstitute $^{15}$N FMN latter, especially for NMR studies in which only $^{15}$N is to be observed and $^{14}$N is silent.

**Preparation of $^{14}$N apo-FD**

Wild-type $^{14}$N FD was overexpressed in LB medium and purified with the procedure described above. The $^{14}$N FMN was dissociated from the FD protein by dilution into 0.1 M CH$_3$COONa, pH=3.9, containing 2 M KBr (125). After centrifugation at 4 °C, 14000 x g for 10 min, the yellow supernatant containing FMN was discarded. The white precipitated $^{14}$N FD protein was restored to a non-denaturing buffer by gentle resuspension in dialysis buffer and dialysis against dialysis buffer at 4 °C overnight.
The concentration of $^{14}$N apo-FD was determined based on OD$_{280}$ by using a HP8452B diode array spectrophotometer (Figure 4.11) ($\varepsilon_{280} = 22,400$ M$^{-1}$ cm$^{-1}$) (107).

**Preparation of $^{15}$N FMN**

$^{15}$N uniformly-labeled FD was overexpressed and purified with the same procedure described above. The $^{15}$N FMN was extracted from $^{15}$N FD from the $^{15}$N apo-FD by the same procedure as was used for preparation of $^{14}$N apo-FD except that the yellow 0.1 M acetate supernatant was retained. We obtained $^{15}$N FMN in sufficiently high concentration that it could be diluted into solutions of $^{14}$N apo-FD under native conditions to achieve FMN reconstitution.

**Reconstitution of the $^{14}$N apo-FD with $^{15}$N FMN**

The $^{14}$N apo-FD solution in dialysis buffer was mixed with a solution containing an equi-molar ratio of $^{15}$N FMN to prepare the $^{15}$N (>95%) FMN reconstituted FD (Figure 4.12). Excess FMN was removed by ultra filtration. The product was monitored by using a HP8452B diode array spectrophotometer (Figure 4.11). The A$_{445}$ and $\varepsilon_{445} = 10,200$ M$^{-1}$ cm$^{-1}$ were used to calculate the FMN concentration and A$_{272}$ with $\varepsilon_{272} = 47,600$ M$^{-1}$ cm$^{-1}$ were used to calculate the protein concentration (124).

**4.2.7 Solution NMR of $^{15}$N FD**

Solution NMR was used to assess the structural integrity of FD and the environment of the FMN. $^{15}$N FD was dissolved in 50 mM NaH$_2$PO$_4$, pH=7.30. D$_2$O was added to 10%.

All the NMR spectra were recorded on a Varian Inova 600 MHz spectrometer at 20 °C. For each spectrum, the probe was retuned, the $^1$H was recalibrated and the shimming was re-optimized.

For $^1$H 1d spectra, solvent suppression was achieved using the wet1d (water suppression enhanced through T1 effects) sequence to rapidly diphase the signal of water (Figure 4.13) (126).
The $^1$H-$^{15}$N heteronuclear single quantum coherence (HSQC) spectrum was collected to reveal the folding state of the protein (Figure 4.14). The recycle delay between scans was set to 1 second and the spectral widths were 12000 Hz and 2400 Hz in the $^{15}$N (F1) and $^1$H (F2) dimensions, respectively. The HSQC spectrum was processed using Varian’s vnmr software. The chemical shifts were referenced to water at 4.77 ppm for $^1$H and external $^{15}$N urea in DMSO at 77 ppm for $^{15}$N. Data were processed using linear prediction to triple the number of $^{15}$N points before Fourier transformed.

$^{15}$N 1d NMR spectra were collected at 20 °C to show the nitrogens in the FMN cofactor (Figure 4.15). There was 1 second recovery time between 0.2 seconds acquisition. 7 μs pulse width and 10526 Hz decoupling field were used.

### 4.2.8 Solid state NMR of $^{15}$N FD (1d CP-MAS)

$^{15}$N uniformly labeled FD was studied by SSNMR. Sample was prepared by dialysis against 5 mM HCO$_2$NH$_2$, pH=7.0, containing 5 mM Glucose and lyophilized. About 80 mg FD powder was packed in a 5 mm zirconia rotor for SSNMR.

$^{15}$N MAS spectra were collected at 40 MHz for $^{15}$N with signal enhancement via ramped Cross Polarization (CP) from $^1$H in 5 mm Zirconia rotors in an HX Chemagnetics-type probe using a Varian Inova™ spectrometer at -50 °C (Figure 4.16). $^{15}$N spectra were collected with interscan delays of 5 sec, CP contact times of 1 ms at a CP field of 50 kHz for $^1$H and $^1$H TPPM2 decoupling at 50 kHz during 20 ms acquisitions (47). Spinning speeds of 3000, 4000 and 5000 Hz were used for the $^{15}$N 1d CP-MAS spectra. All $^{15}$N chemical shifts are quoted relative to liquid ammonia, and spectra were referenced indirectly based on the $^{15}$NO$_3^-$ signal of NH$_4$NO$_3$ at 376 ppm (48).

### 4.2.9 His-tagged TEV-OYE1 cloning into pET32a vector

DNA sequence encoding the thioredoxin domain, a His-tag and a TEV cleavage site were added at the N terminus of the gene for Old Yellow Enzyme1 (OYE1) using the polymerase chain reaction (PCR). This insert was successfully
cloned into pET32a vector using the Ncol and Xhol endonuclease site (Figure 4.17).

The forward primer Nterm_\text{Ncol TEV OYE1} (5' - gat acc cat gcc aga aac ctt gta ttt tca ggg aat gtc att tgt aaa ag - 3') and reverse primer Cterm_\text{XhoI OYE1} (5' – ctg atc tcg agt tta ctt ttt gtc cc a gcc - 3') were used to add TEV site at the N terminus of the OYE1 gene. After amplifying the insert gene by PCR (Figure 4.18), the product was identified by 1% agarose gel electrophoresis and purified using the Gel Extraction Kit (Qiagen) (Figure 4.19). Both TEV-OYE1 insert and pET32a vector were digested with Ncol and Xhol. Digestion products were confirmed by 1% agarose gel electrophoresis and purified using the Gel Extraction Kit (Qiagen).

After ligating the above together using T4 DNA ligase, the resulting Trx-His-TEV-OYE1-pET32a was transformed into XL1-Blue supercompetent cells. After spreading on an LB plate containing 50 μg/mL ampicillin and incubating at 37 °C overnight, different single colonies were selected and each used to inoculate added to 3 mL LB culture containing 50 μg/mL ampicillin. After shaking incubation at 37 °C, 250 rpm for 12 hours, Trx-His-TEV-OYE1-pET32a DNA plasmids were each purified using the QIAPrep Spin miniprep kit (Qiagen). After the Trx-His-TEV-OYE1-pET32a plasmids from different colonies had been digested by Ncol and Xhol, those producing the expected fragments were identified by 1% agarose gel electrophoresis (Figure 4.20).

Strains producing the correct sized restriction products were grown up in 100 mL LB cultures containing 50 μg/mL ampicillin. After shaking incubation at 37 °C, 250 rpm for 14 hours, putative the Trx-His-TEV-OYE1-pET32a plasmids were purified using HiSpeed plasmid midi and maxi kit (Qiagen). Trx-His-TEV-OYE1-pET32a plasmid DNA was sent to Davis Sequencing INC for sequencing.

4.2.10 Overexpression and purification of Trx-His-TEV-OYE1

Trx-His-TEV-OYE1-pET32a plasmid was transformed into BL21 (DE3) competent cells. Trx-His-TEV-OYE1 was overexpressed in \textit{E. coli} in Luria-
Bertani (LB) medium and successfully purified by Ni-NTA agarose column chromatography.

