Stony Brook University

Novel Spectroscopic Probes to Study Protein Folding

A Dissertation Presented

by

Humeyra Taskent

to

The Graduate School

in Partial fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Chemistry

Stony Brook University

August 2010

Stony Brook University

The Graduate School

Humeyra Taskent

We, the dissertation committee for the above candidate for the Doctor of Philosophy degree, hereby recommend acceptance of this dissertation.

Daniel P. Raleigh, Ph. D., Dissertation Advisor
Professor, Department of Chemistry

Isaac Carrico, Ph. D., Chairperson of Defense
Professor, Department of Chemistry

Erwin London, Ph. D., Third Member
Professor, Department of Chemistry

Suzanne Scarlata, Ph. D., Outside Member
Stony Brook University, Professor, Department of Physiology and Biophysics

This dissertation is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the Dissertation

Novel Spectroscopic Probes to Study Protein Folding

by

Humeyra Taskent

Doctor of Philosophy

in

Chemistry

Stony Brook University

2010

Protein folding is one of the most important unsolved questions of structural biology because of the desire to understand the link between the primary sequence and structure. Proteins can fold on a microsecond to millisecond time scale. Fluorescence and IR spectroscopy are very important tools to follow these fast kinetic events, however, natural fluorescent groups in proteins, Tyr and Trp, are not perfect substitutions for most amino acids and new fluorescent probes are needed. In IR spectroscopy the signal from the protein backbone is used to follow protein dynamics. In an IR spectrum, broad and overlapping peaks are generally observed and site-specific IR probes would represent a significant advance. With the advances in molecular biology, it is now possible to introduce new spectroscopic probes into proteins recombinantly.

In this dissertation the N-terminal domain of ribosomal protein L9 (NTL9) was used as a model system to investigate the unnatural amino acids, p-cyanophenylalanine (F_{CN}) as a fluorescent and azidohomoalanine (Aha) as an IR-active probe for folding studies. Recently, F_{CN} was shown as a fluorescent and IR probe to study protein dynamics. The fluorescence quantum yield of the probe increases dramatically when it is hydrogen bonded. NTL9-F5F_{CN} is generated by both peptide synthesis and recombinantly by an orthogonal tRNA/tRNA synthetase pair. The folding kinetics of NTL9 is studied with stopped-flow fluorescence. In addition to this, the effect of amino acid side chains on F_{CN} fluorescence is investigated with model peptides in order to use this popular probe accurately.

The azido stretching vibration is in a transparent region of the protein IR spectrum and is sensitive to solvation. Aha also has a high extinction coefficient. The azido analog of methionine, Aha is incorporated into proteins by solid-phase peptide synthesis and recombinantly in high yield using methionine auxotrophic strains. Aha was incorporated into two sites in NTL9. The mutations did not perturb the overall fold of the protein. The frequency of the azido mode is observed to undergo a significant blue shift in the thermally unfolded state, indicating that the group provides a sensitive probe of protein folding and sidechain burial.

Table of Contents

List of Figures

List of Tables

List of Schemes

List of Abbreviations and Symbols

Aha	azidohomoalanine
ATR-FTIR	Attenuated total reflection fourier transform infrared
CD	circular dichroism
Cα	Alpha carbon
D	the denatured state
ΔG°	standard Gibbs free energy of unfolding
ΔΔG°	unfolding free energy change upon mutation
ΔH°	standard enthalpy of unfolding
ΔS°	standard entropy of unfolding
DIPEA	N,N-diisopropylethylamine
DMF	N,N-dimethylformamide
ESI	electro-spray ionization
F	the folded state
F_{CN}	*p*-cyanophenylalanine
FRET	Fluorescence resonance energy transfer
FTIR	Fourier transform infrared
FMOC	9-fluoromethoxycarbonyl
GuHCl	guanidine hydrochloride
HBTU	O-Benzotriazol-1-yl-N,N,N',N'-tetramethyluronium hexafluorophosphate
HOBT	1-Hydroxybenzotriazole
HPLC	high performance liquid chromatography
k_f	rate constant for folding
k_u	rate constant for unfolding
MALDI-TOF	matrix-assisted laser desorption/ionization time-of-flight
m	slope of ΔG° vs. denaturant concentration
m_f	slope of $\ln(k_f)$ vs. denaturant concentration
MS	mass spectrometry
m_u	slope of $\ln(k_u)$ vs. denaturant concentration

N	the native state
NMP	N-methylpyrrolidone
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
NTL9	N-terminal domain of ribosomal protein L9 from *Bacillus Stearothermophilus*
PAL	5-(4'-Fmoc-aminomethyl-3', 5-dimethoxyphenoxyl) valeric acid
ϕ-value	$\Delta\Delta G^{\circ\ddagger}/ \Delta\Delta G^{\circ}$
$[\theta]$	mean residue ellipticity
θ_m	Tanford value, m_f/m
RMSD	Root mean square deviation
R_0	Förster distance
TFA	trifluoroacetic acid
TOCSY	total correlation spectroscopy
TSP	sodium 3-trimethylsilyl (2,2,3,3,-d4) propionate
U	the unfolded state
1D	one-dimensional
2D	two-dimensional

Acknowledgements

I thank my advisor, Prof. Daniel Raleigh, for his support, encouragement and enthusiasm throughout my Ph.D. His knowledge not only in protein folding but also in many other areas of science has always made it enjoyable to talk to him.

Prof. Isaac Carrico was my second advisor. He introduced me to the world of unnatural amino acids. It was great to discuss science with him, his enthusiasm and encouragement was really invaluable to get to this point.

I thank the Raleigh group members, past and present: Dr. Burcu Anıl, Dr. Jae-Hyun Cho, Dr. Sylvia Tracz, Dr. Yuan Bi, Dr. Benben Song, Dr. Ruchi Gupta, Dr. Ying Li, Dr. Yuefeng Tang, and Dr. Bing Shan, Wenli Meng, Trisha Barua, Ping Cao, Fanling Meng, Shifeng Xiao, Peter Marek, Bowu Luan, Ivan Peran, Hui Wang and Cynthia Li. It has been a great experience to work with and learn from these wonderful people many of whom have become friends to me. Special thanks go to Burcu, Cho and Sylvia for answering all my questions and training me patiently in my first year and Wenli for being a very good friend and his help with NMR experiments.

Juah Chung from Carrico group was one of the undergraduate students I worked with. She stayed with me for the longest time. It was really a lot of fun to work with Juah. She was always cheerful which motivated me a lot. She deserves a special thanks. Also, in azidohomoalanine project I collaborated with Partha Banerjee from Carrico group. Thank you Partha!

Two and a half years ago Vadim Patsalo from David Green's lab joined the lab for the experimental part of his work. We became really good friends. We discussed science, politics, and life in general, it was great to have someone to talk to in the lab. Thanks Vadim!!

I collaborated lots of brilliant scientists from various laboratories for my projects. Everybody was really helpful trying to solve a problem, understanding a concept or a technique. Thank you all: Dr. Scott Brewer from Franklin and Marshall College, Dr. Brian Dyer and Dr. Sureshbabu Nagarajan from Emory University, Dr. Lisa Miller and Ms. Megan Bourassa from BNL with FTIR spectroscopy; Dr. Ryan Mehl from Franklin and Marshall College with 21st pair methodology; Dr. Feng Gai and Dr. Smita Mukherjee from University of Pennsylvania for discussion on F_{CN} fluorescence and temperature jump IR studies; Dr. Miquel Vila-Perello from Rockefeller University with peptide synthesis; and recently Dr. Osman Bilsel and Dr. C. Robert Matthews for time-resolved fluorescence experiments.

I would like to thank two pairs of parents, Masume-Arif Taşkent and Gülsevin-Mehmet Sezgin, and my siblings, Gül Nihal, Esma, Hasan Can and Akif Can for their love and support. I am grateful for their encouragement throughout my PhD.

Lastly, I would like to thank my husband, Dr. Efe Sezgin, for his love, support and everything that he has done for me. Even though he was away, he was always with me and believed in me even when I was not sure. I would not have done it without his support.

Publications

1. Aprilakis, K.N.*, **Taskent, H.*,** Raleigh, D.P. (2007) "Use of the novel fluorescent amino acid p-cyanophenylalanine offers a direct probe of hydrophobic core formation during the folding of the N-terminal domain of the ribosomal protein L9 and provides evidence for two-state folding" *Biochemistry* 46, 12308-12312. (* Joint first author)

2. **Taskent, H.,** Cho, J.H., Raleigh, D.P. (2008) "Temperature-dependent Hammond behavior in a protein-folding reaction: analysis of transition-state movement and ground-state effects" *J. Mol. Biol.* 378, 699-706.

3. **Taskent-Sezgin, H.,** Chung, J., Patsalo, V., Miyake-Stoner, S. J., Miller, A. M., Brewer, S.H., Mehl, R.A., Green, D.F., Raleigh, D.P., Carrico I. (2009) "Interpretation of p-cyanophenylalanine fluorescence in proteins in terms of solvent exposure and contribution of side-chain quenchers: a combined fluorescence, IR and molecular dynamics study." *Biochemistry* 48, 9040-9046.

4. Abedini, A., Gupta, R., Marek, P., Meng, F., Raleigh, D.P., Tracz, S. and **Taskent, H.** (2009) "Post Translational Modifications and Amyloid Formation" In: Protein Misfolding Diseases: Current and Emerging Principles and Therapies. Edited by Christopher M. Dobson, Jeffery W. Kelly and Marina Ramirez-Alvarado. Published by John Wiley and Sons.

5. **Taskent-Sezgin, H.**, Marek, P., Thomas, R, Goldberg, D, Chung, J, Carrico, I, Raleigh, D.P. (2010) "Modulation of p-Cyanophenylalanine Fluorescence by Amino Acid Side Chains and Rational Design of Fluorescence Probes of alpha-Helix Formation." *Biochemistry* 49, 6290-6295.

6. **Taskent-Sezgin, H.,** Chung, J., Banerjee, P.S., Nagarajan, S., Dyer, R.B., Carrico, I., and Raleigh, D.P. (2010) "Azidohomoalanine Provides a Conformationally Sensitive IR Probe of Protein Folding, Protein Structure and Electrostatics Which Can Be Readily Incorporated into Proteins" **Accepted.**

1. Background

1.1. Protein Folding

Proteins need to reach their native conformation to be biologically active in cells. How polypeptide chains achieve this is one of the largest unanswered questions in structural biology. Understanding the mechanism of how proteins start from a linear stretch of amino acids and reach their functional three-dimensional structure will complement the knowledge of the flow of genetic information from DNA, to RNA, to protein, and finally to functional conformation. One of the goals of solving the protein folding problem is to design new proteins with specific functions. In addition, knowledge of the protein folding process can lead to treatments for diseases such as type II diabetes, Alzheimer's and Parkinson's diseases which are related to protein misfolding.

Anfinsen's experiments showed that protein sequence carries the necessary information to reach the functional, native structure (*1*). Proteins fold within submicrosecond to seconds time scale. Thus, protein folding cannot be a random search of all possible conformations (*2*). There should be folding pathways that proteins follow to reach their three-dimensional structure. In the funnel representation of the energy

landscape of protein folding, the denatured state of proteins is an ensemble of structures at the top of the funnel. The protein chain can fold by different pathways, possibly through low energy intermediate state(s) to the native structure with the lowest energy (Figure1.1) (*3*).

There are several models proposed for protein folding mechanisms. The hydrophobic collapse model states that when a protein chain is transferred from a good solvent to a poor one, it collapses just like a polymer chain and forms a compact structure. In the hydrophobic collapse model it is hypothesized that collapse and secondary structure formation occur concomitantly. In this model, non-local hydrophobic interactions are thought to be the driving force of compaction which reduces the conformational freedom of the chain (*4*). On the other hand, the framework model suggests that secondary structure formation is followed by collapse (*5-8*). The framework model states that the unfolded state of a protein has both a compact and an expanded form. The nucleation-condensation model proposes that the collapsed state has native hydrophobic interactions. This collapsed state is the nucleus and initiates the rest of the protein to fold (*9*). Theoretical studies have been done on these questions for over couple decades. With the advances in the time resolution of experimental kinetic methods, experimentalists have been investigating the early events in protein folding reactions and these experiments are testing the theoretical predictions for validation (*10*).

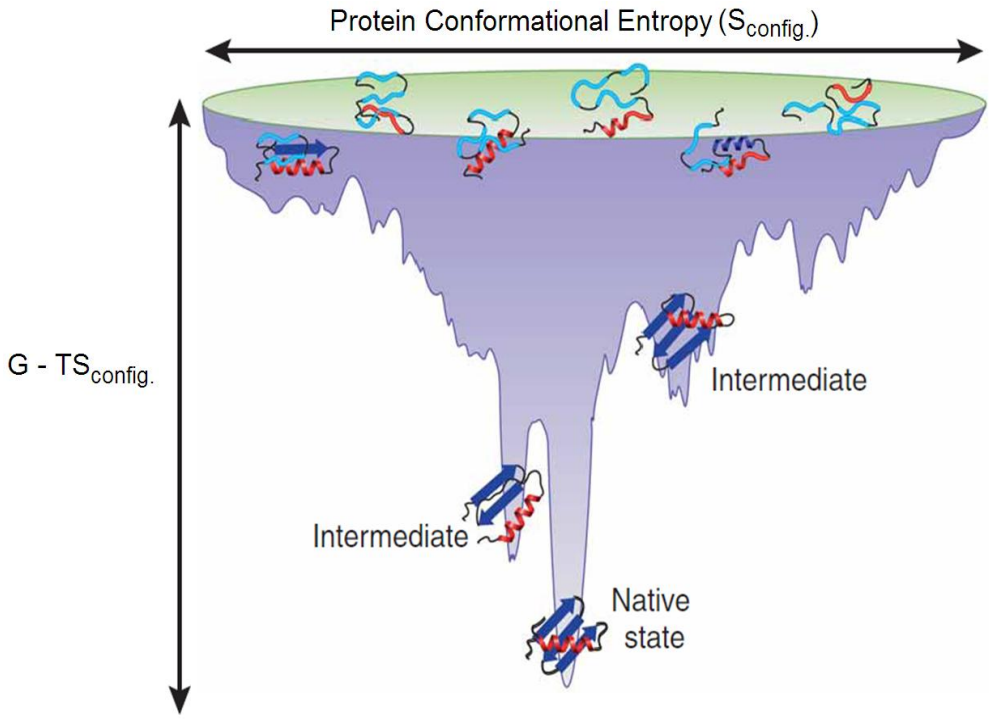


Figure 1.1. The protein folding funnel. The unfolded state ensemble is high free energy state at the top of the funnel. The y-axis represents the free energy of the system with the contribution due to the protein configurational entropy subtracted. The x-axis is proportional to chain entropy. Proteins fold through possible intermediate state(s) and reach the low energy native state. Figure adopted from reference (*3*).

1.2. Thermodynamics and Kinetics of Protein Folding

The stability of the folded protein structures can be perturbed thermally by heat or chemically by chaotropic agents such as urea or guanidinium hydrochloride or by changing the pH of the solution. Most small proteins are capable of folding back to the native structure when the perturbation is removed. Folding might follow a thermodynamic two-state model where only the folded and unfolded states are populated

at all times, or might follow a multi-state thermodynamic model where there are intermediate states along with the folded and unfolded states.

