COMPUTATION OF TRYPTOPHAN FLUORESCENCE QUENCHING BY AMIDE AND HISTIDINE

by

José Ramón Tusell

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Chemistry

MONTANA STATE UNIVERSITY
Bozeman, Montana

November 2011

`

©COPYRIGHT

by

José Ramón Tusell

2011

All Rights Reserved

`

APPROVAL

of a dissertation submitted by

José Ramón Tusell

This dissertation has been read by each member of the dissertation committee and has been found to be satisfactory regarding content, English usage, format, citation, bibliographic style, and consistency and is ready for submission to The Graduate School.

Patrik R. Callis

Approved for the Department of Chemistry and Biochemistry

Bern Kohler

Approved for The Graduate School

Dr. Carl A. Fox

STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this dissertation should be referred to ProQuest Information and Learning, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

José Ramón Tusell

November 2011

DEDICATION

I dedicate this thesis to my daughter Natalia Tusell, and my parents José Ramón Tusell Vila and Sonia Graciela Angulo de Tusell, without their constant support this thesis would have never been written.

ACKNOWLEDGEMENTS

I would like to acknowledge first and foremost, Patrik Callis for his mentoring and patience during the time I spent working with him. Also I would like to acknowledge Matt Prissel, Alex Grigoropoulos, Ken Mousseau, Cameron Elliot, Karl Sebby, Ryan Hutchenson, Nathan Scott and Anupam Goel for their friendship and support. Finally I acknowledge Erin and Maggie Koester for their companionship, support and patience over the past few months.

TABLE OF CONTENTS

1. INTRODUCTION 1

Variability of Tryptophan Fluorescence Quantum Yields in Proteins 1
Electron Transfer Theory Classical and Quantum Mechanical Approximations 2
Tryptophan Fluorescence Quenching via Electron Transfer to the Peptide Backbone 5
An Ultra Fast Folding Protein: Villin Headpiece Subdomain 7
Probing Helix Formation Kinetics of Ac-W-A-A-A-H^+-NH_2 10
Motivation and Goals 11
Tables and Figures 12

2. COMPUTATION OF THE ELECTRON TRANSFER MATRIX ELEMENT 14

Introduction 14
Methods 16
Molecules Used to Calculate V_{et} 16
Selection Method 18
Linear Combination Method 18
Crossing Method 19
Purifying Mixed Orbitals by Maximizing the Electron Density 19
Results 23
Amide V_{et} 23
Results in Proteins 24
Histidine V_{et} 26
Conclusions 28
Tables and Figures 29

3. EXTENDING MOLECULAR DYNAMICS SIMULATIONS TO 100 NS FOR PREDICTION OF FLUORESCENCE QUANTUM YIELD 39

Introduction 39
Methods 41
Results 44
Ground to Excited State Transition Energies and V_{et} for OPLS-AA 44
Ground to Excited State transition Energies and V_{et} for AMBER99SB-ILDN 46

TABLE OF CONTENTS – CONTINUED

Conclusions 49
Tables and Figures 51

4. CALCULATION OF FLUORESCENCE QUANTUM YIELDS CONTROLLED BY HETEROGENEITY 69

Introduction 69
Model of Tryptophan Fluorescence Quantum Yield for Thermal Unfolding of HP-35 72
Conclusions 74
Tables and Figures 75

5. HISTIDINE AND AMIDE ELECTRON TRANSFER QUENCHING AND PROTEIN FOLDING: Ac-W-A-A-A-H^+-NH_2 AND THE VILLIN HEADPIECE SUBDOMAIN 79

WH5 Simulation 79
Amide Quenching 79
Histidine Quenching 81
HP-35 Simulation 82
Folded Simulations 82
Unfolded Simulation 83
Factors that Determine Histidine Quenching 84
Conclusions 87
Tables and Figures 88

6. SUMMARY AND CONLUSIONS 94

REFERENCES CITED 98

LIST OF TABLES

Table Page

3.1 Summary of experimental quantum yields for the 20 simulated proteins and their PDB code 51

4.1 Parameters used to calculate the quantum yield as a function of temperature. 78

LIST OF FIGURES

Figure Page

1.1 Summarized results as of Callis et al. 42 QM-MM simulations for 24 tryptophan residues from 17 proteins. 12

1.2 Crystal structure of HP-35 as determined by Chiu et al. at a resolution of 1.07 Å 12

1.3 Folded structure of Ac-W-$(A)_3$-H^+-NH_2. 13

2.1 Model system for calculating V_{et} for the histidine molecule 29

2.2 Model system for calculating V_{et} for the closest amide backbone from tryptophan molecule 29

2.3 The singly excited determinants that were used in the Gaussian03 calculation. 30

2.4 Subspace of molecular orbitals in a model donor acceptor electron transfer system 30

2.5 Subspace of molecular orbitals in a model donor acceptor electron transfer system 30

2.6 Contour plot for amide Vet using the amide model system χ_2 set at 90. 31

2.7 Contour plot for amide Vet using the amide model system χ_2 set at -90 31

2.8 Calculation of V_{et} using the crossing method 32

2.9 Comparison between V_{et} calculated using the crossing method and the linear combination method 32

2.10 Time traces for V_{et} for azb, b8r, barn35, bpp, cpl and ctx 33

2.11 Time traces of Vet for d6o, dsb76, dsb126, hp35hse, gcn and lyd126 34

LIST OF FIGURES – CONTINUED

Figure Page

2.12 Time traces of Vet for lyd138, mlt, nscp, rnt, sbc and stn. 35

2.13 Time trace of the V_{et} for trpcage 36

2.14 χ_1 rotamer flips for d6o 36

2.15 χ_1 and ψ rotamer flips for gcn 37

2.16 Average V_{et} for each individual protein 37

2.17 The $\log_{10}$ of the square of V_{et} as a function of distance for imidazolium/indole molecule in WH5 38

2.18 Contour diagrams of V_{et} in cm^{-1} as a function of the two tumble angles θ_1 and θ_2 at a fixed center of mass distance 38

3.1 ZINDO vertical transition energies for the ground to ^1L_a state and ground to CT state during a 25 ns simulation for villin, azb and b8r 51

3.2 ZINDO vertical transition energies for the ground to ^1L_a state and ground to CT state during a 25 ns simulation for b8r, barn35, barn71, bpp, cpl and ctx 52

3.3 ZINDO vertical transition energies for the ground to ^1L_a state and ground to CT state during a 25 ns simulation for d6o, dsb76, dsb126, gcn, lyd126 and lyd138 53

3.4 ZINDO vertical transition energies for the ground to ^1L_a state and ground to CT state during a 25 ns simulation for mlt, nscp, rnt, sbc, stn and trpcage 54

3.5 Average V_{et} for the 20 simulated proteins 55

3.6 Theoretical versus experimental quantum yields for the 20 simulated proteins 55

LIST OF FIGURES – CONTINUED

Figure Page

3.7 ZINDO vertical transition energies for the ground to ^1L_a state and ground to CT state during a 100 ns simulation for azb, b8r, barn35, barn71, bpp and cpl using the AMBER99SB-ILDN force field 56

3.8 ZINDO vertical transition energies for the ground to ^1L_a state and ground to CT state during a 100 ns simulation for ctx, d6o, dsb126, dsb76, hp35 and gcn using the AMBER99SB-ILDN force field 57

3.9 ZINDO vertical transition energies for the ground to ^1L_a state and ground to CT state during a 100 ns simulation for lyd126, lyd138, mlt, nscp, rnt and sbc using the AMBER99SB-ILDN force field..... 58

3.10 ZINDO vertical transition energies for the ground to ^1L_a state and ground to CT state during a 100 ns simulation for stn and trpcage using the AMBER99SB-ILDN force field 59

3.11 χ_1, χ_2, ψ, and the calculated V_{et} as a function of time for azb, b8r, barn35 and barn71..... 59

3.12 χ_1, χ_2, ψ, and the calculated V_{et} as a function of time for bpp, cpl ctx, d6o, dsb126 and dsb76..... 60

3.13 χ_1, χ_2, ψ, and the calculated V_{et} as a function of time for ga1yrf, gcn, lyd126, lyd138, mlt and nscp 61

3.14 χ_1, χ_2, ψ, and the calculated V_{et} as a function of time for rnt, sbc, stn and trpcage 62

3.15 Time traces of the sum of all decay constants, $k_r + k_{nr} + k_{et,amide}$, as a function of time for azb, b8r, barn35, barn71, bpp and cpl 63

LIST OF FIGURES – CONTINUED

Figure Page

3.16 Time traces of the sum of all decay constants, $k_r + k_{nr} + k_{et,amide}$, as a function of time for ctx, d6o, dsb126, dsb76, hp35 and gcn 64

3.17 Time traces of the sum of all decay constants, $k_r + k_{nr} + k_{et,amide}$, as a function of time for lyd126, lyd138, mlt, nscp, rnt and sbc 65

3.18 Time traces of the sum of all decay constants, $k_r + k_{nr} + k_{et,amide}$, as a function of time for stn and trpcage 66

3.19 The average survival curve calculated using the time traces for the sum of the rate constants for azb, b8r, barn35, barn71 and bpp 66

3.20 The average survival curve calculated using the time traces for the sum of the rate constants for cpl, ctx, d6o, dsb76 and dsb126 67

3.21 The average survival curve calculated using the time traces for the sum of the rate constants for hp35, gcn, lyd126, lyd138 and mlt 67

3.22 The average survival curve calculated using the time traces for the sum of the rate constants for nscp, rnt, sbc, stn, and trpcage 68

3.23 Fit obtained for theoretical versus experimental quantum yield using the AMBER99SB-ILDN and the survival curve procedure 68

4.1 Three state kinetic model to calculate tryptophan fluorescence quantum yield in the presence of heterogeneity 75

4.2 Fluorescence quantum yield Φ as a function of k_{ex} 75

LIST OF FIGURES – CONTINUED

Figure Page

4.3 Coefficients c_{GS+} and c_{GS-} as a function of k_{ex} 76

4.4 Contour graph of quantum yield as a function of $\log_{10}(k_{relax}) = k_f + k_b$ and the equilibrium constant K_{eq}. 76

4.5 Kinetic model for calculating tryptophan fluorescence quantum yields as HP-35 thermally unfolds 77

4.6 Change of tryptophan fluorescence as a function of thermal unfolding of HP-35 77

4.7 Time trace for the fluorescence quantum yield after a temperature jump from 328 to 337 K kinetics simulation 78

5.1 Amide energy gap for WH5, the value of the sum of the radiative, nonradiative and electron transfer rate constant (kr and knr are constants set at 0.04 ns^{-1} and 0.08 ns^{-1}) and the Cα RMSD as a function of time using the OPLS-AA force field 88

5.2 Results for histidine electron transfer quenching WH5 and the OPLS-AA force field 89

5.3 Amide energy gap for HP-35, the value of the sum of the radiative, nonradiative and electron transfer rate constant (kr and knr are constants set at 0.04 ns^{-1} and 0.08 ns^{-1}) and the Cα RMSD as a function of time using the OPLS-AA force field 89

5.4 Results for histidine electron transfer quenching HP-35 and the OPLS-AA force field 90

5.5 Center of mass distance between tryptophan and histidine and the Cα root mean square deviation for a villin subdomain simulation starting from the crystal structure (1yrf) at 300K using the AMBER99SB-ILDN force field. 90

LIST OF FIGURES – CONTINUED

Figure Page

5.6 Time traces of the center of mass distance between tryptophan and histidine, χ_1 and χ_2 of tryptophan for HP-35 and the AMBER99SB-ILDN force field..... 91

5.7 Histidine energy gap and the distance between tryptophan and histidine as a function of time for a simulation at 400K starting at the crystal structure of 1yrf using the OPLS-AA force field. 91

5.8 Distance dependence of the Frank-Condon density of states, the rate of electron transfer to histidine and V_{et} between the fluorescent state of tryptophan and the charge transfer state to histidine..... 92

5.9 Protein, Solvent and total electrostatic stabilization of the charge transfer state as a function of the distance separation between tryptophan and histidine residues. 92

5.10 Histidine and tryptophan close conformation that shows water that stabilize or destabilize the charge transfer state to histidine 93

5.11 Histidine and tryptophan far conformation that shows water that stabilize or destabilize the charge transfer state to histidine..... 93

ABSTRACT

Tryptophan fluorescence quantum yield is widely used to follow protein folding for the villin headpiece subdomain (HP-35) and a synthetic peptide Ac-W-$(A)_3$-H^+-NH_2 (WH5). These biopolymers have a histidine residue, which is a potent quencher of tryptophan fluorescence, positioned four amino acids away from tryptophan. Experiments assumed that when folding occurs the fluorescence of tryptophan will be quenched by histidine due to the formation of an α helix. The reliability of folding and unfolding rate constants determined by tryptophan fluorescence has been called into question by several computational studies.

A method to calculate the electron transfer matrix element was developed for different donor/acceptor systems. This method shows that the electron transfer matrix element is sensitive to orientation at close distances and that it does not follow a simple exponential decay with distance. This thesis improved the methods developed by Callis and coworker by conducting 100 ns long simulations for single tryptophan proteins and by modifying the calculation of the fluorescence quantum yield to account for heterogeneity in the calculated electron transfer rates. In addition the method was extended to calculate electron transfer rate constants for histidine quenching by conducting 1 μs long simulations of HP-35 and WH5.

Calculated tryptophan fluorescence quantum yields for the single tryptophan proteins show better agreement with experiment than was previously reported. Simulations for HP-35 and WH5 indicate that the ability of histidine to quench the fluorescence of tryptophan is surprisingly controlled by the energy gap dependence on the distance that separates them. The energy gap dependence on this distance arises from water solvation around histidine. At large distances this solvation decreases the ability of histidine to accept an electron from tryptophan. Different tryptophan/histidine rotamers control this distance. Even when HP-35 is completely folded much of the time histidine does not quench tryptophan fluorescence contrary to the idea that histidine is only close when HP-35 is folded. The calculated fluorescence quantum yield is sensitive to the distribution of close and far conformations and the rate of exchange between these two conformations. This sensitivity gives credibility to the folding/unfolding rates derived from tryptophan fluorescence quantum yields.

CHAPTER 1

INTRODUCTION

Variability of Tryptophan Fluorescence
Quantum Yields in Proteins

Tryptophan fluorescence quantum yields are extremely sensitive to the local protein and solvent environment, varying from 0.33 down to 0.01[1]. For 40 years this sensitivity remained poorly understood. Even with a crystal structure there was no method to predict whether tryptophan would fluorescence strongly or not. Even so, tryptophan fluorescence is used to monitor protein folding, protein conformational changes and protein/ligand binding[1-21].

Tryptophan fluorescence quantum yields can be calculated using the following formula:

$$\Phi = \frac{k_r}{k_r + k_{nr} + k_{et}}$$

where k_r is the radiative decay constant, k_{nr} is the non-radiative decay constant and k_{et} is the rate of electron transfer to the peptide backbone or other acceptors.

To begin to understand the variability in tryptophan fluorescence we need to start with a model, the simplest being the chromophores of tryptophan: 3-methylindole (3MI). 3MI has a quantum yield of 0.33 that is independent of solvent polarity[22]. This insensitivity to solvent polarity does not explain the variability in fluorescence intensity due to different protein and solvent environments. Even so, this molecule is used as a model to determine the radiative and non-radiative rate constants for tryptophan in

proteins, since these processes involve only the indole ring. The radiative decay constant has a value of 4×10^7 s^{-1} and the non-radiative rate constant has a value of 8×10^7 s^{-1}.[23] 3MI does not have other processes that can alter its quantum yield, such as electron transfer.

A better model molecule for tryptophan fluorescence quantum yield is n-acetyl tryptophan amide (NATA). The fluorescence quantum yield of NATA in water and at 300 K is 0.14[24]. The fluorescence quantum yield of NATA is sensitive to solvent polarity, when solvated in dioxane at 300 K its quantum yield is reported as 0.35. [25] The only other process that can alter the observed quantum yield is electron transfer to the amide backbone of NATA. The amide group has been shown to be able to undergo electron transfer for a series of small tryptophan derivatives[24,26-30]. If amide electron transfer is possible and all proteins contain amides the question still remains why some proteins fluoresce more than other. Since all of the variability in fluorescence intensity is deposited on k_{et} we need to talk about the basic theory of estimating electron transfer rate constants.

Electron Transfer Theory Classical and Quantum Mechanical Approximations

To understand electron transfer we must discuss the most important aspects of the theory of electron transfer as first proposed by Marcus. When a system undergoes electron transfer the structure of the molecule changes due to the new charge distribution[31,32]. The change in energy after this structural rearrangement is the nuclear reorganization energy, λ_N. Additionally, the solvent molecules reorient themselves so as

to accommodate the new charge distribution; this is known as the solvent reorganization energy, λ_S. The sum of these two reorganization energies is equal to the reorganization energy λ. The potential energy curves of the reactants and products depend on a reaction coordinate x as[33-36]:

$$V_R = \frac{1}{2}k(x + x_o)^2$$

$$V_P = \frac{1}{2}k(x - x_o)^2$$

In these equations k is the force constant, x_o is the value of x at the minimum point on the potential energy surface (PES). If we take the difference between these two equations we get an expression that states that the potential energy difference between V_R and V_P is directly proportional to the reaction coordinate x[37-39]. Thus for a reaction involving electron transfer V_R and V_P are parabolas that are a function of the reaction coordinate x or the difference between V_R and V_P.

However if the reaction is carried out in solution the important variable is the free energy difference as a function of $V_R - V_P$. In this case the ordinate is the free energy and the abscissa is the reaction coordinate corresponding to the polarization and vibrational energy difference between D-A and D^+-A^- [33]. The free energy barrier is equal to the difference between the bottom of the reactant parabola and the intersection region. If entropy changes are ignored, the free energies become potential energies. And the activation energy is given by[40]:

$$\Delta G^{\mp} = \frac{1}{4\lambda}\left(\lambda + \Delta G^0\right)^2$$

$$\lambda = \frac{1}{2}\left(x_R - x_P\right)^2$$

In the previous expression λ is the reorganization energy and ΔG_0 is the free energy difference between products and reactants.

The rate of electron transfer has an exponential dependence on the activation free energy for electron transfer, i.e. the place where both parabolas intercept. This is best set forth by the following expression[40]:

$$k_{ET} = A\exp\left(-\frac{\Delta G^{\mp}}{k_B T}\right) = A\exp\left(-\frac{\left(\Delta G^0 + \lambda\right)^2}{4\lambda k_B T}\right)$$

This equation is the high temperature classical expression for the rate of electron transfer. The quantum mechanical expression for the rate of transition between an initial and a continuum of final states can be expressed in terms of Fermi's Golden Rule[35]:

$$rate = \frac{2\pi}{\hbar}\left|V_{et}^2\right|\rho(E)$$

V_{et} is the electron transfer matrix element between the initial and final states and ρ(E) is the density of final states. When the golden rule is applied to the process of electron transfer in molecules the previous expression becomes[41]

$$k_{ET}\left(\Delta E_{00}, V_{el}\right) = \frac{2\pi}{\hbar}\left|V_{et}^2\right|\rho_{FC}\left(\Delta E_{00}\right)$$

The new term, ρ_{FC}, is the Franck-Condon weighted density of final vibronic states that are quasi-degenerate with the initial state. This expression was applied by Callis[42-44] and coworkers to explain the variability of tryptophan fluorescence in proteins by electron transfer to the amide backbone.

Tryptophan Fluorescence Quenching via Electron Transfer to the Peptide Backbone

Callis and et al. have worked on predicting fluorescence quantum yields of single tryptophan proteins by studying how the protein electric field enables electron transfer to the amide group of tryptophan. In their papers there are two important states whose energies are monitored as a function of time. The first is the fluorescent state of tryptophan which is called the ^1L_a state and it involves a linear combination of mostly HOMO to LUMO and a smaller contribution of HOMO-1 to LUMO+1 singly excited configurations[45]. The second is the charge transfer state and it involves excitation of an electron from the tryptophan ring to an amide π^* molecular orbital.

Callis and Vivian explained the variability of quantum yields on the basis that charges near tryptophan could stabilize or destabilize the charge transfer state (CT state)[44]. This stabilization or destabilization influences the number of states in quasi-resonance with the fluorescent state of tryptophan (it alters the Frank-Condon density of states). In particular Callis and Vivian found that in general a negative charge near the indole ring will decrease the quantum yield, and that a positive charge will increase the quantum yield. The opposite is true if the charge was localized near the acceptor group.

In another paper[42], Callis and Liu were able to quantitatively predict the fluorescence quantum yields of 24 tryptophan residues in 17 proteins solely on the basis of the energy gap between the fluorescing ^1L_a state and the CT state.

One observation that was made by Callis and Liu was that the average energy gap and quantum yields are directly proportional, i.e. a small energy gap gives a low quantum

yield. Their approach to obtain a reasonable fit assumed a V_{et} that was constant (see figure 1.1). Since V_{et} is roughly proportional to the overlap between the initial and final states[32], this can lead to errors in estimating the quantum yield. This is not a good approximation because the overlap is very sensitive to small changes in orientation. At the time, Callis and coworkers assumed a constant V_{et} because the amide group is at a constant distance from the indole ring, and based on the fact that V_{et} decays exponentially with distance, this assumption seemed feasible.

A particular case of this variation and proportionality of V_{et} to the overlap matrix element was studied by Mikkelssen and Ratner[46]. In their paper they studied the electron transfer between $C_6H_6^-$ and C_6H_6. They showed that V_{et} and the overlap matrix element are proportional to each other. Therefore nodes of the molecular orbitals involved in the electron transfer process should be important at close distances.

An improvement in the calculation of quantum yields was made by Callis et al. by calculating V_{et} for 20 tryptophan residues in 17 proteins[43]. Some of the results in Chapter 2 of this thesis were used in that paper and others gave support to the calculations presented in that paper. It showed in particular that V_{et} is sensitive to the dihedral angles between indole and the amide backbone and that extending the simulations to 150 ps didn't improve the predictive power of the method.

This thesis will use the methods we developed to calculate the V_{et} for 20 different proteins and extend the molecular dynamics simulation time scale to 100 ns. The longer time scale of these simulations should improve sampling of the configurational space of proteins. This is important because some of these protein configurations can have

different electric fields that could alter the calculated fluorescence quantum yield. Furthermore the methods will be applied to discriminate whether the tryptophan fluorescence quantum yield can report the folding and unfolding rates of an ultrafast folding protein: the villin head piece subdomain.

An Ultra Fast Folding Protein: Villin Headpiece Subdomain

The villin headpiece subdomain (HP-35) has been extensively studied both computational and theoretical groups. This protein has the ability to fold in the μs regime, and therefore it can serve as a bridge between experiment and simulation since new methodologies make microsecond long simulation easily obtainable. The structure of the villin head piece has been determined by x-ray crystallography to a 1.07 Å resolution. This structure includes hydrogen atoms, and its PDB code is 1YRF [10]. Figure 1.2 shows the crystal structure for HP-35.

This protein was designed to monitor folding via tryptophan fluorescence[8] since in its structure it contains a tryptophan residue that is four amino acids away from a protonated histidine residue. Histidine cation is a potent quencher of fluoresce via electron transfer[47]. These two amino acids are contained within an α helix, and therefore histidine was assumed to quench only after folding of the helix occurs and the distance between donor and acceptor is about 7.5 Å.

Kubelka et al. conducted a series of experiments in which the folding of HP-35 is probed via the fluorescence quantum yield of tryptophan at different temperatures[8]. In addition, Kubelka et al. did temperature jump experiments to probe the folding dynamics

of HP-35[8]. In these experiments they monitor the time dependence of tryptophan fluorescence after a temperature jump as well as the steady state fluorescence as a function of temperature. They observe two time constants in the relaxation kinetics after a temperature jump experiment. There is a fast 70 ns phase, which is attributed to conformational relaxation in the folded side of the well. The slower 5 μs relaxation is assigned to overall folding relaxation.

The crystal structure of HP-35 pointed to some possible mutations that could stabilize the native state and therefore increase the folding rate of the protein[10]. Kubelka et al. then constructed a mutant of HP-35 in which two lysine residues where mutated to norleucine residues (Lys24Nle and Lys29Nle). This villin headpiece mutant has a high resolution crystal structure and it also shows bi-exponential relaxation kinetics, with a 70 ns and a faster 1 μs long slower phase. The folding time for this mutant is less than one microsecond, being 0.7 μs[9].

Computationally there are a number of studies that have probed the folding of the villin headpiece subdomain. The first attempt was made by Kollman[48] in which they simulated the villin headpiece subdomain in explicit water for 200 ns. During this simulation Kollman et al. only observed partial folding to a molten globule.

After the creation of the folding at home project[49] Pande and coworkers ran hundreds of simulations starting from a number of thermally denatured states constituting 354 μs of total simulation[50]. In this study Pande and coworkers found that certain unfolded structures fold rapidly exhibiting bi-exponential kinetics during the folding process, the fast relaxation rate being about 50 ns and a slow relaxation of about 600 ns.

However, they also found that there are a number of structures that exhibit folding lifetimes that are much longer that the experimentally determined folding rate. Pande and coworkers suggest that tryptophan fluorescence monitors folding lifetimes between structures that are similar to the unfolded structures that actually fold in the simulations and it does not report on folding lifetimes that are much longer. Additionally Pande and coworkers suggest that there might be possible discrepancies between folding and the tryptophan fluorescence signal, pointing to the possibility that tryptophan fluorescence only reports on the formation of the helix that contains the tryptophan residue.

Freddolino et al. ran microsecond long simulation of the both 1yrf and 2f4k[51]. They found that for 1yrf there is a long lived flipped state in which all three helices are formed but helix I is flipped 180° from the native state, raising the question as to whether tryptophan fluorescence can accurately report folding rates. These authors suggest that folding rates reported by tryptophan fluorescence can be faster than actual folding because the tryptophan/histidine distance reaches near native distance quicker than the overall folding of the protein.

Shaw et al. used a specialized computer called Anton to simulate many folding events in single continuous molecular dynamics simulations [52,53]. They find that the folding time according to the average lifetimes in the folded and unfolded states is about 1 μs and is consistent with experimental data. They support the idea that tryptophan fluorescence quantum yield can accurately predict folding rates. They use the decay of the autocorrelation function of tryptophan's solvent accessible surface area to calculate a decay constant of 0.9 μs.

Eaton and coworkers used two different independent methods: one was tryptophan fluorescence quantum yield which monitors the formation of the folded[8-10] and the other was triplet quenching experiments[54]. The triplet quenching experiments were designed to monitor contact formation between tryptophan and a cytokine which can only occur in the unfolded state, this method probes the formation of the unfolded state. Eaton et al. found that the folding rate is the same whether probed by fluorescence or by triplet quenching, giving credit to the validity of the rates determined by fluorescence quantum yield.

Finally, Pande and coworkers constructed a Markov State Model[55] of the villin headpiece subdomain in which they predict the lifetime of contact formation between all possible combinations for triplet-triplet energy transfer experiments[56]. They find that the folding unfolding rates they calculate using contact formation agree well with experiment.

Probing Helix Formation Kinetics of Ac-W-A-A-A-H^+-NH_2

Ac-W-A-A-A-H^+-NH_2 (WH5) is a peptide that was designed to probe the helix formation kinetics by using tryptophan fluorescence quantum yield[7,57]. Figure 1.3 shows a folded structure of WH5. Eaton et al. show that a similar peptide, Ac-W-A-A-A-H^+-(A-A-A-R^+A)$_3$A-NH_2 shows bi-exponential kinetics after a T-jump experiment when followed by tryptophan fluoresce quantum yield[7].

Other groups have measured the relaxation kinetics of WH5 by using tryptophan fluorescence but at a high resolution[58,59]. They also show bi-exponential kinetics with a

< 1ns fast lifetime and a slower 4 ns lifetime. Additionally this peptide has been studied by computational methods[60] that show that kinetics and thermodynamics of protein folding are force field dependent.

Motivation and Goals

The use of tryptophan fluorescence quantum yield temperature jump experiments and in the steady state experiments and the disagreement of simulation with experiments is the main motivation of this thesis, both for HP-35 and WH5. This thesis will be divided into four chapters. Chapter 2 will present how the V_{et} between indole to amide and indole to histidine is calculated and what factors determine its magnitude. Chapter 3 will extend, to the nanosecond timescale, and improve with the application of the methodology developed by Callis and coworkers. Chapter 4 will talk about heterogeneity in tryptophan fluorescence and how the lifetime of individual species can affect the observed quantum yield for different scenarios. Using this methodology we developed a four state kinetics model to calculate both the steady state fluorescence quantum yield and the T-jump experiments. In chapter 5 we will calculate the rate of electron transfer to both the histidine cation and the amide backbone. We will show that amide does a lot of the observed quenching but the histidine cation quenching gives the fluorescence quantum yield its ability to discern between the folded and unfolded states for both the villin headpiece subdomain and WH5.

Tables and Figures

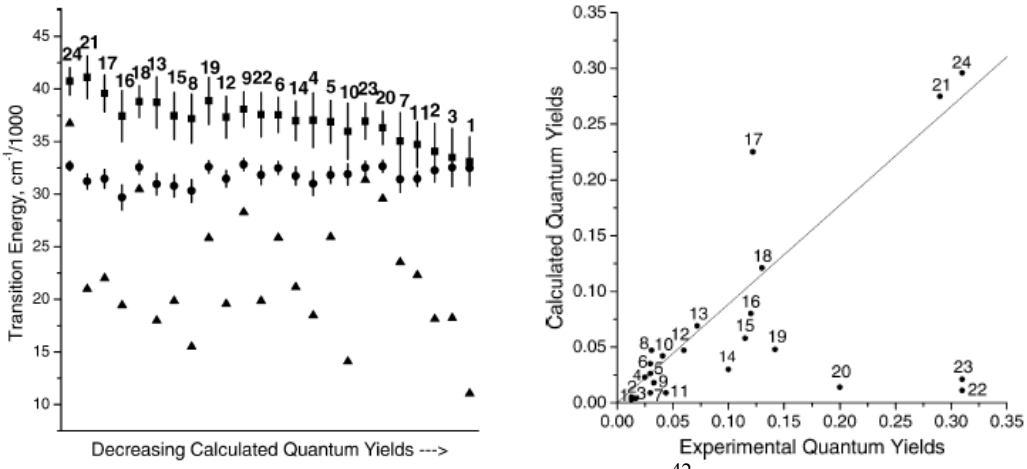

Figure 1.1 Summarized results as of Callis et al. [42] QM-MM simulations for 24 tryptophan residues from 17 proteins. The left figure contains the average CT energy (squares), the average ^1L_a energy (circles) and the energy of the relaxed CT state (triangles). The error bars represent two times the standard deviation. The right figure compares the calculated and experimental yields. The line is a best fit of the data that does not include the three outlying numbers (20,22 and 23), this gave a correlation coefficient of 0.9.

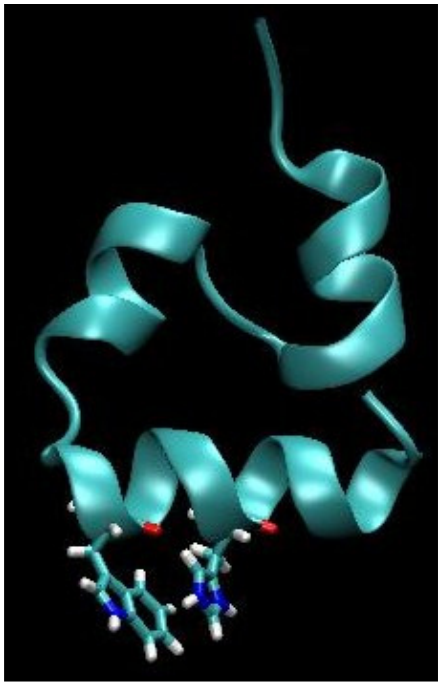

Figure 1.2 Crystal Structure of HP-35 as determined by Chiu et al. at a resolution of 1.07 Å. This figure was made using VMD[61].