**Overexpression of Trx-His-TEV-OYE1**

Trx-His-TEV-OYE1-pET32a plasmid constructed as described above was transformed into BL21 (DE3) competent cells. After spreading on an LB plate containing 50 μg/mL ampicillin and incubated at 37 °C overnight, a single colony was picked and used to inoculate a 3 mL LB culture containing 50 μg/mL ampicillin. After shaking incubation at 32 °C, 220 rpm for 12 hours, 100 μL were added to 100 mL of LB culture containing 50 μg/mL ampicillin. After shaking incubation at 32 °C, 220 rpm for additional 12 hours, the 100 mL LB culture was added to 2 L LB containing 50 μg/mL ampicillin and then shaken at room temperature. When OD$_{600}$ reached to about 0.5, 0.5 M IPTG was added to a final concentration of 0.4 mM to induce overexpression. After shaking incubation at 220 rpm for an additional 6 hours, the cells were harvested by centrifugation at 4 °C, 9500 rpm for 15 min. The cells were suspended in cell wash buffer and collected by centrifuge at 4 °C, 9500 rpm for 15 min. Cell pellets were stored at -20 °C.

**Purification of Trx-His-TEV-OYE1**

Frozen cells were suspended in lysis buffer with 0.5 mg/mL 4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF) and 0.1 mmol/L FMN. After homogenization by hand, cells were broken by passage through a French Press twice at 15k psi. The crude Trx-His-TEV-OYE1 and the cell debris were separated by centrifugation at 4 °C, 15000 g for 45 min.

The supernatant was applied to a 10 mL Ni-NTA agarose column which had been pre-equilibrated with 200 mL of lysis buffer. The flow though was collected and reloaded three more times. The column was washed with 500 mL of wash buffer to remove non-specifically bounded protein. His-tagged OYE1 was eluted using high imidazole elution buffer. Each 5 mL fractions were collected and their OD$_{280}$ was monitored to identify fractions containing protein.
(presumably Trx-His-TEV-OYE1). The eluted protein’s identity was determined on the basis of molecular weight by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 4.21). Fractions containing pure Trx-His-TEV-OYE1 were combined and dialyzed against dialysis buffer.

The Trx-His-TEV-OYE1 was concentrated by centrifugation in an Amicon Ultra-15 centrifugal filter device (10 k NMWL) at 4 °C, 4000 g for 15 to 45 min to a total volume of less than 10 mL. After adding 1mL of 1.0 mg/mL His-tagged TEV protease glycerol stock solution, the mixture was rocked at room temperature for at least 12 hours. The above mixture was applied to the Ni-NTA column again. The flow-through was collected and an additional 15 mL dialysis buffer was loaded to elute all cleaved OYE1. The eluate was concentrated using an Amicon Ultra-15 centrifugal filter device (10 k NMWL), to a final volume of less than 10 mL, by centrifuging at 4 °C, 4000 g for 15 to 45 min. The concentration was determined for measured values of OD$_{278}$, OD$_{381}$ and OD$_{462}$ by using a HP8452B diode array spectrophotometer ($\varepsilon_{278} = 102,500$ M$^{-1}$ cm$^{-1}$, $\varepsilon_{381} = 9,600$ M$^{-1}$ cm$^{-1}$, $\varepsilon_{462} = 10,600$ M$^{-1}$ cm$^{-1}$) (128).

4.2.11 His-TEV-OYE1 cloning into pET3b vector

DNA sequence encoding a His-tag, and a TEV cleavage site were added at the N terminus of the gene for Old Yellow Enzyme1 (OYE1) using the polymerase chain reaction (PCR). This insert was successfully cloned into pET3b vector using the Ndel and BamHI endonuclease site (Figure 4.22).

The forward primer Nterm_ Ndel_His_TEV_OYE1 (5' - gat acc ata tgc atc acc atc acc atc acg aaa acc tgt att itc agg gaa tgt cat ttg taa aag - 3') and reverse primer Cterm_ BamHI_OYE1 (5' - ctg atg gat cct tta ctt ttt gtc cca gcc - 3') were used to add His-tag and TEV site at the N terminus of the OYE1 gene. After amplifying the insert gene by PCR (Figure 4.18), the product was identified by 1% agarose gel electrophoresis and purified using the Gel Extraction Kit (Qiagen). Both His-TEV-OYE1 insert and pET3b vector were digested with Ndel and BamHI. Digestion products were confirmed by 1% agarose gel electrophoresis and purified using the Gel Extraction Kit (Qiagen).
After ligating the above together using T4 DNA ligase, the resulting His-TEV-OYE1-pET3b was transformed into XL1-Blue supercompetent cells. After spreading on an LB plate containing 50 μg/mL ampicillin and incubating at 37 °C overnight, different single colonies were selected and each used to inoculate added to 3 mL LB culture containing 50 μg/mL ampicillin. After shaking incubation at 37 °C, 250 rpm for 12 hours, His-TEV-OYE1-pET3b DNA plasmids were each purified using the QIAprep Spin miniprep kit (Qiagen). After the His-TEV-OYE1-pET3b plasmids from different colonies had been digested by NdeI and BamHI, those producing the expected fragments were identified by 1% agarose gel electrophoresis (Figure 4.23).

Strains producing the correct sized restriction products were grown up in 100 mL LB cultures containing 50 μg/mL ampicillin. After shaking incubation at 37 °C, 250 rpm for 14 hours, putative the His-TEV-OYE1-pET3b plasmids were purified using HiSpeed plasmid midi and maxi kit (Qiagen). TEV-OYE1-pET3b plasmid DNA was sent to Davis Sequencing INC for sequencing.

4.2.12 Overexpression and purification of His-TEV-OYE1

His-TEV-OYE1-pET3b plasmid was transformed into BL21 (DE3) competent cells. OYE1 was overexpressed in *E. coli* and successfully purified by Ni-NTA agarose column chromatography. Overexpression and purification was similar as Trx-6xHis-TEV-OYE1. See 4.2.10 for details.

4.2.13 Preparation, activity assay and titration with p-chlorophenol of native OYE1, apo-FMN OYE1 and FMN-reconstituted OYE1

The cofactor FMN was removed from the native OYE1 to get apo-FMN OYE1. The apo-FMN OYE1 could rebind the cofactor FMN. The native OYE1 and the FMN-reconstituted OYE1 showed similar properties with respect to catalytic activity and titration with p-chlorophenol.

*Preparation of apo-FMN OYE1 and FMN-reconstituted OYE1*
Native His-tagged OYE1 was dialyzed against 0.2 M phosphate buffer, pH=5.30, containing 2 M KBr and 6.7 mM EDTA. After three changes of this buffer over a 24 hours period, there was no yellow color visible in the dialysis bag. Then the apo-FMN OYE1 was dialyzed against dialysis buffer to remove KBr. Apo-FMN OYE1 was recovered and stored at 4 °C. FMN was reconstituted to apo-FMN OYE1 by adding FMN to a final concentration of 0.1 mM and incubating at 4 °C overnight. Excess FMN was separated from FMN-reconstituted OYE1 by centrifuging at 4 °C, 4000 g for 30 min in a centricon (Ultracel YM-10). 1mL of dialysis buffer was added to the centricon to dilute the OYE1 again and the centrifugal concentration was repeated to remove more excess FMN until there was no yellow color visible in the flow-through. The OD$_{278}$, OD$_{381}$ and OD$_{462}$ were measured by a HP8452B diode array spectrophotometer for native OYE1, apo-FMN OYE1 and FMN-reconstituted OYE1 (Figure 4.24 and Table 4.3) ($\varepsilon_{278} = 102,500$ M$^{-1}$ cm$^{-1}$, $\varepsilon_{381} = 9,600$ M$^{-1}$ cm$^{-1}$, $\varepsilon_{462} = 10,600$ M$^{-1}$ cm$^{-1}$) (128).