Folding conditions can be thought of as a poor solvent condition for a polypeptide chain. Proteins bury their hydrophobic side chains in folding buffers in order to avoid paying the high penalty of solvent entropy. The unfolding buffers are good solvent conditions for polypeptide chains because chaotropic agents solvate the constituent parts of proteins. The free energy of unfolding is normally proportional to the denaturant concentration according to:

$$\Delta G° = \Delta G°(H_2O) - m[denaturant] \quad (1.1)$$

where ΔG° is the free energy of unfolding, ΔG° (H_2O) is the free energy of unfolding in the absence of denaturant and m is a constant which is related to the solvent accessible surface area upon unfolding (*11*).

The classical kinetic method to study protein folding mechanism is stopped-flow. The dependence of folding and unfolding rates can be investigated as a function of denaturant with a stopped-flow device coupled with a spectroscopic method. The plot of the natural logarithm of folding and unfolding rates as a function of denaturant, the so called Chevron plot, has a V-shape for two-state folders (Figure 1.2). This shape arises from the linear dependence of the folding and unfolding rates to changes in denaturant concentration as shown in the equations:

$$\ln k_u = \ln k_u^{H_2O} + m_u[denaturant] \quad (1.2)$$

$$\ln k_f = \ln k_f^{H_2O} - m_f[denaturant] \quad (1.3)$$

where $k_u^{H_2O}$ and $k_f^{H_2O}$ are the unfolding and folding rates extrapolated to the absence of denaturant, respectively, and m_u-value is proportional to the change in accessible surface area between the folded state and transition state while m_f-value reports on the difference between the unfolded state and the transition state (Figure 1.3). In the presence of an intermediate state during folding, the Chevron plot often deviates from linearity at low denaturant concentration region (Figure 1.2). The observed rate during the experiment is the sum of folding and unfolding rates:

$$\ln k_{obs} = \ln(k_f + k_u) \tag{1.4}$$

When equations 1.2 and 1.3 are substituted in equation 1.4:

$$\ln(k_{obs}) = \ln(k_f^{H_2O} \exp(m_f[denaturant]/RT) + k_u^{H_2O} \exp(m_u[denaturant]/RT)) \tag{1.5}$$

Plot of $\ln(k_{obs})$ as a function of denaturant concentration gives the Chevron plot. The equilibrium m-value can be calculated by taking the difference between m_f and m_u. It is also possible to determine the extent of buried surface area in the transition state from the $m_f/(m_u-m_f)$ which is also known as the Tanford-parameter or β_T (*12*).

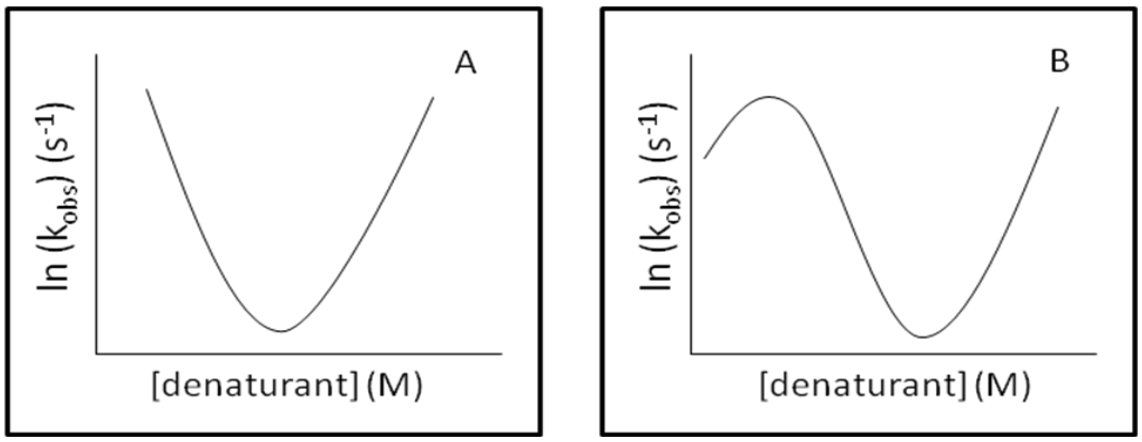


Figure 1.2. Chevron plots for a two-state and a multi-state folding reaction. (A) Classic V-shape Chevron plot indicating a two-state folding mechanism. (B) Deviation from linearity at low concentration of denaturant, the so called roll-over in Chevron plot indicates the presence of an intermediate in the folding reaction.

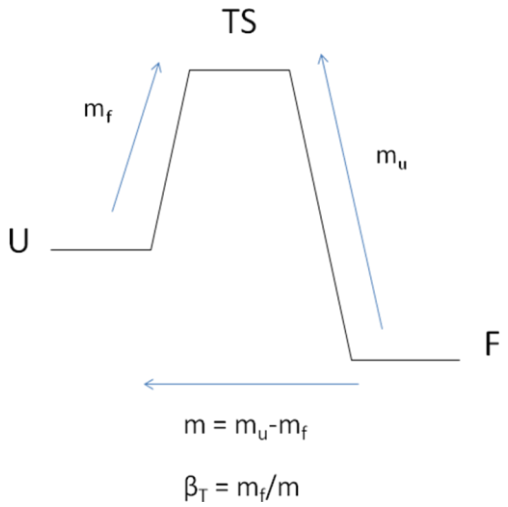


Figure 1.3. Free energy diagram for the folding of a protein. U, F, and TS are unfolded, folded and transition state, respectively. M_f is related to the change in solvent accessible surface area between the unfolded and the transition state; m_u is related to the change between folded and the transition state. M value can be calculated from m_f and m_u. β_T parameter reports on the extent of buried surface area in the transition state.

1.3. Emerging Experimental Kinetic Methods to Study Protein Folding

The stopped-flow technique has a time resolution of milliseconds and longer. However, some of the events during protein folding such as secondary structure formation and chain collapse occur much faster. In addition, experimental validation of theoretical predictions is often limited by the poor time resolution. Thus, there was a great urgency for new techniques with better time resolution. In the last 10 to 15 years, ultrarapid-mixing techniques have advanced the time resolution to submilliseconds (*13, 14*) and laser pulse perturbation techniques have further reduced the time resolution into submicroseconds (*15, 16*). Two of the popular spectroscopic methods, fluorescence and infrared (IR) coupled with these rapid mixing and laser pulse perturbation are helping close the gap between the time resolution of experimental data and theoretical calculations (*10*).

Fluorescence Resonance Energy Transfer (FRET) is one of the methods that has been coupled with continuous-flow or laser temperature jump techniques to investigate protein folding reactions (*16, 17*). The FRET effect is sensitive to distance and can be used to follow chain compaction. FRET requires the presence of a fluorescent donor and a fluorescent acceptor molecule. The donor and acceptor molecules should be chosen in such a way that, the distance to be investigated should fall within $0.5R_0$ to $2R_0$ limits. (R_0, Förster distance, is the distance between the donor and acceptor when there is 50% energy transfer.) Since FRET measurements provide distance between two probes; valuable information can be obtained regarding protein dynamics. At this point in time, the application of FRET is often limited by the available probes. For proteins Trp and Tyr

are the natural fluorescent amino acids, however, they may not be the best replacement for all amino acids because of their sizes and they do not form a FRET pair. There are various dye molecules such as IEADANS, Cy3 and Cy5 that are attached to proteins through Cys side chains. These dyes are either paired with a Trp residue (IEADANS), or another dye molecule to form a FRET pair. One problem can be that the protein fold may be perturbed by the attachment of these hydrophobic dye molecules. If the dye molecule is attached through a Cys residue, all the native Cys residues should be eliminated via mutation to prevent multi-site labeling and one Cys site for the dye attachment should be introduced into the protein. Dye attachment reactions do not generally go to completion, so in the end there will be a mixture of labeled and unlabeled protein solution. To sum up, there is an increasing need for better FRET probes.

IR spectroscopy is another widely used method to study protein dynamics and protein-ligand interactions. Laser induced temperature-jumps coupled with IR detection has a time resolution in nanoseconds, and is the method of choice to investigate protein folding kinetics (*18*). The temperature-jump is induced by the absorption of the pulse energy by the solvent which perturb the system thermally. The backbone amide-I band is the most commonly used probe in IR studies of proteins. However, the amide-I band is often broad due to observed overlapping bands and resolving these bands is not always trivial. Some of the other common IR probes available suffer from poor sensitivity or are difficult to introduce (*19, 20*).

Unnatural amino acids with various desirable properties are becoming more popular probes due in part to the advances in methods for their incorporation. Section 1.4

outlines methods to introduce unnatural amino acids and section 1.5 introduces the unnatural amino acids, *p*-cyanophenylalanine and azidohomoalanine which are utilized for this dissertation research.

1.4. Methods to Incorporate Unnatural Amino Acids into Proteins

Unnatural amino acids have been used as probes of protein structure, folding and dynamics (*21-23*). Different probes have various properties such as being IR-active, fluorescent, being fluorinated for use in ^{19}F-NMR. There have been improvements in techniques to incorporate these probes. Thus, their application is growing. Previously, the only way to put many of the unnatural probes into proteins was by solid-phase peptide synthesis. With peptide synthesis, there is a limit to the size of the protein one can generate because of the decreased yield and purity as the polypeptide chain gets longer (*24*).

Emergence of native chemical ligation (NCL) and expressed protein ligation (EPL) techniques addressed the size limitation of peptide synthesis. With NCL and/or EPL, one can synthesize part of a protein with the unnatural probe and the rest of the protein can be expressed (*25, 26*). Then, these pieces can be ligated using a thioester at the C-terminus of the N-terminal fragment and a Cys residue at the N-terminus of the C-terminal fragment. The ligation site should be chosen carefully to produce stable fragments with high yield. The ligation reaction has a high yield, when the concentration of each fragment is high.

The so called 21st pair technology developed in Schultz's laboratory, is another promising method for the incorporation of unnatural amino acids into proteins. An orthogonal tRNA / aminoacyl-tRNA synthetase pair that will recognize an amber suppresser codon is used for the attachment of the unnatural amino acid to the growing polypeptide chain (*27*). The power of this method is the ability to place the probe at the specific site by using the translational machinery of the expression system. In addition, since the tRNA/ tRNA synthetase pair is orthogonal to the expression system and codes for the amber codon, the pair is selective to the unnatural amino acid of interest. The use of the 21st pair method for the incorporation of *p*-cyanophenylalanine into the N-terminal domain of the ribosomal protein L9 (NTL9) to study the folding reaction is described in chapters 2 and 3 of this thesis.

Another method for introducing probes into proteins is to use the native translational machinery of the expression system. Selenomethionine incorporation into proteins with methionine auxotrophic cell lines for X-ray crystallography has been known for some time (*28*). This method was expanded in Tirrell's laboratory for incorporating methionine analogs into proteins with various functionalities as handles for biomolecular labeling via expression of protein in methionine auxotropic cells in methionine depleted media (*29, 30*). Azidohomoalanine is one of the methionine analogs that the native translational machinery of bacteria can incorporate. The method gives high yield, however, every methionine position in the protein will be replaced with azidohomoalanine. The incorporation of azidohomoalanine into NTL9 to study the folding mechanism of the protein by azido IR spectroscopy is described in chapter 5.

1.5. p-Cyanophenylalanine and Azidohomoalanine

The unnatural amino acid, *p*-cyanophenylalanine (F_{CN}) has been recently shown to have unique fluorescent properties (Figure 1.4). It has high quantum yield and its quantum yield changes depending on hydrogen bonding to solvent through the cyano group (*31*). F_{CN} can be selectively excited at 240 nm, and the emission maximum is around 290 nm. Thus, F_{CN} and Trp form a FRET pair with a Förster distance (R_0) of 16 Å (*32*) which is ideal to study protein folding reactions. F_{CN} has already been used as a fluorescent reporter group in protein-protein, protein-membrane interactions, protein dynamics and amyloid formation (*22, 31-35*). F_{CN} is incorporated into proteins with the 21st pair technique developed by the Schultz group (*36*).

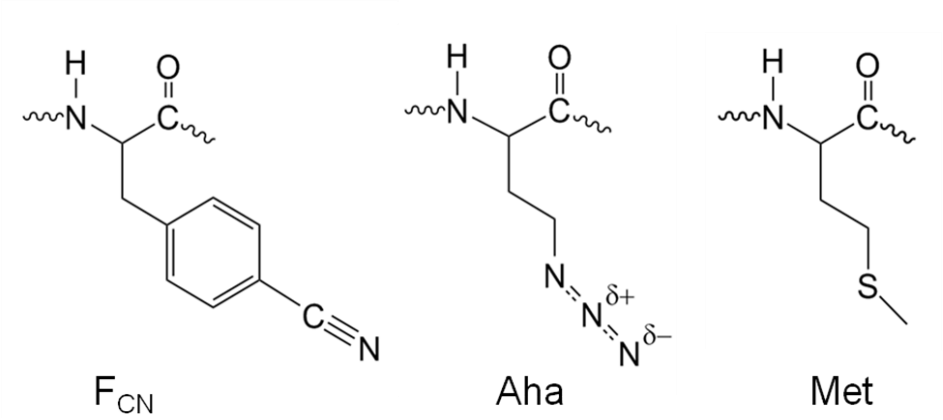


Figure 1.4. Molecular structures of p-cyanophenylalanine (F_{CN}), azidohomoalanine (Aha) and methionine (Met).

The other unnatural amino acid I worked with is azidohomoalanine. The azido IR vibration band of azidohomoalanine is in a transparent region of the protein IR spectrum. Its frequency is also sensitive to the environment. The extinction coefficient of an azido group is approximately twenty fold larger than that of a commonly employed cyano group (*37*). As metioned previously, azidohomoalanine is incorporated into the protein in methionine auxotropic cell lines in methionine depleted media through utilization of the native translational machinery of *E. coli*.

1.6. N-terminal Domain of the Ribosomal Protein L9 (NTL9)

The N-terminal domain of the ribosomal protein L9, NTL9, is the model protein for folding studies in this dissertation. The ribosomal protein L9 is located in the large 50S subunit of ribosome (Figure 1.5) (*38*). The structure of L9 from *Bacillus stearothermophilus* had been determined using X-ray crystallography (*39*). L9 has two domains; the N-terminal domain is connected to the C-terminal domain through a long α-helix (Figure 1.5). Our laboratory has determined the high resolution structure of the isolated N-terminal domain. The biological role of L9 has been suggested to be the restriction of peptidyl-tRNA slippage opposing the stemloop signaling activity in P-site frameshifting in protein translation (*40*).

The N-terminal domain of protein L9 (NTL9) is a 56-residue long protein with three anti-parallel β-strands sandwiched between two α-helices (Figure 1.5). NTL9 does not have any disulfides, does not require any cofactors to fold and has a simple topology. These properties make NTL9 an attractive model system for protein folding studies.

NTL9 folds in a two-state fashion under a wide range of conditions (*41*, *42*). The transition state of the folding reaction of NTL9 has been studied extensively using ϕ-value analysis.