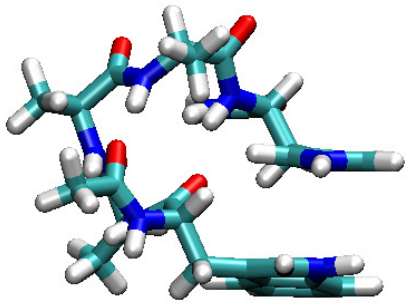

Figure 1.3 Folded structure of Ac-W-A-A-A-H^+-NH_2. This peptide is designed using the same principle as HP-35, tryptophan and histidine are separated by three alanine residues. Therefore, as the protein undergoes folding we expect to see a decrease in tryptophan fluorescence quantum yield. Figure was created using VMD[61].

CHAPTER 2

COMPUTATION OF THE ELECTRON TRANSFER MATRIX ELEMENT

Introduction

Fermi's golden rule[35,41] gives the magnitude of the electron transfer rate constant k_{et} as a function of the electronic matrix element that couples the initial and final states and the Frank-Condon weighted density of states:

$$k_{et}(\Delta E_{00}) = \frac{2\pi}{\hbar} V_{et}^2 \rho_{FC}(\Delta E_{00})$$

where k_{et} is the rate constant for electron transfer, ΔE_{00} is the energy difference between the zero point vibrational energies of the initial and final state, V_{et} is the electron transfer matrix element between the initial and final states and ρ_{FC} is the Frank-Condon weighted density of states.

V_{et} is given from the following expression:

$$V_{et} = \langle f | \hat{H} | i \rangle$$

The final state, $|f\rangle$, when applied to tryptophan fluorescence is the charge transfer state to an acceptor (either the amide backbone or the histidine cation), herein called the CT state. In this state an electron is transferred from the tryptophan ring to an unoccupied molecular orbital of the acceptor.

The initial state, $|i\rangle$, is the fluorescent state of tryptophan which is composed of a linear combination of singly excited configurations. Mostly HOMO to LUMO and a

little of HOMO-1 to LUMO+1[45], this state is herein called the ^1L_a state. Therefore V_{et} is defined as:

$$V_{et} = \left\langle CT \left| \hat{H} \right| {}^1 L_a \right\rangle$$

Many papers have published and used the well known distance dependence of V_{et} with distance[33,62-65]. This dependence is exponential since the orbital overlap decays exponentially with distance[46,65,66]:

$$V_{et}(r) = A \exp\left[-\beta \frac{R}{2} \right]$$

In this expression β determines how fast V_{et} decreases with donor acceptor separation R . The value of β ranges from 1.00 to 3.00 $Å^{-1}$ [65] and it also depends on the relative orientation between donor and acceptor molecules.[62] This β can be different from the one used to describe the exponential decay of electron transfer rates with distance if the Frank-Condon weighted density of states depends on distance. This equation holds for atom to atom electron transfer because there is no orientation dependence; the same cannot be said about molecules. At close distances V_{et} is sensitive to the orientation. This orientation dependence comes from the dependence of V_{et} with the overlap between donor and acceptor orbitals[46]. Even at a set distance we can see large fluctuations in the order of hundreds of wavenumbers. After a certain distance V_{et} decays exponentially as expected and can be modeled by an exponential fit.

Tryptophan fluorescence has been used to monitor the folding of an ultrafast folding protein, the villin headpiece subdomain[8-10,54,67]. In this protein residue 27 was mutated from asparagine to histidine. Residue 27 is four amino acids down from tryptophan in a α-helix. It is known that free histidine quenches the fluorescence of

tryptophan at a collisional rate[47]. These amino acids are contained in a helical section of the protein. Therefore, as the protein folds the tryptophan/histidine distance will become small and the fluorescence of tryptophan will diminish in the folded conformation. Distance dependence of V_{et} will play a crucial role in determining the ability of histidine to quench the fluorescence of tryptophan.

V_{et} has played an important role in the development of methods to determine the fluorescence quantum yield in single tryptophan proteins. Callis and coworkers developed a theory which used a universal V_{et} constant of 10 cm^{-1} and a correction factor of -4000 cm^{-1} that is added to the energy gap difference between the CT state and the ^1L_a state.[42] Subtracting 4000 cm^{-1} from the energies was physically unreasonable because they expected the energy of the CT state to be lower since the amides are in the CT state geometry. Later work showed that the V_{et} is dependent on the protein and that different proteins exhibit different behaviors[43]. V_{et} is controlled mainly by dihedral angles of tryptophan. The most important of these is χ_1 (see figure 2.2 for a definition of this dihedral) since it determines if the V_{et} is large.

Methods

Molecules Used to Calculate V_{et}

To study electron transfer reactions from the excited state of indole to the amide backbone or the imidazolium cation of histidine we constructed two model systems. These are shown in figures 2.1 and 2.2. The geometries were taken directly from MD simulation snapshots unless otherwise noted. We studied three possible methods to

calculate V_{et}: the selection method (which ignores orbital mixing), the linear combination/maximization of electron density method (which takes into account orbital mixing by purifying the orbitals) and the avoided crossing method.

Selection Method

In this method we focus on finding a virtual molecular orbital (MO) that has the highest pi molecular orbital coefficients on carbon, oxygen and nitrogen atoms of amide. As before the CT state is described as a single excitation from the HOMO of the indole ring to the amide orbital. Gaussian03[68] calculates the Hamiltonian matrix on the configuration interaction singles (CIS) basis using single excitations from HOMO to LUMO + 2 (MOs 50 to 53). These excitations are shown schematically on figure 2.3.

The singlet adapted configuration that describes the ^1L_a state is approximately given by:

$$\left|{}^1L_a\right\rangle = \frac{|1\rangle + |4\rangle}{\sqrt{2}} = \frac{\left|...49\overline{49}50\overline{51}...\right| + \left|...49\overline{49}51\overline{50}...\right|}{\sqrt{2}}$$

While the singlet adapted configuration that describes the CT is:

$$\left|CT\right\rangle = \frac{|2\rangle + |5\rangle}{\sqrt{2}} = \frac{\left|...49\overline{49}50\overline{52}...\right| + \left|...49\overline{49}52\overline{50}...\right|}{\sqrt{2}}$$

If we calculate V_{et} between these two states and remembering that the Hamiltonian matrix elements in the CI basis are given by:

$$H_{i,j} = \langle i|H|j\rangle$$

Where |i> and |j> are singly excited configurations, we find in this case that V_{et} is simply given by:

$$V_{et} = \left\langle {}^1L_a \middle| H \middle| CT \right\rangle$$

$$V_{et} = \left\langle \frac{\left|...49\overline{49}50\overline{51}...\right| + \left|...49\overline{49}51\overline{50}...\right|}{\sqrt{2}} \middle| H \middle| \frac{\left|...49\overline{49}50\overline{52}...\right| + \left|...49\overline{49}52\overline{50}...\right|}{\sqrt{2}} \right\rangle$$

$$= \frac{1}{2}\left\langle \left|...49\overline{49}50\overline{51}...\right| \middle| H \middle| \left|...49\overline{49}50\overline{52}...\right| \right\rangle + \frac{1}{2}\left\langle \left|...49\overline{49}50\overline{51}...\right| \middle| H \middle| \left|...49\overline{49}52\overline{50}...\right| \right\rangle$$

$$+ \frac{1}{2}\left\langle \left|...49\overline{49}51\overline{50}...\right| \middle| H \middle| \left|...49\overline{49}50\overline{52}...\right| \right\rangle + \frac{1}{2}\left\langle \left|...49\overline{49}51\overline{50}...\right| \middle| H \middle| \left|...49\overline{49}52\overline{50}...\right| \right\rangle$$

$$= \frac{1}{2}\left[H_{12} + H_{15} + H_{42} + H_{45}\right] = H_{12} + H_{15}$$

In the above expression H_{12} and H_{15} are the matrix elements between the singly excited configurations |1> and |2>, and |1> and |5> respectively of the Hamiltonian in the CI basis of the singly excited determinants.

Linear Combination Method

Since the amide orbital is mixed with orbital 53, a procedure was made in which we maximized the molecular orbital coefficient on the amide carbon by taking a linear combination of orbitals. We took the linear combination which maximizes the amide carbon pi content. Therefore through the use of this method it was possible to localize this orbital on the carbon and calculate V_{et} as the sum of the individual CI matrix elements weighted by their corresponding linear combination coefficients. Additionally this method was extended to calculate V_{et} between the imidazolium ion and indole as will be shown in the next subsection.

Crossing Method

This method has been applied by other groups to calculate V_{et}. In all of this work the system is perturbed with an external field so as to simulate the effect of the

fluctuation of the solvent polarization that brings the two states into resonance. In a particular case, Gao and Marcus[69] applied this method to calculate V_{et} in a series of organic compounds.

In the work by Gao and Marcus they vary the distance of a F^- from the acceptor site. This variation stabilizes the CT state in their system. As in our system there is a singlet excited state as an initial state that is hardly affected by the presence of the F^-. When these two states come into near resonance they have an avoided crossing and it is known that the minimum in energy difference between these two states is $2V_{et}$.

With these ideas in mind we situated a small positive charge equidistant from the carbon, nitrogen and oxygen of the amide. This small charge was varied to perturb the system and lower the CT state energy so as to make an avoided crossing with ^1L_a state. We monitored the singlet eigenvectors of the Hamiltonian in the CI basis, particularly their energy as a function of charge. When this energy difference was a minimum we can calculate V_{et}.

Purifying Mixed Orbitals by Maximizing the Electron Density

In this method we try to get the linear combination of MOs that purify the imidazolium cation LUMO. The procedure starts by obtaining a pure LUMO of the imidazolium cation by running it without indole. This LUMO is then projected onto the LUMOs of the imidazolium and indole molecule, a density matrix is then constructed and diagonalized to produce a set of eigenvalues and eigenvectors. The eigenvector with an eigenvalue of ~1 is the linear combination of molecular orbitals that produce the pure LUMO on the imidazolium cation. V_{et} is simply given by the linear combination

coefficients multiplied by the respective Hamiltonian matrix elements in the CI representation. The derivation of this formula is as follows.

If we have a system in which MO 1 is a pure *HOMO of donor*, MO 2 is a pure *LUMO of donor*, MO 3 is a mixed *LUMO acceptor* and MO 4 is a mixed *LUMO of acceptor* (Figure 2.4), then the singly excited configurations using these MOs are the following:

$$|1\rangle = |2 \quad \bar{1}| \quad |4\rangle = |1 \quad \bar{2}|$$
$$|2\rangle = |3 \quad \bar{1}| \quad |5\rangle = |1 \quad \bar{3}|$$
$$|3\rangle = |4 \quad \bar{1}| \quad |6\rangle = |1 \quad \bar{4}|$$

If we now change the MOs so that MO A is still a pure *HOMO of donor* (ie MO A ≈ MO 1), MO B is still a pure *LUMO of donor* (ie MO B ≈ MO 2), MO C is now a pure *LUMO of acceptor* and MO D is also pure (Figure 2.5). Additionally MO C is a linear combination of MO 3 and MO 4:

$$C(x_i) = c_1 3(x_i) + c_2 4(x_i)$$

were x_i is the electronic coordinates of electron i. In this calculated MO basis set i (i=A,D) we can express the CIS Hamiltonian in this subspace with the following singly excited configurations:

$$|A\rangle = |B \quad \bar{A}| \quad |D\rangle = |A \quad \bar{B}|$$
$$|B\rangle = |C \quad \bar{A}| \quad |E\rangle = |A \quad \bar{C}|$$
$$|C\rangle = |D \quad \bar{A}| \quad |F\rangle = |A \quad \bar{D}|$$

V_{et} is given by:

$$V_{et} = \langle i|H|f\rangle$$

Where $|i\rangle$ is equal to:

$$|i\rangle = \frac{1}{\sqrt{2}}\left[|1\rangle + |4\rangle\right] = \frac{1}{\sqrt{2}}\left[|A\rangle + |D\rangle\right]$$

and $|f\rangle$ is:

$$|f\rangle = \frac{1}{\sqrt{2}}\left[|B\rangle + |E\rangle\right]$$

$$|B\rangle = \left|C \quad \overline{A}\right| = \frac{1}{\sqrt{2}}\begin{vmatrix} C(x_1) & \overline{A}(x_1) \\ C(x_2) & \overline{A}(x_2) \end{vmatrix} = \frac{1}{\sqrt{2}}\begin{vmatrix} c_1 3(x_1) + c_2 4(x_1) & \overline{A}(x_1) \\ c_1 3(x_2) + c_2 4(x_2) & \overline{A}(x_2) \end{vmatrix}$$

$$= \frac{1}{\sqrt{2}}\left[(c_1 3(x_1) + c_2 4(x_1))\overline{A}(x_2) - (c_1 3(x_2) + c_2 4(x_2))\overline{A}(x_1)\right]$$

$$= \frac{1}{\sqrt{2}}\left[c_1 3(x_1)\overline{A}(x_2) + c_2 4(x_1)\overline{A}(x_2) - c_1 3(x_2)\overline{A}(x_1) - c_2 4(x_2)\overline{A}(x_1)\right]$$

$$= \frac{1}{\sqrt{2}}\left[c_1 3(x_1)\overline{A}(x_2) - c_1 3(x_2)\overline{A}(x_1) + c_2 4(x_1)\overline{A}(x_2) - c_2 4(x_2)\overline{A}(x_1)\right]$$

If MO A is equal to MO 1 then:

$$\frac{1}{\sqrt{2}}\left[c_1 3(x_1)\bar{1}(x_2) - c_1 3(x_2)\bar{1}(x_1) + c_2 4(x_1)\bar{1}(x_2) - c_2 4(x_2)\bar{1}(x_1)\right]$$

$$= \left[c_1\left\{\frac{1}{\sqrt{2}}\left[3(x_1)\bar{1}(x_2) - 3(x_2)\bar{1}(x_1)\right]\right\} + c_2\left\{\left[\frac{1}{\sqrt{2}}4(x_1)\bar{1}(x_2) - 4(x_2)\bar{1}(x_1)\right]\right\}\right]$$

$$|B\rangle = c_1\left|3 \quad \bar{1}\right| + c_2\left|4 \quad \bar{1}\right| = c_1|2\rangle + c_2|3\rangle$$

$$|E\rangle = |A \quad \overline{C}| = \frac{1}{\sqrt{2}}\begin{vmatrix} A(x_1) & \overline{C}(x_1) \\ A(x_2) & \overline{C}(x_2) \end{vmatrix} = \frac{1}{\sqrt{2}}\begin{vmatrix} A(x_1) & c_1\overline{3}(x_1)+c_2\overline{4}(x_1) \\ A(x_2) & c_1\overline{3}(x_2)+c_2\overline{4}(x_2) \end{vmatrix}$$

$$= \frac{1}{\sqrt{2}}\left[A(x_1)\left(c_1\overline{3}(x_2)+c_2\overline{4}(x_2)\right)-A(x_2)\left(c_1\overline{3}(x_1)+c_2\overline{4}(x_1)\right)\right]$$

$$= \frac{1}{\sqrt{2}}\left[c_1A(x_1)\overline{3}(x_2)-c_1A(x_2)\overline{3}(x_1)+c_2A(x_1)\overline{4}(x_2)-c_2A(x_2)\overline{4}(x_1)\right]$$

$$= \frac{1}{\sqrt{2}}\left[c_1 1(x_1)\overline{3}(x_2)-c_1 1(x_2)\overline{3}(x_1)+c_2 1(x_1)\overline{4}(x_2)-c_2 1(x_2)\overline{4}(x_1)\right]$$

$$= \left[c_1\left\{\frac{1}{\sqrt{2}}\left[1(x_1)\overline{3}(x_2)-1(x_2)\overline{3}(x_1)\right]\right\}+c_2\left\{\frac{1}{\sqrt{2}}\left[1(x_1)\overline{4}(x_2)-1(x_2)\overline{4}(x_1)\right]\right\}\right]$$

$$|E\rangle = c_1|1 \quad \overline{3}|+c_2|1 \quad \overline{4}| = c_1|5\rangle+c_2|6\rangle$$

Therefore $|f\rangle$ can be expressed as:

$$|f\rangle = \frac{1}{\sqrt{2}}\left[|B\rangle+|E\rangle\right] = \frac{1}{\sqrt{2}}\left[c_1|2\rangle+c_2|3\rangle+c_1|5\rangle+c_2|6\rangle\right]$$

With this we can write V_{et} as:

$$V_{et} = \langle i|H|f\rangle$$

$$\frac{1}{2}\left(\langle 1|+\langle 4|\right)H\left(c_1|2\rangle+c_2|3\rangle+c_1|5\rangle+c_2|6\rangle\right)$$

$$= \frac{1}{2}\left[c_1\langle 1|H|2\rangle+c_2\langle 1|H|3\rangle+c_1\langle 1|H|5\rangle+c_2\langle 1|H|6\rangle+c_1\langle 4|H|2\rangle+c_2\langle 4|H|3\rangle+c_1\langle 4|H|5\rangle+c_2\langle 4|H|6\rangle\right]$$

$$= c_1\left[\frac{1}{2}\left(H_{1,2}+H_{1,5}+H_{4,2}+H_{4,5}\right)\right]+c_2\left[\frac{1}{2}\left(H_{1,3}+H_{1,6}+H_{4,3}+H_{4,6}\right)\right]$$

The terms in parenthesis can be expressed as follows from the mixed MOs:

$$V_{1,2} = \frac{1}{2}\left(\langle 1|+\langle 4|\right)H\left(|2\rangle+|5\rangle\right) = \frac{1}{2}\left(H_{1,2}+H_{1,5}+H_{4,2}+H_{4,5}\right)$$

$$V_{1,3} = \frac{1}{2}\left(\langle 1|+\langle 4|\right)H\left(|3\rangle+|6\rangle\right) = \frac{1}{2}\left(H_{1,3}+H_{1,6}+H_{4,3}+H_{4,6}\right)$$

From this the final expression for V_{et} is:

$$V_{et} = c_1 V_{1,2} + c_2 V_{1,3}$$

Results

Amide V_{et}

Amide V_{et} depends on the dihedral angles χ_1, χ_2 and ψ (please see figure 2.2 for a definition of these dihedral angles). The dihedrals ψ and χ_1 where scanned from -180° to +180° in 5° increments. χ_2 was scanned in either the +90° or -90° angle. At each of these structures a calculation of V_{et} was carried out, using the maximization of electron density on the carbonyl carbon of amide. Results for -90° is shown in figure 2.6 and results for +90° is shown in figure 2.7. It is interesting that V_{et} is sensitive to both χ_1 and ψ. This dependence of V_{et} with χ_1, χ_2 and ψ is due to large non-Condon behavior[46,70-72]. This dependence has been observed previously and is due to through-bond and through-space interaction. In these two contour plots the region near χ_1 = -60° exhibits the highest values for V_{et} and this is due mostly to through bond interaction since calculations done on a space-separated amide and indole ring shows only a small V_{et} (0-25cm^{-1}). This means that the large interaction at -60° is due to through-bond V_{et}. The region in which ψ is -40° and +140° and χ_1 is -60° maximize the through bond "π conjugation" between donor and acceptor therefore increasing the through-bond component in this area[73,74].

In order to corroborate our linear combination method we also conducted a study similar to Gao and Marcus[69]. We added a charge that was equidistant between C, N and O of carbonyl amide and then varied the charge slowly and monitored the exited state energies according to CIS/3-21G. Figure 2.8 shows an example calculation. The

energies of ^1L_a (the excited fluorescent state of tryptophan) and S_2 (another singlet excited state) do not vary significantly with charge since these states have small dipoles. The CT state on the other hand varies quite rapidly with the charge since it has a large dipole. The positive charge applied near the carbonyl amide stabilizes the CT state, therefore decreasing its energy as a function of charge. As the CT state decreases its energy, it first reaches the S_2 state; an avoided crossing occurs between these two states and the curve that was the CT state becomes S_2, while the curve that was the S_2 state becomes the CT state. As the charge is increased there is a crossing between the CT state and the ^1L_a state. The minimum energy difference at the crossing region between the CT state and 1La state is equal to 2 V_{et}.

This procedure was applied to the model system using the following set of the dihedrals, χ_1 = -60°,χ_2 = -90°, and ψ was varied from -180° to +180° in 5° increments. The results of these calculations are shown in figure 2.9. This figure shows the comparison between the crossing method and maximization of amide density. It shows that both methods give the same qualitative behavior as a function of ψ.

Results in Proteins

V_{et} for each of the proteins was calculated using the linear combination method that maximizes the density on the amide carbonyl. Each protein was simulated for 25 ns using Gromacs[75] (for more details about the simulation parameters see chapter 3). For each protein a calculation was carried out every 3.125 ns giving a total 4000 snapshots. The coordinates for the molecule were taken directly from each snapshot and capped with hydrogen. The V_{et} as a function of time for azb (PDB: 1AZB), b8r (PDB: 1B8R), barn35

(PDB: 1A2P), bpp (PDB: 2BPP), cpl (PDB: 2CPL) and ctx (PDB: 1CTX) are shown in figure 2.10. d6o (PDB: 1D6O), dsb76 (PDB: 1FVK), dsb126 (PDB: 1DSB), hp-35hse (PDB: 1YRF), gcn (PDB: 1GCN) and lyd126 (PDB: 1LYD) V_{et} time traces are shown in figure 2.11. In figure 2.12 one can see the time trace of V_{et} for lyd138 (PDB: 1LYD), mlt (PDB: 2MLT), nscp (PDB: 2SCP), rnt (PDB: 9RNT), sbc (PDB: 1SBC) and stn (PDB: 1STN). Finally figure 2.13 shows V_{et} time trace for trpcage1 (PDB: 1L2Y). From some of the time traces we see that there are some proteins that always have a large V_{et}: stn, lyd126 and d6o for part of the time. The behavior of d6o is unique since it shows heterogeneity in V_{et} it goes from a region of very high V_{et} ~800 cm^{-1} to a region were V_{et} is low ~200 cm^{-1}. This is due to a rotamer flip in the χ_1 angle, see figure 2.14. It goes from a region in which χ_1 is mainly -60° and has a high V_{et} due to through bond V_{et}, to a region in which V_{et} is more variable in time. The average ψ angle for tryptophan in this protein is ~ -45° and χ_2 = ~90°. If we look at figure 2.6 we see that for the χ_1 = -60 region V_{et} is strong and never becomes small. As we alter χ_1 to -180° in the contour plot we see that this is a region in which V_{et} fluctuates from 0 cm^{-1} to about 300 cm^{-1}. This is consistent with the behavior observed in the time traces of V_{et} for d6o.

Gcn is another protein that shows heterogeneity in V_{et} as a function of time. In the last 5 ns of the simulation gcn show a marked increase in V_{et}, the previous 20 ns shows a region in which V_{et} fluctuates between 0 and 200 cm^{-1}. In contrast the last 5 ns of the simulation V_{et} shows a V_{et} that never goes to zero and V_{et} goes from ~200 cm^{-1} to about 1100 cm^{-1}. This change in behavior is also explained through a change in the χ_1 and ψ angles see figure 2.15. As χ_1 changes to +60°, ψ also changes its value to +30°.

This set V_{et} in a region in which we expect to see fluctuations between 100 cm^{-1} to about 700 cm^{-1}, which is roughly what we see in the time trace for gcn.

Figure 2.16 shows the average value of V_{et} ± associated standard deviation. This is the value that we chose to use in the actual computation of quantum yields for each of the twenty proteins. However there are some limitations with such a procedure because as we have shown above there are some proteins that show heterogeneity in V_{et} due to rotamer flips. In the future we could possibly improve our predictions of quantum yields by assigning V_{et} via a table that has V_{et} for amide as a function of χ_1, χ_2 and ψ.

Histidine V_{et}

The electron transfer matrix element is one of the factors that determine the ability of the histidine cation to quench the fluorescence of tryptophan via electron transfer from the indole ring to the imidazolium ring. V_{et} decreases exponentially with distance and has played a crucial role in the interpretation of protein folding experiments of HP-35[8,9,50,51,67]. To understand how V_{et} changes with distance we used a small peptide Ac-Trp-$(Ala)_3$-His^+-NH_2, hereby called WH5. This peptide has tryptophan and histidine separated by three alanine residues and are therefore four amino acids away, so it is going to be a good model for HP-35 since tryptophan and histidine are also separated by four amino acids[7,57-59]. We conducted a 1μs simulation of WH5 and calculated V_{et} between tryptophan and histidine using the coordinates of each frame and cropping the protein so that the resulting system is that of figure 2.2. The calculation of the matrix element was done using the linear combination method procedure to account for mixing of the orbitals.

Figure 2.17 shows how the $\log_{10}$ of V_{et} squared varies with the center of mass distance between the two rings. It shows two regions one between 3.5 Å and 6.0 Å that show that the magnitude of V_{et} doesn't change significantly within this distance. After 6.0 Å the exponential decay with distance begins as expected. This figure was used to assign values for V_{et} as a function of distance for HP-35 and WH5 without the need of conducting expensive quantum mechanical calculations. The model assumes a constant V_{et} of 250 cm^{-1} for any snapshoot that has a distance between 3.5 Å and 6.0 Å. After 6.0 Å V_{et} was assigned an exponential fall off and after 15 Å it was assumed to have a value of 0 cm^{-1}.

$$V(x)=\begin{cases} 250cm^{-1} & 3\le x\le 6 \\ 7x10^{11}cm^{-1}\exp\left(-1.8\dfrac{x}{2}\right) & 6\le x\le 15 \\ 0cm^{-1} & x\ge 15 \end{cases}$$

Another study was conducted to show how orientation is important when tryptophan and histidine are close together. In this study tryptophan was set on the x-y plane so that all of its atoms are on the plane. Imidazolium was placed on top of indole in a sandwich at a distance of 5 Å, 6 Å, 7 Å and 8 Å (both ring planes are parallel to each other). Imidazolium was then rotated while holding indole position constant. The rotation was defined by two angles θ_1 and θ_2. Each of these angles were rotated from 0° to 360°, thereby scanning the entire space in which the center of mass distance between indole and imidazolium would be 5 Å, 6 Å 7 Å or 8 Å. Results for each of these distances can be seen in figure 2.18.

Figure 2.18 shows that orientation dependence is very strong at 5 Å showing regions in which V_{et} can be large (close to 1200 cm^{-1}) and regions that have small values

for V_{et}. The regions that correspond to small V_{et} have structures in which the rings are perpendicular to one another mainly $\theta_2 = 90°$ and 270°. As we increase the distance the same pattern remains (see figure 2.18 B when the distance is 6.0 Å) but the magnitude of V_{et} decreases and so does its range. Once we get to 8 Å we see that the contour graph is about the same but the magnitude of V_{et} has decreased by a factor of 100.

Conclusions

We have shown a method of calculating from first principles the value of V_{et} for amide and histidine. Results show that the value of V_{et} is an order of magnitude higher than previously expected. Since V_{et} is larger, Callis and coworkers had to add 4500 cm^{-1} to the energy gap difference between the CT and ^1L_a state[43]. This leads to a more sensitive Frank-Condon factor that gives rise to larger variation in the calculated electron transfer rates. If the electron transfer rate constants can vary significantly they can give rise to heterogeneity and can change the observed quantum yield. Different proteins show different values for V_{et} that are controlled by the dihedral angles that connect amide to the tryptophan ring. Mainly if through bond conjugation exists V_{et} is large and never zero, some examples include d6o, stn and lyd126. Other proteins have dihedral angle combination which sets them in a region where V_{et} fluctuates and is occasionally found to be zero. To calculate V_{et} for histidine we constructed a model of how we expect to decay with distance. Also we have shown that orientation determines the value of V_{et} when the distance between tryptophan and histidine is between 5 to 8 Å.

The exponential decay constant, $\beta = 1.80$ Å^{-1} is comparable to the values obtained by Cave and coworkers for a series of compounds.

Tables and Figures

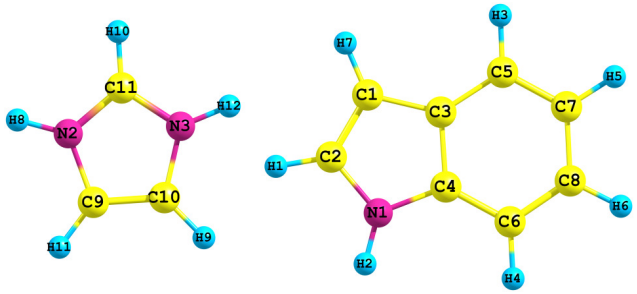

Figure 2.1 Model system for calculating V_{et} for the histidine molecule. The molecule is composed of an indole and an imidazolium cation. The coordinates are taken directly from the MD simulations. A series of calculations to find V_{et} were also carried out at different sandwich while varying the relative orientation of the rings. This figure was made using Chemcraft.[76]

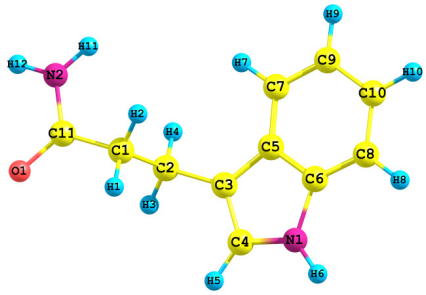

Figure 2.2 Model system for calculating V_{et} for the closest amide backbone from tryptophan. The geometry is taken from the MD simulations snapshots. A series of calculations were carried out to find V_{et} varying three different dihedrals, χ_1, χ_2 and ψ of the tryptophan backbone. χ_1 is defined by this set of atoms H1, C1, C2 and C3. χ_2 is defined by this set of atoms C1, C2 and C3 and C4. ψ is defined by this set of atoms H1, C1, C11 and N2. A series of calculation were carried out in which angles χ_1 and ψ were varied from -180 to +180 at 5 degree increments while holding χ_2 at either +90 or -90. This figure was made using Chemcraft.[76]

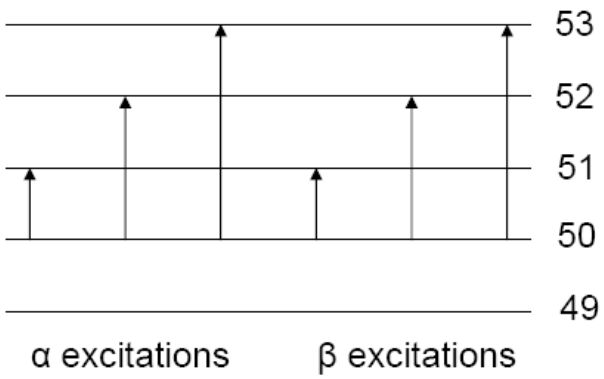

Figure 2.3 The singly excited determinants that were used in the Gaussian03 calculation. The LUMO is orbital 49 while orbital 51 is the HOMO of the indole ring. MOs 52 and 53 are the orbitals which contain the amide pi* orbital.

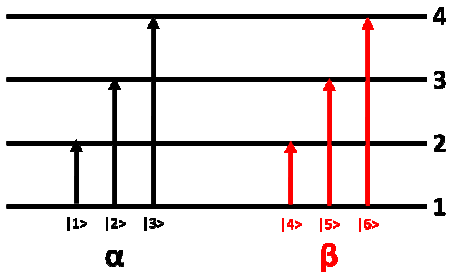

Figure 2.4 Subspace of molecular orbitals in a model donor acceptor electron transfer system. MO 1 pure HOMO of the donor, MO 2 is a pure LUMO of the donor. MO 3 and MO 4 are a mixed LUMOS of the acceptor. The black and red arrows indicate alpha and beta excitations from the HOMO to the appropriate LUMO.