**Activity assay of native OYE1, apo-FMN OYE1 and FMN-reconstituted OYE1**

The activity of OYE1 was assayed by measuring the rate of NADPH oxidation at 340 nm ($\varepsilon_{340} = 6300$ L mol$^{-1}$ cm$^{-1}$) at constant temperature of 25 °C (5, 129). The blank buffer contained 100 μM potassium phosphate, pH=7.0. After 120 nmol of NADPH was added, the absorbance was monitored. The enzyme was added at the 50 second time point. Enzyme activity and specific activity were determined according to the initial rate of NADPH oxidation (Figure 4.25).

**Titration with p-chlorophenol of native OYE1, apo-FMN OYE1 and FMN-reconstituted OYE1**

Spectral changes in OYE1 upon titration with p-chlorophenol were monitored optically at 382 nm, 462 nm and 645 nm using a HP8452B diode array spectrophotometer. The blank buffer contained 0.1 M potassium phosphate, pH=6.45. The enzyme concentration was 2.6 x 10$^{-2}$ mM. The concentrations of p-
chlorophenol used were 0.972 x 10^{-2} \text{ mM}, 1.94 \times 10^{-2} \text{ mM}, 2.91 \times 10^{-2} \text{ mM}, 3.98 \times 10^{-2} \text{ mM}, 5.80 \times 10^{-2} \text{ mM}, 64.0 \times 10^{-2} \text{ mM} \text{ respectively.}

4.3. Results

4.3.1 TEV-6xHis-TEV-FD cloning into pET32a vector

I was successful in producing the forward and reverse primers described above (4.2.2), amplifying the TEV-FD insert while adding the TEV site at the N terminus of FD. Subcloning the TEV-FD gene into the pET32a vector, caused FD to be expressed following a 6xHis tag which itself follows a thioredoxin sequence.

The sequencing result of the clone which displayed the correct bands on the 1% agarose gel (one band at about 6 kb and another one at about 500 bp) showed 100% identity with the FD sequence, without gaps or mutations. Due to the last residue of the TEV targeting sequence, the first amino acid at FD’s N-terminal is Gly instead of Met.

4.3.2 Expression and purification Trx-6xHis-TEV-FD

Using the above procedures, His-tagged FD was overexpressed (Figure 4.5) and purified to 95% purity. The pure FD was contained in the first 15 mL to elute from the Ni-NTA column in 250 mM imidazole, and showed a strong band at about 20 kD in SDS-PAGE. According to the concentration and total volume of the FD, about 30 mg FD can be obtained from a 2 L MM9 growth.

Although the pET32a construct supports an adequate total yield of purified protein, only 1/3 of the resulting Trx-His-TEV-FD could be cleaved by TEV protease. Therefore, the net yield of pure FD was only 30 mg / 2L culture.

4.3.3 FD-6xHis cloning into pET23d vector

Using both the forward and reverse primers described above (4.2.4), I succeeded in amplifying the FD insert (Figure 4.9). The FD gene was then subcloned into the pET23d vector, which supplies code for a His-tag at the C terminus of the inserted gene. Also, there is no thioredoxin module and TEV site
at the N terminus. Cleavage with TEV protease with TEV protease is no longer necessary for our work.

The DNA sequence of the clone which produced the right bands on the 1% agarose gel (one band at about 3.6 kb and another one at about 500 bp). Yield of protein sequencing. Yield of protein upon overexpression showed 100% identity with the FD sequence. Due to the XhoI site and His-tag at the C-terminal of FD, the cloned FD has 8 more amino acids (Leu and Glu from XhoI site, and 6 His for the His-tag) at the C-terminal.

4.3.4 Overexpression and purification of FD-6xHis

Using the above procedures, FD-6xHis was overexpressed and purified to 95% purity. The pure FD was contained in the first 15 mL to elute from the Ni-NTA column in 250 mM imidazole, and showed a strong band at about 16 kD in SDS-PAGE.

4.3.5 Preparation of $^{15}$N FMN reconstituted FD

The optical spectra of native $^{14}$N FD and the $^{15}$N FMN reconstituted $^{14}$N FD were very similar (Figure 4.11). FMN, FD monomer ratios of 0.95 were typical. $^{14}$N FD shows the expected FMN absorbances at 376 and 457 nm, $^{15}$N FMN reconstituted $^{14}$N FD shows the FMN absorbances at 379 and 460 nm. However, apo-FD had none of the optical signature of FMN. Thus, the cofactor FMN could be removed from native FD to generate apo-FD and then added back to apo-FD to generate holo-FD.

4.3.6 Solution NMR of $^{15}$N FD

$^{15}$N FD was monitored by solution NMR to assess the protein structure’s stability. Both wet1d and HSQC showed that $^{15}$N FD was well folded, based on the good dispersion and sharp peaks as well as comparison with literature spectra. All the four nitrogens from FMN were evident in the $^{15}$N1d spectra.

Figure 4.13 shows the proton wet1d $^1$H spectrum of $^{15}$N uniformly-labeled FD. Detection of the solute signal in the presence of the solvent signal was
achieved by using water suppression techniques. The 1d ¹H NMR spectra showed sharp peaks and significant dispersion in the amide proton region (downfield, 6-9 ppm) as well as in aliphatic side-chain proton region (upfield, 0-3 ppm). This indicates that the protein must be in a native, functional conformation. Although proton signals overlapped, each proton has their own specific environment and shows the specific signal.

The ¹⁵N HSQC experiment is probably the most frequently recorded experiment in protein NMR. Each residue of the protein (except proline) has an amide proton attached to a nitrogen in the peptide bond. The ¹H-¹⁵N HSQC of uniformly ¹⁵N-labeled FD (Figure 4.14) was very similar to the spectrum published before (109), and the ¹H resonances are readily resolved by the ¹⁵N dimension. Because the peaks are well dispersed, and most of the individual peaks can be distinguished, the protein is folded. Thus, HSQC is a relatively cheap and quick experiment, and useful to determine whether protein is well folded or not. The number of peaks in the spectrum match the number of residues in the protein (though side chains with nitrogen-bound protons will add additional peaks). Although it is difficult to solve the structure of the protein, obtaining a good HSQC is always the primary step and it is essential for a meaningful interpretation of more advanced NMR experiments.

¹⁵N 1d NMR spectra of ¹⁵N uniformly-labeled FD (Figure 4.15) show all four nitrogens in the FMN cofactor, both with and without proton decoupling spectra. This spectrum shows how little background we have from natural abundance ¹⁵N in ¹⁴N-FD. The signals of the 6xHis-tag are evident in the spectral region shown. The chemical shifts of N1, N3, N5 and N10 were 188.0, 160.0, 341.4 and 165.2 ppm vs urea in DMSO at 77 ppm. The chemical shifts of these four nitrogens were 188.0, 159.9, 341.5, and 165.6 ppm in the literature (34). Thus our resonances are similar and easy identified within 0.4 ppm of the literature chemical shifts. The small difference between our experimental and literature values maybe due to the small difference of pH of the samples. Thus, FD sample containing stoichiometric FMN and the FMN was correctly bounded.
The above NMR studies of FD provide good controls for our upcoming solid state NMR study. The chemical shifts of all four nitrogens in the FMN cofactor should be close to the isotropic chemical shifts obtained SSNMR, but maybe affected by the different temperature and physical states under which these two types of spectra are obtained.

4.3.7 Solid state NMR of $^{15}$N FD (1d CP-MAS)

SSNMR was used to study $^{15}$N uniformly labeled FD. The backbone shows signal centered at about 120 ppm, and the sidebands span over 600 ppm. The signal corresponding to the nitrogens from flavin ring showed up in couple of hours. Although $^{15}$N uniformly labeled FD shows very little background in solution state NMR $^{15}$N 1d spectra, it is not good enough for SSNMR detection. Among three different spinning speed used, only the spectrum detected at 4000 Hz had strong potential peaks corresponding to the signals of nitrogens from flavin ring (Figure 4.16). The sidebands from backbone of protein are strong and broad, although the potential peaks of N5 showed up, they were overlapped with the sidebands from backbone of protein. Thus, it is very hard to get intensity to calculate CSPVs.