Previously conducted research showed that the unfolded state of NTL9 contains significant non-native electrostatic interactions and possibly contains significant secondary structure. The mutation of the surface exposed Lys 12 to Met stabilized the protein significantly and the folding rate was increased (*43*). Mutational analysis demonstrates that Lys 12 is involved in specific non-native electrostatic interactions in the unfolded state. Removal of this interaction by K12M mutation is the reason for the stabilization of the protein (*43, 44*). In addition, a Phe 5 to Ala mutant has been used as a model of the unfolded state of NTL9. Analysis of the mutant was interpreted to indicate that the unfolded ensemble was compact (*45*). These data give information about how compact the unfolded state, but not necessarily about the dynamics of the folding reaction.

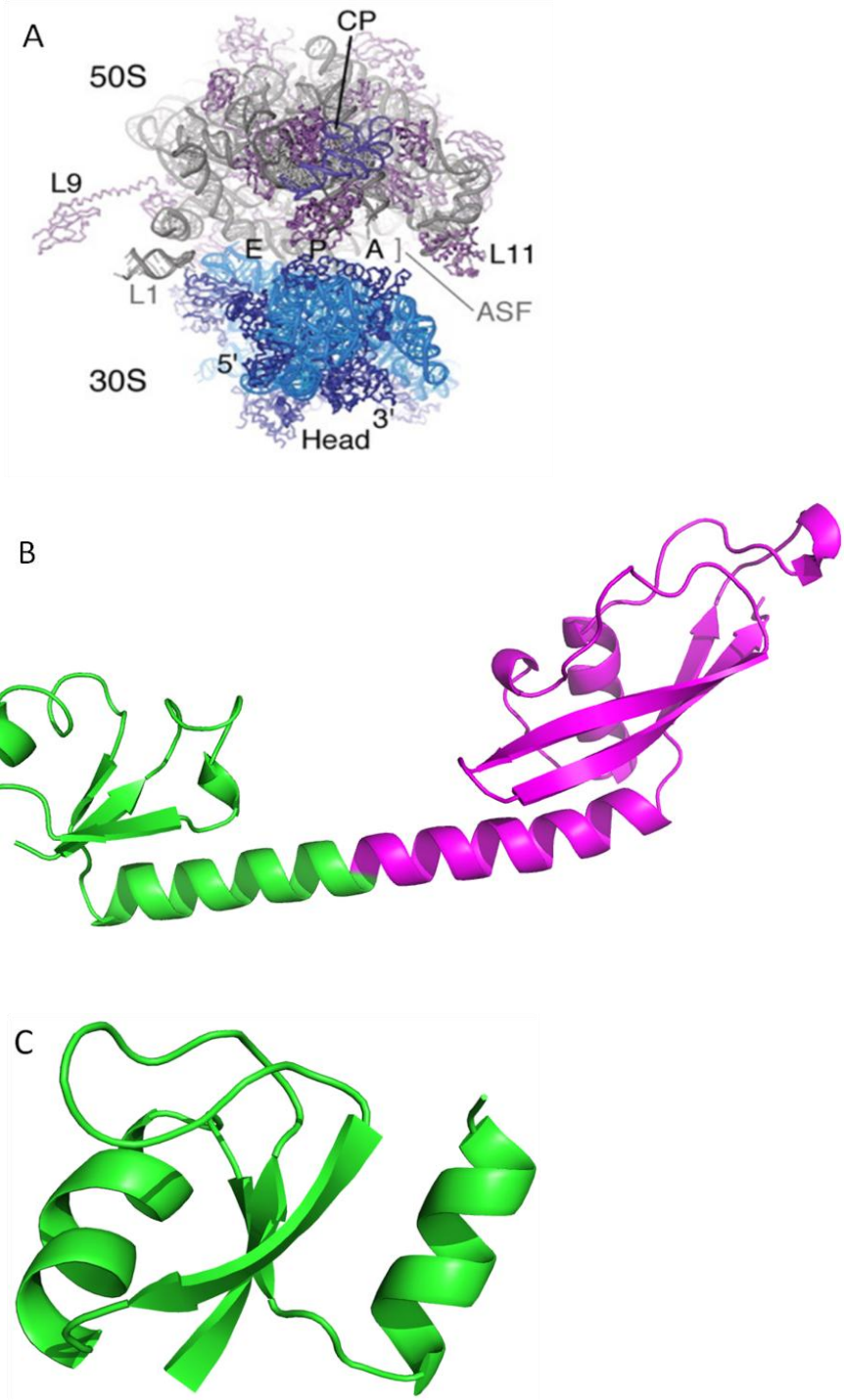


Figure 1.5. (A) Structure of the ribosome with large 50S and small 30S subunits. Figure is adopted from reference (*38*). (B) Ribosomal protein L9, N-terminal domain is green and the C-terminal domain is pink. The diagram is constructed using the pdb code

1DIV and the program Pymol. (C) N-terminal domain of L9.The diagram is generated using the pdb file 2HBA and the program Pymol.

1.7. The aims of this dissertation

This dissertation describes research that is directed at the incorporation of the non-coded amino acids, *p*-cyanophenylalanine (F_{CN}) and azidohomoalanine (Aha) into proteins in the desire to help understand the early folding events in NTL9. The fluorescent unnatural amino acid F_{CN} was incorporated into a core position, F5 in NTL9 by both solid-phase peptide synthesis and recombinantly by an orthogonal tRNA/tRNA synthetase pair. This work is described in Chapter 2 and 3. Interpretation of F_{CN} fluorescence in the context of NTL9 is described in Chapter 2. The folding kinetics of NTL9 monitored by F_{CN} fluorescence are examined in Chapter 3. The modulation of F_{CN} fluorescence by amino acid side-chains is discussed in Chapter 4. This is important for developing a detailed understanding of the factors which control F_{CN} fluorescence and is required if the probe is to be used to study conformational changes. Finally, the IR-active unnatural amino acid, azidohomoalanine, was introduced into NTL9 recombinantly and an analysis of how the change in the azido vibrational properties can be used as a probe of protein folding is described in Chapter 5. It is demonstrated that the azido mode is sensitive to folding and to electrostatics.

1.8. References

1. Anfinsen, C. B. (1973) Principles that govern folding of protein chains, *Science 181*, 223-230.
2. Levinthal, C. (1968) Are there pathways for protein folding?, *J. Chim. Phys. Pcb. 65*, 44-&.
3. Bartlett, A. I., and Radford, S. E. (2009) An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms, *Nat. Struct. Mol. Biol. 16*, 582-588.
4. Dill, K. A., Bromberg, S., Yue, K. Z., Fiebig, K. M., Yee, D. P., Thomas, P. D., and Chan, H. S. (1995) Principles of protein-folding - A perspective from simple exact models, *Protein Sci. 4*, 561-602.
5. Magg, C., and Schmid, F. X. (2004) Rapid collapse precedes the fast two-state folding of the cold shock protein, *J. Mol. Biol. 335*, 1309-1323.
6. Ratner, V., Amir, D., Kahana, E., and Haas, E. (2005) Fast collapse but slow formation of secondary structure elements in the refolding transition of E. coli adenylate kinase, *J. Mol. Biol. 352*, 683-699.
7. Magg, C., Kubelka, J., Holtermann, G., Haas, E., and Schmid, F. X. (2006) Specificity of the initial collapse in the folding of the cold shock protein, *J. Mol. Biol. 360*, 1067-1080.
8. Arai, M., Kondrashkina, E., Kayatekin, C., Matthews, C. R., Iwakura, M., and Bilsel, O. (2007) Microsecond hydrophobic collapse in the folding of Escherichia

coli dihydrofolate reductase, an alpha/beta-type protein, *J. Mol. Biol. 368*, 219-229.

9. Daggett, V., and Fersht, A. R. (2003) Is there a unifying mechanism for protein folding?, *Trends Biochem. Sci. 28*, 18-25.

10. Snow, C. D., Nguyen, N., Pande, V. S., and Gruebele, M. (2002) Absolute comparison of simulated and experimental protein-folding dynamics, *Nature 420*, 102-106.

11. Santoro, M. M., and Bolen, D. W. (1988) Unfolding free-energy changes determined by the linear extrapolation method .1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants, *Biochemistry 27*, 8063-8068.

12. Tanford, C. (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation., *Advan. Protein Chem. 24*, 1-95.

13. Chan, C. K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A., and Hofrichter, J. (1997) Submillisecond protein folding kinetics studied by ultrarapid mixing, *Proc. Natl. Acad. Sci. U.S.A. 94*, 1779-1784.

14. Hagen, S. J., and Eaton, W. A. (2000) Two-state expansion and collapse of a polypeptide, *J. Mol. Biol. 301*, 1019-1027.

15. Jones, C. M., Henry, E. R., Hu, Y., Chan, C. K., Luck, S. D., Bhuyan, A., Roder, H., Hofrichter, J., and Eaton, W. A. (1993) Fast events in protein-folding initiated by nanosecond laser photolysis, *Proc. Natl. Acad. Sci. U.S.A. 90*, 11860-11864.

16. Yang, W. Y., and Gruebele, M. (2003) Folding at the speed limit, *Nature 423*, 193-197.

17. Wu, Y., Kondrashkina, E., Kayatekin, C., Matthews, C. R., and Bilsel, O. (2008) Microsecond acquisition of heterogeneous structure in the folding of a TIM barrel protein, *Proc. Natl. Acad. Sci. U.S.A. 105*, 13367-13372.

18. Vu, D. M., Myers, J. K., Oas, T. G., and Dyer, R. B. (2004) Probing the folding and unfolding dynamics of secondary and tertiary structures in a three-helix bundle protein, *Biochemistry 43*, 3582-3589.

19. Fafarman, A. T., Webb, L. J., Chuang, J. I., and Boxer, S. G. (2006) Site-specific conversion of cysteine thiols into thiocyanate creates an IR probe for electric fields in proteins, *J. Am. Chem. Soc. 128*, 13356-13357.

20. Mukherjee, S., Chowdhury, P., DeGrado, W. F., and Gai, F. (2007) Site-specific hydration status of an amphipathic peptide in AOT reverse micelles, *Langmuir 23*, 11174-11179.

21. Muralidharan, V., Cho, J. H., Trester-Zedlitz, M., Kowalik, L., Chait, B. T., Raleigh, D. P., and Muir, T. W. (2004) Domain-specific incorporation of noninvasive optical probes into recombinant proteins, *J. Am. Chem. Soc. 126*, 14004-14012.

22. Miyake-Stoner, S. J., Miller, A. M., Hammill, J. T., Peeler, J. C., Hess, K. R., Mehl, R. A., and Brewer, S. H. (2009) Probing protein folding using site-specifically encoded unnatural amino acids as FRET donors with tryptophan, *Biochemistry 48*, 5953-5962.

23. Ye, S. X., Zaitseva, E., Caltabiano, G., Schertler, G. F. X., Sakmar, T. P., Deupi, X., and Vogel, R. (2010) Tracking G-protein-coupled receptor activation using genetically encoded infrared probes, *Nature 464*, 1386-U1314.

24. Kent, S. B. H. (1988) Chemical synthesis of peptides and proteins, *Annu. Rev. Biochem. 57*, 957-989.

25. Dawson, P. E., and Kent, S. B. H. (2000) Synthesis of native proteins by chemical ligation, *Annu. Rev. Biochem. 69*, 923-960.

26. Muralidharan, V., and Muir, T. W. (2006) Protein ligation: an enabling technology for the biophysical analysis of proteins, *Nat. Methods 3*, 429-438.

27. Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code, *Annu. Rev. Biophys. Biomol. Struc. 35*, 225-249.

28. Hendrickson, W. A., Horton, J. R., and Lemaster, D. M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD) - A vehicle for direct determination of 3-dimensional structure, *EMBO J. 9*, 1665-1672.

29. Kiick, K. L., and Tirrell, D. A. (2000) Protein engineering by in vivo incorporation of non-natural amino acids: Control of incorporation of methionine analogues by methionyl-tRNA synthetase, *Tetrahedron 56*, 9487-9493.

30. Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation, *Proc. Natl. Acad. Sci. U.S.A. 99*, 19-24.

31. Tucker, M. J., Oyola, R., and Gai, F. (2006) A novel fluorescent probe for protein binding and folding studies: p-cyano-phenylalanine, *Biopolymers 83*, 571-576.

32. Tucker, M. J., Oyola, R., and Gai, F. (2005) Conformational distribution of a 14-residue peptide in solution: A fluorescence resonance energy transfer study, *J. Phys. Chem. B 109*, 4788-4795.

33. Tang, J., Yin, H., Qiu, J. D., Tucker, M. J., DeGrado, W. F., and Gai, F. (2009) Using two fluorescent probes to dissect the binding, insertion, and dimerization kinetics of a model membrane peptide, *J. Am. Chem. Soc. 131*, 3816.

34. Tucker, M. J., Tang, J., and Gai, F. (2006) Probing the kinetics of membrane-mediated helix folding, *J. Phys. Chem. B 110*, 8105-8109.

35. Marek, P., Gupta, R., and Raleigh, D. P. (2008) The fluorescent amino acid p-cyanophenylalanine provides an intrinsic probe of amyloid formation, *Chembiochem 9*, 1372-1374.

36. Schultz, K. C., Supekova, L., Ryu, Y. H., Xie, J. M., Perera, R., and Schultz, P. G. (2006) A genetically encoded infrared probe, *J. Am. Chem. Soc. 128*, 13984-13985.

37. Oh, K. I., Lee, J. H., Joo, C., Han, H., and Cho, M. (2008) beta-Azidoalanine as an IR probe: Application to amyloid A beta(16-22) aggregation, *J. Phys. Chem. B 112*, 10352-10357.

38. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. D. (2005) Structures of the bacterial ribosome at 3.5 angstrom resolution, *Science 310*, 827-834.

39. Hoffman, D. W., Cameron, C. S., Davies, C., White, S. W., and Ramakrishnan, V. (1996) Ribosomal protein L9: A structure determination by the combined use of X-ray crystallography and NMR spectroscopy, *J. Mol. Biol. 264*, 1058-1071.

40. Herr, A. J., Nelson, C. C., Wills, N. M., Gesteland, R. F., and Atkins, J. F. (2001) Analysis of the roles of tRNA structure, ribosomal protein L9, and the bacteriophage T4 gene 60 bypassing signals during ribosome slippage on mRNA, *J. Mol. Biol. 309*, 1029-1048.

41. Kuhlman, B., Boice, J. A., Fairman, R., and Raleigh, D. P. (1998) Structure and stability of the N-terminal domain of the ribosomal protein L9: Evidence for rapid two-state folding, *Biochemistry 37*, 1025-1032.

42. Kuhlman, B., and Raleigh, D. P. (1998) Global analysis of the thermal and chemical denaturation of the N-terminal domain of the ribosomal protein L9 in H2O and D2O. Determination of the thermodynamic parameters, Delta H degrees, Delta S degrees, and Delta C degrees(p), and evaluation of solvent isotope effects, *Protein Sci. 7*, 2405-2412.

43. Cho, J. H., Sato, S., and Raleigh, D. P. (2004) Thermodynamics and kinetics of non-native interactions in protein folding: A single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state, *J. Mol. Biol. 338*, 827-837.

44. Cho, J. H., and Raleigh, D. P. (2005) Mutational analysis demonstrates that specific electrostatic interactions can play a key role in the denatured state ensemble of proteins, *J. Mol. Biol. 353*, 174-185.