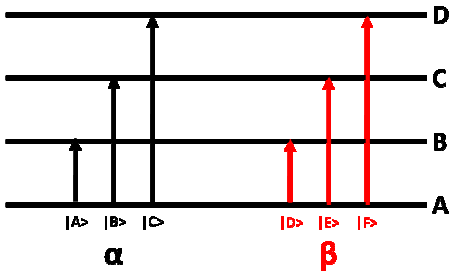

Figure 2.5 Subspace of molecular orbitals in a model donor acceptor electron transfer system. MO 1 pure HOMO of the donor, MO 2 is a pure LUMO of the donor. MO 3 and MO 4 are pure LUMOS of the acceptor. The black and red arrows indicate alpha and beta excitations from the HOMO to the appropriate LUMO.

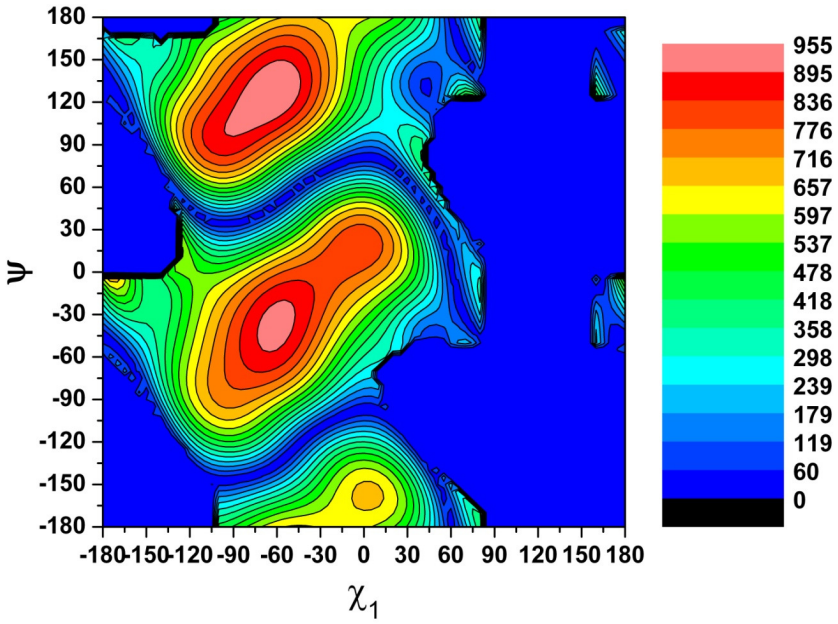

Figure 2.6 Contour plot for amide V_{et} using the model system in Figure 1.2. χ_2 is set at 90^0. V_{et} is plotted unit of cm^{-1}. The blue sections of the plot are areas in which there was significant overlap between atoms. Because of this, these structures were not used and were assigned an arbitrary value of zero.

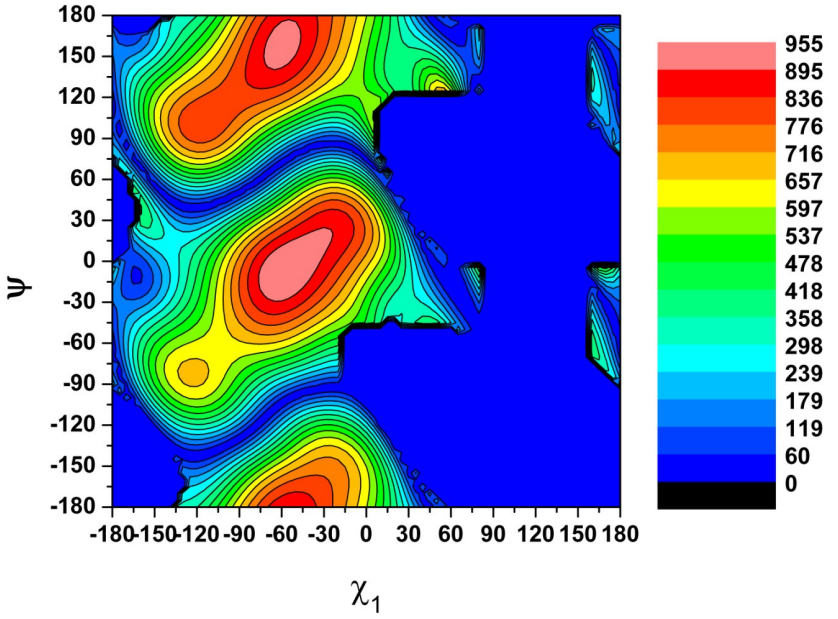

Figure 2.7 Contour plot for amide V_{et} using the model system in figure 1.2. χ_2 is set at -90^0. V_{et} is plotted unit of cm^{-1}. The blue sections of the plot are areas in which there was significant clashing between atoms because of this these structures were not used and they were assigned an arbitrary value of zero.

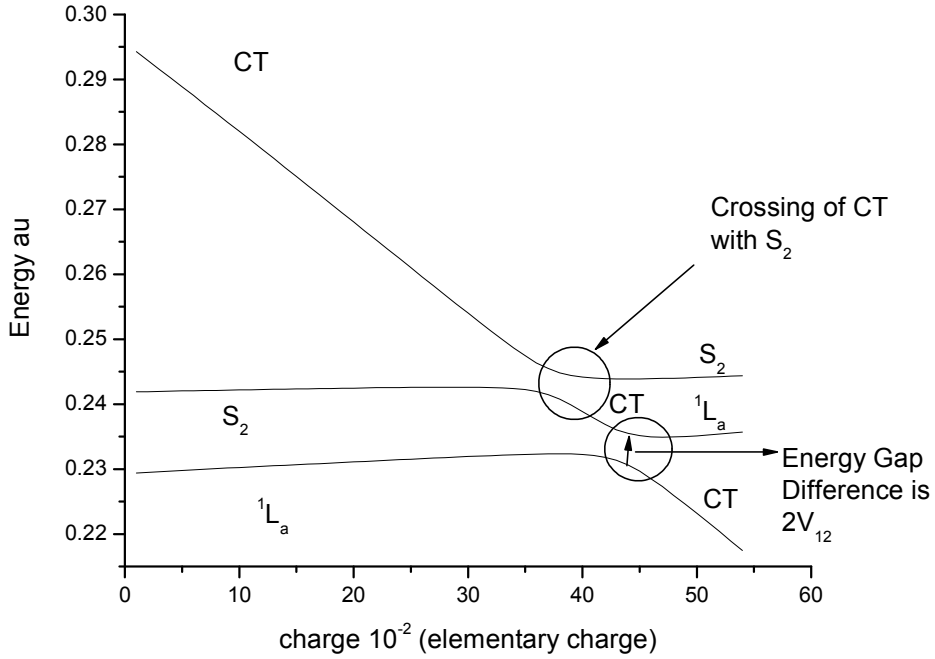

Figure 2.8 Calculation of V_{et} using the crossing method. It is noted that the CT state first experiences an avoided crossing with the second lowest singlet state denoted by S_2 and after crossing this energy surface it crosses the ^1L_a state and the minimum in energy gap difference is taken as 2 V_{et}.

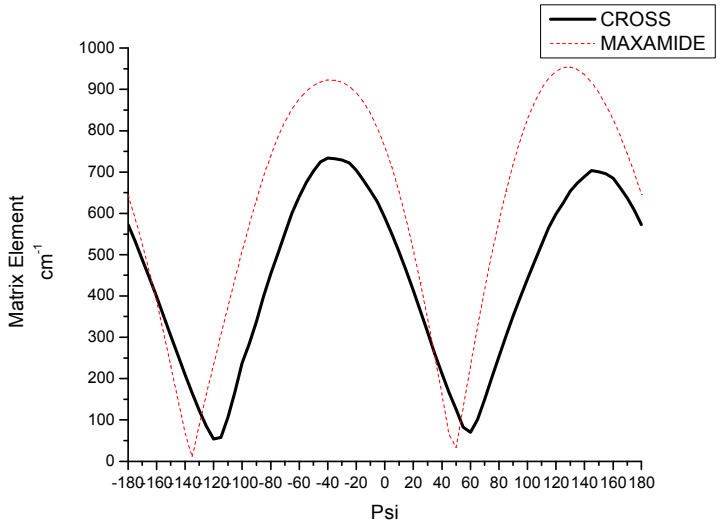

Figure 2.9 This figure compares the values of V_{et} between the CT state and the ^1L_a state for two methods, the Crossing Method and the Linear Combination method.

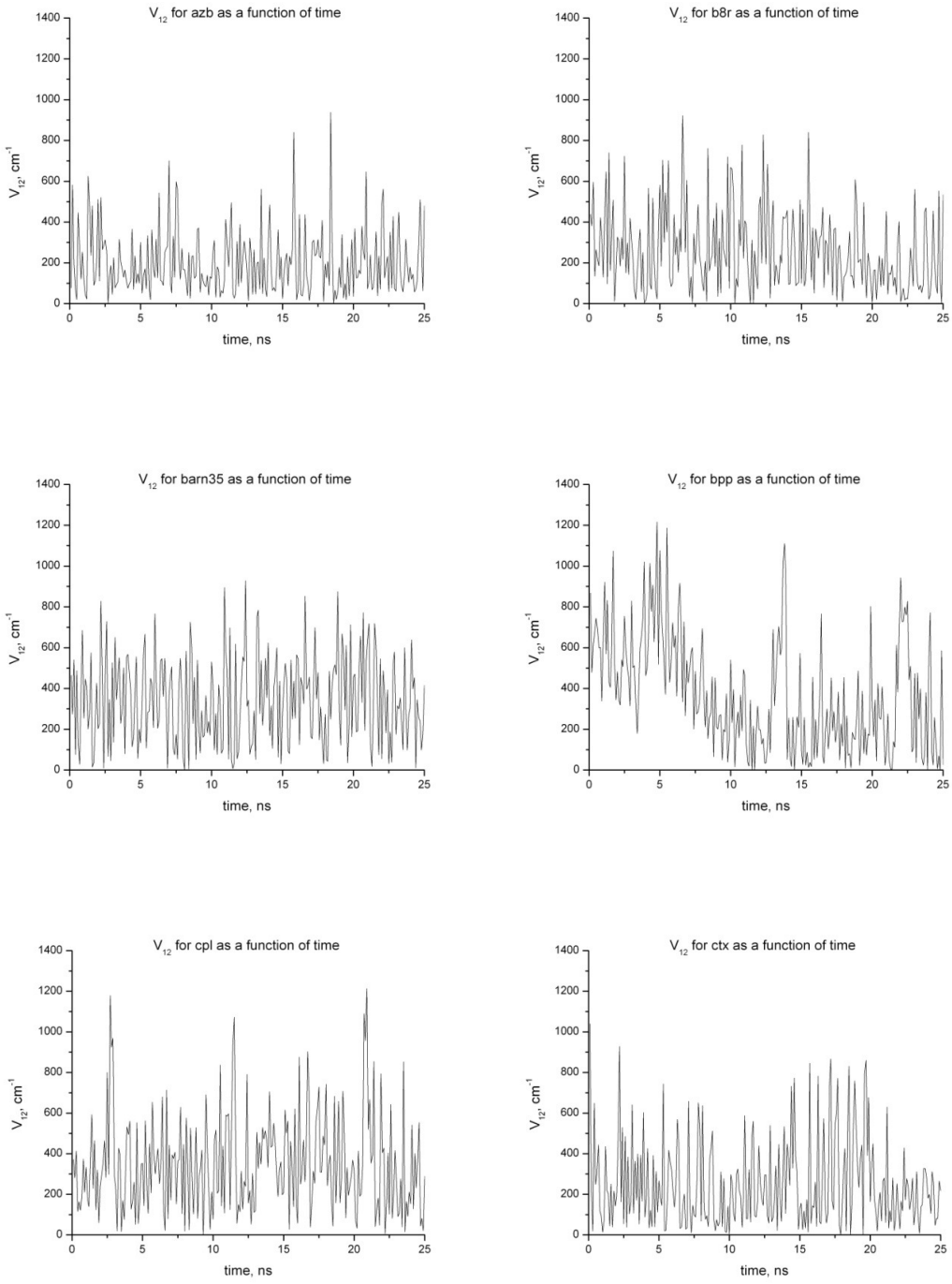

Figure 2.10 Time traces of V_{et} for amide electron transfer for the following proteins: azb, b8r, barn35, bpp, cpl and ctx. Notice the variability in V_{et} depending on the protein. There are a total of 4000 snapshots, making the time step equal to 3.125 ns for each calculation of V_{et}. V_{et} was calculated using the linear combination method.

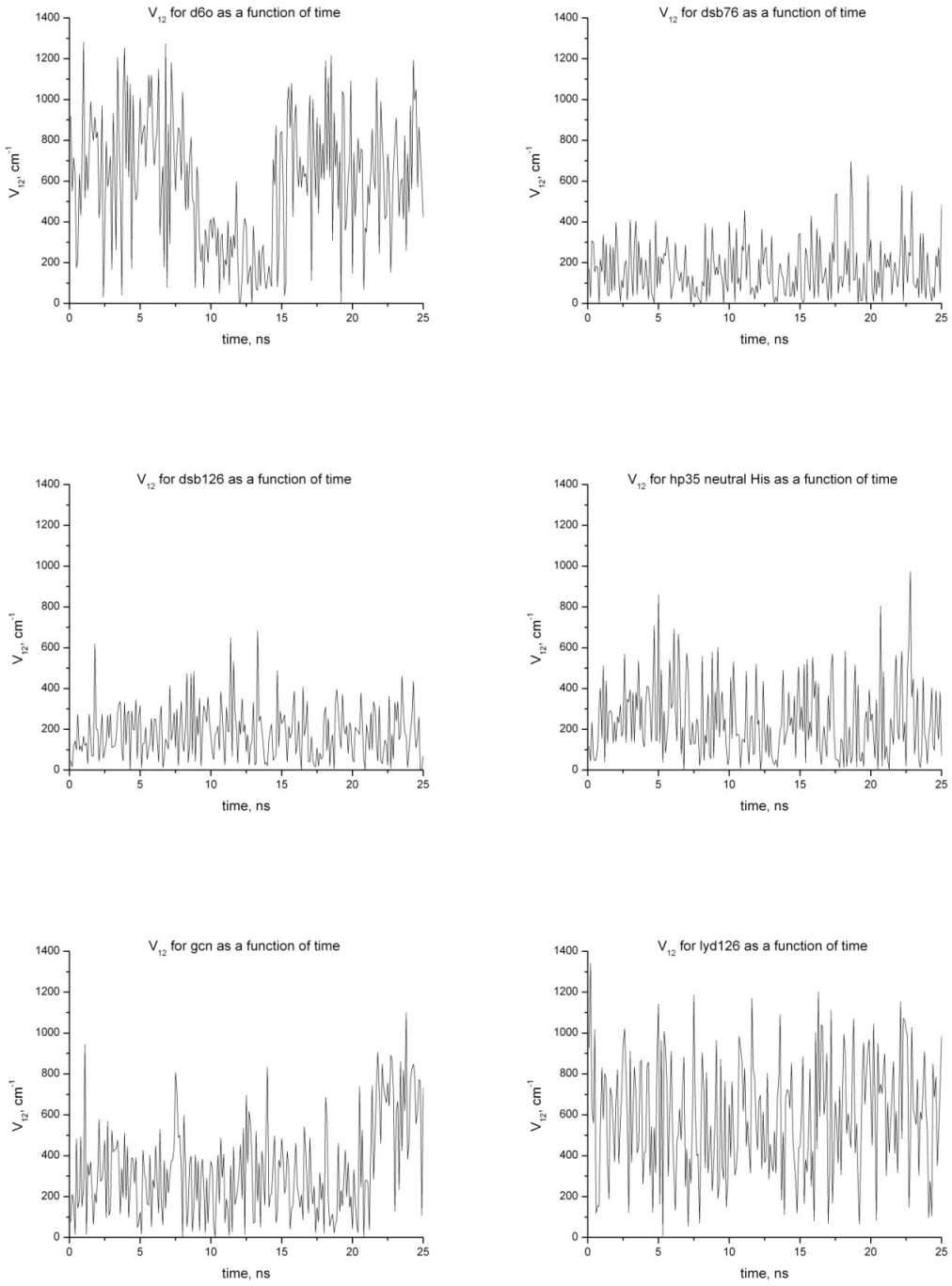

Figure 2.11 Time traces of V_{et} for amide electron transfer for the following proteins: d6o, dsb76, dsb126, hp35, gcn and lyd126. Notice the variability in V_{et} depending on the protein. There are a total of 4000 snapshots, making the time step equal to 3.125 ns for each calculation of V_{et}. V_{et} was calculated using the linear combination method.

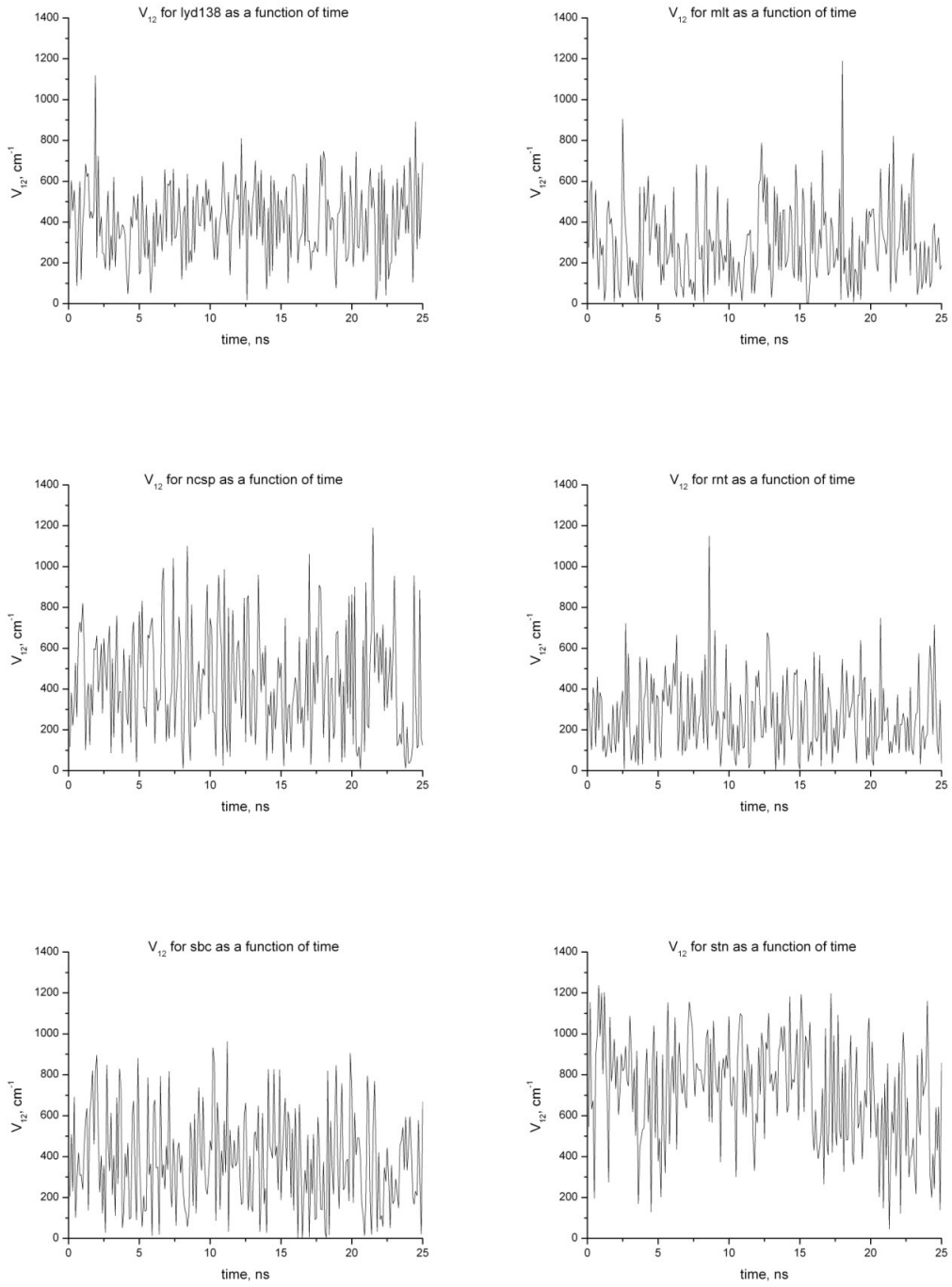

Figure 2.12 Time traces of V_{et} for amide electron transfer for the following proteins: lyd138, mlt, nscp, rnt, sbc and stn. Notice the variability in V_{et} depending on the protein. There are a total of 4000 snapshots, making the time step equal to 3.125 ns for each calculation of V_{et}. V_{et} was calculated using the linear combination method.

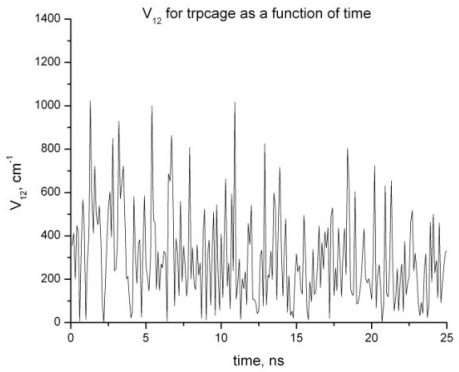

Figure 2.13 Time traces of V_{et} for amide electron transfer for trpcage. There are a total of 4000 snapshots, making the time step equal to 3.125 ns for each calculation of V_{et}. V_{et} was calculated using the linear combination method.

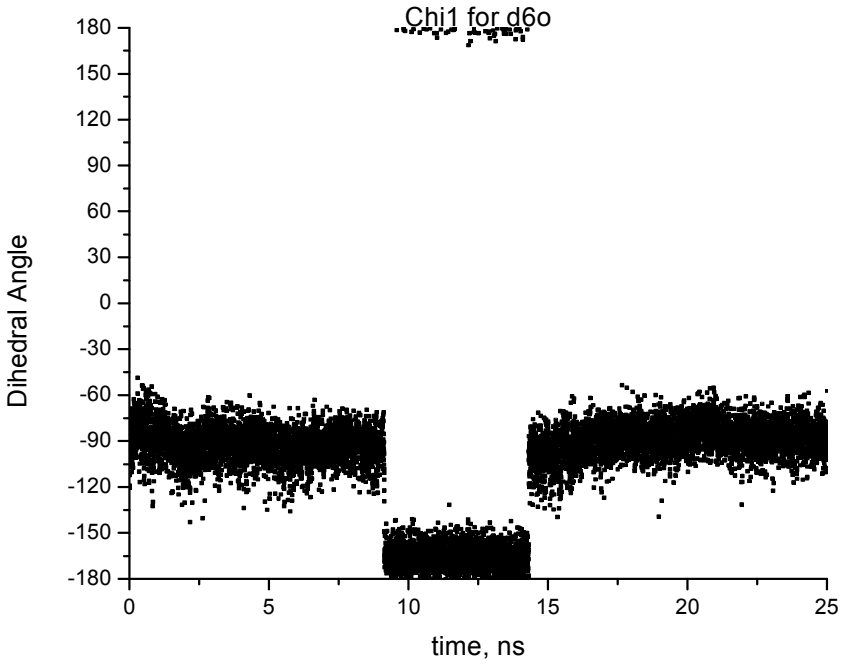

Figure 2.14 χ_1 rotamer flips for d6o. These rotamer flips are important in determining the magnitude of V_{et} for amide. In the $\chi_1 = -60°$ state through bond V_{et} is enhanced via "π conjugation." The value of the average V_{et} in this region is ~800 cm^{-1}. In the $\chi_1 = -180°$ the average value of V_{et} is ~ 200cm^{-1}. These observations are consistent with the contour graphs of V_{et} as a function of χ_1, χ_2 and ψ.

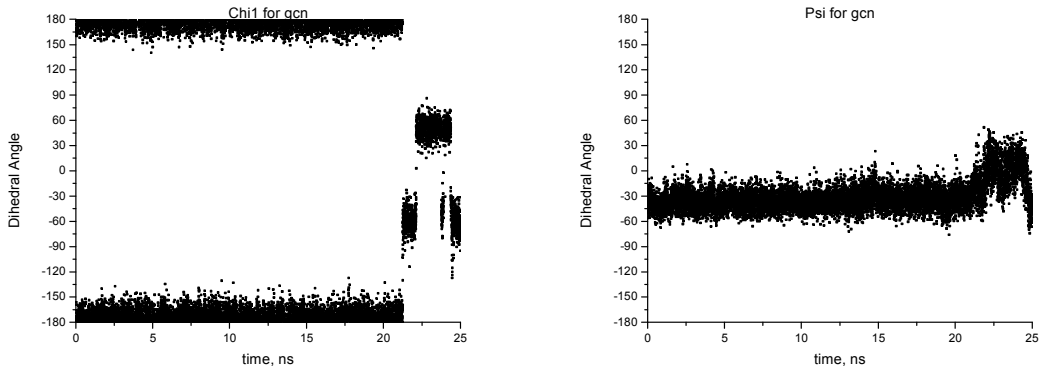

Figure 2.15 χ_1 and ψ rotamer flips for gcn. As before in the χ_1 = -60° state through bond V_{et} is enhanced via "π conjugation." The value of the average V_{et} in this region is ~800 cm^{-1}. In the χ_1 = -180° the average value of V_{et} is ~ 200cm^{-1}. Additionally the region with χ_1 = +60, ψ = +30 also shows high values of V_{et}. These observations are consistent with the contour graphs of V_{et} as a function of χ_1, χ_2 and ψ.

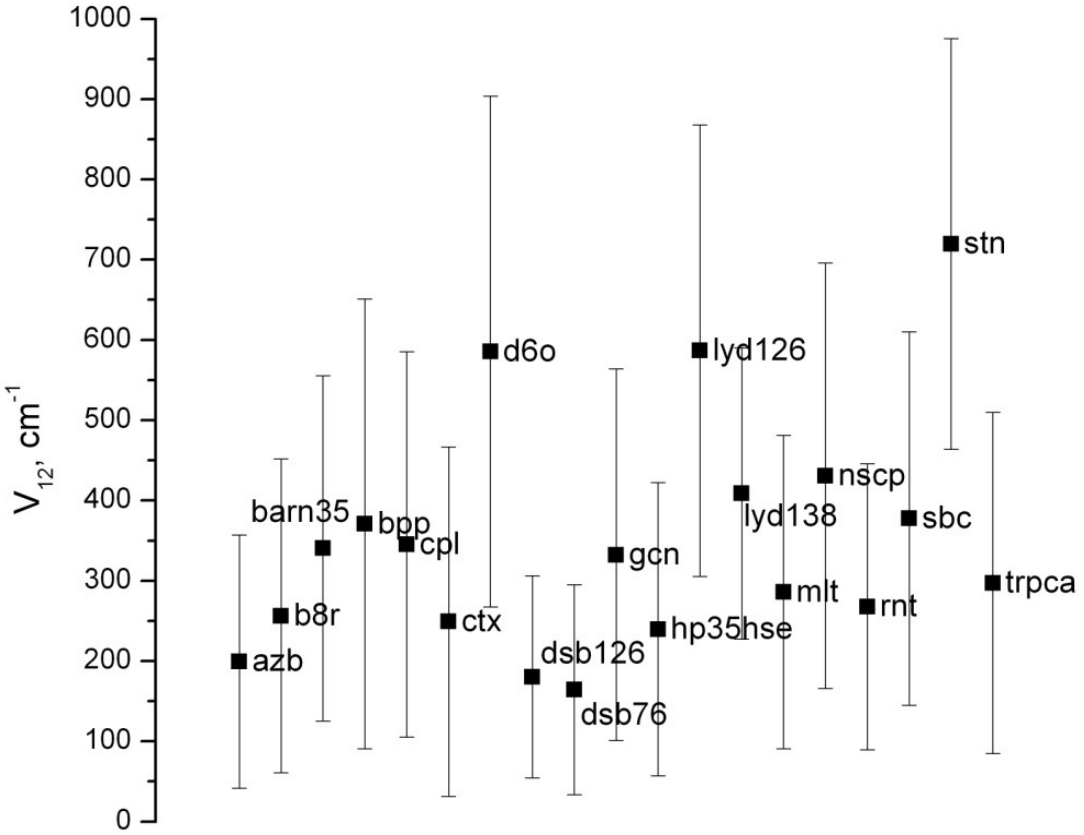

Figure 2.16 Average V_{et} for each individual protein, the bars are plus/minus one standard deviation from the average.

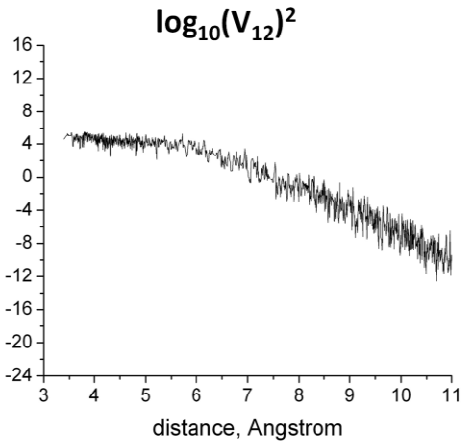

Figure 2.17 The $\log_{10}$ of square of V_{et} as a function of distance for imidazolium/indole molecule in WH5. Notice that there are two regions one from 3.5 Å to about 6.0 Å in which the V_{et} remains constant, after 6.0 Å that the coupling starts to decay exponentially with distance.

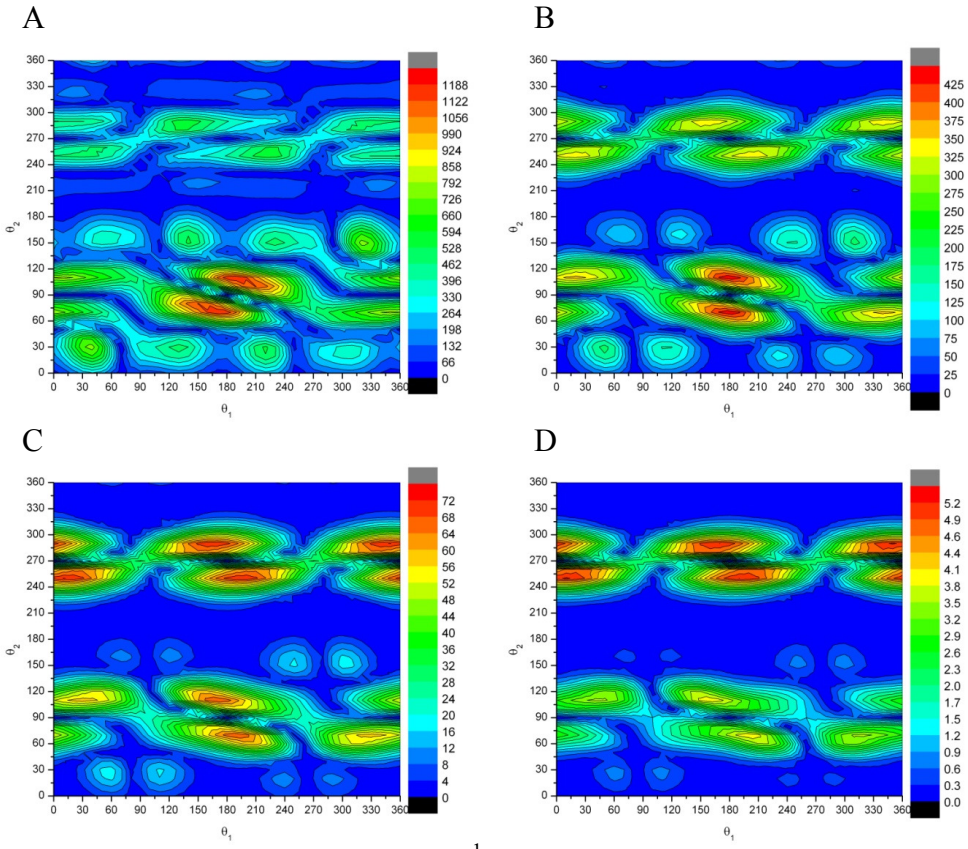

Figure 2.18 Contour diagrams of V_{et} in cm^{-1} as a function of the two tumble angles θ_1 and θ_2 at a fixed center of mass distance, set at 5 Å (A), 6 Å (B), 7 Å (C) and 8 (D) Å. Notice the sensitive of V_{et} at close distances and its insensitivity at far distances.