This SSNMR study illustrates that reconstituted sample is necessary if we are looking for the signal corresponding the nitrogen from flavin ring.

4.3.8 His-tagged TEV-OYE1 cloning into pET32a vector

By using both the forward and reverse primer I succeeded in adding bases encoding a TEV cleavage site at the N terminus of the OYE1 gene and amplifying the TEV-OYE1 insert (Figure 4.19). The TEV-OYE1 gene was then subcloned into the pET32a vector, which supplies code for a His-tag and a thioredoxin module at the N terminus of inserted gene. This His-tag greatly facilitates protein purification by Ni-NTA agarose column chromatography.

The DNA sequence of the one which produced the right bands on the 1% agarose gel (one band at about 6 kb and another one at about 1.2 kb) (Figure 4.20) showed 100% identity with the OYE1 sequence.
4.3.9 Overexpression and purification of Trx-His-TEV-OYE1

Using the above procedures, Trx-His-TEV-OYE1 was overexpressed and purified to 95%. The pure Trx-His-TEV-OYE1 was contained in the first 15 mL elute and showed a strong band at about 64 kDa (49 kDa for OYE1 and 15 kDa for Trx) (Figure 4.21). According to the concentration and total volume of the Trx-His-OYE1, about 30 mg His-tagged OYE1 can be obtained from a 2 L LB growth.

However, Trx-His-TEV protease could not cleave His-tag and OYE1 through TEV site. When I applied the overnight cleavage mixture to the Ni-NTA column again, the OYE1 could still bind the column. So, for the protein after above purification, the thioredoxin domain, His-tag, and a TEV cleavage site still at the N terminus of OYE1.

To solve this problem, I used pET3b vector to replace pET32a vector to add His-tag and TEV site at the N terminus of the gene for OYE1 only (the next part of this dissertation). Although I could not cleave His-tag through TEV site, the small amount amino acids at the N terminus will not affect the properties and structure of OYE1.

4.3.10 His-TEV-OYE1 cloning into pET3b vector

By using both the forward and reverse primer I succeeded in adding bases encoding a His-tag and a TEV cleavage site at the N terminus of the OYE1 gene and amplifying the His-TEV-OYE1 insert. The His-TEV-OYE1 gene was then subcloned into the pET3b vector. The His-tag greatly facilitates protein purification by Ni-NTA agarose column chromatography.

The DNA sequence of the one which produced the right bands on the 1% agarose gel (one band at about 6 kb and another one at about 1.2 kb) showed 100% identity with the OYE1 sequence.

4.3.11 Overexpression and purification of His-TEV-OYE1
Using the above procedures, His-TEV-OYE1 was overexpressed and purified to 95%. The pure His-TEV-OYE1 was contained in the first 15 mL elute and showed a strong band at about 49 kDa.

4.3.12 Preparation, activity assay and titration with p-chlorophenol of native OYE1, apo-FMN OYE1 and FMN-reconstituted OYE1

The optical spectra of native OYE1 and the FMN-reconstituted OYE1 were very similar (Figure 4.24), both of them displayed the FMN absorbance bands at 381 and 462 nm. However, apo-FMN OYE1 had none of the optical signature of FMN. The FMN content of FMN-reconstituted OYE1 was about 20% higher than that of native OYE1, suggesting that as isolated OYE1 was slightly FMN deficient. Thus, the cofactor FMN could be removed from native OYE1 to generate apo-FMN OYE1 and then add back to apo-FMN OYE1.

Removal of FMN from OYE1 abolished NADPH oxidizing activity, as expected (Figure 4.25). Upon reconstitution of FMN to apo-FMN OYE1, activity was recovered to a 40% higher level than was observed for native OYE1 as-isolated. Thus, the NADPH oxidizing activity is increased when increasing the FMN content and we confirm that our procedures for removing and reconstituting FMN actually do more good for activity than harm.

When titrated with p-chlorophenol, both native OYE1 and FMN-reconstituted OYE1 showed a long-wavelength absorption band centered at 645 nm where the free enzyme has no absorbance. This indicated a charge-transfer transition reflecting p-chlorophenol binding and interaction with FMN (Figure 4.26 (A) and (B)). We can also see the decreases in FMN absorption at 381 and 462 nm that were also reported by Massey group (121). After fitting the titration data to a binding curve for each of native OYE1 and FMN-reconstituted OYE1 (Figure 4.26 (C) and (D)), the dissociation constants calculated (2.7 x 10^-5 M) are the same.

The similar properties of native OYE1 and the FMN-reconstituted OYE1 with respect to catalytic activity and titration with p-chlorophenol indicated the cofactor FMN was removed and added back successfully. Thus, we can remove
14N FMN from 14N OYE and reconstitute with 15N FMN to produce a protein-bound FMN that can be observed by NMR with relatively little background, and can be considered to be fully native.

4.4 Discussion

Both FD and OYE1 were cloned into a vector chosen for the convenience it offers for overexpression and purification of the encoded protein. Thus, FD and OYE1 were in His-tag constructs, to expedite purification. The N signals of the side chains can be seen by 15N NMR at 178-214 ppm in Figure 4.15. The lack of chemical shift dispersion demonstrates that the tag is not structured. 14N proteins with 15N flavin were produced by overexpressing the protein in both LB and MM9, removing the flavins and reconstituting the 15N flavin into unlabeled protein. The precipitate lost during unfolding / folding could cause loss of impurity and increase the purity of protein. Indeed, the flavin to protein ratio increased from 0.78 to 0.96 for old yellow enzyme. We could also have had some apoproteins upon purification that acquired a flavin upon reconstitution. Nonetheless, total protein losses were small. The method we used has been characterized in detail by Mayhew (125), and 90% of the flavin was found to be incorporated into proteins. Thus, the valuable isotopically labeled material is efficiently incorporated into the NMR sample molecule. Deflavinated proteins were reconstituted in over 88% yield and the unflavinated portion is NMR-silent in our experiments and therefore not problematic. Thus, both reconstituted FD and OYE could be produced in high quantity and quality. Given our high expression yield of 50 mg pure protein per 2 liters, in combination with the requirement for 80 mg of protein per SSNMR sample, we can produce a sample with two 2 liters growths.

The spectroscopic signatures of the flavin in both FD and OYE were characterized to verify that our reconstituted proteins have similar flavin binding modes and environments to the native enzymes. However, our reconstituted proteins have higher flavin content than the as-isolated proteins. Moreover, the activity assays enable us to make sure that reconstitution does no harm to the
activity of proteins. The apparent turnover number for NADPH oxidation of native OYE1 was 28 min\(^{-1}\) and that of the FMN-reconstituted OYE1 was 41 min\(^{-1}\). Both of them are very close to the range of previously reported values, 30 min\(^{-1}\) to 51 min\(^{-1}\) (130). Thus, I have generated protein samples that will provide foundation for SSNMR studies of flavins bound in proteins. The amount of \(^{15}\)N from protein backbone is the number of amino acids times \(^{15}\)N natural abundance, 399 (for OYE1) x 0.37\%, which is about 1.5 nitrogens. The number of nitrogens in FMN is 4. However, protein signals occur in the range of 100 to 120 ppm. The location of FMN nitrogens are from 160 to 320 ppm. Thus, in a spectrum of \(^{15}\)N FMN-reconstituted OYE1, the N signals from protein backbone has similar intensity as the ones from FMN but different position.