45. Anil, B., Li, Y., Cho, J. H., and Raleigh, D. P. (2006) The unfolded state of NTL9 is compact in the absence of denaturant, *Biochemistry 45*, 10110-10116.

2. Interpretation of *p*-Cyanophenylalanine Fluorescence in Proteins in Terms of Solvent Exposure and Contribution of Side-chain Quenchers: A Combined Fluorescence, IR and Molecular Dynamics Study

Abstract

The use of non-coded amino acids as spectroscopic probes of protein folding and function is growing rapidly, in large part because of advances in the methodology for their incorporation. Recently *p*-cyanophenylalanine has been employed as a fluorescence and IR probe, as well as a FRET probe to study protein folding, protein–membrane interactions, protein–protein interactions and amyloid formation. The probe has been shown to be exquisitely sensitive to hydrogen bonding interactions involving the cyano group, and its fluorescence quantum yield increases dramatically when it is hydrogen bonded. However, a detailed understanding of the factors which influence its fluorescence is required to be able to use this popular probe accurately. Here we demonstrate the recombinant incorporation of *p*-cyanophenylalanine in the N-terminal domain of the ribosomal protein L9. Native state fluorescence is very low which suggests that the group is sequestered from solvent, however, IR measurements and molecular dynamics simulations show that the cyano group is exposed to solvent and forms hydrogen bonds to water. Analysis of mutant proteins and model peptides demonstrates that the reduced native state fluorescence is caused by the effective quenching of *p*-cyanophenylalanine fluorescence via FRET to tyrosine side-chains. The implications for

the interpretation of *p*-cyanophenylalanine fluorescence measurements and FRET studies are discussed.

Note: The material presented in this chapter has been published (H. Taskent-Sezgin, J. Chung, V. Patsalo, S.J. Miyake-Stoner, A.M. Miller, S.H. Brewer, R.A. Mehl, D.F. Green, D.P. Raleigh and I. Carrico. *Biochemistry.* **2009**, 48, 9040-9046). This chapter contains direct excerpts from the manuscript that was written by me with suggestions and revisions by Professors Raleigh and Carrico. V. Patsalo ran the molecular dynamics simulations and the FTIR spectra were collected at Franklin and Marshall College.

2.1. Introduction

The experimental study of protein folding remains an area of intense interest, owing in part to the realization that protein misfolding can play a major role in certain diseases (*1*). The increasing interplay between experimental and computational investigations of folding has also helped to refocus interest in the area (*2*, *3*). Most experimental studies of folding rely on changes in the fluorescence of intrinsic fluorophores such as Trp or Tyr (*4*, *5*). Recently, *p*-cyanophenylalanine (F_{CN}) has been shown to have unique fluorescence properties. F_{CN} can be selectively excited at 240 nm and its emission maximum is 290 nm; its fluorescence quantum yield depends on hydrogen bonding interactions involving the cyano (CN) group (*6*). F_{CN} fluorescence can be selectively detected in the presence of Trp and Tyr and the group can be tolerated in both a hydrophobic environment and in a hydrophilic environment, making it a conservative substitution for Tyr, Phe, and Trp.

Gai and co-workers have pioneered the use of F_{CN} as a fluorescence probe and have also demonstrated that F_{CN}-Trp can be used as a FRET pair with a Förster distance (R_0) of 16 Å (*6*). This is a convenient R_0 for studies of protein folding. Gai and co-workers have utilized F_{CN} fluorescence and F_{CN}-Trp FRET to study protein–membrane (*7*, *8*) and protein–protein interactions (*6*), while we have used F_{CN} fluorescence to probe amyloid formation (*9*) and protein folding (*10*, *11*). These studies have led to increasing interest in the use of F_{CN}. The CN group also provides a useful IR probe (*12*). Due to the sensitivity of its stretching frequency to solvation and electric field effects, the CN group can serve as a reporter of the local environment, and Boxer and coworkers have used CN groups in Stark tuning experiments to map protein electric fields (*13*).

F_{CN} fluorescence intensity is traditionally interpreted to reflect the presence or absence of H-bonding interactions with the CN group but little is known about other factors that can influence the quantum yield such as quenching by amino-acid side chains. A deeper understanding of F_{CN} fluorescence is clearly required given the utility of the group as a probe of protein structure, protein–ligand interactions and protein–membrane interactions. Here we use a combination of fluorescence, infrared (IR) spectroscopy and molecular dynamics (MD) simulations to explore the factors which control F_{CN} fluorescence. As a model system we choose the N-terminal domain of the ribosomal protein L9 (NTL9). NTL9 is a 56-residue, mixed α-β protein and is the simplest example of the split β-α-β motif. Folding of the domain is reversible and two-state (*14, 15*). Almost all of the published folding studies of NTL9 have made use of Tyr fluorescence since Tyr25 is the only fluorescent group in the domain. Tyr25 is on the solvent exposed face of the first helix in NTL9, however, the ring packs against the body of the protein and there is a measurable change in fluorescence upon unfolding (Figure-1). NTL9 can tolerate a wide range of substitutions in its hydrophobic core (*14, 15*), and we recently prepared a derivative containing a Phe5 to F_{CN} substitution (NTL9-F5F_{CN}) by solid-phase peptide synthesis and have shown that the substitution does not perturb the structure (*10*). Thus NTL9 is an excellent model system to probe the origins of changes in F_{CN} fluorescence.

Here we show that the analysis of F_{CN} fluorescence is more complicated than initially thought and demonstrate that the fluorescence intensity cannot be interpreted solely in terms of hydrogen bonding to the CN group owing to complications caused by quenching from amino acid side-chains.

2.2. Materials and Methods

2.2.1. Protein Expression, Purification and Characterization

The NTL9-WT gene was inserted into a pBAD vector. The codon for F5 in the NTL9 gene was replaced with the amber stop codon to allow coding for F_{CN} (*16*). The vector that is generated is denoted pBAD-NTL9-F5F_{CN}. A pDule vector carrying the gene for the aminoacyl-tRNA synthetase for the incorporation of F_{CN} into the growing protein chain was used (*11, 17*). The two vectors, pBAD and pDule, were co-transformed into electrocompetent *E.coli* DH10B cells. A 4mL LB culture grown overnight was used as a seed for a sample of 750mL of main auto-induction media containing 1mM F_{CN}. The culture was grown for 24 hours at 37˚C at 250 rpm (*18*). The proteins were purified with ion exchange chromatography followed by reverse-phase HPLC on a Vydac C8 preparative column as described previously (*14*). The protein, NTL9-F5F_{CN} was characterized by LTQ-Orbitrap XL Mass Spectrometer. The observed monoisotopic MW was 6241.448 Da and calculated monoisotopic MW was 6241.443 Da. NTL9-F5F_{CN}Y25F was characterized by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). The observed average MW was 6228.70 Da and calculated average MW was 6228.30 Da.

2.2.2. Peptide Synthesis and Characterization

The control peptides, GGF$_{CN}$AA and GGF$_{CN}$YA were synthesized by standard Fmoc solid-phase peptide synthesis on an Applied Biosystems 433A peptide synthesizer. Both peptides have a free N-terminus and the PAL-PEG-PS resin leads to an amidated C-terminus. The peptides were cleaved from the resin using 91% (v/v) trifluoroacetic acid

(TFA), 3% (v/v) anisole, 3% (v/v) thioanisole, 3% (v/v) 1,2-ethanedithiol. The peptides were precipitated using cold ethyl ether and scavengers were removed under vacuum. The crude peptide was then purified using reverse-phase HPLC on a Vydac C18 semi-preparative column. An A-B gradient system was used with buffer A composed of 0.1% (v/v) solution of TFA and buffer B composed of 90% (v/v) isopropanol, 9.9% (v/v) water and 0.1% (v/v) TFA. The gradient was 10 min of buffer A wash followed by 0-90% B in 120 min. The peptides were characterized by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). For $GGF_{CN}AA$, Observed MW: 446.10 Da and calculated MW: 446.45 Da and for $GGF_{CN}YA$, Observed MW: 538.57 Da and calculated MW: 538.55 Da.

2.2.3. Equilibrium Denaturation Experiments

Circular Dichroism (CD) wavelength scans and CD-monitored denaturation experiments were conducted as previously described for NTL9 (*14*). Fluorescence-monitored denaturation experiments were conducted on an Applied Photon Technologies fluorimeter. The protein and peptide concentrations were determined by measuring the absorbance of the sample at 280 nm. The extinction coefficient for F_{CN} is 850 $M^{-1}cm^{-1}$ (*6*) and for Tyr, it is 1490 $M^{-1}cm^{-1}$. Fluorescence- and CD-monitored experiments were performed at pH 5.4 and 25°C. These experimental conditions were chosen to match previous studies. Samples were dissolved in 20 mM sodium acetate and 100 mM NaCl. For the fluorescence emission spectra, the F_{CN} group was excited at 240 nm and the emission signal was recorded in the range 250-450 nm. For the fluorescence-monitored urea denaturation experiments, the excitation wavelength for the F_{CN} group was 240 nm,

and the emission signal was followed at 291 nm. The concentration of urea solutions was determined by measuring the refractive index. Denaturation curves were fit to the following equation:

$$f = \frac{a_n + b_n[denaturant] + (a_d + b_d[denaturant])\exp(-\Delta G_u^o([denaturant]/RT))}{1+\exp(-\Delta G_u^o([denaturant]/RT))} \quad (2.1)$$

where:

$$\Delta G_u^o([denaturant]) = \Delta G_u^o(H_2O) - m[denaturant] \quad (2.2)$$

f is the measured signal, (ellipticity in CD, and fluorescence signal in fluorescence-monitored denaturation experiments), a_n is the intercept of the extrapolated curve in the pre-transition region, b_n is the slope of this same curve. a_d is the intercept of the curve in the post-transition region, b_d is the slope in this region. ΔG_U^o is the change in free energy of unfolding, *T* is the absolute temperature and *R* is the gas constant.

2.2.4. Equilibrium FTIR Measurements

Equilibrium FTIR absorbance spectra were recorded on a Bruker Vertex 70 FTIR spectrometer equipped with a globar source, KBr beamsplitter, a liquid nitrogen cooled mercury cadmium telluride (MCT) detector, and a Harrick BioATRcell II accessory. The spectra were the result of 512 scans and were recorded at a resolution of 1.0 cm^{-1}. The NTL9-F5F_{CN} protein was dissolved in a buffer containing 20 mM sodium acetate and 100 mM NaCl at a pH of 5.4, with or without 9.8 M urea. The protein concentration for the IR experiments was ~1.5 mM. The spectra were measured at 25˚C. The nitrile IR absorbance band was fit with a Gaussian line shape using Igor Pro (Wavemetrics, Inc.).

2.2.5. Molecular Dynamics Simulation System Setup

NTL9-WT coordinates (PDB: 2HBB) were retrieved from the Protein Data Bank (The Research Collaboratory for Structural Bioinformatics (RCSB), http://www.rcsb.org/pdb/). The last five residues of NTL9 are partially disordered and are not listed in the PDB file. Thus, the simulations were conducted on residues 1 to 51. Hydrogens were added with the HBUILD (*19*) module of CHARMM (*20*), and the system was minimized briefly using the PARAM22 force field (*21*). An orthorhombic box of pre-equilibrated TIP3P solvent was constructed 10 Å around the protein, and the system was charge-neutralized by random placement of sodium and chloride ions to a final concentration of approximately 145 mM. The final system contained 3607 TIP3P water molecules.

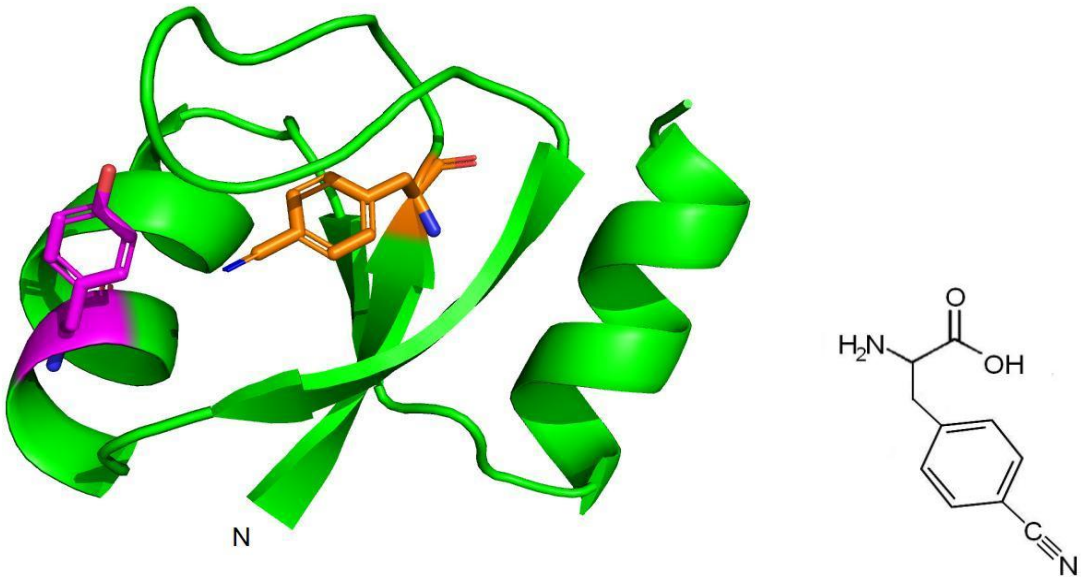
For NTL9-F5F_{CN}, Phe5 was mutated to F_{CN} by replacing the *para*-hydrogen with a nitrile, keeping all other atoms fixed. The F_{CN} side-chain was minimized while keeping the protein fixed, followed by unconstrained minimization. A box of solvent was then constructed and the system was neutralized by the addition of ions, as above.

Molecular dynamics calculations were run with NAMD (*22*). Briefly, the fully solvated system was minimized for 5000 steps, and equilibrated by heating to 300K over 200 ps using Particle Mesh Ewald electrostatics and a time step of 2 fs. Production dynamics were run at 300K for 90 ns at constant temperature and pressure (1 atm), with full electrostatics evaluated every time step.

2.3. Results and Discussion

A ribbon diagram of NTL9 is displayed in Figure 2.1 and shows the location of F5, the site chosen for F_{CN} substitution, and Y25 which is the dominant natural fluorophore. F5 is a core residue and is located on the first β-strand. F_{CN} was introduced by expression in *E.coli* using a modified version of the methodology pioneered by the Schultz group *(11, 16, 18, 23, 24).* The plasmid (pDule), containing the gene for the aminoacyl t-RNA synthetase used to incorporate the unnatural phenylalanine variants, and the protein expression conditions were originally optimized in the Mehl laboratory (*11, 18*). We modified these conditions to provide optimum expression of NTL9-F5F_{CN}. A yield of 7.5 mg of purified NTL9-F5F_{CN} was obtained from 0.75 liters of culture supplemented with 1mM F_{CN}. The incorporation of F_{CN} into the mutant, NTL9-F5F_{CN} was confirmed with the LTQ Orbitrap XL Mass Spectroscopy. The monoisotopic mass distribution is consistent with essentially complete incorporation (Figure 2.2).