CHAPTER 3

EXTENDING MOLECULAR DYNAMICS SIMULATIONS TO 100 NANOSECONDS FOR PREDICTION OF FLUORESCENCE QUANTUM YIELDS

Introduction

Tryptophan fluorescence has been used extensively to study proteins[1-21]. Tryptophan's sensitivity to protein changes arises from the local protein environment. The work of many groups has suggested that amide is a quencher of tryptophan fluorescence by electron transfer[24,27-30,77]. Based on this, simulation studies have been carried out in the past to calculate the variation in quantum yield in a set of single tryptophan proteins by Callis and coworkers[42-44]. In this work it was found that there is a correlation between the fluorescence quantum yield and the energy gap. This energy gap is the energy difference between the fluorescing state of tryptophan and the CT transfer state from the HOMO of the tryptophan ring to a π^* orbital of the amide. This energy gap is controlled by the protein environment, specifically the orientation and charge of amino acids, and the formation of hydrogen bonds to amide. Electron transfer happens only when the energy gap is low. Predictions of the fluorescence intensity have been carried out using the following equation:

$$\Phi = \frac{k_r}{k_r + k_{nr} + k_{et}}$$

Where Φ is the fluorescence quantum yield, k_r is the radiative decay rate set at 0.04 ns^{-1}, k_{nr} is the nonradiative decay rate set at 0.08 ns^{-1} and k_{et} is the rate of electron

transfer to the amide[47]. The variation in fluorescence intensity is only due to the variation in k_{et} from protein to protein. k_{et} is estimated by using Fermi's golden rule[35,41]:

$$k_{et}(\Delta E_{00}) = \frac{2\pi}{\hbar} V_{et}^2 \rho(\Delta E_{00})$$

In this equation ΔE_{00} is the energy difference between the vibrational zero points of the CT and ^1L_a states[42], V_{et} is the electron transfer matrix element between the ^1L_a state and the CT state and ρ_{FC} is the Franck-Condon weighted density of states. ρ_{fc} is calculated using the same procedure as Callis and coworkers[42,43]. In short, ρ_{fc}, is calculated by integrating over the overlap between the Frank-Condon factors for ionization of indole and electron capture for formamide.

Previous work had been carried out to a maximum time of 150 ps[43]. An open question still remained after this work was complete. Is 150 ps a long enough time scale to capture all the fluctuations and basins that are possible in proteins? There are fluctuations that occur in the native states of proteins that can significantly alter the energy gap which will therefore change the observed quantum yield. These possible fluctuations can be due to rotamer states of tryptophan or charged side chains and could give credibility that rotamer transitions are the cause of lifetime heterogeneity. If these conformations have lifetimes comparable to the fluorescence lifetime they cause the observed non-exponential decay of tryptophan fluorescence[78-82]. In this work we capture a small glimpse of such possible configurations in some proteins. In chapter 4 we discuss how to calculate quantum yields using a kinetics model to simulate the fluorescence decay rates from two different basins that exchange at a rate comparable to the fluorescence lifetime.

Methods

All proteins were simulated using GROMACS 4.0.7[75,83-85]. The crystal structure of each protein was taken directly from the PDB file according to table 3.1. Each protein was solvated so that at least 1 nm of tip3p water separated the surface of the protein and the edge of the water cube. Either sodium cations or chloride anions were added to neutralize the overall charge of the system. An energy minimization was carried out so as to have a maximum force on all atoms of less than 1000 kJ/mol·nm. NVT and NPT equilibrations where then carried out each for 100 ps. In these simulations, the time step was set to 2 fs and all atom bonds were constrained using the P-LINCS algorithm[86]. Electrostatics were calculated using particle mesh Ewald[87]. The temperature was set at 300 K, the protein and solvent were coupled separately to a velocity rescale thermostat[88] with a time constant of 0.1 ps. Pressure was set at 1 bar using a Parrinello-Rahman barostat[89,90] with a 2 ps time constant and a compressibility of 4.5×10^{-5} bar^{-1}. Periodic boundary conditions were used with a cubic box as the unit cell. The OPLS-AA[91-95] force field was used in all simulations. After equilibration, a 25 ns production simulation was carried out under NPT conditions, pressure and temperature were again set at 1.0 bar and 300K respectively, using velocity rescaling and a Parrinello-Rahman barostat using the same time constants and compressibility. Coordinates were saved every 2 ps for a total of 12500 snapshots per simulation. All of the amino acids had ground state charges including tryptophan. Zerner's INDO/S CI[96], modified to include the local electric field and potential[97,98] as well as modified parameters for oxygen[99], was used to calculate the

transition energies for the ^1L_a and amide charge transfer states, using a capped version of tryptophan, making the QM molecule N-formyltryptophanamide (NFTA). V_{et} between tryptophan and amide were calculated from the CIS Hamiltonian, as in Chapter 2, using Gaussian03[68] with a read window from HOMO to LUMO+2 for tryptophan/amide. V_{et} was corrected to account for extensive mixing of orbitals using the maximization of the electron density carbonyl carbon method described in the previous chapter. For results on the average V_{et} and time traces of the V_{et} see chapter 2.

To calculate the electron transfer rate for each of these proteins the following expression for Fermi's golden rule was used[43]:

$$ket_i = \frac{2\pi}{\hbar}\left\langle V_{et,i}\right\rangle^2 \left\langle \rho_{fc,i}\left(\Delta E_{00}\right)\right\rangle$$

Where $k_{et,i}$ is the electron transfer rate for the ith protein, $\langle V_{et,i}\rangle$ is the average V_{et} and $\langle \rho_{fc}\rangle$ is the average Frank-Condon density of states. This rate was then used to calculate the quantum yield.

Additionally a set of 100 ns simulations where carried out using an improved version of the AMBER99SB[100] force field called AMBER99SB-ILD004[101]. The protocol was exactly the same except for three points. Gromacs-4.5.4 with AMBER99SB-ILDN force fields was used instead of Gromacs-4.0.7 with OPLS-AA force field. No counter ions were added to the simulation box. The last difference is in the way we analyze the data obtained in the simulations. Instead of having a constant V_{et} for all the proteins, figures 2.7 and 2.8 were used to create a table in which V_{et} was listed as a function of the dihedral angles χ_1, χ_2, and ψ. The GROMACS utility g_angle was used to calculate the dihedral angles χ_1, χ_2 and ψ as a function of time, these were estimated to nearest set of

dihedral angles in the table and a V_{et} was assigned for that particular point. Now instead of having an average V_{et} and an average Frank-Condon density of states for each protein we have instantaneous V_{et} and Frank-Condon densities. With these instantaneous values we can therefore easily calculate the instantaneous rate as a function of time.

The time series of instantaneous rates was used to construct a series of survival curves by integration of the first order rate equation:

$$\frac{d^1L_a(t)}{dt} = -k(t)^1L_a(t)$$

$$\int_{^1L_a(0)}^{^1L_a(t)} \frac{d^1L_a(t')}{^1L_a(t')} = -\int_0^t k(t')dt' = -\int_0^t (k_{et}(t') + k_r + k_{nr})dt' = -\left(k_r t + k_{nr} t + \int_0^t k_{et}(t')dt' \right)$$

$$\ln\left(\frac{^1L_a(t)}{^1L_a(0)} \right) = -\left(k_r t + k_{nr} t + \int_0^t k_{et}(t')dt' \right)$$

$$^1L_a(t) = ^1L_a(0)\exp\left(-\left(k_r t + k_{nr} t + \int_0^t k_{et}(t')dt' \right) \right)$$

$$^1L_a(0) = 1$$

$$^1L_a(t) = \exp\left(-\left(k_r t + k_{nr} t + \int_0^t k_{et}(t')dt' \right) \right)$$

Where k_r and k_{nr} are the usual radiative and non-radiative decay constants for 3-methylindole and $k_{et}(t)$ is the instantaneous value of k_{et} at time t. The window for each survival curve was 25 ns which is three times the lifetime of a completely unquenched tryptophan. After calculating and saving a survival curve the window was moved 0.05 ns in time and a new survival curve was calculated and saved. This was done until the last frame of the window reached the end of the simulation. With this procedure a total of 1601 survival curves were calculated. All of these survival curves were used to calculate

an average survival curve by averaging all the populations at time t. The quantum yield is then calculated using the average survival curve using the following:

$$\Phi = \frac{\int_0^{t'} \langle {}^1L_a(t) \rangle dt}{\int_0^{t'} \exp(-k_r) dt} \approx \frac{\int_0^{t'} \langle {}^1L_a(t) \rangle dt}{\frac{1}{k_r}} \approx k_r \int_0^{t'} \langle {}^1L_a(t) \rangle dt$$

Results

Ground to Excited State Transition Energies and V_{et} for OPLS-AA

Figures 3.1 to 3.4 show the Zerner's INDO/S CIS calculated ground state to ^1L_a state transition energies and the ground state to the first amide charge transfer state transition energies. A number of these proteins show long lived basins for the CT state energy, which include: azb, b8r, cpl, d6o, gcn, lyd138, mlt, rnt and trpcage1. If the lifetime of these basins is comparable to the fluorescence lifetime these simulations could be a sign of heterogeneity. For example gcn (which is a protein that is known to have bi-exponential fluorescence decay) has two possible basins one at ~46 kcm^{-1} and by the end of the simulation there is a second basin that has an energy of ~50 kcm^{-1}.

In this treatment we follow the procedure that was followed by Callis and coworkers in 2007[43]. In this treatment the calculated quantum yield is mostly sensitive to the energy gap, which is the energy difference between the ^1L_a state and the CT state transition energy. The energy gap determines the number of vibrational states that are in resonance between the final and initial states. This term is called the Frank-Condon density of states and it is very sensitive to the energy gap. The energy gap depends on

the electric field that the surrounding amino acids create on the tryptophan moiety[43]. For example a protein configuration that contains a positive charge near the amide of tryptophan and a negative charge near the tryptophan ring is going to have a smaller energy gap than the opposite case (a negative charge near the amide and a positive charge near the ring). Therefore hydrogen bonding to the oxygen of the amide can play a central role in the lowering the energy gap (either by formation of a helix or beta sheet or by water accessibility). Additionally positively charged groups that can get near the amide can also significantly decrease the gap. For the ring any negatively charged amino acid that is close to the ring and has the right orientation will decrease the energy gap. It seems plausible therefore that rotamer flips of the tryptophan dihedral angles χ_1, χ_2 and ψ that change the relative distance of charged amino acids to the ring can change the energy gap and the observed fluorescence.

Figure 3.5 shows the average V_{et} for the 20 different proteins in different environments. The magnitude and fluctuations of the V_{et} is determined for the most part by the χ_1 dihedral angle as mentioned in chapter 2 of this thesis. Proteins that contain $\chi_1 \approx -60°$ will have a V_{et} that fluctuates at about 900 cm^{-1} and never goes to zero. On the other hand proteins that have a χ_1 of either 60° or ± 180° will have a V_{et} that on the average is about 200 cm^{-1} whose value frequently goes to zero.

Using Fermi's golden rule with the average V_{et} and the average Frank-Condon density of states the average electron transfer rate constant to the amide was calculated. In order to make agreement with experiment an offset had to be added to the energy gap. The offset was chosen so as to make the best fit, the value of this was 240 cm^{-1}. Figure

3.6 shows the theoretically calculated versus experimental quantum yield for the 20 single tryptophan proteins. There are a number of proteins that deviate from the experimental value, some of these include lyd126, barn71, stn, gcn and mlt. Gcn and mlt are expected to have non-exponential fluorescence decay because of the different distinct energy gaps which have a lifetime close to the fluorescence lifetime. Therefore treating each of these with a single Gaussian distribution is not correct, which is what is assumed in the calculation of the electron transfer rate. Stn, barn71 and lyd126 seem to have the wrong electric fields this is due to, as we will see in our second study, differences in the atomic partial charges. These charges can make the energy gap change significantly.

Ground to Excited State Transition Energies
and V_{et} for AMBER99SB-ILDN

The second set of simulations was carried out using the AMBER99SB-ILDN force field. The simulation length was 100 ns and there were no counter ions to make the protein neutral. Figures 3.7 to 3.10 show the results for the transition energies from the ground state to the 1La state and the amide CT state. There are some proteins different basins, some of these include gcn, b8r, bpp, ctx, d6o and mlt. Just by inspection we should expect that these proteins are going to have heterogeneity in their lifetimes because each individual basin will have a different energy leading to a different value for the Frank-Condon density of states. However since, the energy gap is not the only variable that determines the rate of electron transfer we also need to consider what V_{et} is doing in time too. A comparison between AMBER99SB-ILDN transition energies and OPLS-AA transition energies show that the transition energies to the ^1L_a state are not

significantly different between the force fields. However, when the charge transfer states are compared there are significant deviations. These deviations are due to differences in assigned charges of each residue in each force field. It is interesting to note that when comparing the average charge transfer state energy that OPLS-AA is consistently higher in energy than the AMBER99SB-ILDN.

Instead of averaging V_{et} over each trajectory we calculated them by calculating the three dihedrals χ_1, χ_2, and ψ and then mapping these onto figures 2.7 and 2.8 (previous chapter). With this an estimation of V_{et} was achieved for all 20 proteins as a function of time (yielding a total of 50,000 values for V_{et}). The results are presented in figures 3.11 to 3.14. As expected proteins that have a χ_1 value of ~-60° maintain a large V_{et} that does not go to zero, whereas proteins that have nominal values of χ_1 equal to ±180° or 60° have V_{et} values that are much smaller and routinely go to zero.

Some examples of heterogeneity in V_{et} are worth mentioning. Bpp shows transitions between χ_1 from -60° to + 60° that change the behavior of V_{et}. Additionally χ_2 undergoes transitions from +90° to =90° but χ_1 is the dihedral that controls the behavior of V_{et}. Others that show this behavior are d6o, gcn and mlt.

Using V_{et} and the transition energy differences as a function of time the instantaneous electron transfer rate constant can be calculated. Instead of presenting the instantaneous electron transfer rate constant as a function of time we are going to present the inverse of the lifetime as a function of time. This is defined as:

$$\tau^{-1} = k_r + k_{nr} + k_{et,amide}(t)$$

The results of this calculation are shown in figures 3.15 to 3.18. Besides the time traces, a smoothed average is presented to be able to follow the average behavior of τ^{-1}. The window in which the averaging was 5 ns (this was chosen because in this time the fraction left of a completely unquenched tryptophan would be about 0.5). There are some remarkable examples of heterogeneity in the expected τ^{-1}. For example b8r, bpp, ctx, gcn and mlt show that in these proteins τ^{-1} can persistently vary by two orders of magnitude. With this in mind we can expect that for b8r, bpp, ctx, gcn and mlt the fluorescence decay should be multi-exponential.

Finally, by using these time traces for τ^{-1} we can integrate the kinetic equation for the decay of the fluorescent state and come up with individual survival curves for each protein. By averaging over these survival curves we can get an estimate of the ensemble average survival curve. The results are presented for all 20 proteins in figures 3.19 to 3.22. It is good to note that many of the proteins that are expected to have a low quantum yield show survival curves that decay quickly, where are proteins that have a high quantum yield decay more slowly. From the survival curves we can notice that there are a few proteins that indeed show non-exponential decay in the survival curves these include: b8r, ctx, bpp, gcn and mlt as expected from the instantaneous electron transfer rate constants.

Integrating these survival curves and multiplying them by k_r we can estimate the fluorescence quantum yield for each of the proteins. In order to reach agreement with experiment we added 1940 cm^{-1} to the energy gap (difference between the ^1L_a and CT state transition energies). The energy offset was calculated by minimizing the residuals

squared of each calculated quantum yield with respect to its experimental quantum yield. This means that most CT state energies in the AMBER99SB-ILDN are on average higher than the OPLS-AA. The obtained fit is shown in figure 3.23. It is interesting to note that stn and ctx are two proteins that still don't fit in this curve. The precise physical reason for this is still being investigated but preliminary results seem to indicate that these energies are dependent on which force field is used. The protein stn was simulated by using CHARMM27 and the CT state energy seems to fall exactly where it should be whereas for AMBER99SB-ILDN and OPLS-AA this CT state is too low. Ctx was ran using both OPLS-AA and AMBER99SB-ILDN in OPLS-AA the CT state energy is just right to give ctx the correct quantum yield but in AMBER99SB-ILDN the CT state is too high and therefore it gives a high quantum yield.

Conclusions

This study shows that GROMACS4 can be used in combination with either the OPLS-AA force field or the AMBER99SB-ILDN force field to estimate the fluorescence quantum yield by first principles. The time scale of each of these simulations was augmented by about three orders of magnitude than previous studies (from ps to ns). The longer times scales have revealed heterogeneous behavior in the fluorescence decay of single tryptophan proteins, a behavior which had not been noted before in previous work. Two different schemes were used to calculate the fluorescence quantum yield.

The first scheme was that of Callis and coworkers, in which each individual trajectory is assigned a single gaussian distribution for the energy gap. This scheme

might be applicable for some proteins that do not contain multiple energetic basins for the charge transfer state, however significant deviation might be expected if multiple energy basins exist. This procedure cannot treat non-exponential fluorescence decay since it assigns a single gaussian distribution for each protein. Another possible artifact of this procedure is that V_{et} is assigned a single value, the average value from a set number of points. In view of the results of chapter 2 of this thesis this assumption might not hold for all proteins, specifically those who undergo tryptophan χ_1 χ_2 or ψ dihedral transitions.

These two artifacts are taken into account by using the second procedure, in which the time dependence of both the energy gap and V_{et} are taken into consideration. An instantaneous τ^{-1} is calculated as a function of time for all the proteins and heterogeneity is taken into account by creating an average survival curve. In view of these improvements the experimental versus theoretical quantum yields shown in figures 3.6 and 3.23 show significant improvements for the AMBER99SB-ILDN\survival curve procedure.

Future work in this field should involve incorporation of a sphere around tryptophan that is treated quantum mechanically. This will probably take care of the energy gap difference that exists between OPLS-AA and AMBER99SB-ILDN force fields.

Tables and Figures

Table 3.1 Summary of experimental fluorescence quantum yields for the 20 simulated proteins and the initial pdb structure for the initial steps in the MD simulation of each of these proteins.

Name	**PDB Code**	**Experimental φ**
Villin N27H pH 7	1YRF	0.055[8]
azb	1AZB	0.31[1]
b8r	1B8R	0.13[1]
barn35	1A2P	0.142[102]
barn71	1A2P	0.122[102]
bpp	2BPP	0.031[1]
cpl	2CPL	0.025[103]
ctx	1CTX	0.1[1]
d6o	1D6O	0.03[104]
dsb76	1DSB	0.2[105]
dsb126	1DSB	0.041[105]
gcn	1GCN	0.12[1]
lyd126	1LYD	0.06[106]
lyd138	1LYD	0.044[106]
mlt	2MLT	0.115[1]
nscp	2SCP	0.033[107]
rnt	9RNT	0.31[1]
sbc	1SBC	0.03[1]
stn	1STN	0.29[1]
trpcage	1L2Y	0.03[108]

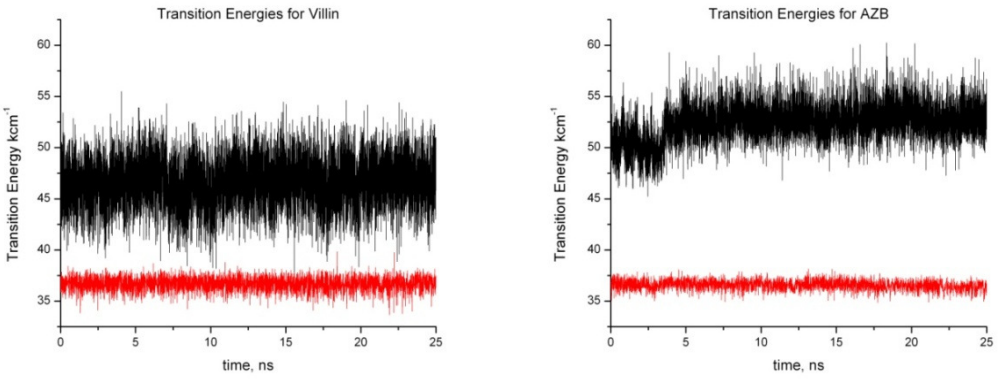

Figure 3.1 ZINDO vertical transition energies for ground to 1La state (red) and ground to amide CT state during a 25 ns simulation for villin and azb using the OPLS-AA force field.

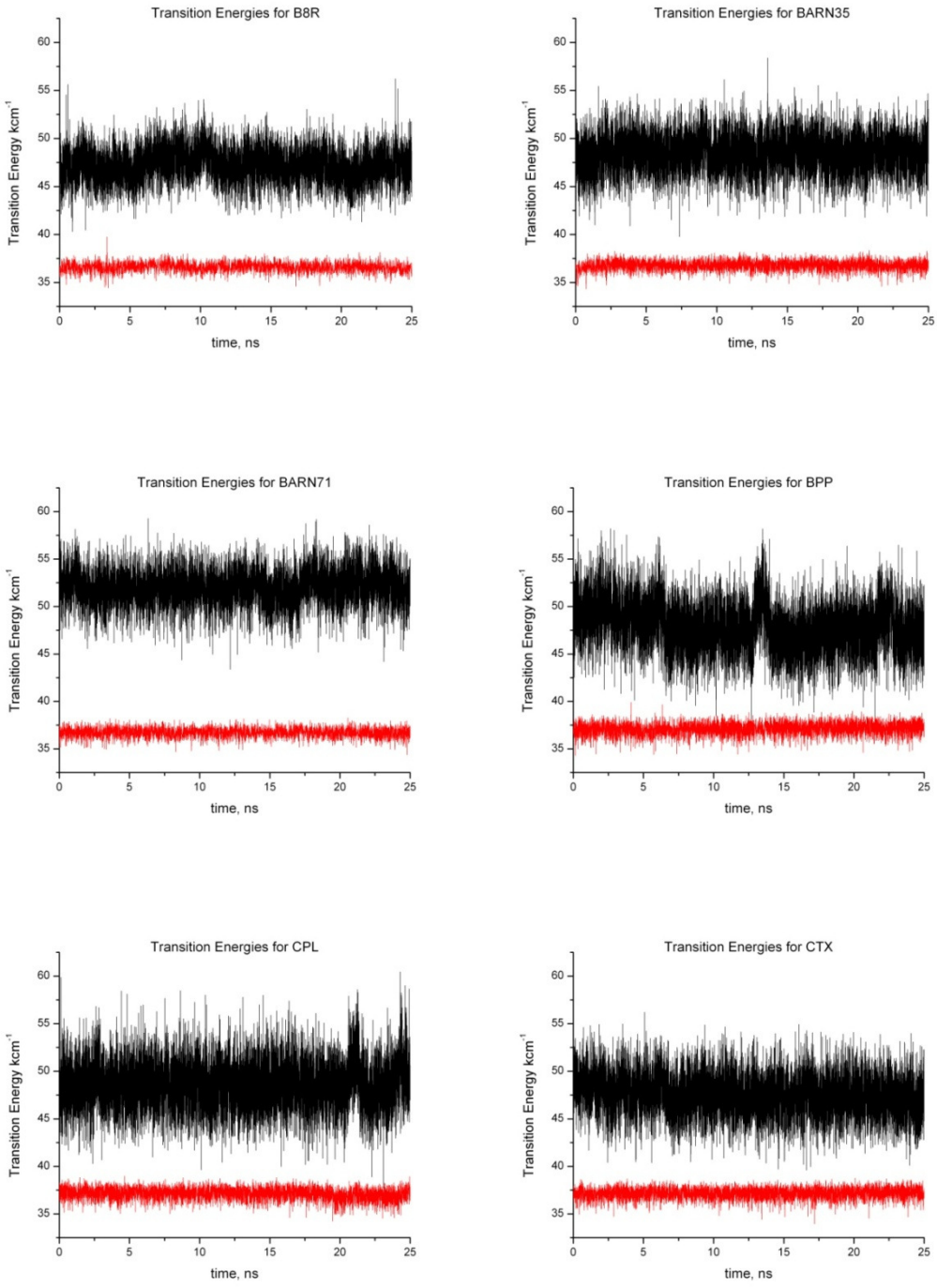

Figure 3.2 ZINDO vertical transition energies for ground to 1La state (red) and ground to amide CT state during a 25 ns simulation for b8r, barn35, barn71, bpp, cpl and ctx using the OPLS-AA force field.

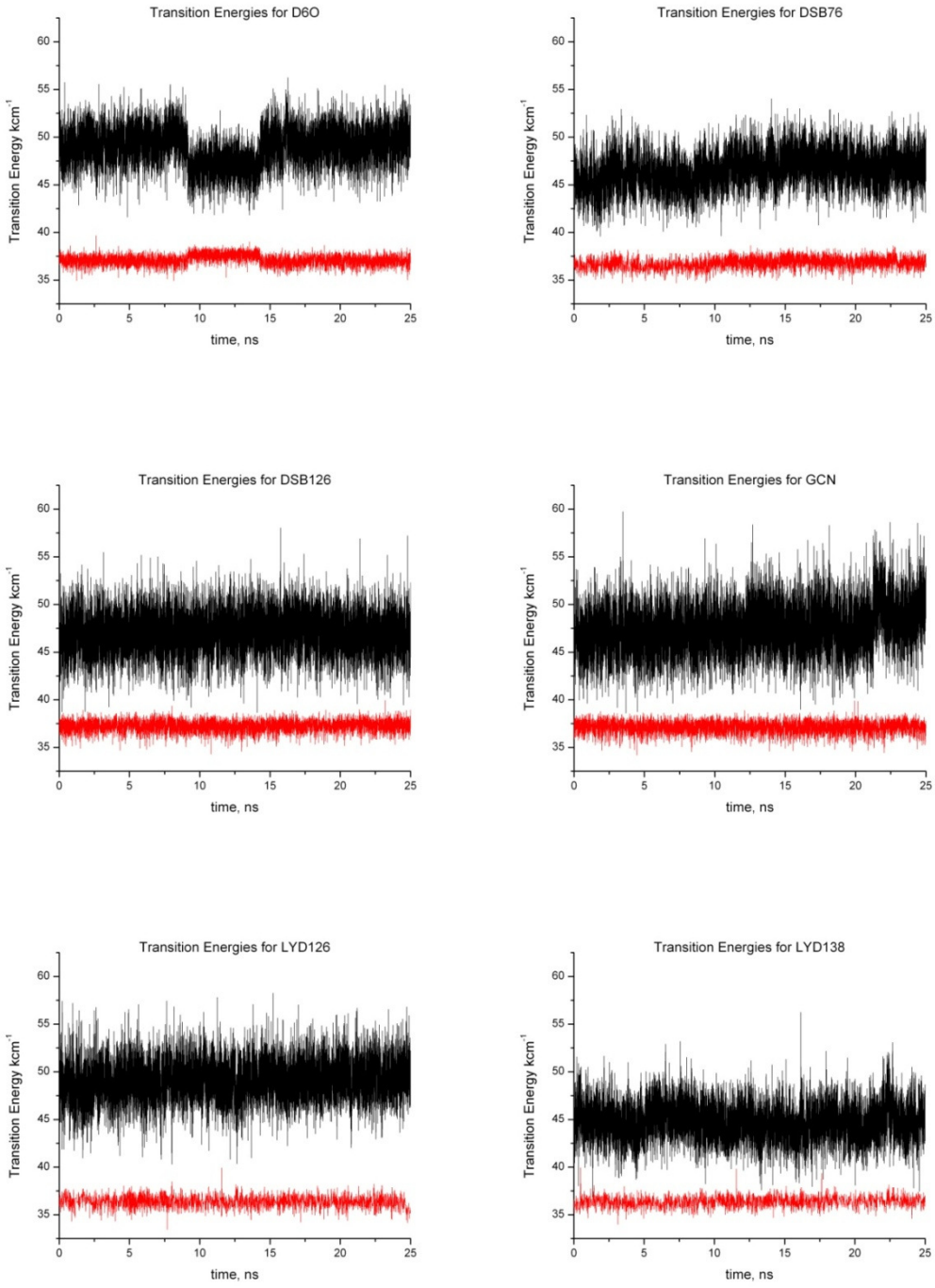

Figure 3.3 ZINDO vertical transition energies for ground to 1La state (red) and ground to amide CT state during a 25 ns simulation for d6o, dsb76, dsb126, gcn, lyd126 and lyd138 using the OPLS-AA force field.

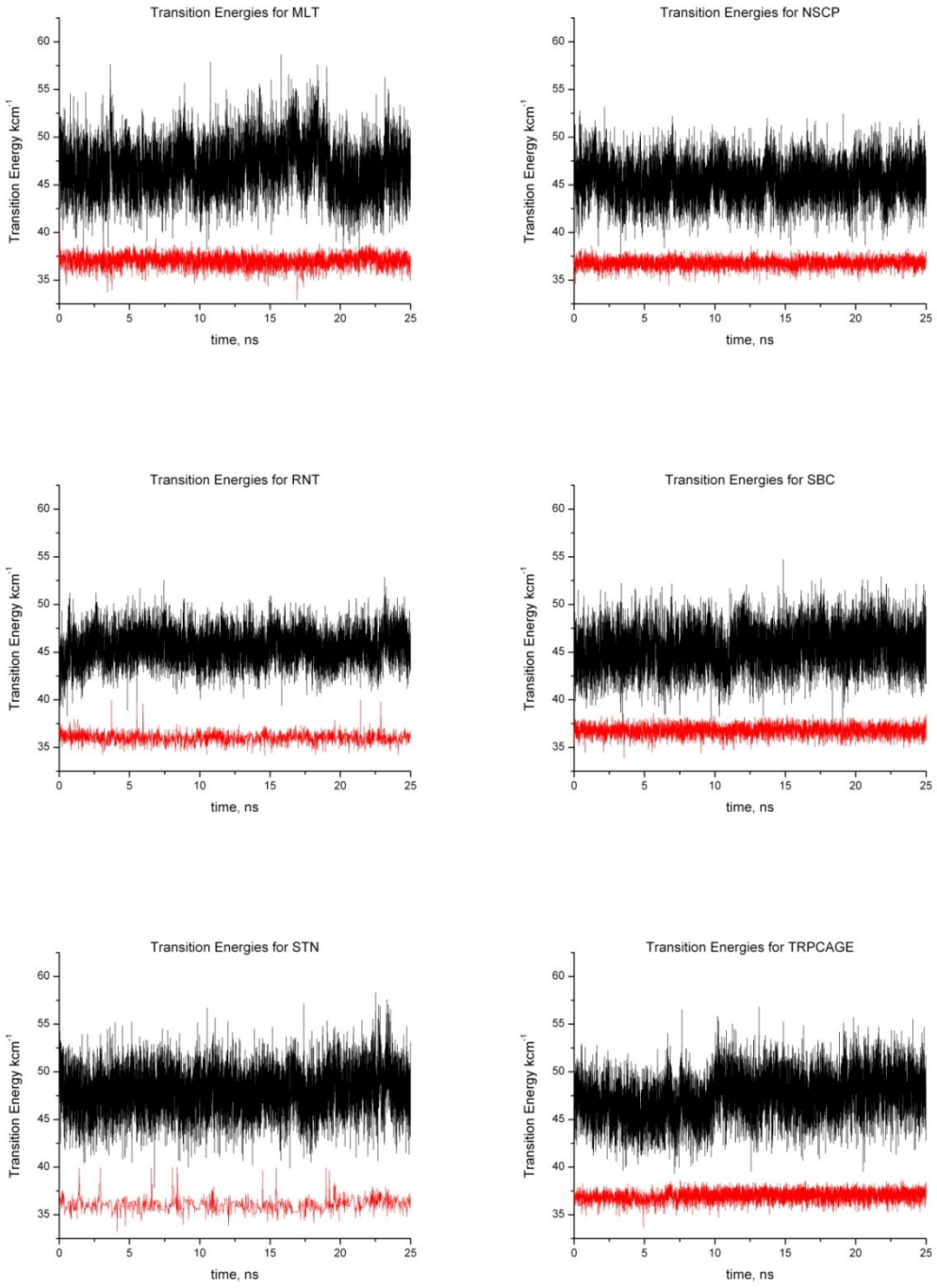

Figure 3.4 ZINDO vertical transition energies for ground to 1La state (red) and ground to amide CT state during a 25 ns simulation for mlt, nscp, rnt, sbc, stn and trpcage using the OPLS-AA force field.