SSNMR has been applied to study a flavin in bound protein for the first time. Initial observation was based simply on the \(^{15}\)N NMR signals of the four nitrogens in the FMN of uniformly \(^{15}\)N labeled FD. Relatively heavy background from protein backbone made the signals of flavin nitrogens overlap with the signals of the protein backbone. Thus, \(^{14}\)N protein samples are necessary to decrease the protein’s contribution to the spectrum. As calculated above, the amount of \(^{15}\)N natural abundance expected in \(^{14}\)N protein is only about 3 times smaller than the number of \(^{15}\)N in \(^{15}\)N FMN. However, the nitrogens of FMN, especially N5 has very broad tensor, 700 ppm. Thus, it is very easy to tell the signal from FMN nitrogen and the \(^{15}\)N natural abundance of protein backbone.

We would like to test the ability of SSNMR to produce chemical tensor components with sufficient sensitivity to reveal differences in the electronic structure of flavins in different sites. This will pave the way towards understanding flavin reactivity and its manipulation, at a fundamental level.
<table>
<thead>
<tr>
<th>Protein/Position</th>
<th>N1</th>
<th>N3</th>
<th>N5</th>
<th>N10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconstituted OYE1 in solution (pH = 8.5)</td>
<td>194.3</td>
<td>164.1</td>
<td>319.4</td>
<td>162</td>
</tr>
<tr>
<td>Flavodoxin (Desulfovibrio vulgaris) (pH = 8.0)</td>
<td>188.0</td>
<td>159.9</td>
<td>341.1</td>
<td>165.6</td>
</tr>
<tr>
<td>FMN in water (pH = 7.0)</td>
<td>190.8</td>
<td>160.5</td>
<td>334.7</td>
<td>163.5</td>
</tr>
<tr>
<td>TARF in CDCl₃</td>
<td>200.1</td>
<td>159.6</td>
<td>346.0</td>
<td>151.9</td>
</tr>
</tbody>
</table>

**Table 4.1:** Isotopic chemical shifts of $^{15}$N-FMN in various different environments in ppm vs. liquid NH₃ (13).
<table>
<thead>
<tr>
<th>Protein</th>
<th>Vector</th>
<th>Resistant Drug</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>OYE1</td>
<td>pET32a</td>
<td>Ampicillin</td>
<td>Trx-6xHis-NcoI-TEV-OYE1-XhoI</td>
</tr>
<tr>
<td>OYE1</td>
<td>pET3b</td>
<td>Ampicillin</td>
<td>NdeI-6xHis-TEV-OYE1-BamHI</td>
</tr>
<tr>
<td>FD</td>
<td>pET32a</td>
<td>Ampicillin</td>
<td>Trx-6xHis-NcoI-TEV-FD-XhoI</td>
</tr>
<tr>
<td>FD</td>
<td>pET23d</td>
<td>Ampicillin</td>
<td>NcoI-FD-6xHis-XhoI</td>
</tr>
<tr>
<td>NR</td>
<td>pET32a</td>
<td>Ampicillin</td>
<td>Trx-6xHis-NcoI-TEV-NR-XhoI</td>
</tr>
<tr>
<td>NR C216S</td>
<td>pET32a</td>
<td>Ampicillin</td>
<td>Trx-6xHis-NcoI-TEV-NR-C216S-XhoI</td>
</tr>
</tbody>
</table>

**Table 4.2:** List of Plasmids.
<table>
<thead>
<tr>
<th></th>
<th>Native OYE1</th>
<th>Apo-FMN OYE1</th>
<th>FMN-reconstituted</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_{278} )</td>
<td>0.31</td>
<td>0.19</td>
<td>0.37</td>
</tr>
<tr>
<td>( A_{381} )</td>
<td>0.020</td>
<td>0.00081</td>
<td>0.036</td>
</tr>
<tr>
<td>( A_{462} )</td>
<td>0.025</td>
<td>0.0020</td>
<td>0.037</td>
</tr>
<tr>
<td>( \varepsilon_{278} (\text{M}^{-1} \text{cm}^{-1}) )</td>
<td>102,500</td>
<td>80,600</td>
<td>102,500</td>
</tr>
<tr>
<td>( \varepsilon_{462} (\text{M}^{-1} \text{cm}^{-1}) )</td>
<td>10,600</td>
<td>N/A</td>
<td>10,600</td>
</tr>
<tr>
<td>Protein concentration ( \frac{A_{278}}{\varepsilon_{278} \times b} ) (mM)</td>
<td>0.30</td>
<td>0.14</td>
<td>0.22</td>
</tr>
<tr>
<td>FMN content</td>
<td>0.78</td>
<td>0.08</td>
<td>0.96</td>
</tr>
</tbody>
</table>

**Table 4.3:** Comparison of the FMN content of native OYE1, apo-FMN OYE1 and FMN-reconstituted OYE1 (128).
<table>
<thead>
<tr>
<th></th>
<th>Native OYE1</th>
<th>Rebind OYE1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial rate (Au/s)</td>
<td>-1.7364 x 10^{-2}</td>
<td>-2.7755 x 10^{-2}</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>9.5396 x 10^{-5}</td>
<td>2.2213 x 10^{-4}</td>
</tr>
<tr>
<td>Enzyme activity (μmol min⁻¹)</td>
<td>0.1688</td>
<td>0.2725</td>
</tr>
<tr>
<td>Specific activity (mol g⁻¹ min⁻¹)</td>
<td>5.660 x 10^{-4}</td>
<td>8.380 x 10^{-4}</td>
</tr>
<tr>
<td>Turnover number (s⁻¹)</td>
<td>0.46</td>
<td>0.68</td>
</tr>
</tbody>
</table>