Position 5 was chosen based on our results with chemically synthesized NTL9 variant which demonstrated that replacement of F5 with the F_{CN} group is a conservative mutation (*10*). The far-UV CD wavelength spectrum of the mutant, denoted here NTL9-F5F_{CN}, indicates that the secondary structure of this variant is similar to the wild-type protein (Figure 2.3). We determined the stability of NTL9-F5F_{CN} by both CD and fluorescence monitored urea denaturation experiments (Figure 2.4). The stability and the m-values determined by the two methods are identical confirming that the modified domain folds cooperatively. Importantly, the stability and m-value are very close to those for wild-type NTL9 providing additional evidence that the F_{CN} substitution does not perturb the fold (Table 2.1).



MKVI<u>F</u>LKDVKGKGKKGEIKNVADG<u>Y</u>ANNFLFKQGLAIEATPANLKALEAQ KQKEQR

Figure 2.1. Ribbon diagram of NTL9 showing the location of the F5F_{CN} (orange) mutation and Y25 (magenta). The amino acid sequence of NTL9 is shown, F5 and Y25 are underlined. Insert: The structure of p-cyanophenylalanine (F_{CN}).

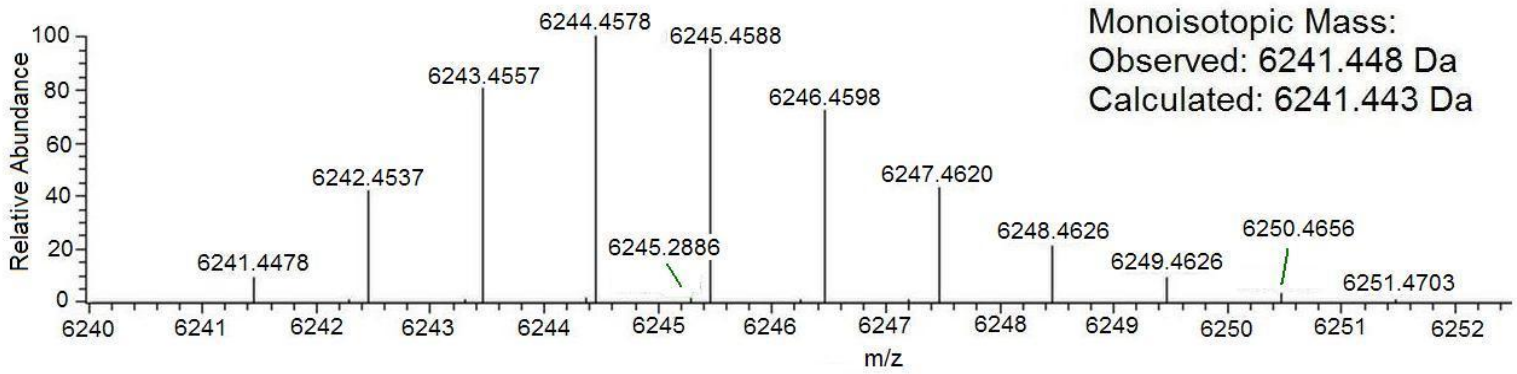


Figure 2.2. Orbitrap MS of NTL9-F5F$_{CN}$. Observed (6241.448 Da) and calculated (6241.443 Da) monoisotopic masses are in excellent agreement.

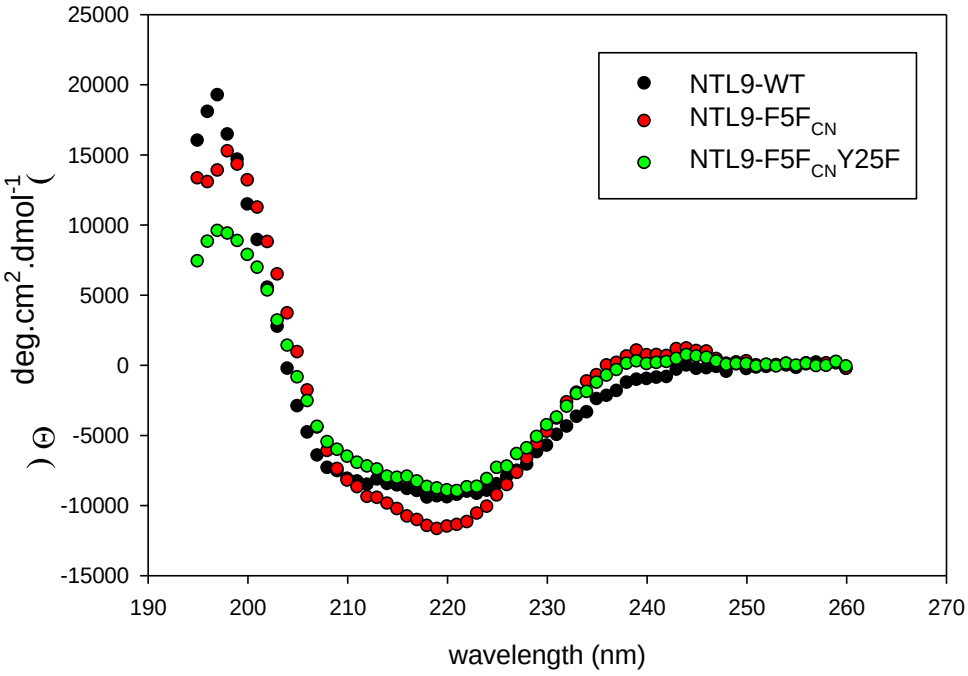


Figure 2.3. Far-UV CD wavelength spectra of NTL9-WT (black) and F5F$_{CN}$ (red) and F5F$_{CN}$Y25F (green) mutants. The spectra were recorded at 25°C, pH 5.4 in 20mM sodium acetate, 100mM NaCl.

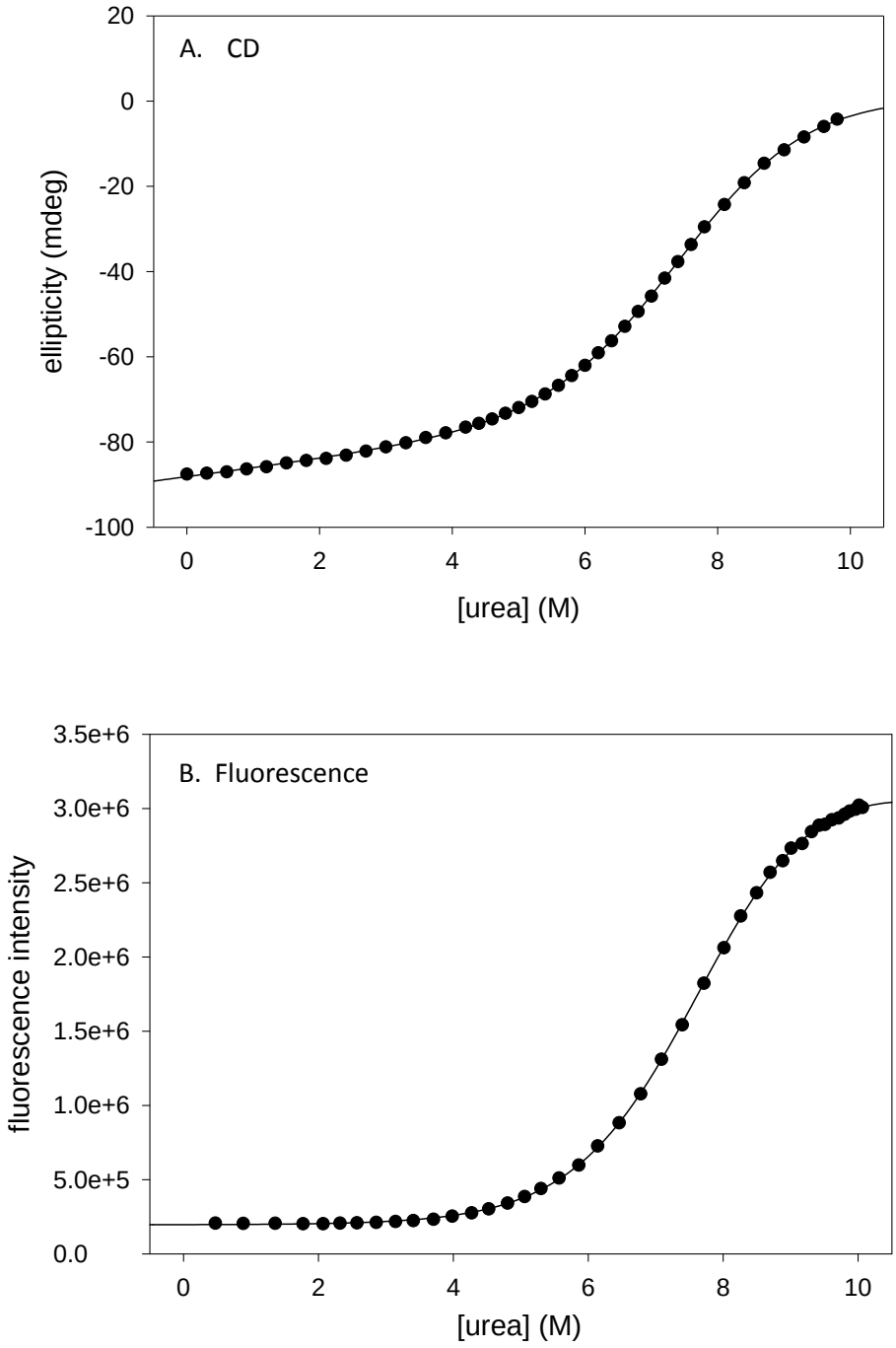


Figure 2.4. (A) CD- and (B) Fluorescence-monitored urea denaturation of NTL9-F5F$_{CN}$. The CD signal was monitored at 222 nm. The F_{CN} group is excited at 240 nm and the emission signal is followed at 291 nm. Experiments were conducted at pH 5.4 and 25°C.

		ΔG^o (kcal.mol^{-1})	m (kcal.mol^{-1}.M^{-1})
NTL9-WT		4.30 ± 0.36[a]	0.66 ± 0.07[a]
NTL9-F5F_{CN}	CD	4.48 ± 0.05	0.59 ± 0.01
	Fluorescence	4.98 ± 0.06	0.64 ± 0.01
NTL9-F5F_{CN}Y25F	CD	4.24 ± 0.04	0.62 ± 0.01
	Fluorescence	4.54 ± 0.08	0.66 ± 0.01

Table 2.1.Thermodynamic parameters for NTL9-WT, F5F_{CN} and F5F_{CN}Y25F variants. All experiments were conducted in 20 mM sodium acetate, 100 mM NaCl at 25°C, pH 5.4. The numbers after the ± sign represent the standard errors to the fit. [a]Values taken from Cho et al.(*25*). The uncertainty for wild-type represents the standard deviation from multiple measurements.

The fluorescence emission spectra of NTL9-F5F_{CN} in the folded and in the urea-induced unfolded state are shown in Figure 3.5. The spectrum of wild-type NTL9 in the folded state is included for comparison. An excitation wavelength of 240 nm was used for all three samples. The fluorescence intensity is drastically reduced in the folded state of NTL9-F5F_{CN} relative to the unfolded state and the ratio is similar to what was previously reported for the chemically synthesized variant (*10*). The low fluorescence intensity of the F_{CN} group in the native state is, by the traditional interpretation, consistent with the burial of the CN group (*6*). However, there are some peculiarities which suggest that the situation may be more complex. For example, the fluorescence emission spectrum of the wild-type protein recorded using an excitation wavelength of 240 nm is actually more intense than the spectrum of the F5F_{CN} mutant (Figure 2.5). This is surprising since the mutant includes both F_{CN} and Tyr fluorophores.

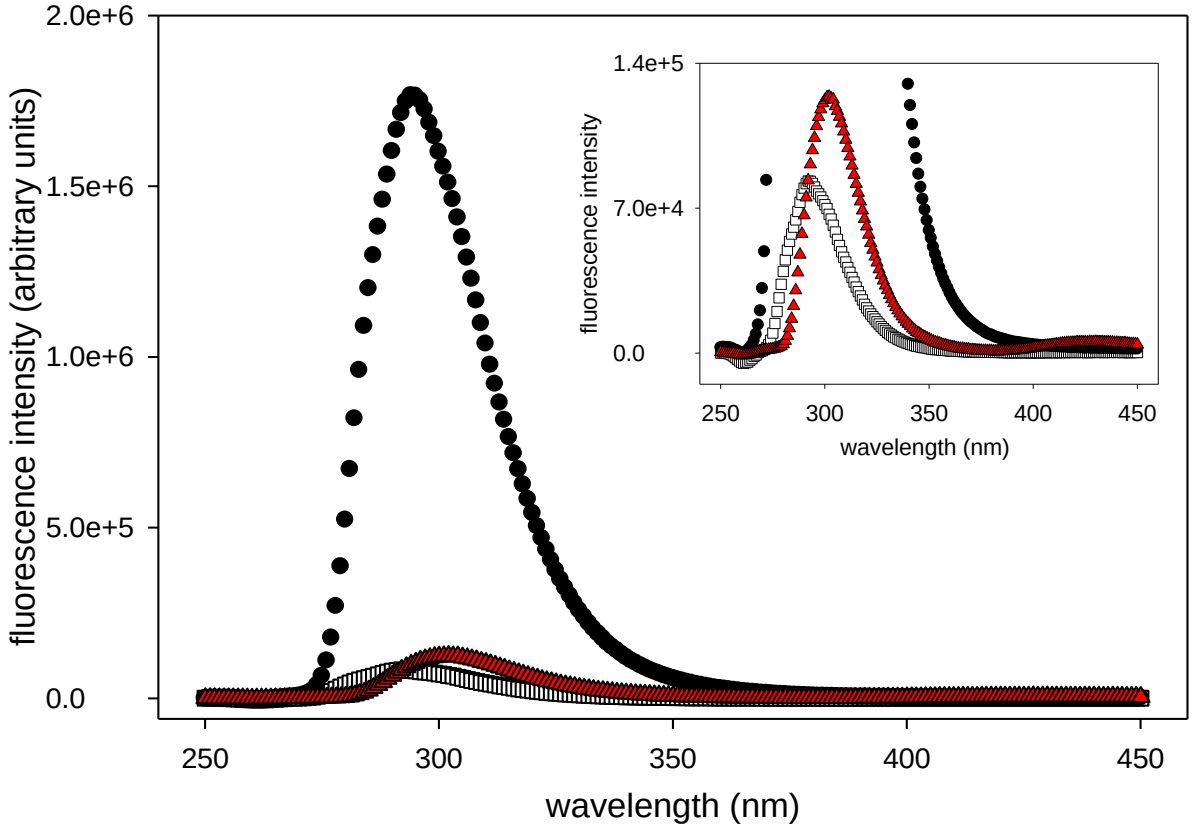


Figure 2.5. Fluorescence emission spectra of NTL9-F5F$_{CN}$ in the folded (open squares), in the urea-induced unfolded state (filled circles) and wild type NTL9 in the folded state (red triangles). The increase in the F_{CN} fluorescence signal upon going from the folded state to the unfolded state in NTL9-F5F$_{CN}$ is traditionally thought to reflect increased solvation of the F_{CN} group. Insert: Expansion of the emission spectra of NTL9-F5F$_{CN}$ and NTL9-WT in the folded state. Protein concentration was 15uM. Experiments were conducted at pH 5.4 and 25°C in 20mM sodium acetate, 100mM NaCl. The excitation wavelength was 240nm for all experiments.