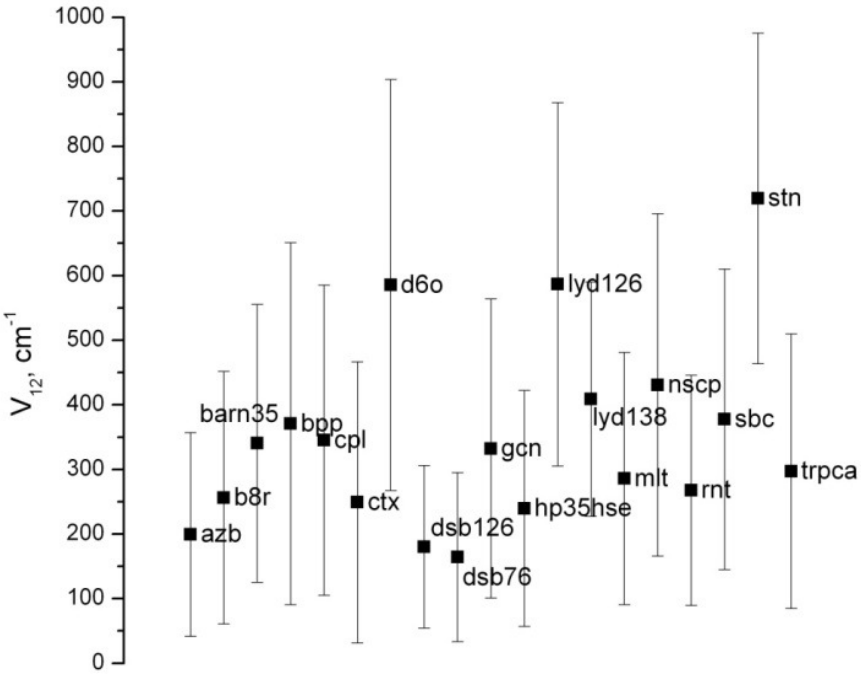

Figure 3.5 Average V_{et} for the 20 proteins; here hp35hse is villin headpiece subdomain at pH7 (neutral His).

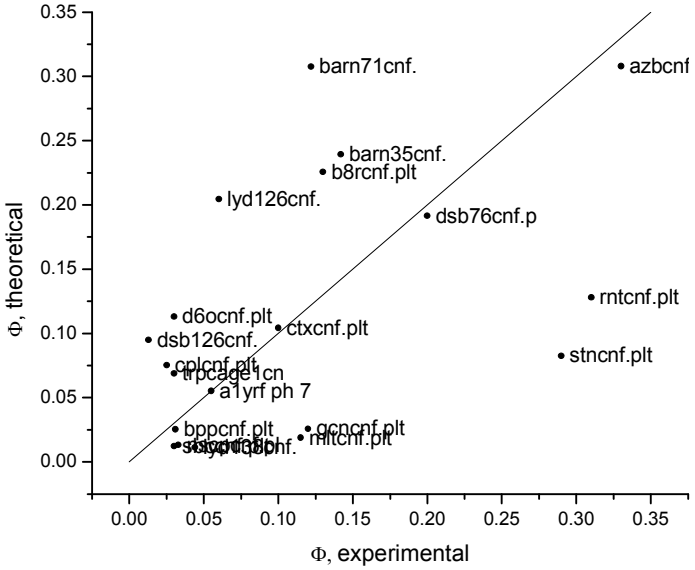

Figure 3.6 Theoretical versus experimental fit using equations [3] and [1]. The added constant to obtain the fit was 240 cm^{-1}.

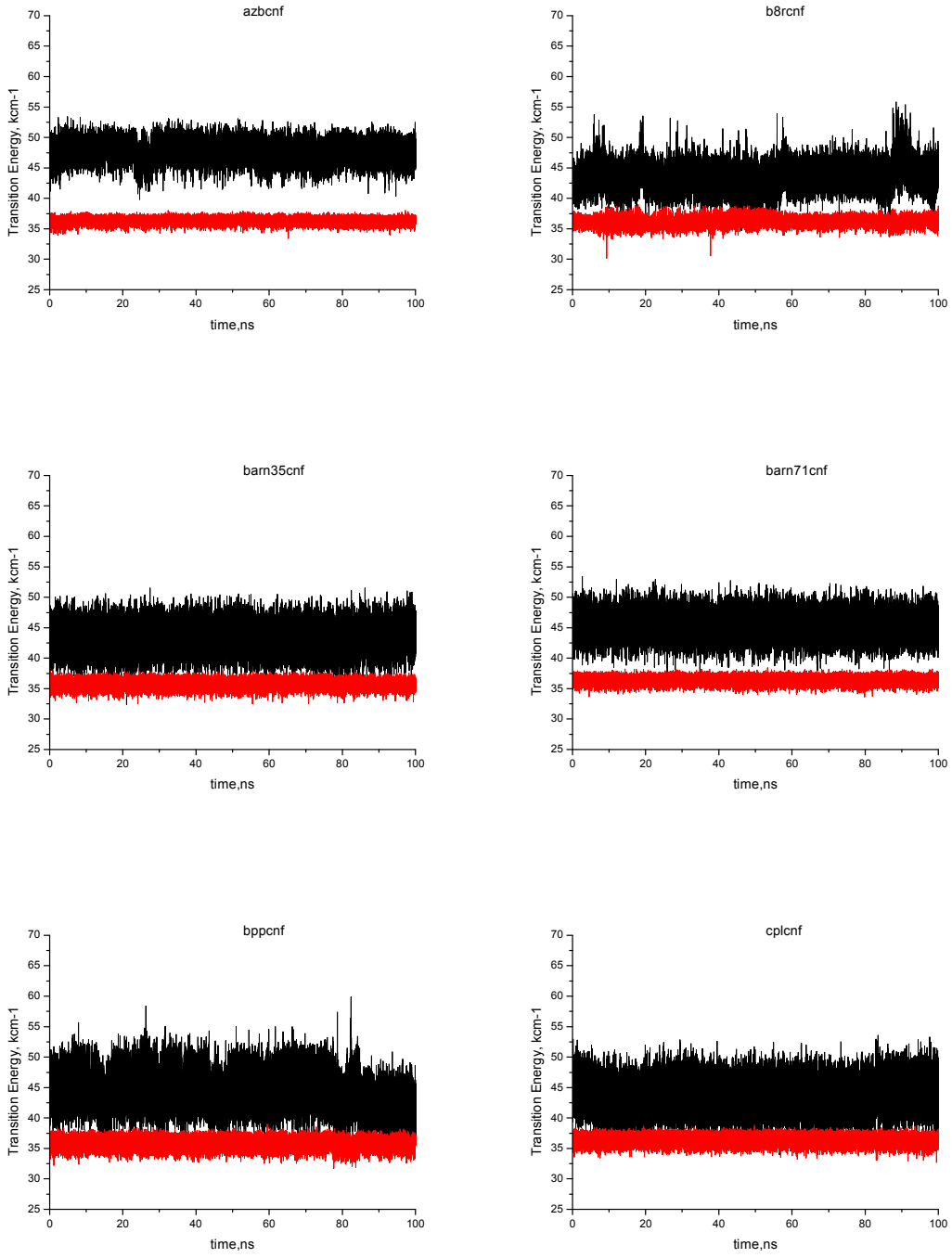

Figure 3.7 ZINDO vertical transition energies for ground to ^1L_a state (red) and ground to amide CT state during a 100 ns simulation for azb, b8r, barn35, barn71, bpp and cpl using the AMBER99SB-ILDN.

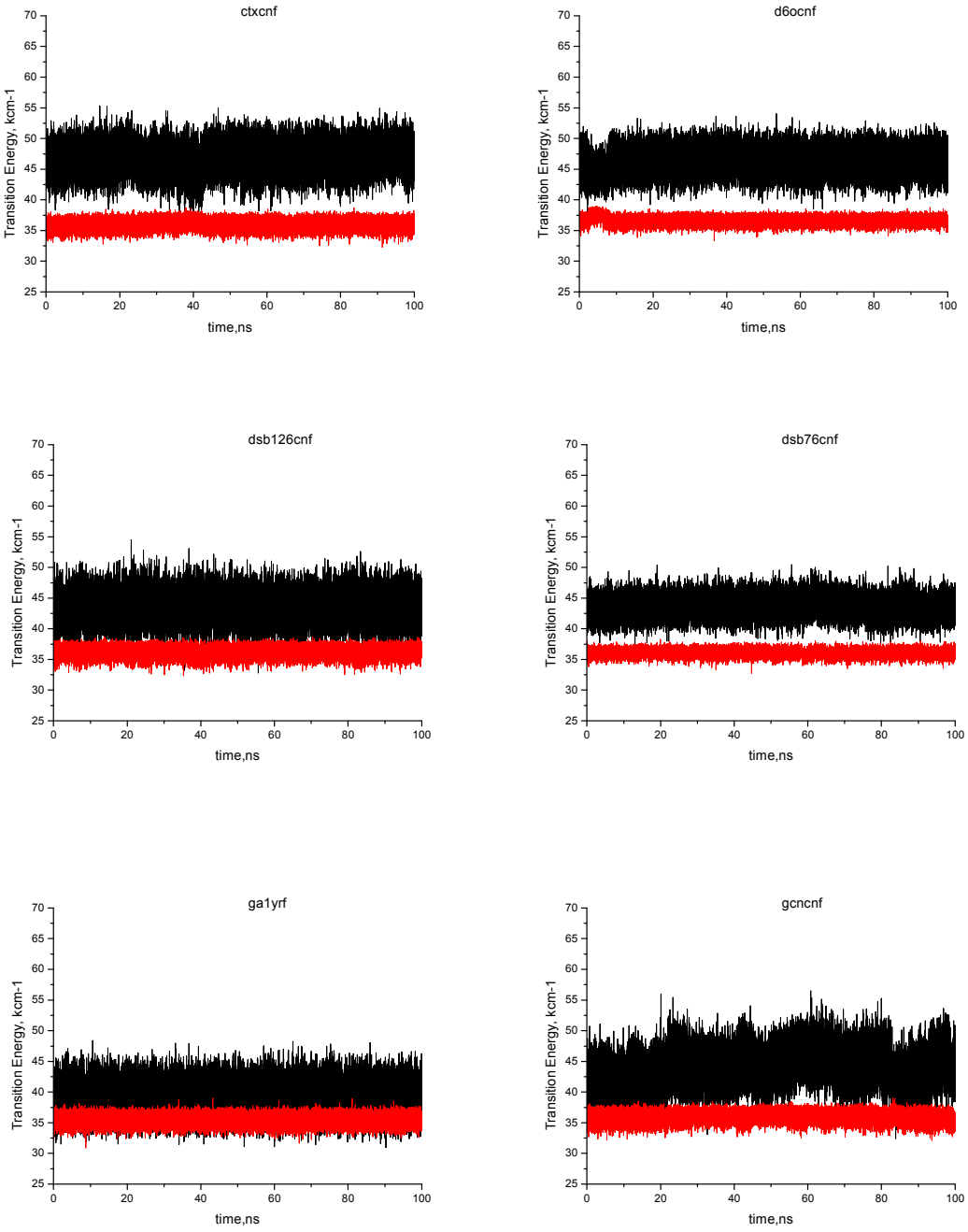

Figure 3.8 ZINDO vertical transition energies for ground to ^1L_a state (red) and ground to amide CT state during a 100 ns simulation for ctx, d6o, dsb126, dsb76, hp35 and gcn using the AMBER99SB-ILDN.

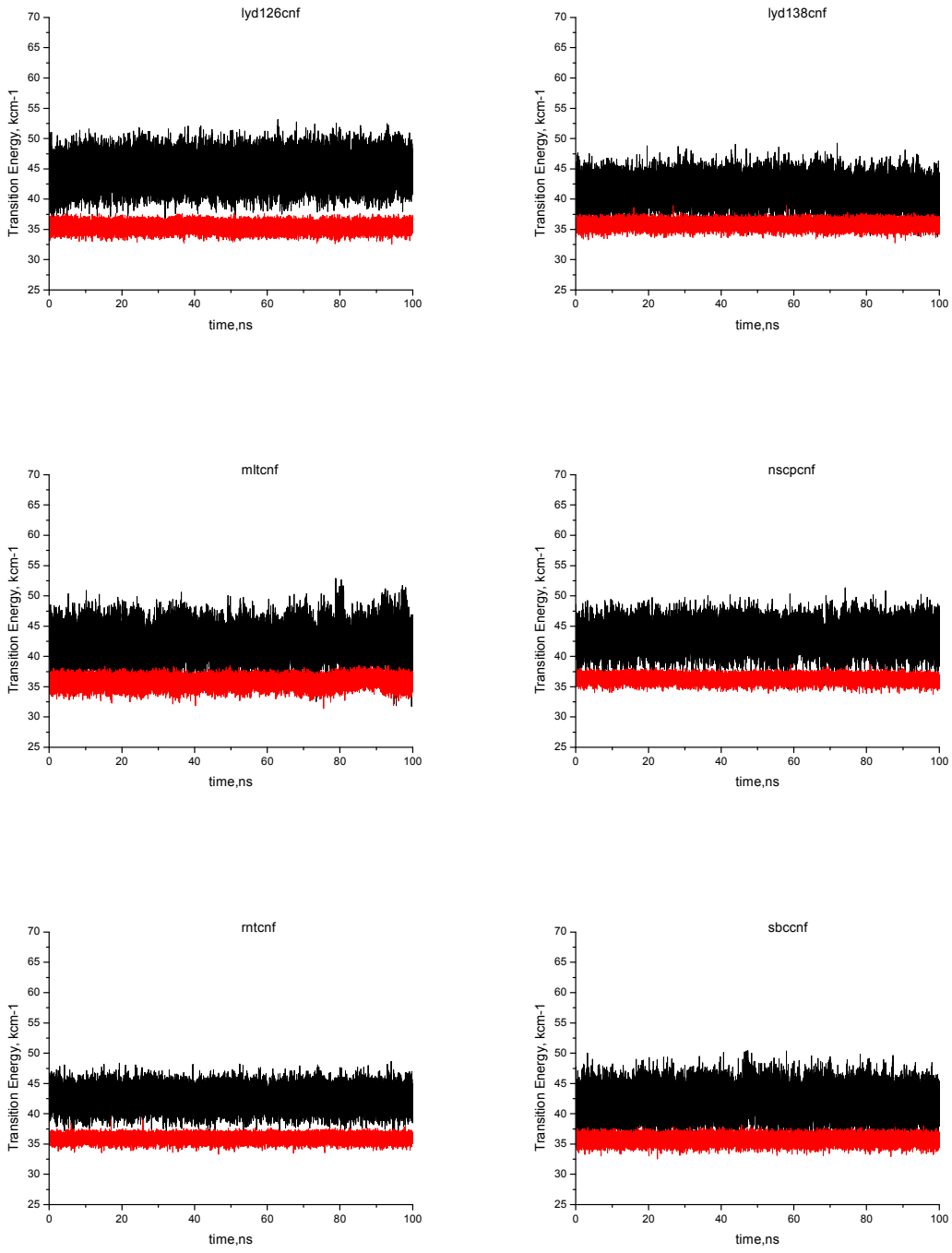

Figure 3.9 ZINDO vertical transition energies for ground to ^1L_a state (red) and ground to amide CT state during a 100 ns simulation for lyd126, lyd138, mlt, nscp, rnt and sbc using the AMBER99SB-ILDN.

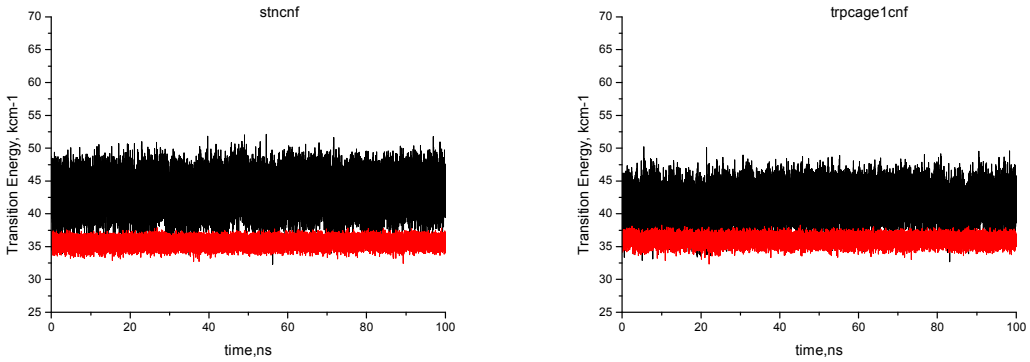

Figure 3.10 ZINDO vertical transition energies for ground to ^1L_a state (red) and ground to amide CT state during a 100 ns simulation for stn and trpcage1 using the AMBER99SB-ILDN.

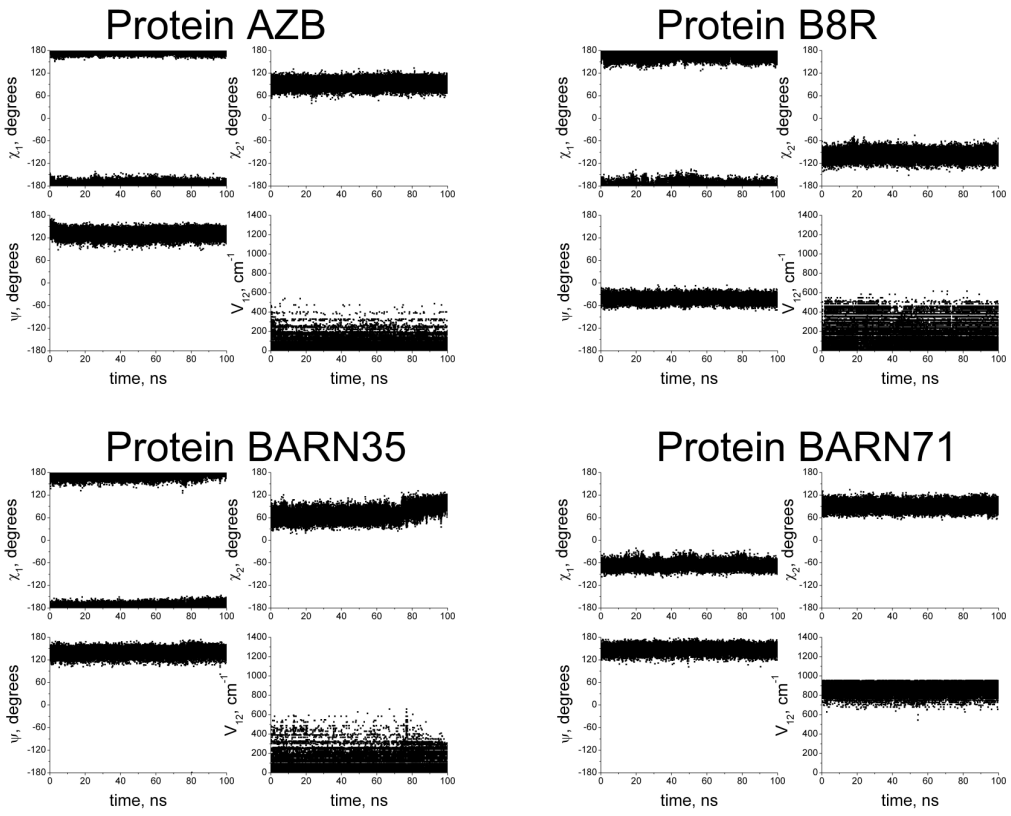

Figure 3.11 χ_1, χ_2, ψ and the calculated V_{et} as a function of time for the proteins azb, b8r, barn35 and barn71 using the AMBER99SB-ILDN 100 ns simulation.

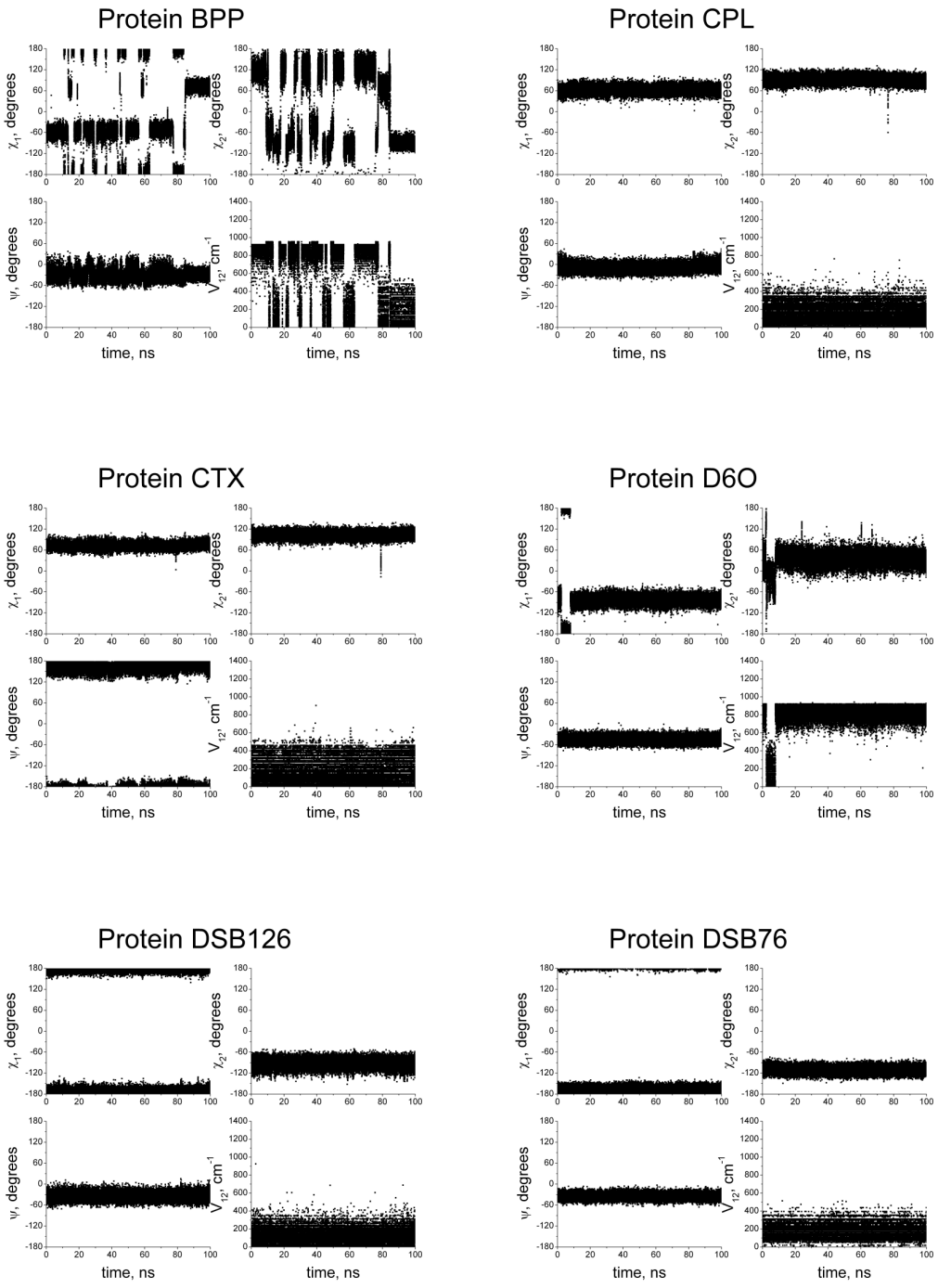

Figure 3.12 χ_1, χ_2, ψ and the calculated V_{et} as a function of time for the proteins bpp, cpl, ctx, d6o, dsb126 and dsb76 using the AMBER99SB-ILDN 100 ns simulation.

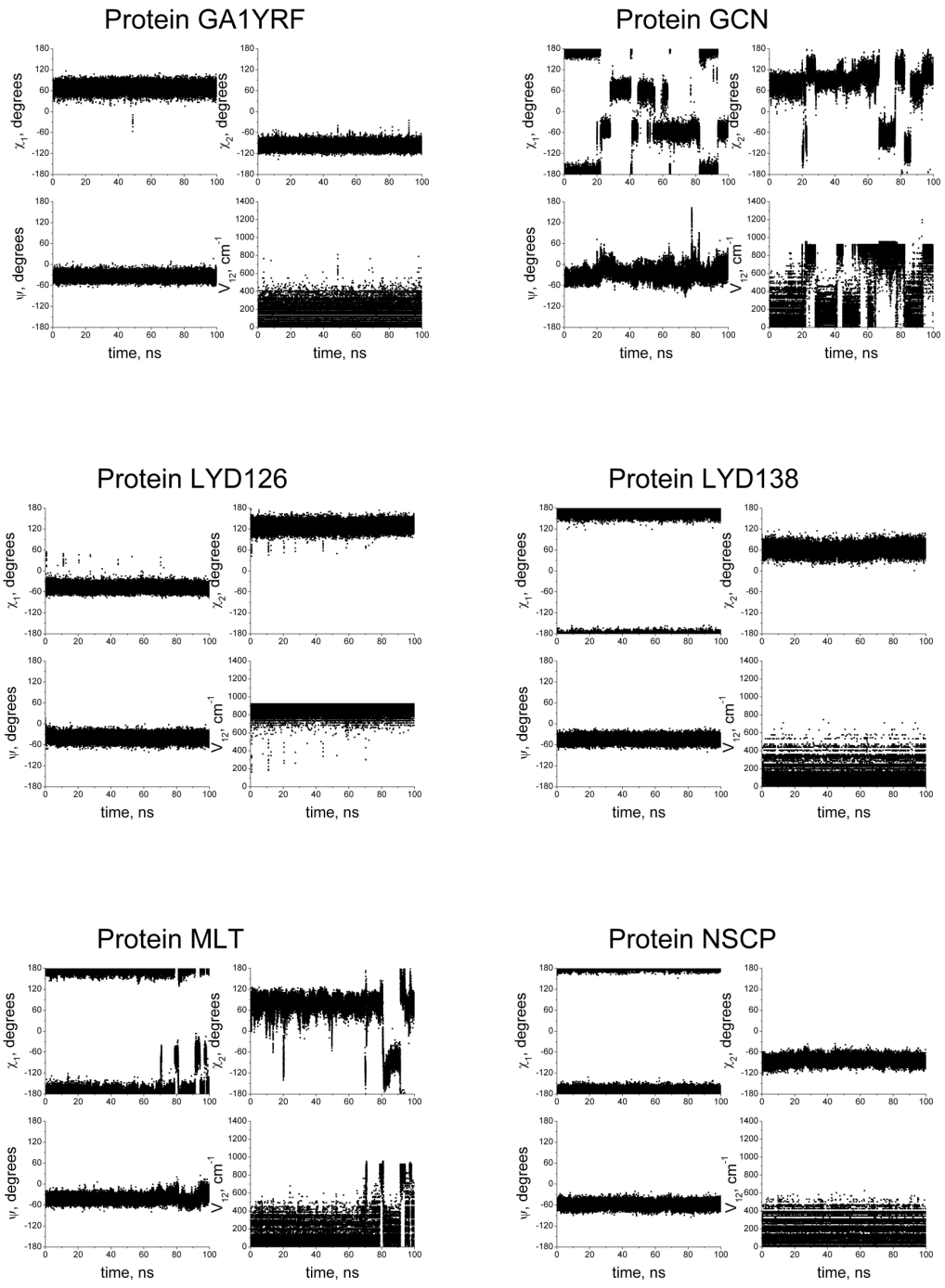

Figure 3.13 χ_1, χ_2, ψ and the calculated V_{et} as a function of time for the proteins ga1ryf, gcn, lyd126, lyd138, mlt and nscp using the AMBER99SB-ILDN 100 ns simulation.

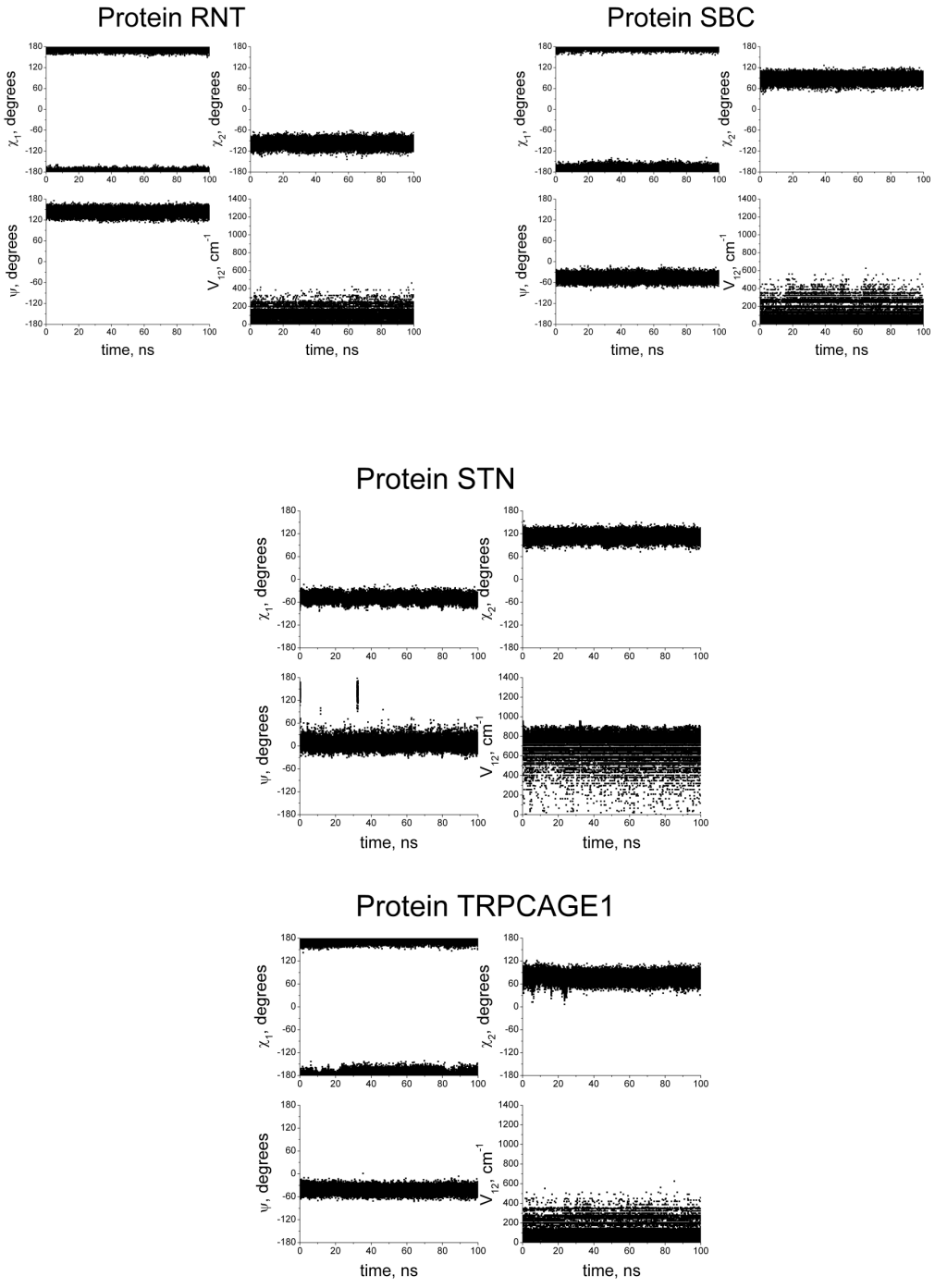

Figure 3.14 χ_1, χ_2, ψ and the calculated V_{et} as a function of time for the proteins rnt, sbc, stn and trpcage1 using the AMBER99SB-ILDN 100 ns simulation.