**Table 4.4**: comparison of enzyme activity assay of native OYE1, and FMN-reconstituted OYE1 by measuring the rate of NADPH oxidation at 340 nm (129).
Figure 4.1: The ribbon diagram of flavodoxin. It is generated by interpolating a smooth curve through the polypeptide backbone. α-helices are shown as coiled ribbons, β-strands as arrows, and lines for random coils. The direction of the polypeptide chain is indicated by a color ramp along the length of the ribbon. The cofactor, FMN, is shown in stick model. Carbon, oxygen, nitrogen, and phosphorus atoms are colored in gray, red, blue, and orange, respectively. Image is courtesy of Dr. Anne-Frances Miller. (1J8Q.pdb)
Figure 4.2: The ribbon diagram of old yellow enzyme. It is generated by interpolating a smooth curve through the polypeptide backbone. α-helices are shown as coiled ribbons, β-strands as arrows, and lines for random coils. The direction of the polypeptide chain is indicated by a color ramp along the length of the ribbon. The cofactor, FMN, is shown in stick model. Carbon, oxygen, nitrogen, and phosphorus atoms are colored in green, red, blue, and orange, respectively. Image is courtesy of Dr. Anne-Frances Miller. (1OYA.pdb)
Figure 4.3: Plasmid map of Trx-6xHis-TEV-FD-pET32a (base number addresses are provided in parentheses).
Figure 4.4: PCR program for amplifying the TEV-FD insert.
Figure 4.5: SDS-PAGE gel (stained with Coomassie blue) showing overexpression of FD-His in MM9. Lane 1, prestained SDS-PAGE standards, low range (Bio-Rad); lane 2-7, cells before addition of IPTG (0 hours, 1 hour, 2 hours, 3 hours and 4 hours after inoculation, respectively); lane 6-7, cells after addition of IPTG (1 hour and 4 hours after addition of IPTG, respectively); lane 8, kaleidoscope polypeptide standards (Bio-Rad).
Figure 4.6: The Ni-NTA agarose column shows the dark purple semiquinone produced when reduced FD released from the anaerobic interior of *E. coli* becomes oxidized by air.
Figure 4.7: SDS-PAGE gel (stained with Coomassie blue) showing purity of FD-6xHis. Lane 1, kaleidoscope polypeptide standards (Bio-Rad); lane 2-4, purified FD growth in LB at different concentration, 9x, 3x, 1x; lane 5-7, purified FD growth in MM9 at different concentration, 9x, 3x, 1x.
Figure 4.8: Plasmid map of FD-His-pET23d (base number addresses are provided in parentheses).
Figure 4.9: 1% agarose gel analysis of the products of PCR (amplification of FD insert). Lane 1, DNA ladder (pBR322 DNA-BstN I Digest); lane 2, FD insert (452bp).
Figure 4.10: 1% agarose gel analysis of pET23d vector and FD insert after digestion with Ncol and Xhol. Lane 1, DNA ladder (Hind III digest); lane 2, pET23d vector after Ncol and Xhol digestion; lane 3, DNA ladder (pBR322 DNA-BstN I Digest); lane 4, FD insert after Ncol and Xhol digestion.
Figure 4.11: Comparison of the optical spectra of $^{14}$N FD, $^{14}$N apo-FD and $^{15}$N FMN-reconstituted $^{14}$N FD. (The native FD sample was considerably more concentrated than the reconstituted FD.)
Figure 4.12: Flow chart for preparing $^{15}$N FMN reconstituted $^{14}$N FD for SSNMR.
Figure 4.13: Proton wet1d NMR spectrum of FD.
Figure 4.14: $^1$H-$^{15}$N HSQC spectrum of uniformly $^{15}$N-labeled FD at 20 °C.
Figure 4.15: $^{15}$N 1d NMR spectrum of uniformly $^{15}$N-labeled FD.
Figure 4.16: SSNMR 1d CP-MAS spectrum for $^{15}$N FD at 4000 Hz spinning speed. Blue line indicates the potential signal of N5 on flavin ring and the red line on protein backbone.
Figure 4.17: Plasmid map of Trx-His-TEV-OYE1-pET32a (base number addresses are provided in parentheses).
Figure 4.18: PCR program for amplifying TEV-OYE1 insert (His-TEV-OYE1 insert).
Figure 4.19: 1% agarose gel analysis of the products of PCR (amplification of TEV-OYE1 insert). Lane 1, DNA ladder (pBR322 DNA-BstNI Digest); lane 2, TEV-OYE1 insert (1204 bp).
**Figure 4.20:** 1% agarose gel analysis of the His-tagged TEV-OYE1-pET32a DNA plasmid. Lane 1, DNA ladder (Hind III digest); lane 2, specific DNA standard (1100 bp); lane 3, TEV-OYE1-pET32a DNA plasmid after Ncol and Xhol digestion; lane 4, TEV-OYE1-pET32a DNA plasmid before Ncol and Xhol digestion.
Figure 4.21: SDS-PAGE electrophoresis (stained with Coomassie blue) showing overexpression and protein fractions at different stages of purification of Trx-His-TEV-OYE1. Lane 1, prestained SDS-PAGE standards, low range; lane 2, cells before add IPTG; lane 3, cells after add IPTG; lane 4, supernatant after French Press; lane 5, final flow through after bind to Ni-NTA agarose column; lane 6, final wash-through after washing off impurities; lane 7 to lane 9, elution fractions of Trx-His-TEV-OYE1 (5 mL fractions, 10 μL samples loaded of each).
**Figure 4.22:** Plasmid map of His-TEV-OYE1-pET3b (base number addresses are provided in parentheses).
Figure 4.23: 1% agarose gel analysis of the His-TEV-OYE1-pET3b DNA plasmid. Lane 1, His-TEV-OYE1-pET3b DNA plasmid after NdeI and BamHI digestion; lane 2, DNA ladder (Hind III digest); lane 3, His-TEV-OYE1-pET3b DNA plasmid, before NdeI and BamHI digestion.
Figure 4.24: Comparison of the optical spectra of native OYE1, apo-FMN OYE1 and FMN-reconstituted OYE1.
Figure 4.25: Enzyme activity assay of native OYE1, apo-FMN OYE1 and FMN-reconstituted OYE1 by measuring the rate of NADPH oxidation at 340 nm.
Titration of native OYE1 with p-chlorophenol

Titration of FMN-reconstituted OYE1 with p-chlorophenol
Figure 4.26: Spectral changes of titration of (A) native OYE1 and (B) FMN-reconstituted OYE1 with p-chlorophenol. The enzyme concentration was 2.6 x 10^{-2} mM in 0.1 M potassium phosphate, pH=6.45. The concentration of p-chlorophenol was 1 x 10^{-2} mM, 2 x 10^{-2} mM, 3 x 10^{-2} mM, 4 x 10^{-2} mM, 6 x 10^{-2} mM, 60 x 10^{-2} mM respectively. Binding curves of titration of (C) native OYE1 and (D) FMN-reconstituted OYE1 with p-chlorophenol using the absorbance at 645 nm at 25 °C.
Chapter Five: Conclusions and Future Perspectives

5.1 Conclusions

Enzymes not only greatly accelerate the rates of reactions, but they are also generally held to do so with high specificity with respect to the reaction. Acting as biological catalysts, enzymes bring molecules together in such a way that they can react. However, in the absence of enzymes, these reactions could not occur at a rate sufficient to sustain life. Enzymes are also amenable to regulation and they are highly specific, so they support life’s need to constantly evade entropy. Enzymes also mediate electron transfer and the synthesis of electron transfer cofactors (such as quinones, cytochromes etc.). For enzymes that employ chemically versatile cofactors, control over the reaction is all the more important.

Information obtained by solution NMR has had important impact in enzymology. Solution NMR has enabled researchers to determine protein structures and probe interactions between DNA and protein or cofactor and protein. Solution NMR has also enabled researchers to study conformational dynamics, monitor binding equilibrium and pKs, and determine structures of intermediates. SSNMR is also becoming increasingly important in the study of enzymes, providing new insights of structures and interactions (14).

Here, SSNMR has been applied to study a flavin in bound protein for the first time. Based on the study of flavoprotein models and flavoprotein enzyme active sites by SSNMR, we are developing our ability to perceive changes in flavin valence electronics. We also combine SSNMR data with DFT calculation and use DFT calculation to interpret SSNMR data in terms of electron density changes and energy spacing between HOMO and LUMO. Thus we hope to improve our understanding of how the protein specifies flavin reactivity.

In studying TPARF and its complex with the H-bonding partners DBAP or H₂O, SSNMR could detect the changes in CSPVs of N5 on the flavin ring. For a specific H-bonding model, TPARF and DBAP, there was a 10 ppm change upon H-bonding with DBAP for N5 although none of the H-bonds directly engage N5.
In comparison with the 4 ppm change observed by solution state NMR, SSNMR was more sensitive to changes. For observing the weak H-bonding model of TPARF and H$_2$O, there was a -6 ppm change in $\delta_{11}$ upon H-bonding with H$_2$O, although the $\delta_{iso}$ changed by less than 1 ppm. Thus, SSNMR could detect a change that was invisible by solution NMR. SSNMR is a new and sensitive tool studying the flavin ring system, and it could be much more specific and informative than solution state NMR.

SSNMR signatures could be interpreted in terms of flavin electronics and reactivity. No one has done this yet. Based on geometry optimization and NMR GIAO calculations of flavin models, the calculated results replicate the observed changes in CSPVs observed by SSNMR experiments. Thus, the computed electronic structure and molecular orbitals can be used as a means of understanding how a given flavin binding site produces its characteristic activity.