To further investigate the origins of this effect, we used IR spectroscopy to probe the molecular environment of the CN group in order to obtain an independent assessment of the environment of the CN group. The CN stretching mode is a convenient spectroscopic marker since it appears in a transparent region of protein IR spectra, and its frequency is sensitive to hydrogen bonding and electric field effects. A blue-shifted CN stretching vibration is observed when the group is transferred from a hydrophobic

environment to water (*12*). The vibrational frequency reported for the CN group of F_{CN} was 2237.2 cm^{-1} in water and the CN group of Fmoc-F_{CN} was 2228.8 cm^{-1} in THF (*12*). Figure 2.6A shows the ATR-FTIR (Attenuated Total Reflection Fourier Transform Infrared) spectra of NTL9-F5F_{CN} in the absence and presence of urea. The nitrile stretching frequency of NTL9-F5F_{CN} in the absence of urea is 2234.2 cm^{-1} suggesting a partially solvated nitrile group in the folded state of NTL9. The nitrile stretching frequency of NTL9- F5F_{CN} in the urea-induced unfolded state (9.8 M urea) is 2232.6 cm^{-1}, which is only a 1.6 cm^{-1} red shift relative to the folded state. A blue shift was expected upon unfolding since the nitrile moiety should become more solvated. However, the high concentration of urea present in the aqueous buffer lowers the dielectric of the buffer and may alter water structure. The net effect is a red shift in the frequency of an exposed nitrile group in 9.8M urea relative to that observed in the native buffer. This effect is illustrated in Figure 2.6B which shows the ATR-FTIR spectra of a control peptide, GGF_{CN}AA in the absence and presence of urea. The nitrile stretching frequency of the control peptide shifted from 2236.0 cm^{-1} to 2233.7 cm^{-1} upon addition of 9.8 M urea, a red shift of 2.3 cm^{-1}. This peptide is disordered and thus the CN group should be fully solvated in both the native and urea containing buffers. The 2.3 cm^{-1} shift represents a reasonable correction for the effects of 9.8M urea. After correcting for this intrinsic urea dependence of the CN stretch, an estimated blue shift of ~0.7 cm^{-1} for the nitrile stretching frequency is obtained upon unfolding of NTL9-F5F_{CN}. Irrespective of the details of the urea induced correction, the key result is that the CN stretching frequency is very similar in the folded and unfolded states and is close to that observed for model compounds in water.

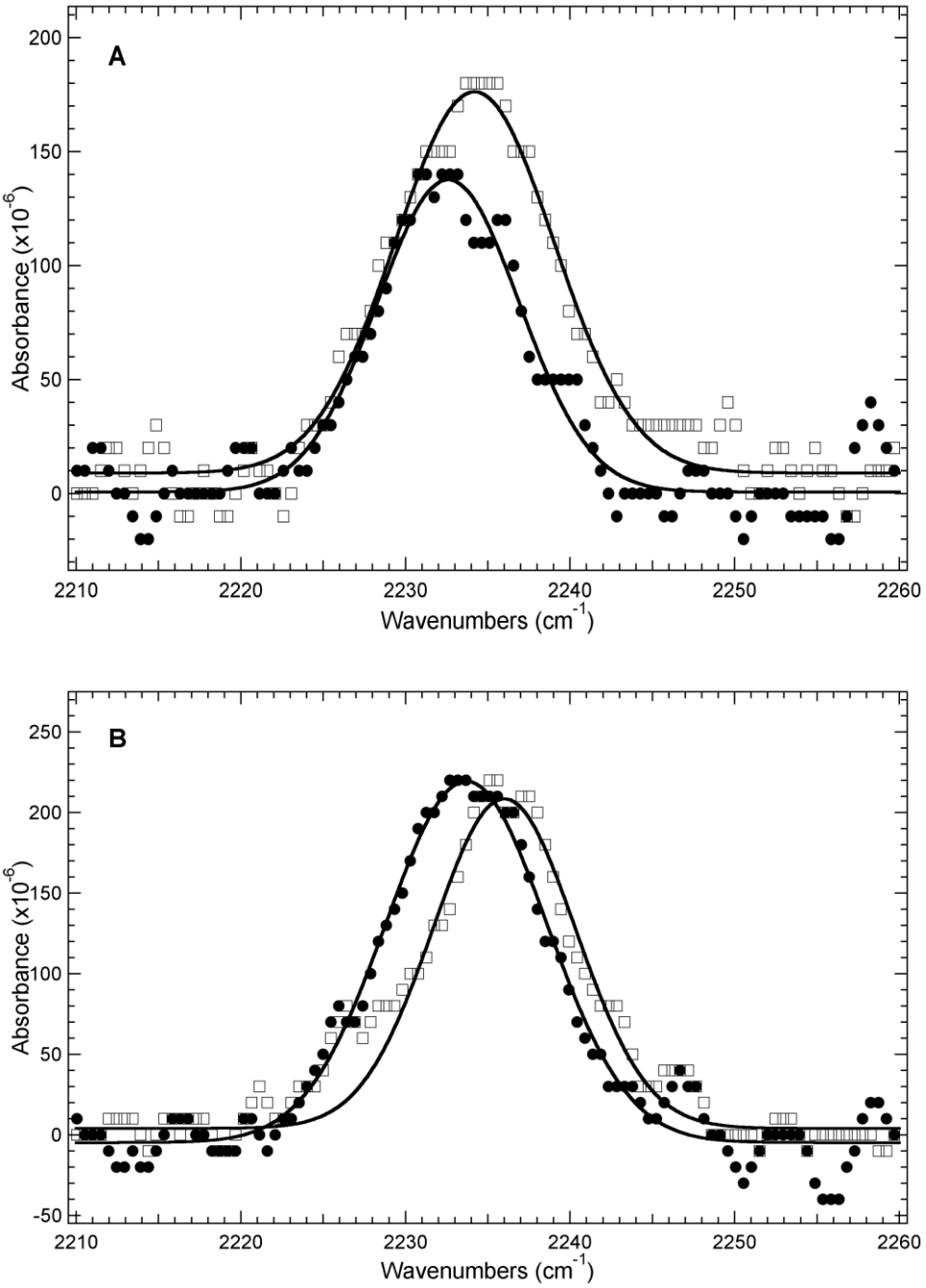


Figure 2.6. (A) FTIR spectra of NTL9-F5F_{CN} in the folded (open squares) and in the urea-induced unfolded state (filled circles). The CN stretching frequency was observed at 2234.2 cm^{-1} and 2232.6 cm^{-1} in the folded and in the unfolded state, respectively. This indicates that the CN group is in a similar environment in both states. The protein concentration was around 1.5 mM. (B) FTIR spectra of the control peptide, GGF_{CN}AA in the absence (open squares) and presence (filled circles) of urea. Samples were at pH 5.4 and 25°C. Each spectrum was fit to a Gaussian line shape (solid curves).

The fluorescence and IR studies lead to apparently incompatible conclusions. The very low fluorescence intensity in the folded state argues that the F_{CN} group is buried, while the IR data indicates that the CN group is exposed. Comparison of the fluorescence emission spectrum of wild-type NTL9 with NTL9-F5F_{CN} (recorded with an excitation wavelength of 240 nm) suggests, as noted above, that the tyrosine side chain may quench the fluorescence of the F_{CN} group in the native state. In the crystal structure of NTL9, the distance between the center of the Tyr25 ring and the center of the Phe5 ring is only 7.4Å. The emission maximum of F_{CN} in NTL9-F5F_{CN}Y25F is 290nm and the full width of the emission spectrum at 50% intensity is about 30nm, hence there is considerable overlap between the tyrosine absorption band and the NTL9-F5F_{CN}Y25F emission spectrum. Given the significant spectral overlap, the relatively high fluorescence quantum yield of F_{CN} (comparable to Trp) and the close approach of F5F_{CN} and Y25, the most likely cause for the unexpectedly weak native state F_{CN} fluorescence is quenching due to FRET to Y25.

To more rigorously test the possibility that Tyr25 quenches the F_{CN} fluorescence, we prepared an F5F_{CN}Y25F double mutant. Far-UV CD indicated that the double mutant has the same fold as the wild-type (Figure 2.3). The stabilities and m-values calculated from the urea-induced unfolding experiments monitored by fluorescence and CD were very similar to the wild-type values (Figure 2.7) indicating that the substitution is well tolerated. The fluorescence emission spectra of NTL9-F5F_{CN}Y25F in the native and urea-induced unfolded state are shown in Figure 2.8. An approximately 40-fold increase in the fluorescence emission signal of NTL9-F5F_{CN}Y25F in the native state is detected relative to that observed for NTL9-F5F_{CN}. This directly validates the hypothesis that Tyr25

quenches the fluorescence signal of the F_{CN} group in NTL9-F5F_{CN} via FRET. The intense fluorescence of F5F_{CN} in NTL9-F5F_{CN}Y25F in the folded state is consistent with solvent exposure of the CN group and is fully consistent with the FT-IR results. To further test the ability of Tyr to quench F_{CN} fluorescence, we synthesized a pair of small F_{CN} containing peptides. Figure 2.9 compares the fluorescence emission spectra of F_{CN} in the GGF_{CN}AA and GGF_{CN}YA peptides. The F_{CN} fluorescence is considerably reduced in the Tyr containing peptide confirming that Tyr is an effective quencher of F_{CN} fluorescence.

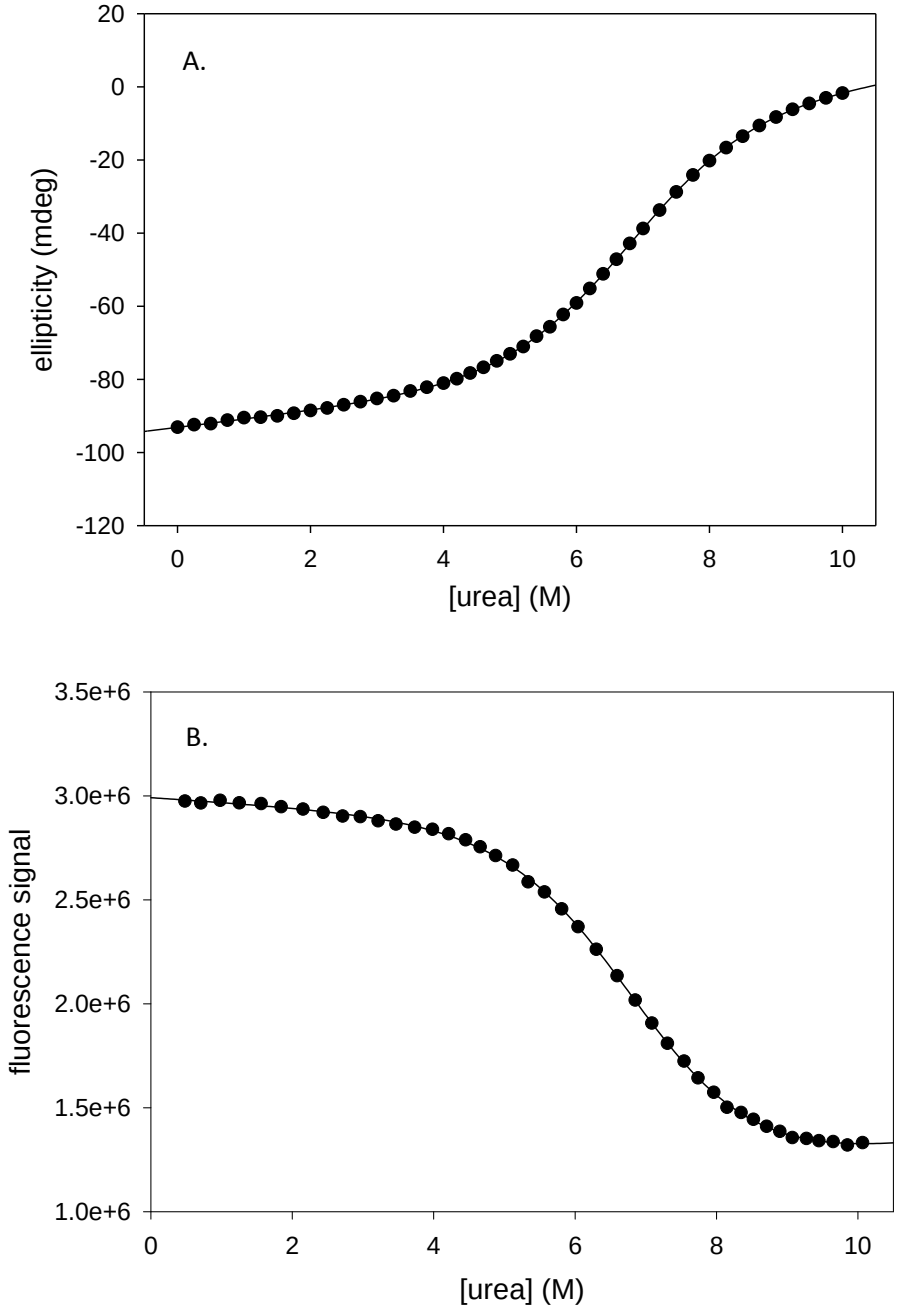


Figure 2.7. (A) CD- and (B) Fluorescence-monitored urea denaturation of NTL9-F5F$_{CN}$Y25F. The CD signal was monitored at 222 nm. The F_{CN} group was excited at 240 nm and the emission signal was followed at 291 nm. Samples were in 20mM sodium acetate, 100mM NaCl at pH 5.4, 25°C.

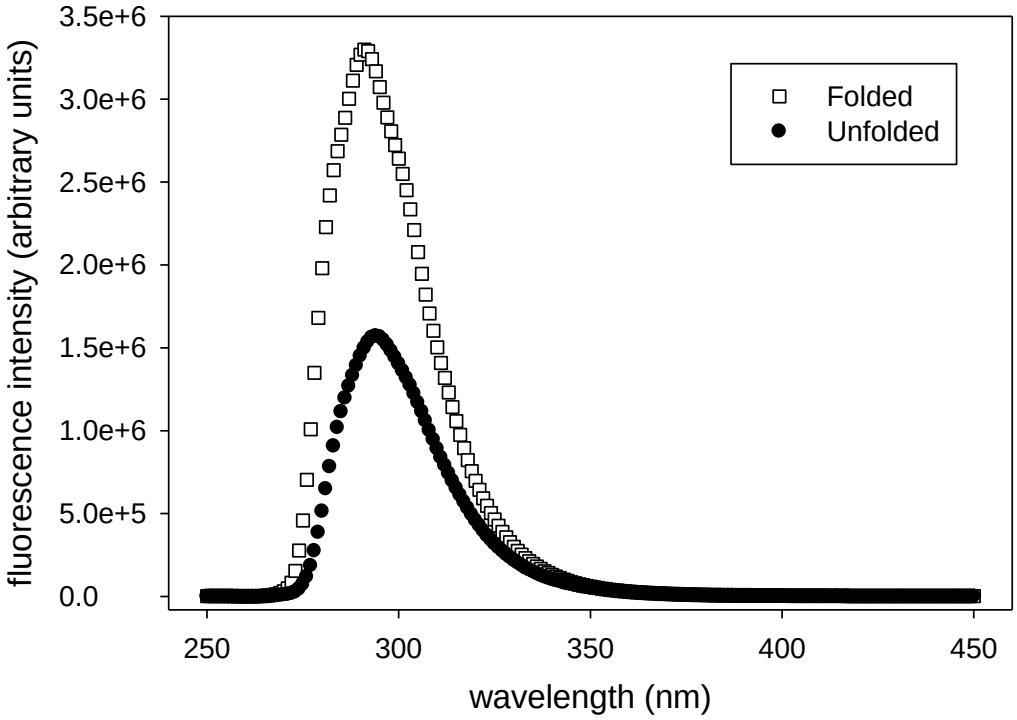


Figure 2.8. Fluorescence emission spectra of NTL9-F5F$_{CN}$Y25F in the folded (open squares) and in the urea-induced unfolded state (filled circles). The high fluorescence intensity in the folded state indicates that the F_{CN} group is not completely buried in NTL9-F5F$_{CN}$Y25F and forms H-bonds with the solvent. Protein concentration was 15uM. Protein samples were in 20mM sodium acetate, 100mM NaCl, pH 5.4 at 25°C.