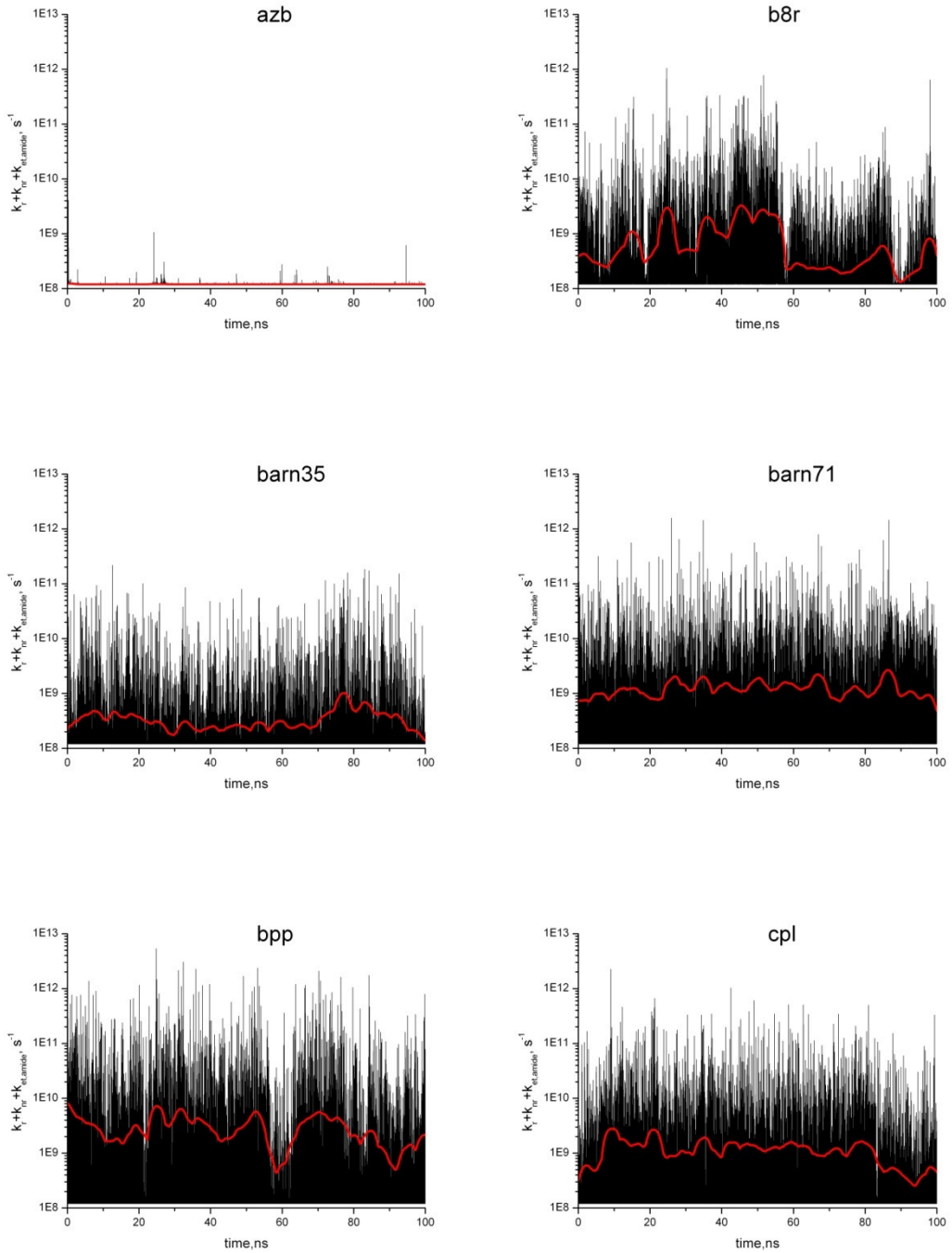

Figure 3.15 Time traces of the sum of all decay constants, $k_r + k_{nr} + k_{et,amide}$, as a function of time for azb, b8r, barn35, barn71, bpp and cpl. The red line shows the smoothed curved using an average window of 5 ns.

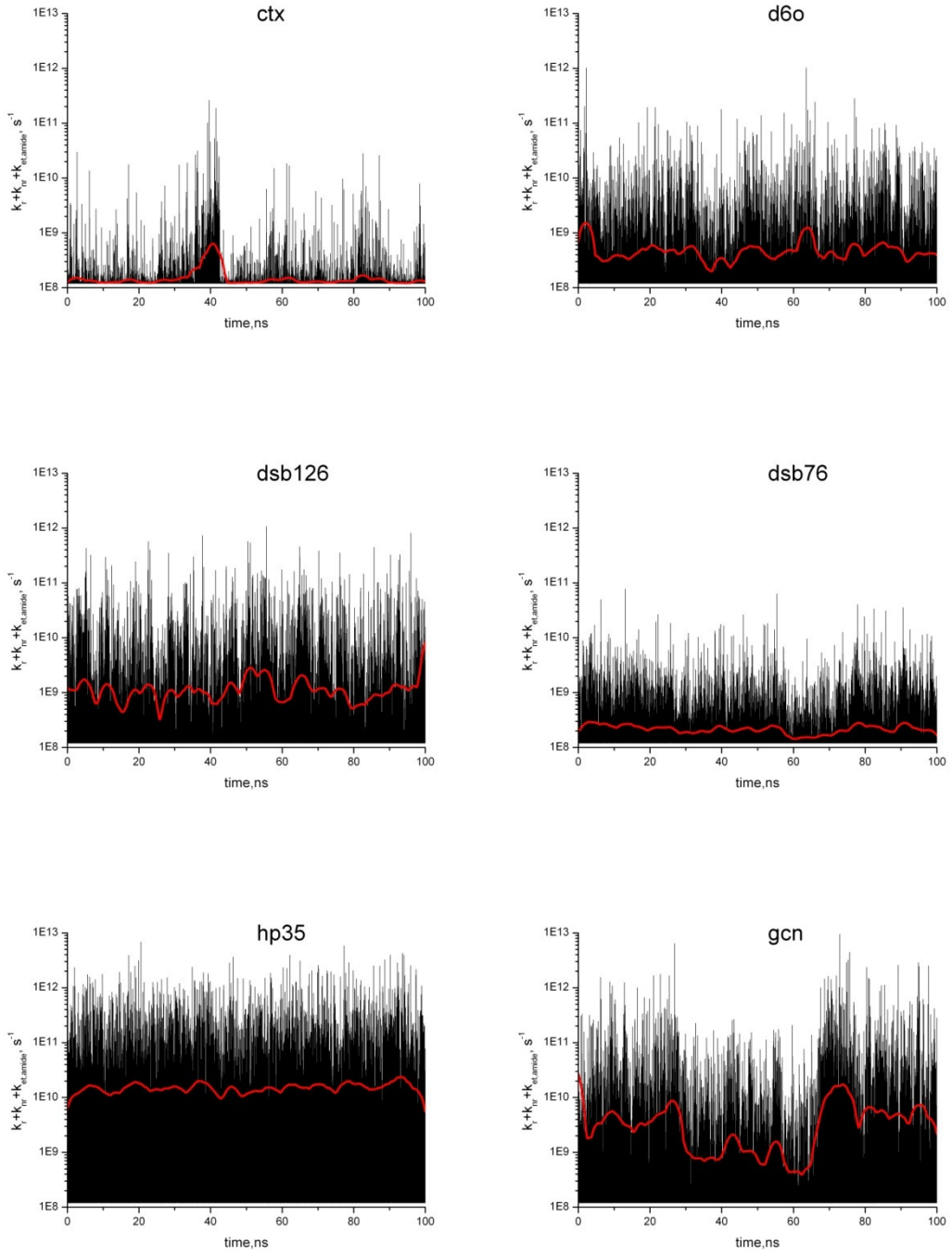

Figure 3.16 Time traces of the sum of all decay constants, $k_r + k_{nr} + k_{et,amide}$, as a function of time for ctx, d6o, dsb126, dsb76, hp35 and gnc. The red line shows the smoothed curved using an average window of 5 ns.

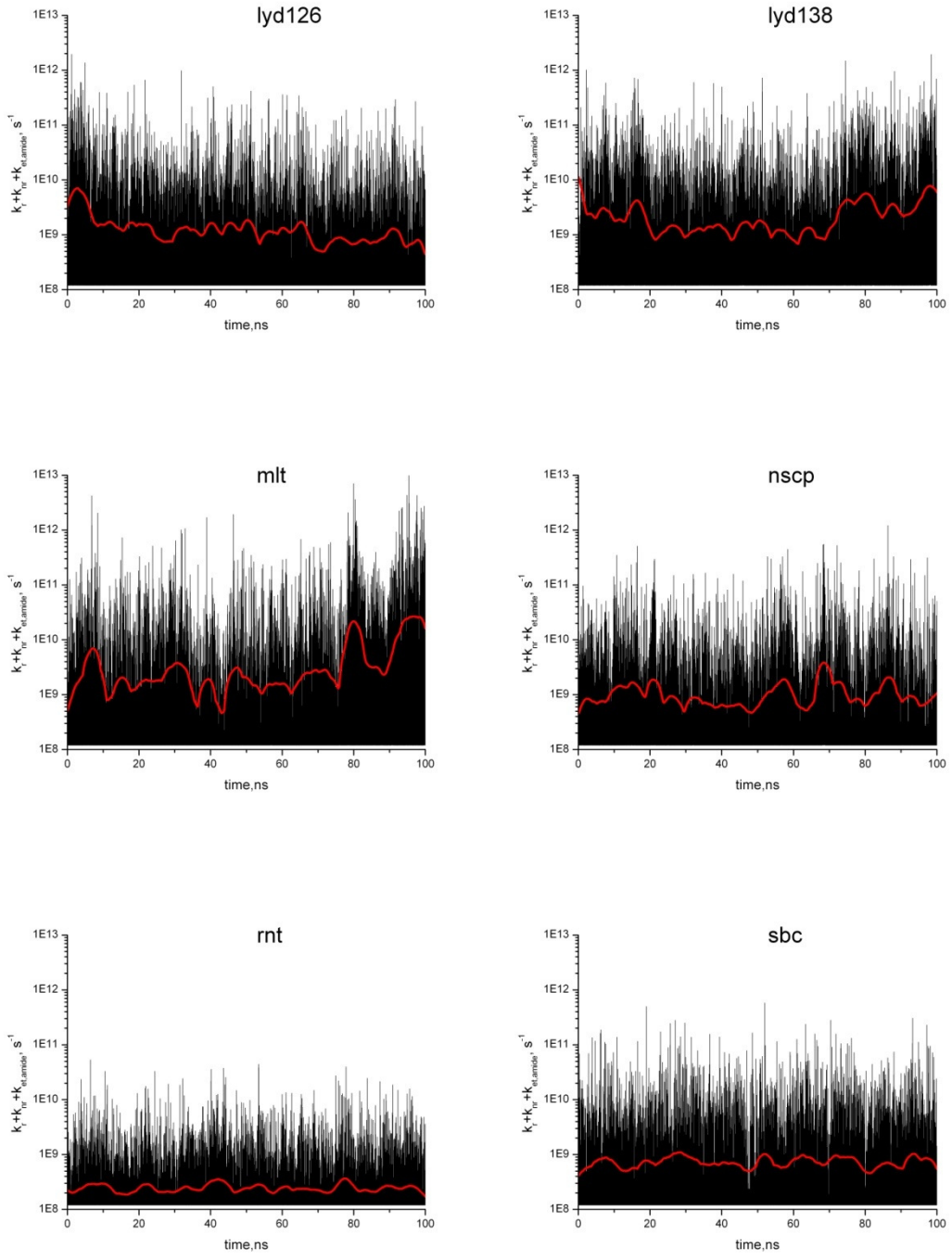

Figure 3.17 Time traces of the sum of all decay constants, $k_r + k_{nr} + k_{et,amide}$, as a function of time for lyd126, lyd138, mlt, nscp, rnt and sbc. The red line shows the smoothed curved using an average window of 5 ns.

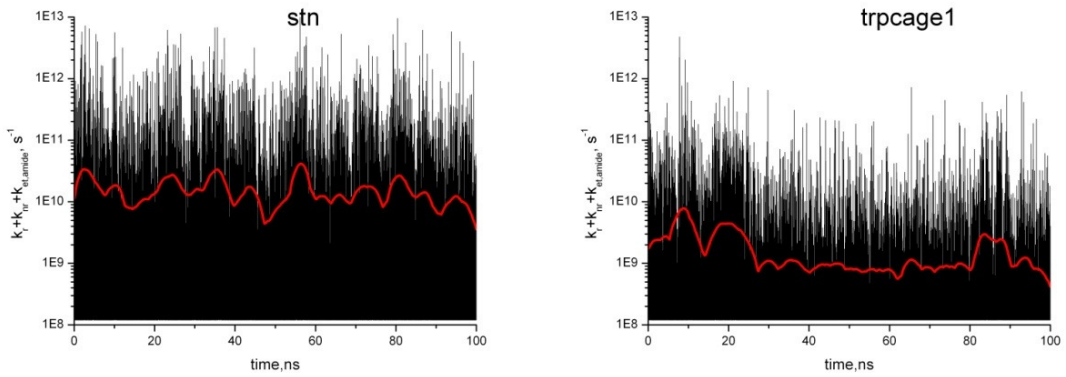

Figure 3.18 Time traces of the sum of all decay constants, $k_r + k_{nr} + k_{et,amide}$, as a function of time for stn and trpcage1. The red line shows the smoothed curved using an average window of 5 ns.

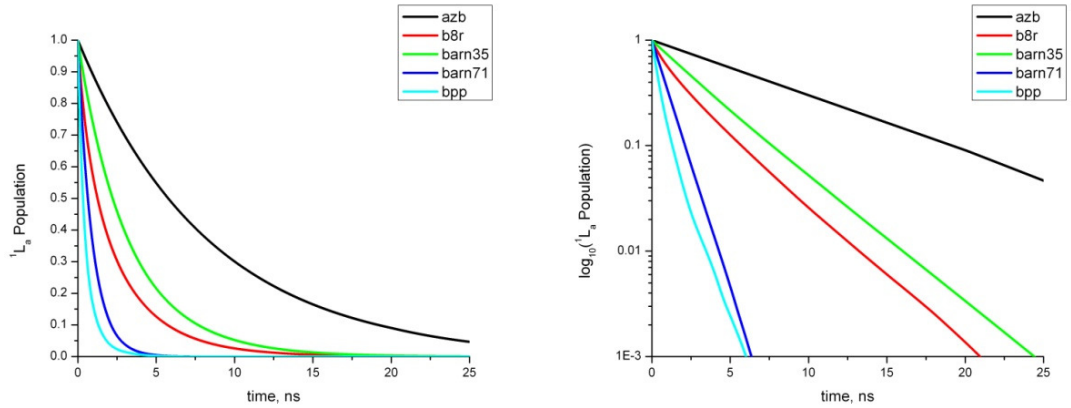

Figure 3.19 The average survival curves calculated using the time traces for the sum of the rate constants for each protein. A 25 ns window was used to calculate a survival curve, this window was then moved 0.05 forward in time and a new survival curve was calculated. All of these were saved for a total of 1601 survival curves. Each point in time was then averaged using the 1601 survival curves. Notice that bpp shows multi-exponential kinetics as expected from its heterogeneous behavior.

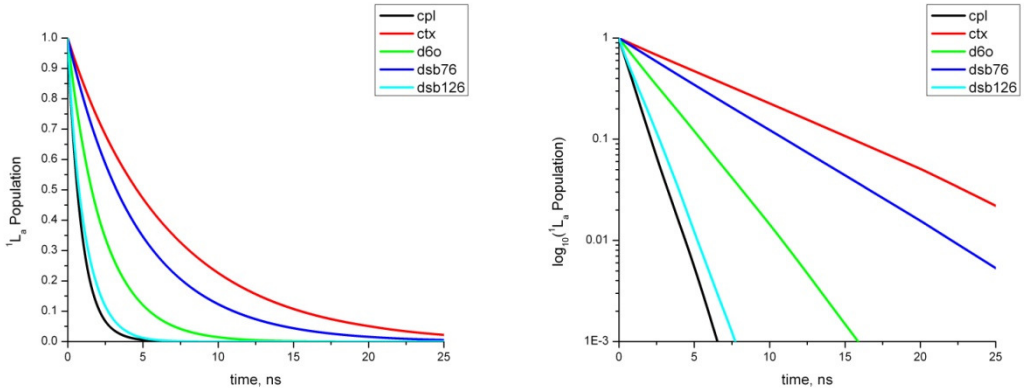

Figure 3.20 The average survival curves calculated using the time traces for the sum of the rate constants for each protein. A 25 ns window was used to calculate a survival curve, this window was then moved 0.05 forward in time and a new survival curve was calculated. All of these were saved for a total of 1601 survival curves. Each point in time was then averaged using the 1601 survival curves. Notice that cpl and dsb126 show multi-exponential kinetics as expected from its kinetic heterogeneous behavior.

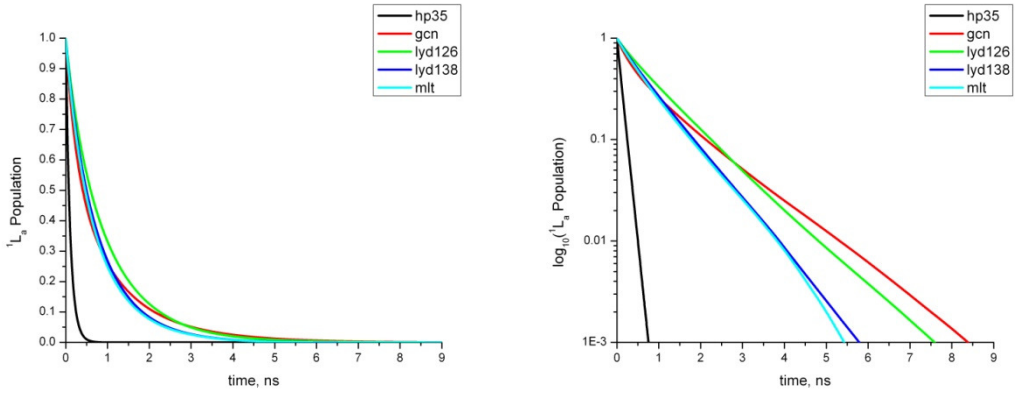

Figure 3.21 The average survival curves calculated using the time traces for the sum of the rate constants for each protein. A 25 ns window was used to calculate a survival curve, this window was then moved 0.05 forward in time and a new survival curve was calculated. All of these were saved for a total of 1601 survival curves. Each point in time was then averaged using the 1601 survival curves. Notice that gcn, mlt and lyd138 show multi-exponential kinetics as expected from its kinetic heterogeneous behavior.

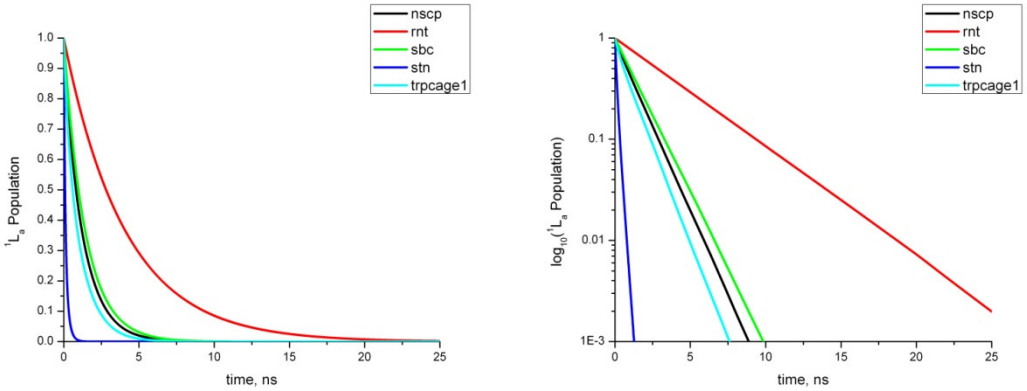

Figure 3.22 The average survival curves calculated using the time traces for the sum of the rate constants for each protein. A 25 ns window was used to calculate a survival curve, this window was then moved 0.05 forward in time and a new survival curve was calculated. All of these were saved for a total of 1601 survival curves. Each point in time was then averaged using the 1601 survival curves. Notice that none of these proteins show multi-exponential kinetics as expected from its kinetic heterogeneous behavior.

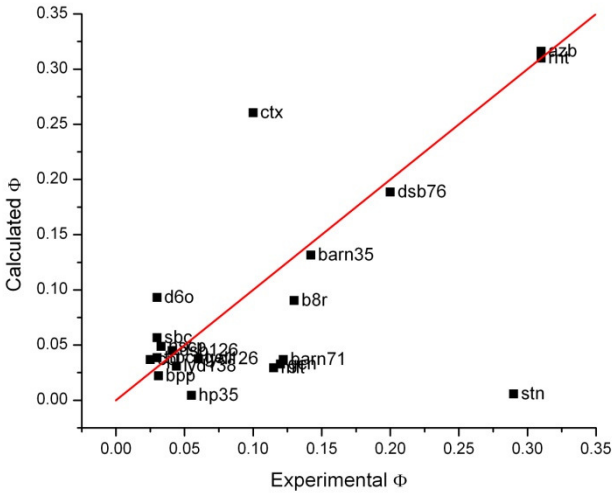

Figure 3.23 Fit obtained using AMBER99SB-ILDN and the survival curve procedure. The constant added to the energy gap in this case was 1940 cm^{-1}.

CHAPTER 4

CALCULATION OF FLUORESCENCE QUANTUM YIELDS CONTROLLED BY HETEROGENEITY

Introduction

The calculation of tryptophan fluorescence quantum yield is dependent on whether there is heterogeneity in the lifetimes. This heterogeneity has two limits: the fast exchange limit, in which the observed lifetime is equal to the average lifetime and the slow exchange limit in which we observe two lifetimes, corresponding to the pure lifetimes of each the species present. Another region of this heterogeneity is one in which the fluorescence decay rates are close to the exchange rates between different species. In this region we expect to see two lifetimes that are not the pure lifetimes of each of the species. We constructed a model that is composed of three states (Figure 4.1) to study how heterogeneity determines the fluorescence quantum yield. The equations that govern this three state model are the following:

$$\begin{bmatrix} \frac{dA}{dt} \\ \frac{dB}{dt} \\ \frac{dGS}{dt} \end{bmatrix} = \begin{bmatrix} -(k_f + k_{dA}) & k_b & 0 \\ k_f & -(k_b + k_{dB}) & 0 \\ k_{dA} & k_{dB} & 0 \end{bmatrix} \begin{bmatrix} A \\ B \\ GS \end{bmatrix}$$

In order to solve these coupled equations we first find the eigenvalues of the matrix. There are three possible eigenvalues: the eigenvalue for equilibrium i.e. $\lambda_{eq} = 0$

and the eigenvalues for non-equilibrium conditions are given by the following expression:

$$\lambda_{\pm} = -\frac{k_f + k_b + k_{dA} + k_{dB}}{2} \pm \frac{\sqrt{\left(k_f + k_b + k_{dA} + k_{dB}\right)^2 - 4\left(k_f k_{dB} + k_b k_{dA} + k_{dA} k_{dB}\right)}}{2}$$

We are going to examine the case in which k_{etA} is 0.5 ns^{-1} and k_{etB} is 2 ns^{-1} since it is pertinent to the study of fluorescence of tryptophan in HP-35. The behavior of λ_+ is the following: at slow exchange rates its value is equal to the fast decay component (k_{dB}); at the fast exchange rate its value is approaches the sum of k_f and k_b. The behavior of λ_+ is the following: at slow exchange rates its value is equal to the slow decay component (k_{dA}); at the fast exchange limit its value approaches the average of k_{dA} and k_{dB}. The population of each state in time is the following:

$$A(t) = c_{A+}\exp(\lambda_+ t) + c_{A-}\exp(\lambda_- t) + 1$$
$$B(t) = c_{B+}\exp(\lambda_+ t) + c_{B-}\exp(\lambda_- t) + 1$$
$$GS(t) = c_{GS+}\exp(\lambda_+ t) + c_{GS-}\exp(\lambda_- t) + 1$$

Where c_{A+}, c_{A-}, c_{B+}, c_{B-}, c_{GS+} and c_{GS-} are determined by the initial conditions; we assume that species A and B are excited with equal probability and their populations in the excited state are determined by k_f and k_b. GS is the initially excited state population that has returned to the ground state.

$$A(0) = \frac{1}{1+K_{eq}} \quad B(0) = \frac{K_{eq}}{1+K_{eq}} \quad GS(0) = 0$$
$$K_{eq} = \frac{k_f}{k_b}$$

The decay of the excited state can therefore be expressed as:

$$^1L_a(t) = 1 - GS(t) = c_{GS+}\exp(\lambda_+ t) + c_{GS-}\exp(\lambda_- t)$$

The quantum yield is then equal to:

$$\Phi = k_r \int_0^\infty \left[c_{GS+} \exp(\lambda_+ t) + c_{GS-} \exp(\lambda_- t) \right] dt = -k_r \left\{ \frac{c_{GS+}}{\lambda_+} + \frac{c_{GS-}}{\lambda_-} \right\}$$

Now we can track the behavior of φ, c_{GS+}, c_{GS-}, λ_+ and λ_-. Figure 4.2 shows how the quantum yield changes when we change the exchange rate without altering the equilibrium between species A and B(assuming K_{eq}=1, k_f=k_b=k_{ex}). At slow exchange rates the following formula applies:

$$\Phi_{slow,exchange} = c_{GS+}\Phi_A + c_{GS-}\Phi_B = c_{GS+}\left[\frac{k_r}{k_r + k_{nr} + k_{dA}} \right] + c_{GS-}\left[\frac{k_r}{k_r + k_{nr} + k_{dB}} \right]$$

At the other extreme, at fast exchange rates the following formula applies:

$$\Phi_{fast,exchange} = \frac{k_r}{k_r + k_{nr} + c_{GS+}k_{dA} + c_{GS-}k_{dB}}$$

The interesting region is where the exchange rate is in between k_{etA} and k_{etB}. In this region the equations to calculate the quantum yield, for fast exchange and the slow exchange, do not apply and instead we use the equation determined by integration of the exited state decay. Figure 4.3 shows how the coefficients and the fluorescence lifetime for each of the species change as a function of the rate of exchange.

We made a plot of quantum yield φ as a function of $\log_{10}(k_f+k_b)=\log_{10}(k_{relax})$ and $\log_{10}(K_{eq})=\log_{10}([C]/[F])$ (Figure 4.4). This graph shows how the quantum yield can change by varying the relaxation time without changing K_{eq}. As well as how sensitive the quantum yield is as function of $K_{eq.}$ A change in K_{eq} or k_{relax} can change the observed quantum yield.

The procedure to calculate quantum yields was important in our past calculations of fluorescence intensity because in some of our proteins we have noticed the existence of heterogeneity. In some cases we have seen discrete fluorescent states that have a lifetime similar to the fluorescence lifetime and therefore the fast exchange approximation will not necessarily hold. It is also important in the calculation of the fluorescence of HP35 since as shown in figure 5.4 there are also discrete energetic states for histidine CT state. The calculation of the fluorescence quantum yield in this case is therefore not straight forward and should be calculated by the procedure described here.

Model of Tryptophan Fluorescence Quantum Yield for Thermal Unfolding of HP-35

The model we use to model equilibrium fluorescence is shown in Figure 4.5. In this model we have a folded state that has two populations FF and FC, which stands for Folded Far and Folded Close (the far and close refer to the distance between tryptophan and histidine) and an unfolded state that also has a far and close population, UF and UC. The values of k_{fold} and k_{unfold} as a function of temperature are taken directly from kinetic data fitting by Kubelka et al.[8]. The kinetic constants for decay to the ground state are k_{df} and k_{dc} (the decay constant from the far conformation and the decay constant in the close conformation). k_{df} and k_{dc} are set to:

$$k_{df} = k_r + k_{nr} + k_{etAm}$$
$$k_{dc} = k_r + k_{nr} + k_{etAm} + k_{etHis}$$

k_{etAm} is the average of the electron transfer rate due to amide quenching and k_{etHis} is the average electron transfer due to histidine when it is close to tryptophan. In these

equations there is no temperature dependence for the fluorescence quantum yield, therefore we will only see changes in quantum yield due to changes in population.

The sum of $k_{ff,fc}$ and $k_{fc,ff}$ was made to be equal to the fast component seen in temperature jump experiments ie $k_{ff,fc} + k_{fc,ff} = (70^{-1})$ns. Temperature dependence in the folded states was introduced so that when there is a temperature jump we will see a nanosecond lifetime due to depopulation of the close conformation (without it there wouldn't be a fast (ns) increase of fluorescence quantum yield after a temperature jump). Additionally $k_{uf,uc} + k_{uc,uf}$ was also made equal to (70^{-1})ns with no temperature dependence.

$$K_{eqf}(T) = K_{eqf}^{\circ} \exp\left(-\left(\frac{\Delta H^{\circ}}{R}\right)\left(\frac{1}{T} - \frac{1}{T^{\circ}}\right)\right)$$

$$k_{FF,FC}(T) = \left(\frac{1}{70ns}\right)\left(\frac{K_{eqf}(T)}{K_{eqf}(T)+1}\right)$$

$$k_{FC,FF}(T) = \left(\frac{1}{70ns}\right)\left(\frac{1}{K_{eqf}(T)+1}\right)$$

The starting populations of FF, FC, UF and UC were made so as to follow the equilibrium constants defined by the model (without taking into account the ground state). GS was again set to 0. The quantum yield was calculated using the above mentioned procedure (ie building a kinetic model, finding its eigenvalues, calculating the coefficients that satisfy the initial conditions and integrating the result over time). The results are shown in figure 4.6.

Additionally we simulated temperature jump experiment with the model. This was done by eliminating all of the decays to the ground state and let the system

equilibrate (in order to reproduce steady state conditions). For each specific set of concentrations we carried out a quantum yield calculation using the five state model that includes decay to the ground state. The results for this modeling are shown in figure 4.7 for a 9 K temperature jump starting at 328 K. The trace shows the two decay constants: a fast ~70 ns decay, which is equilibration between the far and close configuration in the folded state and a 5 μs decay which is the sum of k_{fold} and k_{unfold}.

Conclusions

Tryptophan fluorescence quantum yield cannot be calculated from an average k_{et} if there are two or more states that have lifetimes that are close to the lifetime of the fluorescent state of tryptophan, instead to calculate the quantum yield under these conditions you can either use an average survival curve approach as in chapter 2 of this thesis or by using a kinetics model approach. This procedure could reconcile previous quantum yield calculations by accounting for states that have a lower or higher quantum yield. On the other hand if the conversions are fast relative to the lifetime the average rate can be used to calculate the quantum yield.