The long term goal is to study FMN in proteins. Because the protein contains nitrogen that is more abundant in the sample than the nitrogen from flavin, producing high quality and quantity samples in which the flavin alone is isotopically labeled is a necessary and important step towards reaching the long term goal. Both FD and OYE had been cloned in a convenient vector wherein the 6xHis-tag greatly facilitates protein purification by Ni-NTA agarose column chromatography. FD and OYE both overexpressed and purified successfully in both unlabeled and labeled media. Thus, by reconstituting $^{15}$N FMN to $^{14}$N protein, we have the means of observing flavin signals while the signals from protein backbone will be relatively silent in the SSNMR study. Preliminary study on the reconstituted samples, such as optical spectra, activity assays and solution NMR studies demonstrated that our protein samples have the native active site, equal or greater FMN content and activity as good or better than that of the native proteins. Moreover, SSNMR study of uniformly $^{15}$N-labeled FD demonstrated that selective labeling of the flavin is necessary in order to be able to observe the flavin $^{15}$N signal despite the presence of the numerous N sites in the protein.
Flavins and flavoproteins represent very important players in biochemistry, but their SSNMR detection remained an unexplored area until the Miller group recently advanced their studies by SSNMR. SSNMR methods can provide atomic-level structural constraints for proteins in forms that cannot be characterized by other high-resolution structural techniques, either because they are not highly soluble or because they do not crystallize. For example, structure of α-spectrin Src-homology 3 (SH3) domain has been determined by solid state magic angle spinning NMR (131). SSNMR also provides insight into chemical properties. This aspect of SSNMR is illustrated by our results on flavins. Although understanding the basis of how the host protein tunes flavin reactivities remains a significant challenge, our SSNMR study on flavin systems has already brought rewards. We can detect changes produced upon H-bonding more sensitively than by solution NMR, and use DFT calculations to interpret our SSNMR data in terms of the electron density distribution, molecular orbitals and energy spacing between orbitals. Studying flavins and flavoproteins is a new application for SSNMR. This application illustrates the potential of SSNMR to address not only flavin structure and interactions, but also function and reactivity.

The range of diagnostic and therapeutic applications for flavins and flavoproteins goes far beyond the use of riboflavin as a vitamin supplement. For example, glutathione reductase and thioredoxin reductase represent important antioxidant principles in the cell and have been identified as promising targets for the design of cytostatic, antiparasitic, and antirheumatic drugs (132). Glutathione reductase serves as a model protein for studying the modifications which result from exposure to endogenous and pharmacologic nitric oxide donors (133). Furthermore, in vitro saturation tests for erythrocytic flavoenzymes, in particular for glutathione reductase, are employed for the diagnosis of riboflavin deficiency (134). The growing knowledge of the flavoproteins has not only had considerable impact in basic science, but will also guide design of novel catalysts and inhibitors including potential therapeutics and diagnostics.

5.2 Future Perspectives
5.2.1 SSNMR comparison of different protein sites and calculation

Flavins in proteins are subject to numerous interactions at once, which may make canceling contributions to chemical shift values. Therefore, we have chosen active sites which resemble each other in many ways, but differ significantly in one. Thus, we would like to compare the SSNMR signal of N5 of FMN bound in FD with that of FMN bound in OYE1 to evaluate the effects of negative charge near N1 in FD vs. positive charge in OYE1. These two proteins produce very different isotropic shifts for N5 (13).

In each protein, future studies will also test the ability of calculations to capture dominant differences between the spectra obtained. The Miller group will calculate NMR CSPVs for highly simplified models of the active sites of FD and OYE1. They will hold heavy atoms fixed at their positions in the relevant crystal structures, neglect the ribityl side chain, but optimize the position of H atoms involved in H-bonds, as well as the flavin before calculation flavin NMR CSPVs. If the calculated CSPVs differ in the same way as they differ in the experimental spectra, then the validated electronic structures of the flavin determined by computation should provide a basis for understanding the different flavin reactivities in different sites.

5.2.2 SSNMR detection of charge transfer complex formation in OYE1

It is also very interesting to test the ability of SSNMR to detect changes in single interactions, with the rest of the active site relatively constant. Because OYE1 binds phenolates and forms striking charge transfer complexes in which the phenolate effectively donates electron density to the flavin. Thus, the electronic structure of flavin is changed. This provides a clear-cut system in which to test SSNMR’s ability to detect changes in electronic structure.

So far, the vast majority of atomic-resolution structural data on proteins come from two techniques, X-ray crystallography and solution state NMR. Both of these techniques are extremely well established and widely practiced. However, X-ray crystallography depends on the availability of high-quality crystals and solution state NMR generally requires high solubility and low
molecular weight. SSNMR techniques do not require crystallinity or solubility and can be applied to proteins with relatively higher molecular weight. Beside these advantages, SSNMR techniques have the unique capability to provide atomic-level information on molecular structures and can provide information that is complementary to the information from X-ray crystallography and solution state NMR. As illustrated by our SSNMR experimental results, SSNMR can provide atomic-level information on flavin interactions. In systems such as other flavoproteins, the atomic-level information from SSNMR can be crucial to understand flavin reactivity.

This dissertation research supplied a foundation for future work on understanding flavin reactivity via SSNMR. It showed that SSNMR has considerable promise of being able to reveal how interactions with protein sites change the flavin electronics. This will pave the way towards understanding flavin reactivity and its manipulation, at a fundamental level.
Appendix A: Setting up and Running Gaussian Jobs

I. Geometry optimization

1. Building molecules

2. **Calculate**=>**Gaussian**…(Gaussian Calculation Setup)

3. Job Type:
   - **Optimization**
     - Optimize to a **Minimum**
     - Calculate Force Constants **Never**

4. Method:
   - Method: **Ground State DFT**… **Default Spin** **B3LYP**
   - Basis Set: **6-311G ++ (2d, 2p)**
   - Charge: **0**
   - Spin: **Singlet**

5. Title:
   - Job Title:

6. Link 0:
   - Link 0 Commands: `%chk=<filename>.chk`

7. General:
   - **Write Connectivity**

8. Guess:
   - Guess Method: **Default**

9. NBO:
   - Type: **None**
   - Checkpoint Save: **Don’t save**

10. Solvation:
    - Model: **Default**
    - Solvent: **Benzene**

11. Submit…
    - Save: `<filename>.com`
II. NMR GIAO calculation
Job Type: NMR GiAO Method

III. Natural Population Analysis
NBO:
Type: NPA only
Checkpoint Save: Don’t save
Additional Keywords: gfprint

IV. Basis set superposition error (BSSE)
Additional Keywords: counterpoise=2
Appendix B: Example .com File

%chk=LUMI-OPT028.chk
# opt b3pw91/6-311++g(2d,2p) geom=connectivity

LumiFlavin- Optimization. Ground State. DFT. Default Spin.
B3PW91/6-311++G(2d,2p)

0 1
C
C 1 B1
C 2 B2 1 A1
C 3 B3 2 A2 1 D1
C 4 B4 3 A3 2 D2
C 5 B5 4 A4 3 D3
C 6 B6 3 A5 2 D4
C 7 B7 4 A6 3 D5
C 8 B8 7 A7 4 D6
H 2 B9 1 A8 6 D7
H 5 B10 4 A9 3 D8
H 9 B11 8 A10 7 D9
N 8 B12 7 A11 4 D10
N 9 B13 8 A12 7 D11
N 7 B14 4 A13 3 D12
N 8 B15 7 A14 4 D13
C 14 B16 9 A15 8 D14
O 9 B17 8 A16 7 D15
O 17 B18 14 A17 9 D16
C 1 B19 2 A18 3 D17
H 20 B20 1 A19 2 D18
H 20 B21 1 A20 2 D19
H 20 B22 1 A21 2 D20
C 6 B23 5 A22 4 D21
H 24 B24 6 A23 5 D22
H 24 B25 6 A24 5 D23
H 24 B26 6 A25 5 D24
C 16 B27 8 A26 7 D25
H 28 B28 16 A27 8 D26
H 28 B29 16 A28 8 D27
H 28 B30 16 A29 8 D28