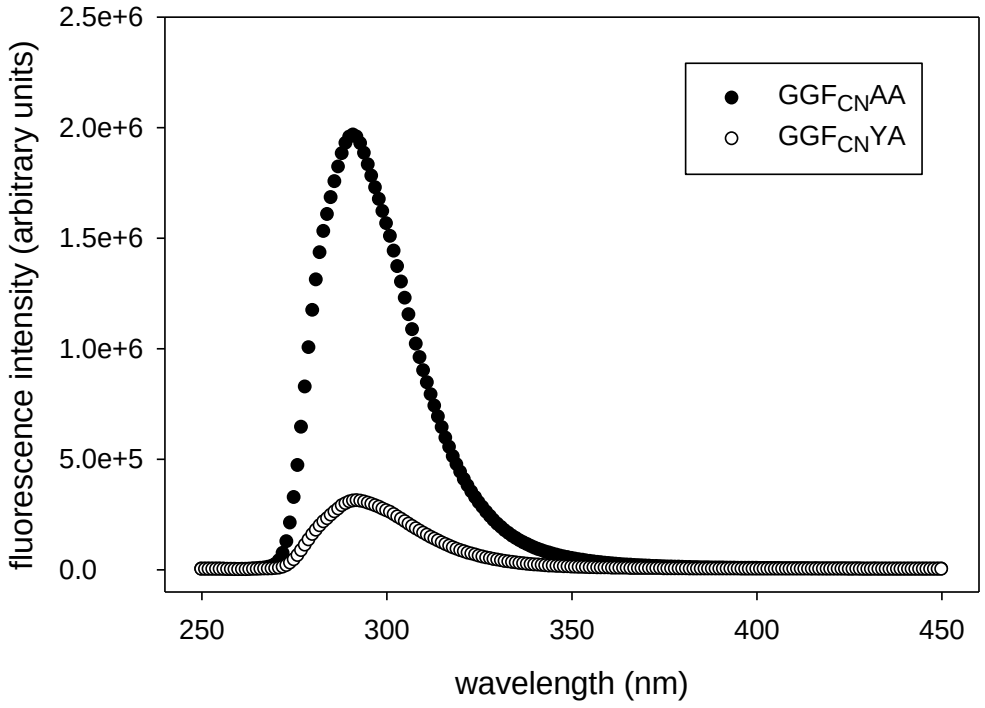


Figure 2.9. Fluorescence emission spectra of model peptides confirm that Tyr is an efficient quencher of F_{CN} fluorescence. GGF$_{CN}$AA (filled circles) and GGF$_{CN}$YA (open circles) in folding buffer. The decrease in the fluorescence signal intensity of F_{CN} (in GGF$_{CN}$AA) with the introduction of a Tyr (in GGF$_{CN}$YA) demonstrates that Tyr is quenching the fluorescence signal of F_{CN}. The peptide concentrations were 20uM.

To assess whether the model suggested by these results was consistent with structural data, we conducted molecular dynamics (MD) simulations of wild-type NTL9 and NTL9-F5F$_{CN}$ using the CHARMM (*20*) PARAM22 force field (*21*) and the TIP3P water model to probe for solvent hydrogen bond interactions with the F_{CN} group. This work was carried out by Mr. Vadim Patsalo in our laboratory. It was first necessary to parameterize the F_{CN} group. Parameters for F_{CN} were modeled by homology, using the Tyr residue of PARAM22 and the cyano group parameters from the 3-cyano-pyridine residue of the CHARMM pyridine force field (*24*). The parametrization placed a partial charge of -0.53e on the cyano nitrogen and +0.40e on the carbon of the cyano group.

The Cα backbone RMSD of the whole protein (residues 1 to 51) and of the N-terminal sub-domain (residues 1 to 39) (Figure 2.10) showed both proteins to be stably folded near the X-ray structure for the course of the 90 ns simulations. The simulations also appeared to be reasonably converged. The RMSD for the 1-39 core unit relative to the starting structure for the simulation, averaged over the last 45 nanoseconds of the MD trajectories, was 0.80 Å for wild-type and 0.79 Å for NTL9-F5F_{CN}. Chi angle analysis for the side chain at position 5 (Phe in wild-type, blue and F_{CN} for NTL9-F5F_{CN}, red) showed that the two side chains adopt the same χ_1 and χ_2 angles (Figure 2.11), further supporting the claim that the F5F_{CN} mutation is conservative, and that the side chain adopts the same orientation in both proteins. Using a hydrogen bond cut-off distance of 2.4 Å, the CN group was found to accept a hydrogen bond from surrounding water molecules in 55% of the trajectory frames during the course of the 90 ns MD trajectory. In addition, we calculated the distance and the relative orientation between the side chains at positions 5 and 25 (Figure 2.12). The observed mean distance between F5F_{CN} and Y25 was 7.43Å. The side chains were also found to adopt a fairly rigid relative orientation, despite the solvent accessibility of Y25. The close proximity of F5F_{CN} to Y25 side-chain and the observed H-bonding interactions between the CN group of the F5F_{CN} mutant with water in the native state are consistent with the IR data and support the hypothesis that the low fluorescence signal intensity in NTL9-F5F_{CN} mutant is not due to the burial of the CN group but rather caused by quenching due to the Tyr.

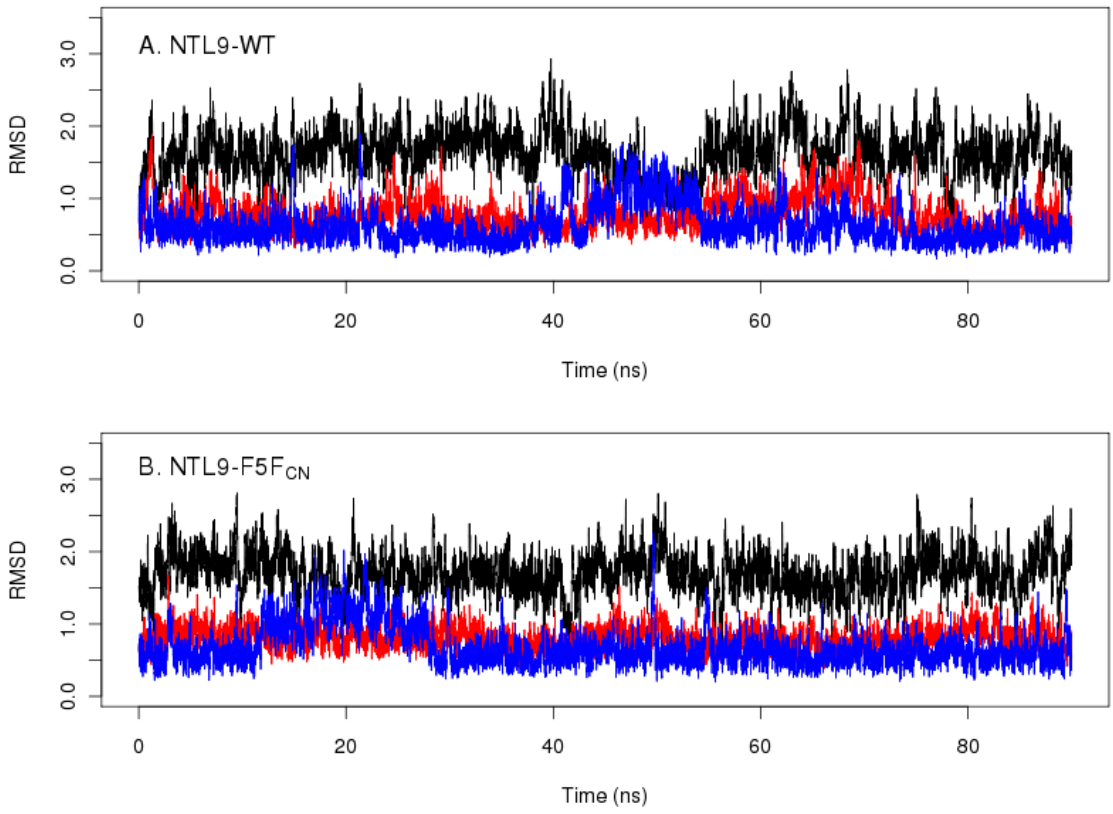


Figure 2.10. Calculated Cα backbone RMSD of the whole protein (residues 1 to 51) (black), the N-terminal sub-domain (residues 1-39) (red), and residues 40-51 (blue). (A) wild-type, (B) NTL9-F5F$_{CN}$.

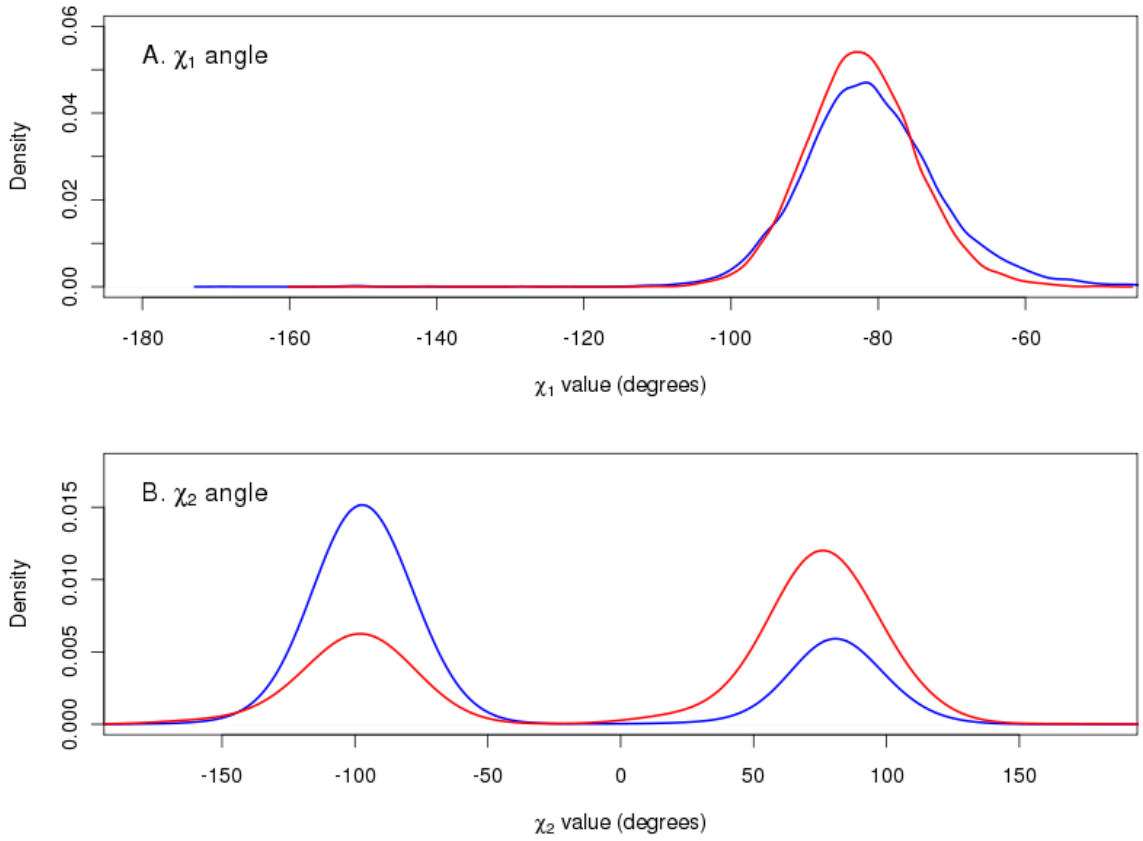


Figure 2.11. Calculated probability density functions for Chi angles (χ_1 and χ_2) for the side chain at position 5, Phe in NTL9-WT (blue) and F_{CN} for NTL9-F5F_{CN} (red). (A) χ_1, (B) χ_2. Note, χ_2 includes rotation of a group with C_2 symmetry, and thus the peaks at $\approx\pm100°$ are equivalent.

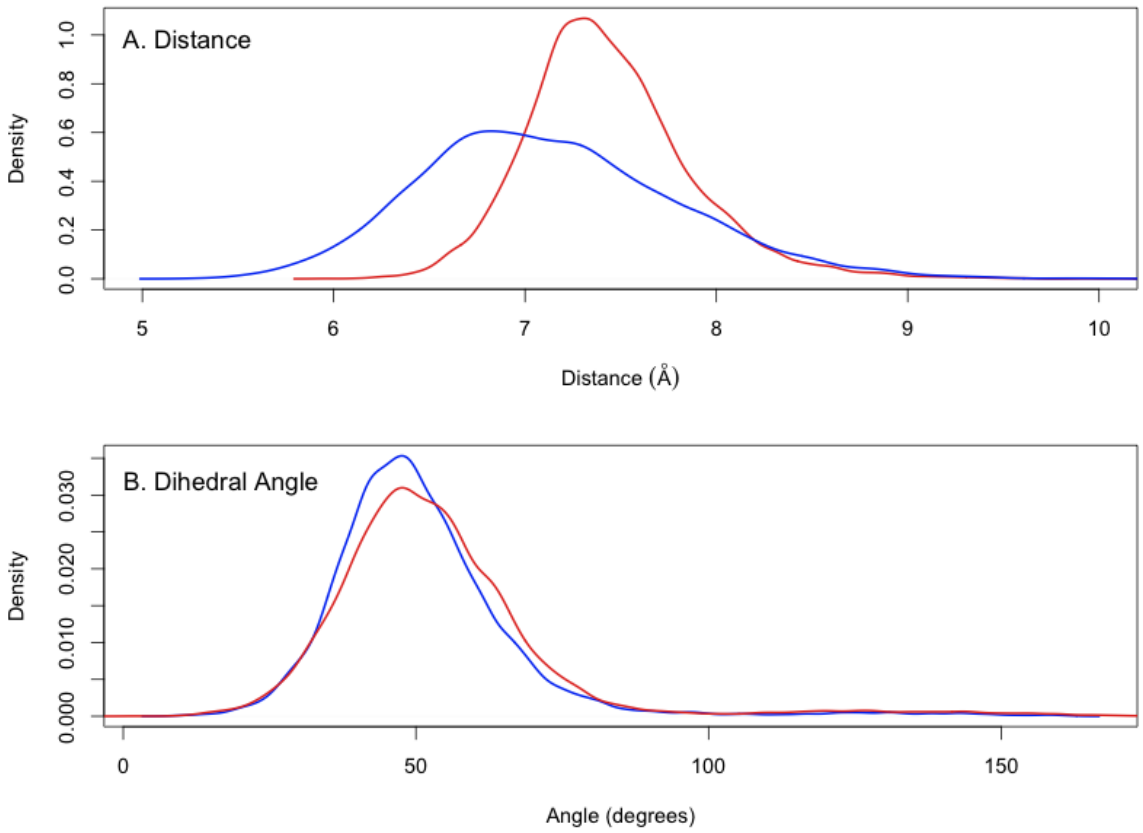


Figure 2.12. Calculated probability density functions for distance and inter-residue dihedral angle for the side chains at positions 5 and 25 as observed by molecular dynamics. (A) Center-of-mass distances between the phenyl rings of the side chains at positions 5 and 25 for NTL9-WT (blue) and NTL9-F5F_{CN} (red). (B) Calculated dihedral angle between geometric planes fit through phenyl rings of the side chains at positions 5 and 25 for NTL9-WT (blue) and NTL9-F5F_{CN} (red).