In the case of HP-35 and WH-5 the fluorescence quantum yield needs to be calculated by using a kinetics model such as figure 4.5. The picture that arises is that the quantum yield for these two proteins can be modified either by the equilibrium constant or by the interconversion rate between the conformations that have histidine close or far.

Tables and Figures

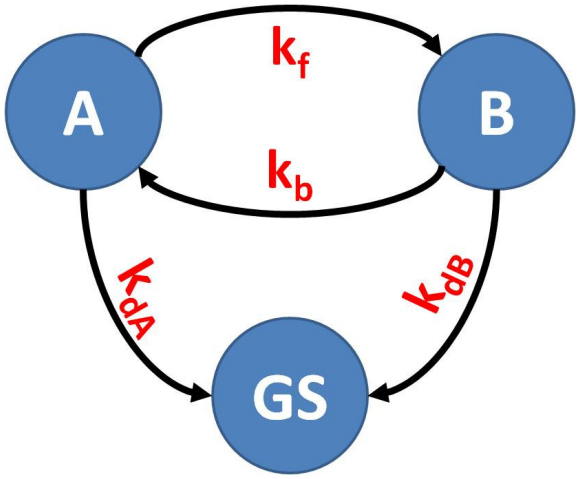

Figure 4.1: Three state kinetic model to calculate tryptophan fluorescence quantum yield. Species A and B are excited states of tryptophan that exchange at rates k_f and k_b. These also decay to the ground state (GS) with two different decay constants, $k_{dA}=k_r+k_{nr}+k_{etA}$ and $k_{dB}=k_r+k_{nr}+k_{etA}+k_{etB}$ (assuming that state B has both electron transfer rates). k_r is the radiative decay constant set at 0.04 ns^{-1}, k_{nr} is the nonradiative decay constant set at 0.08 ns^{-1}, and $k_{etA,B}$ is the rate of electron transfer from the excited tryptophan of species A and B respectively.

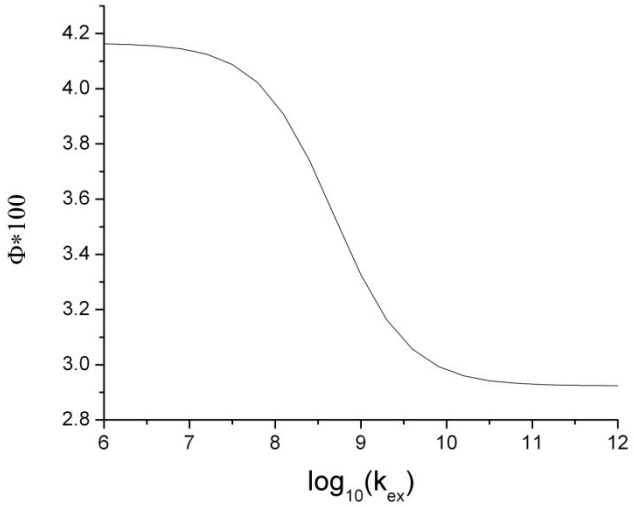

Figure 4.2: Fluorescence quantum yield Φ as a function of k_{ex}. Notice that the quantum yield starts to deviate at exchange rates that are close to the value of ketA.

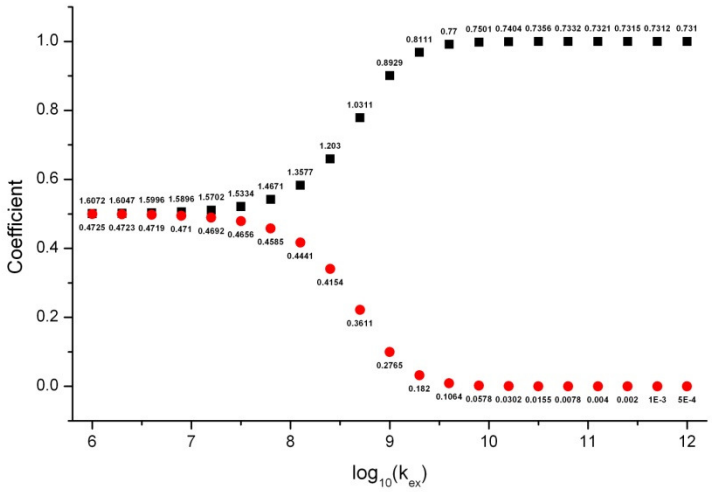

Figure 4.3: Coeffiecients c_{GS+} (black) and c_{GS-} (red) as a function of k_{ex}. As labels we have included the fluorescence lifetimes of each of the species as a function of k_{ex}.

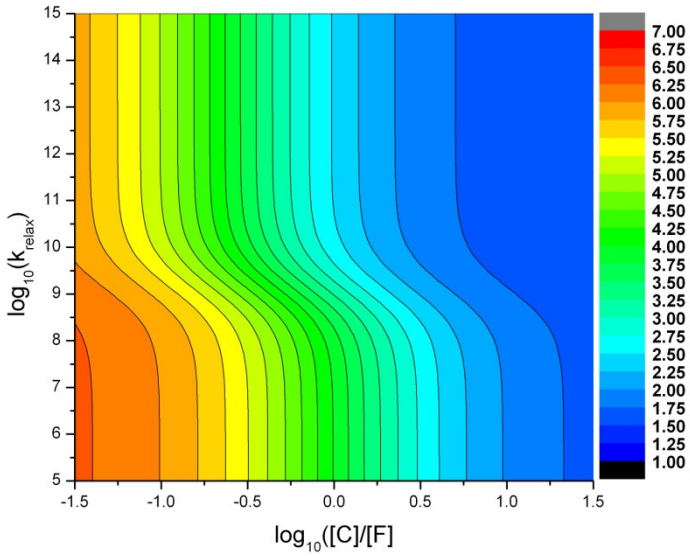

Figure 4.4: Contour graph of quantum yield as a function of $\log_{10}(k_{relax}) = k_f + k_b$ and the equilibrium constant K_{eq}. Observe how a change in either the relaxation rate or the equilibrium constant can change the quantum yield by 0.25.

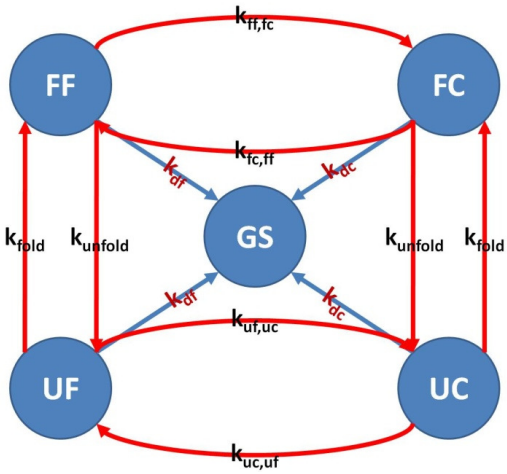

Figure 4.5: Kinetic model for calculating Tryptophan fluorescence quantum yields as a function of temperature. The species are the following: FF is folded far, FC is folded close, UF is unfolded far, UC is unfolded close and GS is ground state. Far and close refers to the distance between tryptophan and histidine. This distance in turn determines the decay of each state as seen in the above equations. $k_{ff,fc}$ is the reaction rate from FF to FC, $k_{fc,ff}$ is the reaction rate from FC to FF, $k_{uf,uc}$ is the reaction rate from UF to UF, $k_{uc,uf}$ is the reaction rate from UC to UF, k_{fold} is the folding rate, k_{unfold} is the unfolding rate, k_{df} is the decay rate to the ground state when histidine is far and k_{dc} is the decay rate to the ground state when histidine is close. k_{df} is equal to $k_r + k_{nr} + k_{et,amide}$ and k_{dc} is equal to $k_r + k_{nr} + k_{et,amide} + k_{et,His}$. k_r is the radiative decay rate, k_{nr} is the non-radiative decay rate, $k_{et,amide}$ is the rate of electron transfer to amide and $k_{et,His}$ is the rate of electron transfer to the histidine cation.

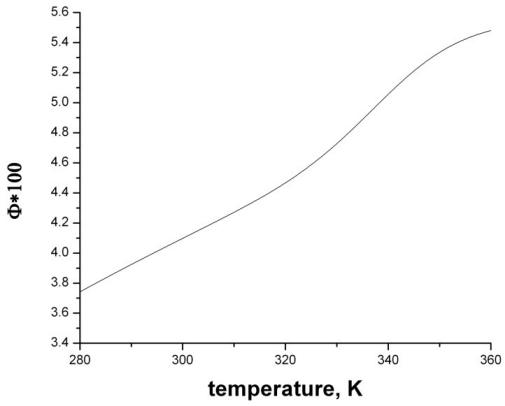

Figure 4.6: Change in tryptophan fluorescence as a function of temperature, when there is no temperature dependence for k_{nr}. The slope from 280 K to 320 K is due to the temperature dependence of the close/far equilibrium in the folded basis, the inflection point at 342 is the transition when the folded and unfolded basins are equally populated.

Table 4.1 Parameters used to calculate the quantum yield as a function of temperature.

Parameter	Value
K_{eq}^{o}	0.9
T^{o}	300 K
ΔH_{fc}	-2.39 kcal/mol
k_{ufuc}	$2.5\mu s^{-1}$
k_{ucuf}	$12\ \mu s^{-1}$
k_{etam}	$0.5\ ns^{-1}$
k_{etHis}	$2.0\ ns\ s^{-1}$
k_r	$0.04\ ns^{-1}$
k_{nr}	$0.08\ ns^{-1}$
T_m	342 K [8]
$\Delta H_{fold/unfold}$	27.00 kcal/mol [8]
k_{0f}	$0.23\ \mu s^{-1}$ [8]
T_0	300 K [a]
ΔH_f^{+}	1.0 cal/mol[8]

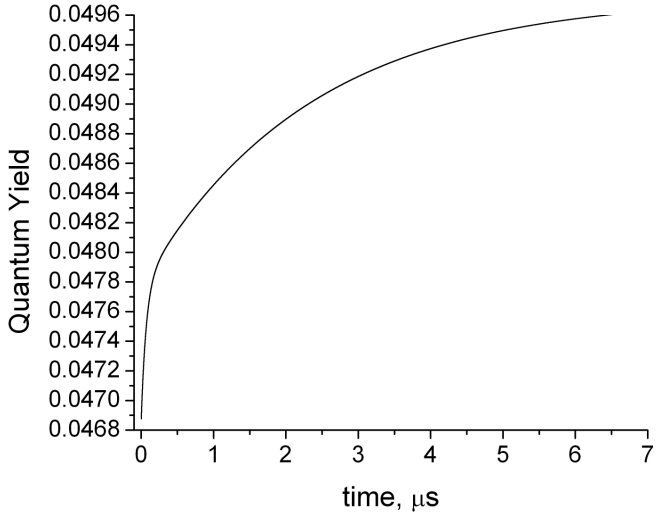

Figure 4.7: Time trace for the fluorescence quantum yield after a temperature jump 328 K to 337 K simulation.

CHAPTER 5

HISTIDINE AND AMIDE ELECTRON TRANSFER QUENCHING AND PROTEIN FOLDING: Ac-W-A-A-A-H^+-NH_2. AND THE VILLIN HEADPIECE SUBDOMAIN

WH5 Simulation

WH5 is a synthetic peptide that has the following sequence: Ac-W-A-A-A-H^+-NH_2.[7,57-59] It is a model to the motif found in the villin headpiece subdomain, since it contains W and H^+ separated by four alanine residues. The fluorescence of WH5 after a temperature jump experiment has been studied by two independent research groups. Eaton and coworkers found that as the temperature increases the fluorescence quantum yield increases[7,57]. Zewail and coworkers did fluorescence measurements on the same system peptide and found a decreasing fluorescence signal after relaxation and a sub-nanosecond relaxation constant[58,59]. CD spectra show that the helical content at 310 K is about 15% according to Zewail et al.. In this peptide there are two possible quenchers: the amide backbone of tryptophan and the histidine cation. Let us first talk about amide quenching.

Amide Quenching

Amide electron transfer rate constant for the peptide backbone near tryptophan was calculated using the same methodology as in chapter 2. In short we calculated the instantaneous energy gap at every point from ZINDO[96-99] calculations and we assigned an

average V_{et} of 240cm^{-1} for the trajectory. This average comes from the average calculated for the villin headpiece subdomain. The results are shown in figures 5.1.

The upper panel shows how the amide energy gap as a function of time. The energy gap is defined as the energy difference between the fluorescent state of tryptophan and the charge transfer state wherein an electron has been transferred from the tryptophan ring to amide backbone. It can be seen that the average energy gap for amide electron transfer is a constant that has fluctuations that are short compared to the fluorescence lifetime. Therefore the correct way to calculate the observed fluorescence quantum yield would be to take the average observed gap and calculate the average Frank-Condon density of states (in other words we expect the fluorescence decay to be mono-exponential).

The middle panel of figure 5.1 shows how τ^{-1} (which is the sum of $k_r + k_{nr} + k_{et,amide}$) fluctuates in time. If its value is close to 1×10^8 s^{-1}, we expect to observe a quantum yield close to that of 3MI; there is no electron transfer mechanism is available. The smoothed curved showing what we expect $<\tau^{-1}>$ in 5 ns is shown in red. From this data we can calculate is 1×10^9 s^{-1} which gives an expected quantum yield of about 0.04.

The bottom of figure 5.1 also shows how the C_α RMSD varies with time for WH5. The C_α RMSD is calculated in reference to the perfect helical structure in which all dihedral angles are set to their ideal α helical structure and the resulting structure is solvated and then energy minimized. From figure 5.1 it is easily seen that amide electron transfer rate constants are not sensitive to the folding. When Eaton and coworkers observed the steady state fluorescence quantum yield for a similar protein,

whose sequence is WAAA, they observed temperature dependence similar to what is expected for n-acetyl tryptophan amide[7]. This gives credit to the notion that amide electron transfer would not be able to report changes in folding.

Histidine Quenching

Figure 5.2 has four different graphs that show, the tryptophan and histidine center of mass distance, the C_α RMSD, τ^{-1}, and the energy gap difference between the fluorescent state of tryptophan and the charge transfer state to the histidine ring. This figure shows that all four are well correlated, showing that histidine can be used to follow folding. This graph shows that the distance between donor and acceptor controls the energy gap and the rate of electron transfer, indicating that V_{et} is not what determines the electron rate constant. A constant was added to the energy so that the close conformation would have the same electron transfer rate as a mutant of barnase[102] that has the histidine cation stacked next to tryptophan. This electron transfer rate is about 1×10^9 s^{-1} which is agreement with the section in which the distance is ~ 4.5 Å.

It is interesting that even though histidine is a potent quencher[47] of the fluorescence of tryptophan it only contributes about the same as amide quenching. And it does so only transiently for a few nanoseconds. We therefore expect that WH5 will have non-exponential fluorescence decay due to the protein structures that have a predisposed conformation in which the rate will be fast and protein structures that will not have that conformation that will decay slower. From looking at this data we can suggest that tryptophan fluorescence quantum yield is a sensitive probe to observe changes and kinetics in this small peptide.

HP-35 Simulations

Folded Simulation

Simulations were carried out starting from the crystal structure of the villin headpiece subdomain crystal structures (PDB: 1yrf or 2f4k). All proteins were simulated using GROMACS 4.0.7[75,83-85] and using the same sequence of steps and parameters as the 25 ns simulations using the OPLS-AA force field (see chapter 2). The final structure after the NPT equilibration was used as the initial structure for the folded simulation.

One simulation was carried out on 1yrf using either the OPLS-AA[91-95] or the AMBER99SB-ILDN[101] force field. The OPLS-AA simulation ran for 1.2 μs, saving every 10 ps time frame, the AMBER99SB-ILDN simulation ran for 1.0 μs and we saved a frame every 2 ps. Additionally another simulation was carried out on 2f4k, which is a double norleucine mutant, this ran for 0.9 μs and we saved a frame every 10 ps. Zerner's INDO/S CIS was used to calculate the transition energies for the ^1L_a and amide or histidine charge transfer states for each of these frames.

Figure 5.3 shows the results for electron transfer do to amide quenching. The overall observed behavior is similar in nature to that of WH5, a rate is independent of the C_α RMSD of the protein. Additionally its fluctuations are faster than the fluorescence lifetime and the average behavior will suffice since we expect that there is heterogeneity in the electron transfer rate due to amide.

Figure 5.4 shows how the center of mass distance between tryptophan and histidine, the C_α RMSD, τ^{-1} and the energy gap to the histidine charge transfer state vary

with time. Again it shows similar behavior to that of WH5 with the exception that electron transfer rate to histidine is not correlated to C_α RMSD. The C_α RMSD of this simulation suggests that we are seeing a little bit of unfolding as the simulation progresses. Even though the C_α RMSD changes significantly the effect on the amide and histidine electron transfer properties because the α helical turn between tryptophan and histidine remains intact throughout.

Figure 5.5 shows the 1 μs simulation using the AMBER99SB-ILDN force field. This simulation shows similar behavior in the center of mass distances between tryptophan and histidine but it shows a more stable C_α RMSD. Additionally the 2F4K simulation, figure 5.6 shows the same behavior for the distance between tryptophan and histidine as a function of time. This figure also shows that there is a very specific rotamer transition that correlates with the decrease in distance. χ_1 and χ_2 simultaneously change to produce a small tryptophan/histidine distance. χ_1 has to go from 60° to +/- 180° where as χ_2 has to change from -90° to +90°. This behavior is observed in all three simulations of 1yrf with the OPLS-AA force field, 1yrf with the AMBER99SB-ILDN force field and 2f4k with the OPLS-AA.

Unfolded Simulation

A simulation was carried in which the same steps were taken as in the folded simulations to set up the system and equilibrate it, except that the temperature was set at 400K. The idea here is to cause the protein to unfold. Results indicate that amide quenching is not sensitive to the protein structure and that histidine will show a significant amount of transitions between far and close conformations. Figure 5.7 shows

the results for the center of mass distance between tryptophan and histidine and the histidine energy gap for 400 ns of simulation. This graph shows similar behavior to the energy observed for WH5 at the melting temperature. This suggests that due to the presence of histidine in the protein we should be able to tell the difference between 1yrf at 300K and 1yrf at 400K due to differences in populations that have a higher electron transfer rate for histidine.

Factors that Determine Histidine Quenching

One of the most important factors that determined if histidine will be able to quench of not is the center of mass distance between tryptophan and histidine. It is good to remember that there are two contributing factors that determine the electron transfer rate:

$$k_{et}(\Delta E_{00}) = \frac{2\pi}{\hbar} V^2 \rho_{FC}(\Delta E_{00})$$

One of them is V_{et}^2 and the second is the Frank-Condon weighted density of states, ρ_{FC}. In order to analyze which of these two determine whether histidine would be able to quench we plotted on a logarithmic scale V_{et}^2 and ρ_{FC} as a function of distance. Figure 5.8 shows the results. In these graphs any electron transfer that is smaller than 1×10^7 s^{-1} will not alter the observed fluorescence quantum yield. As can be seen from the center panel of figure 5.8, within a few tenths of an angstrom (from about 5.2 Å to 5.4 Å) the electron transfer rate to histidine falls quickly by two orders of magnitude. This effect is mirrored by ρ_{FC} whereas V_{et}^2 does not change its value up until 6.0 Å. Therefore what controls the ability of histidine to quench is not by V^2 but rather by ρ_{FC}. ρ_{FC} is

delicately controlled by the energy gap difference between the charge transfer state and the fluorescent state of tryptophan. Therefore we turn our attention onto what makes the energy gap difference change by 30,000 cm^{-1}

Previous work has shown that the electron transfer energy difference is affected by the electrostatic environment around the tryptophan ring and the acceptor group[42-44]. It was shown by Callis[42-44] et al. that a positive charge near the amide backbone and a negative charge near the ring will enhance electron transfer and decrease the observed quantum yield for that particular protein. It is therefore natural for us to try and understand the underlying fluctuations in the energy gap by examining the classical electrostatic energy difference between the charge transfer state and the fluorescent state of tryptophan. We calculate this energy by using the following formula:

$$\Delta E_{env} = \sum_{\alpha} V_{\alpha} \Delta \rho_{\alpha}$$

Where ΔE_{env} is the classical electrostatic energy difference between the charge transfer state and the fluorescent state of tryptophan, V_α is the coulomb potential, $\Delta\rho_\alpha$ is the charge density difference in the αth quantum atom. V_α is defined as:

$$V_{\alpha} = \sum_{k} \frac{q_k}{r_{\alpha,k}}$$

Where q_k is the charge at the kth non-quantum atom and $r_{\alpha,k}$ is the distance between the kth non-quantum atom and the αth quantum atom.

By restricting the k and α indices appropriately, one may estimate the major electrostatic contributions from any part of the environment for any of the possible quenching electron transfers.

We decided to decompose this classical electrostatic energy difference into two parts. First we looked at the protein contribution this is shown in figure 5.9 as the black line. This graph shows that the center of mass distance between histidine and tryptophan does not significantly change the electrostatic energy difference due to the protein. The solvent contribution on the other hand is large and very sensitive to the distance between tryptophan and histidine. This is shown as the blue line in figure 5.9. The electrostatic energy increases rapidly at about the same distance in which the density of states (figure 5.8) decreases by order of magnitude. For completeness the total electrostatic energy difference was included as the red line in figure 5.9. This graph shows that histidine electron transfer is very sensitive with distance and that the solvent is controlling the electron transfer rate constant.

To understand this we analyzed different snapshots in which tryptophan and histidine were either close or very far apart. The analysis for the close conformation is shown in figure 5.10. Blue waters indicate waters that stabilize the charge transfer state and red water indicate waters that destabilize the charge transfer state. There are a number of waters that are on the face of tryptophan that feel the electrostatic charge of histidine on the other side and orient their dipoles to that positive charge. All of these waters therefore stabilize the charge transfer state since after the electron leaves the tryptophan ring it leaves a positive charge behind which is stabilized these waters. On the other hand the waters around histidine destabilize the charge transfer state since their dipoles are pointed so as to stabilize a positive charge rather than a neutral charge.

Figure 5.11 shows a conformation in which tryptophan and histidine are far. When this distance is large the number of waters that do not stabilize the neutral charge left on the histidine ring after electron transfer significantly increase, while the water around tryptophan does not favor or oppose electron transfer.

Conclusions

Results indicate amide electron transfer is always present with a rate of about 0.5 ns^{-1} and that histidine cation can quench transiently at a rate of about 2 ns^{-1}. The ability of histidine to quench the fluorescence of tryptophan is controlled by the dihedral rotamers of tryptophan and histidine. These rotamers are only present about 5% to 40% percent of the simulation time in folded state. This would seem to indicate that the equilibrium between far and close is what determines the overall quantum yield, and that tryptophan fluorescence could possibly distinguish between different structures of the protein, for example the flipped state of HP-35. The observed fluorescence quantum yield can be affected according to chapter 4 by either the rate of exchange between far and close or the equilibrium distribution between far and close. The rate of exchange is too slow to alter the observed quantum yield and the quantum yield is controlled by equilibrium distribution.

Figures 5.10 and 5.11 seem to indicate that the most important factor that determines histidine electron transfer rates is solvation around the rings. Close tryptophan/histidine distances have water contributions that nearly cancel each other

whereas far tryptophan/histidine distances have waters that overwhelming increase the energy gap.

Tables and Figures

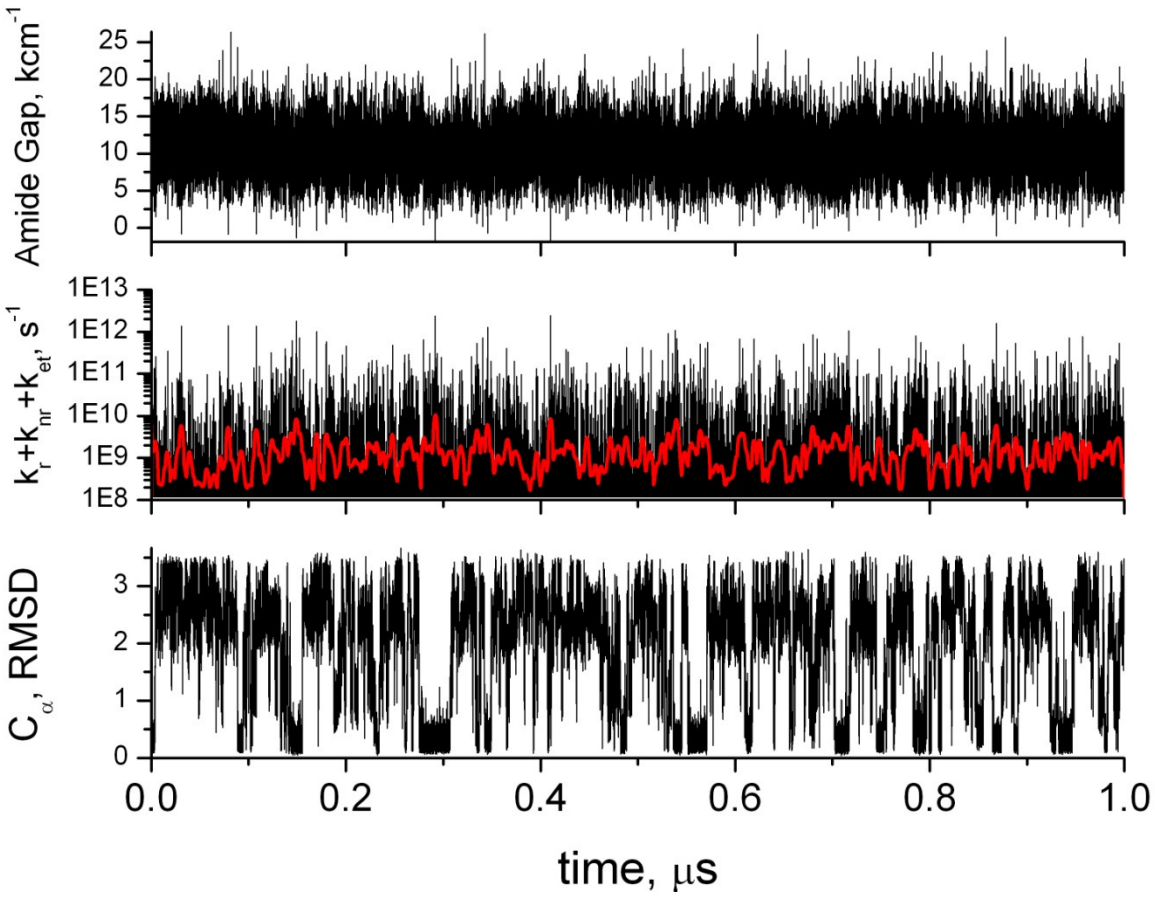

Figure 5.1 Amide energy gap for WH5, the value of the sum of the radiative, nonradiative and electron transfer rate constant (k_r and k_{nr} are constants set at 0.04 ns^{-1} and 0.08 ns^{-1}) and the C_α RMSD as a function of time using the OPLS-AA force field. The red line is the 5 ns average of the sum of the decay constants for WH5.

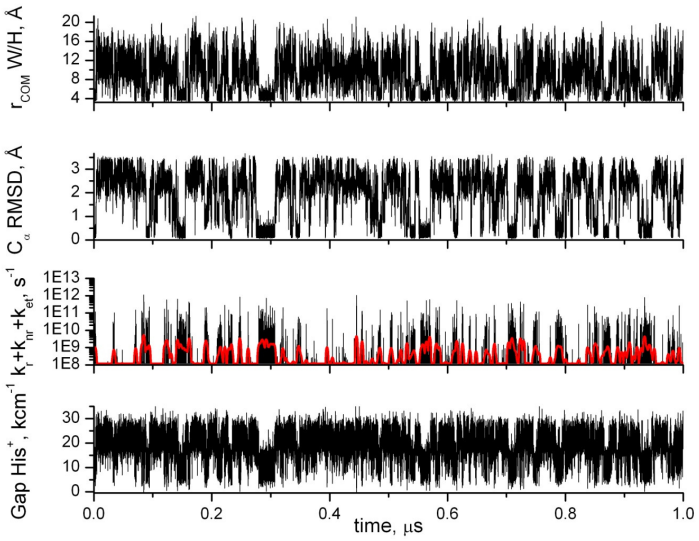

Figure 5.2 Results for WH5 and the OPLS-AA force field. Distance between tryptophan and histidine amino acids (r_{COM} W/H^+), C_α root mean square deviation, the $\log_{10}$ of the sum of the radiative, non-radiative and electron transfer to histidine rate constants, τ^{-1}, (k_r and k_{nr} are constants set at 0.04 ns^{-1} and 0.08 ns^{-1}) and the energy gap for electron transfer to histidine as a function of time. Notice the correlation between the C_α, r_{COM} W/H^+, τ^{-1}, and energy gap.

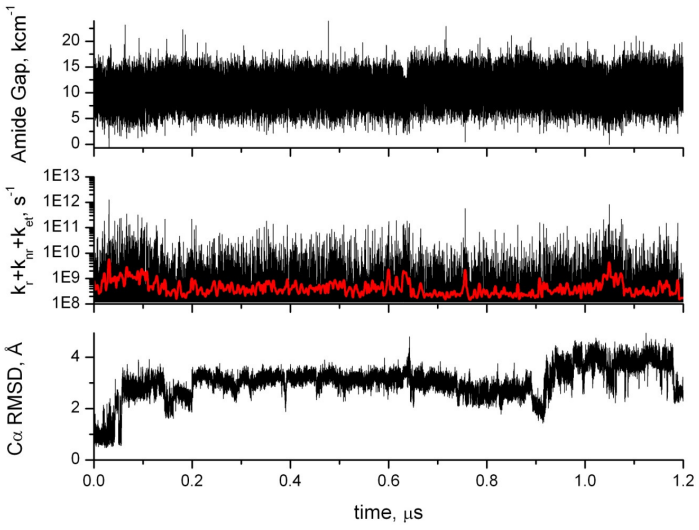

Figure 5.3 Amide energy gap for HP-35, the sum of the radiative, nonradiative and electron transfer rate constant to the amide backbone (k_r and k_{nr} are constants set at 0.04 ns^{-1} and 0.08 ns^{-1}) and the C_α RMSD (for residues 3-33) as a function of time using the OPLS-AA force field. The red line is the 5 ns average of the sum of the decay constants for HP-35.

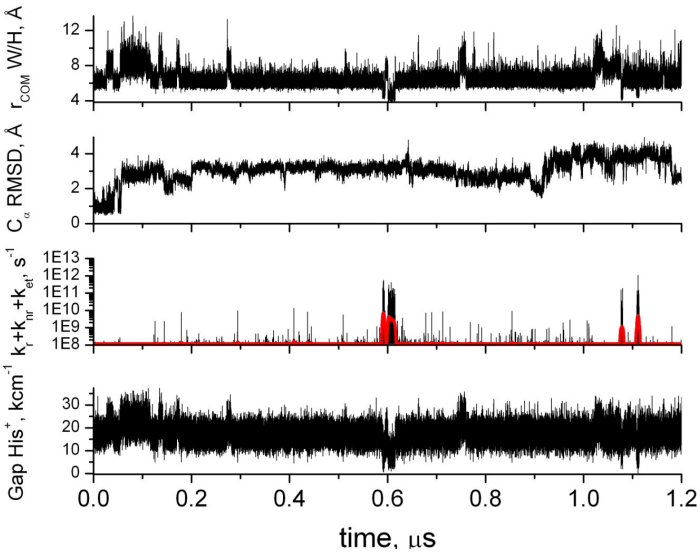

Figure 5.4 Results for HP-35 folded state simulation using the OPLS-AA force field. Distance between tryptophan and histidine amino acids (r_{COM} W/H^+), C_α root mean square deviation (for residues 3-33), the $\log_{10}$ of the sum of the radiative, non-radiative and electron transfer to histidine rate constants, τ^{-1}, (k_r and k_{nr} are constants set at 0.04 ns^{-1} and 0.08 ns^{-1}) and the energy gap for electron transfer to histidine as a function of time for HP-35 and the OPLS-AA force field. Notice the correlation between the C_α, r_{COM} W/H^+, τ^{-1}, and energy gap.