B1  1.38555129
B2  1.39774924
B3  1.41081619
B4  1.40186661
B5  1.37694552
B6  2.27831863
B7  1.45471126
B8  2.31384162
B9  1.08036568
B10  1.08268991
B11  2.05309163
B12  1.29938958
B13  1.40674806
B14  1.29159580
B15  1.37640384
B16  1.37288035
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B17</td>
<td>1.20900924</td>
<td></td>
</tr>
<tr>
<td>B18</td>
<td>1.20789190</td>
<td></td>
</tr>
<tr>
<td>B19</td>
<td>1.49937158</td>
<td></td>
</tr>
<tr>
<td>B20</td>
<td>1.09274011</td>
<td></td>
</tr>
<tr>
<td>B21</td>
<td>1.09274093</td>
<td></td>
</tr>
<tr>
<td>B22</td>
<td>1.08872088</td>
<td></td>
</tr>
<tr>
<td>B23</td>
<td>1.50080794</td>
<td></td>
</tr>
<tr>
<td>B24</td>
<td>1.09279297</td>
<td></td>
</tr>
<tr>
<td>B25</td>
<td>1.09279297</td>
<td></td>
</tr>
<tr>
<td>B26</td>
<td>1.08867764</td>
<td></td>
</tr>
<tr>
<td>B27</td>
<td>1.45766785</td>
<td></td>
</tr>
<tr>
<td>B28</td>
<td>1.08998431</td>
<td></td>
</tr>
<tr>
<td>B29</td>
<td>1.08444167</td>
<td></td>
</tr>
<tr>
<td>B30</td>
<td>1.08998360</td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>121.23160674</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>118.75142151</td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>119.30596773</td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>122.07620952</td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>91.92420514</td>
<td></td>
</tr>
<tr>
<td>A6</td>
<td>92.80247423</td>
<td></td>
</tr>
<tr>
<td>A7</td>
<td>94.14119436</td>
<td></td>
</tr>
<tr>
<td>A8</td>
<td>118.68583416</td>
<td></td>
</tr>
<tr>
<td>A9</td>
<td>116.66031883</td>
<td></td>
</tr>
<tr>
<td>A10</td>
<td>115.01678748</td>
<td></td>
</tr>
<tr>
<td>A11</td>
<td>125.29553043</td>
<td></td>
</tr>
<tr>
<td>A12</td>
<td>88.64154115</td>
<td></td>
</tr>
<tr>
<td>A13</td>
<td>31.81690016</td>
<td></td>
</tr>
<tr>
<td>A14</td>
<td>115.72602591</td>
<td></td>
</tr>
<tr>
<td>A15</td>
<td>127.78981823</td>
<td></td>
</tr>
<tr>
<td>A16</td>
<td>152.48279678</td>
<td></td>
</tr>
<tr>
<td>A17</td>
<td>123.22463202</td>
<td></td>
</tr>
<tr>
<td>A18</td>
<td>119.49306681</td>
<td></td>
</tr>
<tr>
<td>A19</td>
<td>111.24034877</td>
<td></td>
</tr>
<tr>
<td>A20</td>
<td>111.24031099</td>
<td></td>
</tr>
<tr>
<td>A21</td>
<td>111.46332220</td>
<td></td>
</tr>
<tr>
<td>A22</td>
<td>120.66964574</td>
<td></td>
</tr>
<tr>
<td>A23</td>
<td>111.62121927</td>
<td></td>
</tr>
<tr>
<td>A24</td>
<td>111.62121927</td>
<td></td>
</tr>
<tr>
<td>A25</td>
<td>110.86267072</td>
<td></td>
</tr>
<tr>
<td>A26</td>
<td>118.94107374</td>
<td></td>
</tr>
<tr>
<td>A27</td>
<td>109.99122188</td>
<td></td>
</tr>
<tr>
<td>A28</td>
<td>107.47279386</td>
<td></td>
</tr>
<tr>
<td>A29</td>
<td>109.99130249</td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>0.00000000</td>
<td></td>
</tr>
<tr>
<td>D2</td>
<td>0.00000000</td>
<td></td>
</tr>
<tr>
<td>D3</td>
<td>0.00000000</td>
<td></td>
</tr>
<tr>
<td>D4</td>
<td>180.00000000</td>
<td></td>
</tr>
<tr>
<td>D5</td>
<td>0.00000000</td>
<td></td>
</tr>
<tr>
<td>D6</td>
<td>180.00000000</td>
<td></td>
</tr>
<tr>
<td>D7</td>
<td>180.00000000</td>
<td></td>
</tr>
<tr>
<td>D8</td>
<td>180.00000000</td>
<td></td>
</tr>
<tr>
<td>D9</td>
<td>0.00000000</td>
<td></td>
</tr>
<tr>
<td>D10</td>
<td>180.00000000</td>
<td></td>
</tr>
<tr>
<td>D11</td>
<td>0.00000000</td>
<td></td>
</tr>
<tr>
<td>D12</td>
<td>180.00000000</td>
<td></td>
</tr>
<tr>
<td>D13</td>
<td>0.00000000</td>
<td></td>
</tr>
<tr>
<td>D14</td>
<td>0.00000000</td>
<td></td>
</tr>
</tbody>
</table>

145
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>D15</td>
<td>180.00000000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D16</td>
<td>180.00000000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D17</td>
<td>180.00000000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D18</td>
<td>120.63640166</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D19</td>
<td>-120.63627831</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D20</td>
<td>0.00000000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D21</td>
<td>180.00000000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D22</td>
<td>120.24329079</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D23</td>
<td>-120.24329079</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D24</td>
<td>0.00000000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D25</td>
<td>-180.00000000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D26</td>
<td>-119.65928072</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D27</td>
<td>0.00000000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D28</td>
<td>119.65940397</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

```
1 2 2.0 6 1.5 20 1.0  
2 3 1.5 10 1.0  
3 4 1.5 16 1.5  
4 5 1.5 15 1.5  
5 6 2.0 11 1.0  
6 24 1.0  
7 8 1.0 15 2.0 17 1.0  
8 13 2.0 16 1.5  
9 13 1.5 14 1.0 18 2.0  
10  
11  
12 14 1.0  
13  
14 17 1.5  
15  
16 28 1.0  
17 19 2.0  
18  
19  
20 21 1.0 22 1.0 23 1.0  
21  
22  
23  
24 25 1.0 26 1.0 27 1.0  
25  
26  
27  
28 29 1.0 30 1.0 31 1.0  
29  
30  
31  ```
References:


Vita

Personal information
Name   Dongtao Cui
Date of birth  10/06/1978
Place of birth  Beijing, China

Education
Ph.D. candidate,   Biological Chemistry, University of Kentucky,
Lexington, KY, USA, 2005-present
B.S., Chemistry,   Beijing Normal University,
Beijing, China, 1997-2001

Professional experience
2007-2010   Graduate Research Fellow, Department of Chemistry,
University of Kentucky, Lexington, KY, USA
2005-2007   Teaching Assistant, Department of Chemistry,
University of Kentucky, Lexington, KY, USA
2001-2004   Chemistry Teacher, Beijing No.80 High School,
Beijing, China

Scholastic honors
2007-2010   Research Challenge Trust Fellowship   University of Kentucky

Professional publications
• Boatright, W. L., Jahan, M. S., Walters, B. M., Miller, A.-F., Cui, D., Hustedt,
E. J., and Lei, Q. (2008) Carbon-centered radicals in isolated soy proteins,
Journal of Food Science 73, C222-C226.
• Zhang, P., Cui, D., and Miller, A.-F. Rapid conformational averaging of a
protein at 4 °C in the absence of denaturants. Submitted to J. Am. Chem.
Soc.

• Maly, T., Cui, D., Griffin, R. G., and Miller, A.-F. High-field dynamic nuclear polarization with an endogenous radical. In preparation.


• “High-sensitivity solid-state NMR experiments for studying the active sites of flavoenzymes” XXVIII Midwest Enzyme Chemistry Conference. Chicago, IL, Oct 4, 2008.

• “Modulation of flavin electronics by H-bonding models of enzyme active sites”. 35th Annual Naff Symposium on Chemistry and Molecular Biology. Lexington, KY, Apr 24, 2009.

• “$^{15}$N solid-state NMR chemical shift tensor as a tool for studying the active sites of flavoenzymes” 238th ACS National Meeting, Washington, DC, Aug 16-20, 2009.