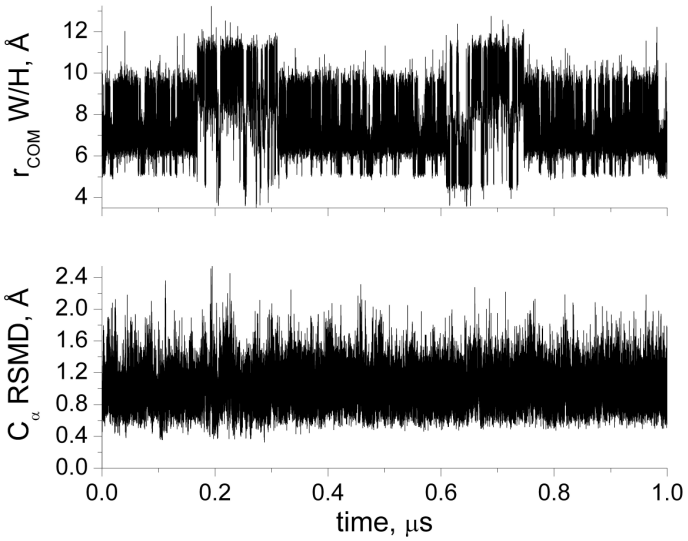

Figure 5.5 Center of mass distance between tryptophan and histidine and the C_α root mean square deviation for a villin subdomain simulation starting from the crystal structure (1yrf) at 300K using the AMBER99SB-ILDN force field.

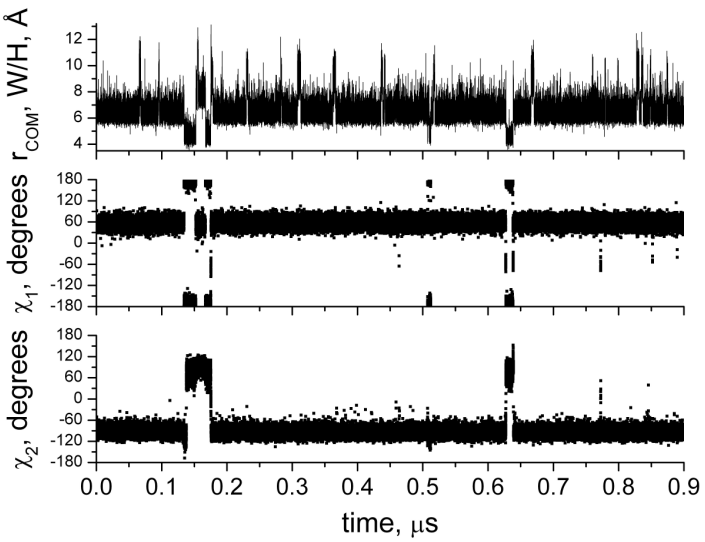

Figure 5.6 Time traces of the center of mass distance between tryptophan and histidine, χ_1 and χ_2 of tryptophan. Notice the correlation between the center of mass distance and the particular combination of χ_1 and χ_2. Only when χ_1 is -180° and χ_2 is +90° can histidine get close enough for it to efficiently quench the fluorescence of tryptophan via electron transfer. This simulation is for the villin headpiece subdomain double norleucine mutant (PDB code: 2F4K) using the OPLS-AA force field.

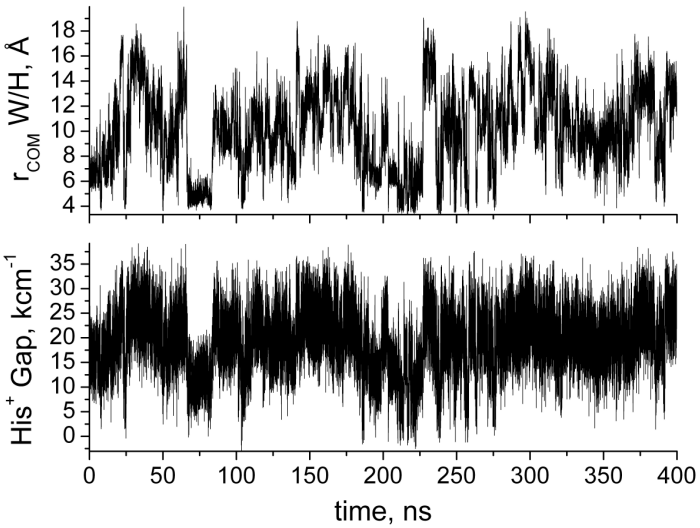

Figure 5.7 Histidine energy gap and the distance between tryptophan and histidine as a function of time for a simulation at 400K starting from the crystal structure of 1yrf. Notice how the distance is starting to behave like the WH5 in figure 5.2.

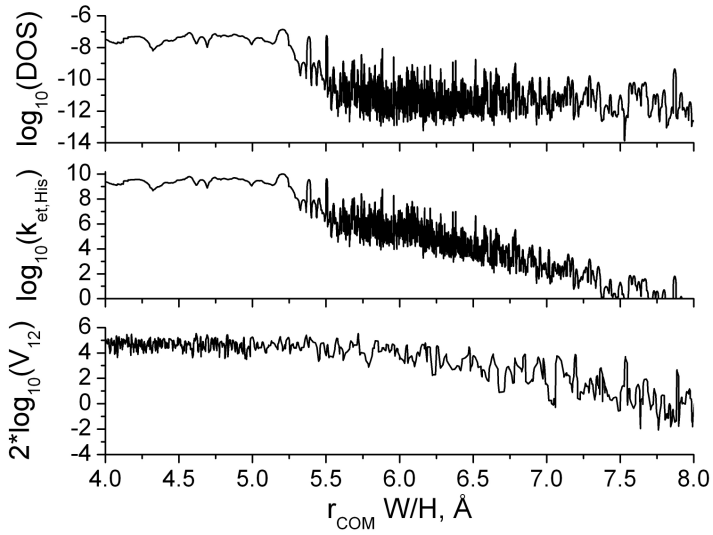

Figure 5.8 Distance dependence of the Frank-Condon density of states, the rate of electron transfer to histidine and V_{et} between the fluorescent state of tryptophan and the charge transfer state to histidine. At about 5.2 Å there is a drop off in the electron transfer rate that is correlated with a drop off in the Frank-Condon density of states. V_{et} is not correlated with this drop off in electron transfer rate.

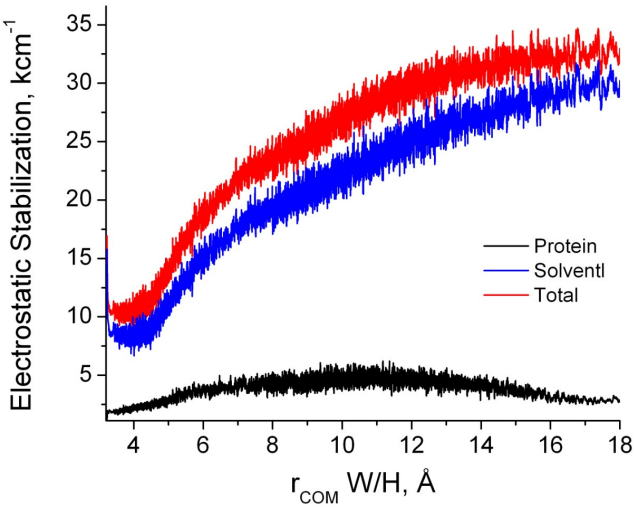

Figure 5.9 Protein, Solvent and total electrostatic stabilization of the charge transfer state as a function of the distance separation between tryptophan and histidine residues. The solvent shows dependence as the Frank-Condon density of states and the electron transfer rate constant to histidine. The protein shows no correlation to the Frank-Condon density of states and the electron transfer rate constant to histidine.

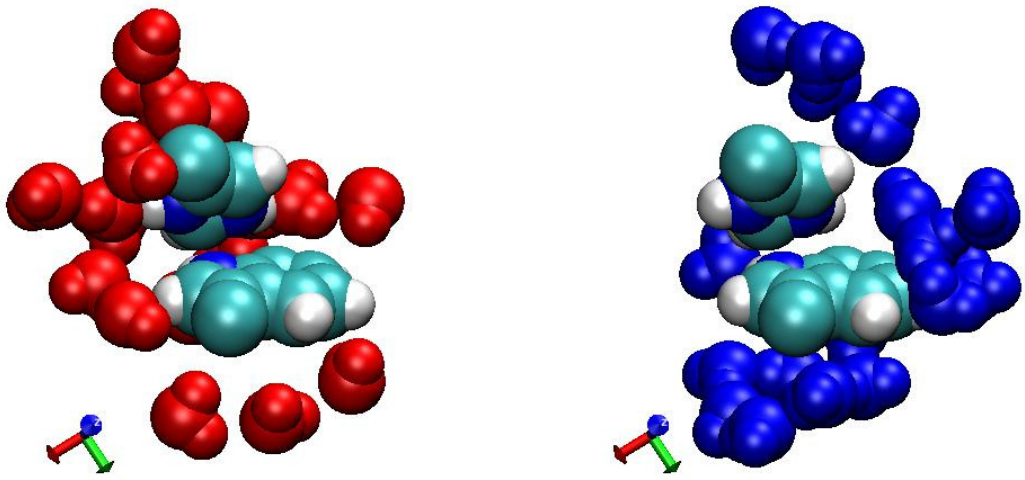

Figure 5.10 Histidine and tryptophan at a close conformation. The left panel shows waters (in red) that destabilize the charge transfer state to histidine. The right hand panel shows waters that stabilize the charge transfer state to histidine. Water molecules that are next to the lower face of tryptophan point their dipoles and stabilize the charge left on the tryptophan ring as the electron leaves the ring to the histidine ring.

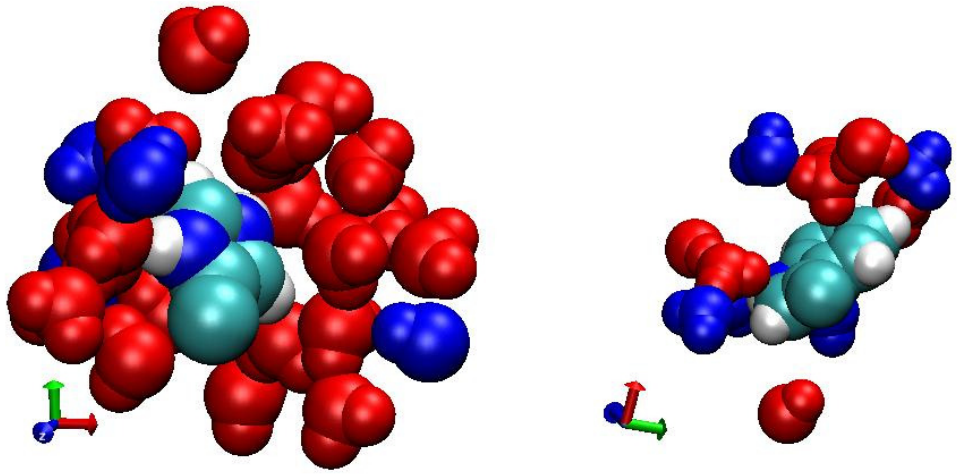

Figure 5.11 The left panel shows the histidine ring and surrounding waters (waters that are colored red are destabilizing and waters that are blue are stabilizing the charge transfer state to the His). The overall effect on the histidine ring is destabilization of the charge transfer state. The right panel shows the tryptophan ring and surrounding waters (same color scheme as the left panel). The overall net effect of the waters around tryptophan is negligible as their dipoles are oriented randomly neither opposing nor favoring electron transfer.

CHAPTER 6

SUMMARY AND CONCLUSIONS

This thesis presented a method to calculate the electron transfer matrix element (V_{et}) used for calculating electron transfer reactions for both amide and the histidine cation. A series of calculations were conducted on model systems that led to creation of either a table (in the case of amide) or a function (in the case of the histidine cation) to estimate the electron transfer matrix element. The use of a model system for calculation of amide and histidine also led to a better understanding of the behavior of V_{et} both at short and long distances. It also helped to understand the trends that are observed from calculations of V_{et} from actual MD simulation snapshots. V_{et} was calculated to be 10 to 100 times larger in magnitude than what was used by Callis and coworkers[42] in 2004. Since V_{et} was larger the correction that is added to the energy gap had to be +4500 cm^{-1}. The consequence of this is that the Frank-Condon factors are now much more sensitive to the energy gap. Therefore, small changes in energy gap difference can lead to very different electron transfer rates for the same protein, which naturally leads to the possibility of heterogeneity.

I conducted molecular dynamics simulations of 20 single tryptophan proteins to 100 nanoseconds and discussed how to calculate fluorescence quantum yields in proteins that exhibit heterogeneity in their electron transfer rates. The question of whether the 150 ps simulations sampled enough conformational space still remained open from previous work. This thought led to the use GROMACS to probe the 100 nanosecond

timescale for all the proteins. From the 100 ns simulations it can be seen that many proteins have different energetic basins that have a residence time close the lifetime of the fluorescent state of tryptophan, which points out the possibility of heterogeneity in fluorescence decay. It was also found that the energy correction that is added to the energy gap to reach agreement with experiment is dependent on force field. To account for the observed heterogeneity in the simulations I averaged thousands of survival curves to calculate an average survival curve for each protein. By numerically integrating this survival curve and multiplying it by the radiative rate constant we could calculate fluorescence quantum yields for each protein. The use of such a method is superior to the method previously used by Callis and coworkers [42,43], since that method cannot account for heterogeneity.

Future studies should simulate each of these proteins for 1 μs using the CHARMM force field as used by Callis and coworkers in 2004 and 2007. The simulation should be at least 1 μs or longer to probe different energetic basins and provide better sampling of these. The fluorescence quantum yields should be calculated using the average survival curve method to account for heterogeneity.

I calculated instantaneous rates of electron transfer to both the amide backbone and the histidine cation for HP-35 and WH5. Amide electron transfer is insensitive to whether HP-35 and WH5 are folded or not, therefore the observed changes in fluorescence quantum yield upon unfolding are due to differences in the electron transfer rate to the histidine cation. In the case of histidine, its rate is controlled by the tryptophan and histidine dihedrals. Therefore any structural difference along the folding pathway that

changes the stability of the close conformations relative to the far conformations will decrease the observed quantum yield. However upon unfolding, Eaton and coworkers reported an increase in quantum yield, therefore unfolding should decrease the stability of the close conformation, favoring instead states that have histidine far from tryptophan. Since this equilibrium is sensitive to changes in the equilibrium distribution between the far and close conformations it seems plausible that tryptophan fluorescence can distinguish between the folded state and a long lived intermediate state found by Freddolino et al. called the flipped state[51]. Because this state has different salt bridges between charged amino acids near histidine, when compared to the crystal structure[51], this state can have a different equilibrium distribution between far and close conformations. This suggests that tryptophan fluorescence might be able to discriminate between this conformation and the folded conformation.

Future studies for the villin headpiece subdomain should include a replica exchange molecular dynamics simulation that expands the entire range of temperatures probed by experiments. This simulation would give us an equilibrium ensemble distribution for the far and close conformation of villin at all the different temperatures. This could help us understand if tryptophan fluorescence can distinguish between folded and unfolded structures. In addition we should use Markov state models to accurately calculate the kinetics of folding as followed by tryptophan fluorescence quantum yield. This combination of simulation studies should point out if tryptophan fluorescence is a reliable marker of folding/unfolding both at the steady state and kinetic levels.

In the simulations of HP-35, WH5 and the twenty single tryptophan proteins different electron transfer rate constants (varying several orders of magnitude) whose residence times are comparable to the lifetime of the fluorescent state of tryptophan are observed. This led me to study how the quantum yield should be calculated when these species exchange at rates that fall in between the long residence time limit (pure heterogeneity) and the very short residence time limit (single exponential decay). I found that the calculated lifetimes depend on the exchange rate. At long residence lifetimes the decays constant are the pure decay constant for each species. As the exchange rate is increased one of the decay constants slowly becomes the average between the two pure decay constants while the second one becomes the sum of the exchange rates. This treatment is important in interpreting the observed lifetimes in a time resolved fluorescence experiment. It was also noted that the calculated quantum yield is sensitive to the lifetimes of each of these species.

For HP-35 and WH5 the fluorescence quantum yield can vary by two mechanisms: either by changing the equilibrium distribution between far and close configurations or by changing the rate of exchange between these two states. The latter seems to be the mechanism by which fluorescence quantum yield changes when HP-35 or WH5 unfolds. It was shown that a simple four state model is sufficient to calculate the fluorescence quantum yield as HP-35 unfolds. This model shows that when there is a difference in the fraction of close conformation between the folded and the unfolded state, or any other state, tryptophan fluorescence can be used to discriminate between these states.

REFERENCES CITED

(1) Eftink, M. R. *Methods of Biochemical Analysis* **1991**, *35*, 127-205.

(2) Beechem, J. M.; Brand, L. *Annual Review of Biochemistry* **1985**, *54*, 43-71.

(3) Demchenko, A. P. *UltraViolet Spectroscopy of Proteins Springer-Verlag: New York***1986**, *1*.

(4) Lakowicz, J. *Principles of Fluorescence Spectroscopy, 2nd ed.;Plenum: New York* **1999**.

(5) Ervin, J.; Sabelko, J.; Gruebele, M. *Journal of Photochemistry and Photobiology B-Biology* **2000**, *54*, 1-15.

(6) Qiu, L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. *Journal of the American Chemical Society* **2002**, *124*, 12952-12953.

(7) Thompson, P. A.; Munoz, V.; Jas, G. S.; Henry, E. R.; Eaton, W. A.; Hofrichter, J. *Journal of Physical Chemistry B* **2000**, *104*, 378-389.

(8) Kubelka, J.; Eaton, W. A.; Hofrichter, J. *Journal of Molecular Biology* **2003**, *329*, 625-630.

(9) Kubelka, J.; Chiu, T. K.; Davies, D. R.; Eaton, W. A.; Hofrichter, J. *Journal of Molecular Biology* **2006**, *359*, 546-553.

(10) Chiu, T. K.; Kubelka, J.; Herbst-Irmer, R.; Eaton, W. A.; Hofrichter, J.; Davies, D. R. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 7517-7522.

(11) May, J. M.; Qu, Z.-C.; Beechem, J. M. *Biochemistry* **1993**, *32*, 9525-9531.

(12) Chadborn, N.; Bryant, J.; Bain, A. J.; O'Shea, P. *Biophysical Journal* **1999**, *76*, 2198-2207.

(13) Chen, S. H.; Burgner, J. W.; Krahn, J. M.; Smith, J. L.; Zalkin, H. *Biochemistry* **1999**, *38*, 11659-11669.

(14) Nair, S. K.; Thomas, T. J.; Greenfield, N. J.; Chen, A.; He, H.; Thomas, T. *Journal of Molecular Endocrinology* **2005**, *35*, 211-223.

(15) Kozachkov, L.; Padan, E. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 15769-15774.

(16) Gasymov, O. K.; Abduragimov, A. R.; Glasgow, B. J. *Biochemistry*, *49*, 582-590.

(17) Zelent, B.; Odili, S.; Buettger, C.; Shiota, C.; Grimsby, J.; Taub, R.; Magnuson, M. A.; Vanderkooi, J. M.; Matschinsky, F. M. *Biochemical Journal* **2008**, *413*, 269-280.

(18) Bian, Q. Q.; Liu, J. Q.; Tian, J. N.; Hu, Z. D. *International Journal of Biological Macromolecules* **2004**, *34*, 333-337.

(19) Liu, J. Q.; Tian, J.; Li, Y.; Yao, X. J.; Hu, Z. D.; Chen, X. G. *Macromolecular Bioscience* **2004**, *4*, 520-525.

(20) Veenstra, M.; Turk, E.; Wright, E. M. *Journal of Membrane Biology* **2002**, *185*, 249-255.

(21) Palmisano, D. V.; Grothvasselli, B.; Farnsworth, P. N.; Reddy, M. C. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* **1995**, *1246*, 91-97.

(22) Meech, S. R.; Phillips, D.; Lee, A. G. *Chemical Physics* **1983**, *80*, 317-328.

(23) Yu, H. T.; Colucci, W. J.; McLaughlin, M. L.; Barkley, M. D. *Journal of the American Chemical Society* **1992**, *114*, 8449-8454.

(24) Eftink, M. R.; Jia, Y. W.; Hu, D.; Ghiron, C. A. *Journal of Physical Chemistry* **1995**, *99*, 5713-5723.

(25) Muino, P. L.; Callis, P. R. *Journal of Physical Chemistry B* **2009**, *113*, 2572-2577.

(26) Chang, M. C.; Petrich, J. W.; McDonald, D. B.; Fleming, G. R. *Journal of the American Chemical Society* **1983**, *105*, 3819-3824.

(27) Cowgill, R. W. *Archives of Biochemistry and Biophysics* **1963**, *100*, 36-&.

(28) Feitelso.J *Israel Journal of Chemistry* **1970**, *8*, 241-&.

(29) Privat, J. P.; Wahl, P.; Auchet, J. C. *Biophysical Chemistry* **1979**, *9*, 223-233.

(30) Colucci, W. J.; Tilstra, L.; Sattler, M. C.; Fronczek, F. R.; Barkley, M. D. *Journal of the American Chemical Society* **1990**, *112*, 9182-9190.

(31) Hynes, J. T. *Journal of Physical Chemistry* **1986**, *90*, 3701-3706.

(32) Mikkelsen, K. V.; Ratner, M. A. *Journal of Chemical Physics* **1989**, *90*, 4237-4247.

(33) Barbara, P. F.; Meyer, T. J.; Ratner, M. A. *Journal of Physical Chemistry* **1996**, *100*, 13148-13168.

(34) Newton, M. D. *Chemical Reviews* **1991**, *91*, 767-792.

(35) Marcus, R. A.; Sutin, N. *Biochimica Et Biophysica Acta* **1985**, *811*, 265-322.

(36) Marcus, R. A. *Reviews of Modern Physics* **1993**, *65*, 599-610.

(37) Marcus, R. A. *Journal of Chemical Physics* **1956**, *24*, 966-978.

(38) Marcus, R. A. *Journal of Chemical Physics* **1956**, *24*, 979-989.

(39) Marcus, R. A. *Annual Review of Physical Chemistry* **1964**, *15*, 155-&.

(40) Marcus, R. A. *Journal of Electroanalytical Chemistry* **2000**, *483*, 2-6.

(41) Newton, M. D. In *Electron Transfer-from Isolated Molecules to Biomolecules, Pt 1*; John Wiley & Sons Inc: New York, 1999; Vol. 106, p 303-375.

(42) Callis, P. R.; Liu, T. Q. *Journal of Physical Chemistry B* **2004**, *108*, 4248-4259.

(43) Callis, P. R.; Petrenko, A.; MuiÃ±o, P. L.; Tusell, J. R. *The Journal of Physical Chemistry B* **2007**, *111*, 10335-10339.

(44) Callis, P. R.; Vivian, J. T. *Chemical Physics Letters* **2003**, *369*, 409-414.

(45) Patrik R, C.; Ludwig Brand, M. L. J. In *Methods in Enzymology*; Academic Press: 1997; Vol. Volume 278, p 113-150.

(46) Mikkelsen, K. V.; Ratner, M. A. *Journal of Physical Chemistry* **1989**, *93*, 1759-1770.

(47) Chen, Y.; Barkley, M. D. *Biochemistry* **1998**, *37*, 9976-9982.

(48) Duan, Y.; Wang, L.; Kollman, P. A. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 9897-9902.

(49) Shirts, M.; Pande, V. S. *Science* **2000**, *290*, 1903-1904.

(50) Ensign, D. L.; Kasson, P. M.; Pande, V. S. *Journal of Molecular Biology* **2007**, *374*, 806-816.

(51) Freddolino, P. L.; Schulten, K. *Biophysical Journal* **2009**, *97*, 2338-2347.

(52) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. *Biophysical Journal*, *100*, L47-L49.

(53) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W. *Science*, *330*, 341-346.

(54) Cellmer, T.; Buscaglia, M.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 6103-6108.

(55) Pande, V. S.; Beauchamp, K.; Bowman, G. R. *Methods*, *52*, 99-105.

(56) Beauchamp, K. A.; Ensign, D. L.; Das, R.; Pande, V. S. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 12734-12739.

(57) Eaton, W. A.; Munoz, V.; Thompson, P. A.; Henry, E. R.; Hofrichter, J. *Accounts of Chemical Research* **1998**, *31*, 745-753.

(58) Mohammed, O. F.; Jas, G. S.; Lin, M. M.; Zewail, A. H. *Angewandte Chemie-International Edition* **2009**, *48*, 5628-5632.

(59) Lin, M. M.; Mohammed, O. F.; Jas, G. S.; Zewail, A. H. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 16622-7.

(60) Hegefeld, W. A.; Chen, S.-E.; DeLeon, K. Y.; Kuczera, K.; Jas, G. S. *Journal of Physical Chemistry A*, *114*, 12391-12402.

(61) Humphrey, W.; Dalke, A.; Schulten, K. *Journal of Molecular Graphics* **1996**, *14*, 33-&.

(62) Smith, D. M. A.; Rosso, K. M.; Dupuis, M.; Valiev, M.; Straatsma, T. P. *Journal of Physical Chemistry B* **2006**, *110*, 15582-15588.

(63) Mallick, G.; Karna, S. P.; He, H. Y.; Pandey, R. *International Journal of Quantum Chemistry* **2009**, *109*, 1302-1310.

(64) Closs, G. L.; Miller, J. R. *Science* **1988**, *240*, 440-447.

(65) Henderson, T. M.; Cave, R. J. *Journal of Chemical Physics* **1998**, *109*, 7414-7423.

(66) Li, X. Y.; Tang, X. S.; Xiao, S. Q.; He, F. C. *Journal of Molecular Structure-Theochem* **2000**, *530*, 49-58.

(67) Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Current Opinion in Structural Biology* **2004**, *14*, 76-88.

(68) M. J. Frisch, G. W. T., H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, . Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc. **2004**.

(69) Gao, Y. Q.; Marcus, R. A. *Journal of Physical Chemistry A* **2002**, *106*, 1956-1960.

(70) Zhang, L. Y.; Friesner, R. A.; Murphy, R. B. *Journal of Chemical Physics* **1997**, *107*, 450-459.

(71) Toutounji, M. M.; Ratner, M. A. *Journal of Physical Chemistry A* **2000**, *104*, 8566-8569.

(72) Daizadeh, I.; Medvedev, E. S.; Stuchebrukhov, A. A. *Proceedings of the National Academy of Sciences of the United States of America* **1997**, *94*, 3703-3708.

(73) Prytkova, T. R.; Kurnikov, I. V.; Beratan, D. N. *Science* **2007**, *315*, 622-625.

(74) Onuchic, J. N.; Kobayashi, C.; Miyashita, O.; Jennings, P.; Baldridge, K. K. *Philosophical Transactions of the Royal Society B-Biological Sciences* **2006**, *361*, 1439-1443.

(75) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *Journal of Chemical Theory and Computation* **2008**, *4*, 435-447.

(76) http://www.chemcraftprog.com.

(77) Petrich, J. W.; Chang, M. C.; McDonald, D. B.; Fleming, G. R. *Journal of the American Chemical Society* **1983**, *105*, 3824-3832.

(78) Moors, S. L. C.; Jonckheer, A.; De Maeyer, M.; Engelborghs, Y.; Ceulemans, A. *Current Protein & Peptide Science* **2008**, *9*, 427-446.

(79) Maglia, G.; Jonckheer, A.; De Maeyer, M.; Frere, J.-M.; Engelborghs, Y. *Protein Science* **2008**, *17*, 352-361.

(80) Rolinski, O. J.; Martin, A.; Birch, D. J. S. *Journal of Biomedical Optics* **2007**, *12*.

(81) Moors, S. L. C.; Hellings, M.; De Maeyer, M.; Engelborghs, Y.; Ceulemans, A. *Biophysical Journal* **2006**, *91*, 816-823.

(82) Hellings, M.; De Maeyer, M.; Verheyden, S.; Hao, Q.; Van Damme, E. J. M.; Peumans, W. J.; Engelborghs, Y. *Biophysical Journal* **2003**, *85*, 1894-1902.

(83) Lindahl, E.; Hess, B.; van der Spoel, D. *Journal of Molecular Modeling* **2001**, *7*, 306-317.

(84) Van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *Journal of Computational Chemistry* **2005**, *26*, 1701-1718.

(85) Berendsen, H. J. C.; Vanderspoel, D.; Vandrunen, R. *Computer Physics Communications* **1995**, *91*, 43-56.

(86) Hess, B. *Journal of Chemical Theory and Computation* **2008**, *4*, 116-122.

(87) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *Journal of Chemical Physics* **1995**, *103*, 8577-8593.

(88) Bussi, G.; Donadio, D.; Parrinello, M. *Journal of Chemical Physics* **2007**, *126*.

(89) Nose, S. *Molecular Physics* **1984**, *52*, 255-268.

(90) Parrinello, M.; Rahman, A. *Journal of Applied Physics* **1981**, *52*, 7182-7190.

(91) Price, M. L. P.; Ostrovsky, D.; Jorgensen, W. L. *Journal of Computational Chemistry* **2001**, *22*, 1340-1352.

(92) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. L. *Journal of Physical Chemistry B* **2001**, *105*, 6474-6487.

(93) Watkins, E. K.; Jorgensen, W. L. *Journal of Physical Chemistry A* **2001**, *105*, 4118-4125.

(94) McDonald, N. A.; Jorgensen, W. L. *Journal of Physical Chemistry B* **1998**, *102*, 8049-8059.

(95) Jorgensen, W. L.; McDonald, N. A. *Theochem-Journal of Molecular Structure* **1998**, *424*, 145-155.

(96) Ridley, J.; Zerner, M. *Theoretica Chimica Acta* **1973**, *32*, 111-134.

(97) Theiste, D.; Callis, P. R.; Woody, R. W. *Journal of the American Chemical Society* **1991**, *113*, 3260-3267.

(98) Sreerama, N.; Woody, R. W.; Callis, P. R. *Journal of Physical Chemistry* **1994**, *98*, 10397-10407.

(99) Li, J. B.; Williams, B.; Cramer, C. J.; Truhlar, D. G. *Journal of Chemical Physics* **1999**, *110*, 724-733.

(100) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins-Structure Function and Bioinformatics* **2006**, *65*, 712-725.

(101) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; shaw, D. E. *Proteins-Structure Function and Bioinformatics*, *78*, 1950-1958.

(102) De Beuckeleer, K.; Volckaert, G.; Engelborghs, Y. *Proteins-Structure Function and Genetics* **1999**, *36*, 42-53.

(103) Gastmans, M.; Volckaert, G.; Engelborghs, Y. *Proteins-Structure Function and Genetics* **1999**, *35*, 464-474.

(104) Egan, D. A.; Logan, T. M.; Liang, H.; Matayoshi, E.; Fesik, S. W.; Holzman, T. F. *Biochemistry* **1993**, *32*, 1920-1927.

(105) Hennecke, J.; Sillen, A.; HuberWunderlich, M.; Engelborghs, Y.; Glockshuber, R. *Biochemistry* **1997**, *36*, 6391-6400.

(106) Harris, D. L.; Hudson, B. S. *Chemical Physics* **1991**, *158*, 353-382.

(107) Sillen, A.; Diaz, J. F.; Engelborghs, Y. *Protein Science* **2000**, *9*, 158-169.

(108) Qiu, L. L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. *Journal of the American Chemical Society* **2002**, *124*, 12952-12953.