Unnatural amino acids are becoming increasingly popular spectroscopic probes as the methodology to incorporate them into proteins improves (*16, 26*). Given its ease of incorporation and convenient spectroscopic properties we anticipate that the use of F_{CN} will continue to grow. Here, we have demonstrated the efficient recombinant incorporation of F_{CN} into a model protein. The combined fluorescence and IR investigations reported here clearly demonstrate that caution must be used when

interpreting F_{CN} fluorescence in terms of solvent exposure and specific hydrogen bonding interactions involving the CN group. This study also demonstrates the power of combining IR and fluorescence to probe the local environment. This is one of the attractive features of an F_{CN} substitution. A second advantage of this probe is that IR spectra can be recorded in high concentrations of urea which facilitates studies of protein folding. Our demonstration that Tyr effectively quenches F_{CN} fluorescence has important implications for F_{CN}–Trp based FRET studies since conformational transitions which modulate the distance of a F_{CN}–Trp FRET pair could also modulate any F_{CN}–Tyr interactions. Thus the apparent F_{CN}-Tpr FRET response could include a contribution from differential F_{CN}-Tyr FRET in the two states. Of course, F_{CN}-Tyr FRET can be used directly to probe conformational changes in proteins. We hope our analysis of F_{CN} quenching will motivate similar studies in other systems and will aid in the interpretation of the spectral response of this very useful probe. One can envisage more sophisticated applications involving proteins which contain Tyr, Trp and F_{CN} in which the probes are used to monitor relative chain motion between three sites. Finally, the parameterization of the F_{CN} group reported here should facilitate MD studies of F_{CN} containing proteins.

2.4. References:

1. Selkoe, D. J. (2003) Folding proteins in fatal ways, *Nature 426*, 900-904.
2. Dyer, R. B. (2007) Ultrafast and downhill protein folding, *Curr. Op. Struc. Biol. 17*, 38-47.
3. Kubelka, J., Hofrichter, J., and Eaton, W. A. (2004) The protein folding 'speed limit', *Curr. Op. Struc. Biol. 14*, 76-88.
4. Fersht, A. R. (2000) *Structure and Mechanism in Protein Science - A Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman & Company, New York.
5. Royer, C. A. (2006) Probing protein folding and conformational transitions with fluorescence, *Chem. Rev. 106*, 1769-1784.
6. Tucker, M. J., Oyola, R., and Gai, F. (2006) A novel fluorescent probe for protein binding and folding studies: p-cyano-phenylalanine, *Biopolymers 83*, 571-576.
7. Tang, J., Signarvic, R. S., DeGrado, W.F., and Gai, F. (2007) Role of helix nucleation in the kinetics of binding of Mastoparan X to phospholipid bilayer, *Biochemistry 46*, 13856-12863.
8. Tucker, M. J., Tang J., and Gai F. (2006) Probing the kinetics of membrane-mediated helix folding, *J. Phys. Chem. B 110*, 8105-8109.
9. Marek, P., Gupta, R., and Raleigh, D.P. (2008) The fluorescent amino acid p-cyanophenylalanine provides an intrinsic probe of amyloid formation, *ChemBioChem 9*, 1372-1374.
10. Aprilakis, K. N., Taskent, H., and Raleigh, D. P. (2007) Use of the novel fluorescent amino acid p-cyanophenylalanine offers a direct probe of hydrophobic core formation during the folding of the N-terminal domain of the ribosomal

protein L9 and provides evidence for two-state folding, *Biochemistry 46*, 12308-12313.

11. Miyake-Stoner, S. J., Miller, A. M., Hammill, J. T., Peeler, J. C., Hess, K. R., Mehl, R. A., and Brewer, S. H. (2009) Probing Protein Folding Using Site-Specifically Encoded Unnatural Amino Acids as FRET Donors with Tryptophan, *Biochemistry 48*, 5953-5962.
12. Getahun, Z., Huang, C. Y., Wang, T., De Leon, B., DeGrado, W. F., and Gai, F. (2003) Using nitrile-derivatized amino acids as infrared probes of local environment, *J. Am. Chem. Soc. 125*, 405-411.
13. Suydam, I. T., Snow, C. D., Pande, V. S., and Boxer, S. G. (2006) Electric fields at the active site of an enzyme: direct comparison of experiment with theory, *Science 313*, 200-204.
14. Anil, B., Sato, S., Cho, J. H., and Raleigh, D. P. (2005) Fine structure analysis of a protein folding transition state; distinguishing between hydrophobic stabilization and specific packing, *J. Mol. Biol. 354*, 693-705.
15. Horng, J. C., and Raleigh, D. P. (2003) Phi-Values beyond the ribosomally encoded amino acids: kinetic and thermodynamic consequences of incorporating trifluoromethyl amino acids in a globular protein, *J. Am. Chem. Soc. 125*, 9286-9287.
16. Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code, *Annu. Rev. of Biophys. & Biomol. Struc. 35*, 225-249.

17. Jackson, J. C., Hammill, J. T., and Mehl, R. A. (2007) Site-specific incorporation of a (19)F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity, *J. Am. Chem. Soc. 129*, 1160-1166.

18. Hammill, J. T., Miyake-Stoner, S., Hazen, J.L., Jackson, J.C. and Mehl, R.A. (2007) Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization, *Nature Prot. 2*, 2601-2607.

19. Brunger, A. T., and Karplus, M. (1988) Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison, *Proteins 4*, 148-156.

20. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations, *J. Comp. Chem. 4*, 187-217.

21. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B 102*, 3586-3616.

22. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD, *J. Comp. Chem. 26*, 1781-1802.

23. Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of Escherichia coli, *Science 292*, 498-500.

24. Yin, D. (1997) Ph.D. Thesis, "Parameterization for Empirical Force Field Calculations and A Theoretical Study of Membrane Permeability of Pyridine Derivative", *Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland*.

25. Cho, J. H., Sato, S., and Raleigh, D. P. (2004) Thermodynamics and kinetics of non-native interactions in protein folding: a single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state, *J. Mol. Biol. 338*, 827-837.

26. Link, A. J., Mock, M. L., and Tirrell, D. A. (2003) Non-canonical amino acids in protein engineering, *Curr. Op. Biotech. 14*, 603-609.

3. Use of the Novel Fluorescent Amino Acid p-Cyano-Phenylalanine Offers a Probe of Folding of the N-Terminal Domain of the Ribosomal Protein L9

Abstract

Fluorescence detected stopped flow measurements are the method of choice for studies of protein folding kinetics. However, the methodology suffers from the limitation that the protein of interest must either contain an intrinsic fluorophore or can tolerate its introduction by mutagenesis. Here we take advantage of this methodology to monitor the folding reaction of the N-terminal domain of L9 (NTL9) with an unnatural fluorescent amino acid, p-cyano-phenylalanine (F_{CN}). Previously, we have shown that F_{CN} fluorescence is quenched by Tyr and Met. Two mutants were generated, NTL9-F5F_{CN} and NTL9-F5F_{CN}Y25F in order to deconvolve the effect of Tyr quenching from the folding rate. The variants adopt the same fold as wild-type NTL9 and are slightly more stable at equilibrium. Refolding and unfolding was monitored using urea jump experiments. Plot of the natural log of the observed relaxation rate vs. denaturant concentration, so called chevron plot, exhibited the characteristic V shape, and no hint of deviation from linearity was observed at low denaturant concentration, however, the observed folding rates for the two mutants were different. NTL9-F5F_{CN} protein kinetics reports on F_{CN} fluorescence quenching by Tyr. The rate was 990 s^{-1}, which was very similar to wild-type protein folding rate. NTL9-F5F_{CN}Y25F protein kinetics report on

unquenching of F_{CN} fluorescence, most likely by Met1. The rate of this reaction was 1680 s^{-1}. The results suggest that either a) there is a non-native hydrophobic collapse which brings Met1 and $F5F_{CN}$ in close proximity in the initial stage of NTL9 folding or b) the non-native hydrophobic collapsed state exists due to the flexible nature of the protein chain in the unfolded state and as the protein folds it dissociates, unquenching F_{CN} fluorescence. This rearrangement that separates Met1 and $F5F_{CN}$ is much faster than the F_{CN} fluorescence quenching by Tyr.

Note: Part of the material presented in this chapter has been published (K.N. Aprilakis*, H.Taskent*, and D. P. Raleigh. *Biochemistry* **2007,** 46, 12308-12313). This chapter contains direct excerpts from the manuscript that was written by me with suggestions and revisions by Professor Raleigh. The experiments were conducted together with Mr. K.N. Aprilakis (* indicates joint first authorship).

3.1. Introduction

Protein folding is an extremely active area of research, motivated both by the desire to decipher the code which links primary sequence and function and by the importance of protein misfolding in biotechnology and human health (*1-7*). Probably the most widely applied experimental method for the direct study of folding and unfolding kinetics is fluorescence-detected stopped flow measurements (*3*). Such studies rely on the presence of a tryptophan or tyrosine residue which experiences a significant change in fluorescence upon unfolding. The method, although very powerful, does suffer from the requirement for an intrinsic fluorophore or depends upon the ability of the protein to tolerate the introduction of Tryptophan or Tyrosine at a suitable position. Recently the cyano (nitrile) analog of phenylalanine (F_{CN}) has been proposed for use as a fluorescence probe of folding and binding (*8*). The fluorescence properties of F_{CN} are discussed in Chapter 2. We have previously shown that F_{CN} fluorescence is quenched by Tyr (Chapter 2). In addition, Met also quenches F_{CN} fluorescence (Chapter 4).

In the present work we use F_{CN} fluorescence as a probe of folding of an α-β protein, the N-terminal domain of the ribosomal protein L9 (NTL9). NTL9 is a relatively small, 56 residue domain and is the simplest example of a common class of structure, the split β-α-β motif (*9-11*). The domain has been widely used for studies of protein folding and stability (*11-15*). Kinetic investigations have relied on the single Tyr found at position 25, since the protein does not contain any Trp residues. Tyr25 is located on the surface of the protein on the solvent exposed face of an α-helix. In the native state of NTL9 the Tyr side-chain adopts a conformation whereby the ring packs against the helix leaving one face exposed to solvent while shielding the other. This is enough to lead to a

change in fluorescence quantum yield between the folded and unfolded states, but the change in fluorescence does not directly report on hydrophobic core formation (*12*).

The structure of NTL9 is shown in ribbon format in Figure 3.1. The topology of the domain is β_1-β_2-α_1-β_3-α_2 where the three β-strands form an antiparallel β-sheet and the second helix extends beyond the globular domain to form the connector to the C-terminal domain. Tyr25 is located on the first helix. The domain is highly soluble, folds and unfolds reversibly over a wide range of conditions and does not require metal ion or cofactor binding to fold (*11*, *16*). Thus, it is an attractive model system for biophysical investigations. Phenylalanine-five, located on the first β-strand was replaced by F_{CN}. In NTL9-F5F_{CN}, the kinetics report on the F_{CN} fluorescence quenching by Tyr25 as the protein folds and unfolds. In NTL9-F5F_{CN}Y25F, Tyr25 is eliminated in order to follow the kinetics without the effect of Tyr quenching. The mutant, NTL9-F5F_{CN}Y25F experiences a change in its fluorescence intensity as the protein folds and unfolds which we speculate is due to the quenching effect of Met1 residue. The location of Met1, F5F_{CN} and Tyr25 are indicated in the ribbon diagram of NTL9 shown in Figure 3.1.

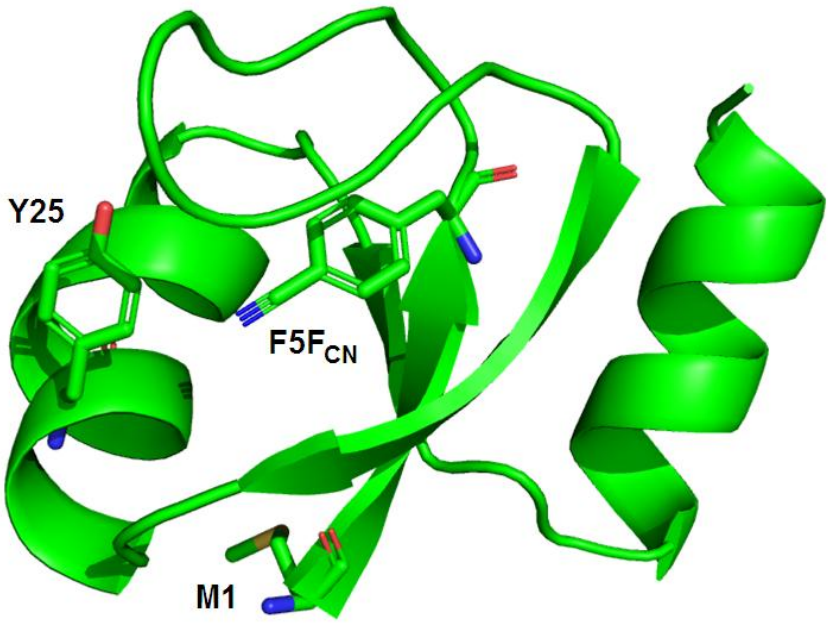


Figure 3.1. A ribbon diagram of NTL9 showing the location of the M1, F5F$_{CN}$, and Y25. Phe5 is completely buried in the native state. The diagram was generated using the pdb file 2HBA and the program Pymol.

3.2.Materials and Methods

3.2.1.Protein Synthesis and Purification

NTL9-F5F$_{CN}$ was synthesized using Fmoc chemistry on an Applied Biosystems 433A peptide synthesizer. Standard Fmoc protocols were used as described elsewhere (*14*). The Fmoc-F$_{CN}$ derivative was purchased from EMD Chemicals. The protein was purified using reverse phase HPLC on a Vydac C4 column as described (*14*).

3.2.2. Protein Expression and Purification

NTL9-F5F$_{CN}$Y25F mutant was expressed by 21st pair methodology as explained in Chapter 2. The protein was purified with ion exchange chromatography followed by

reverse-phase HPLC on a Vydac C8 preparative column as described previously for NTL9 (*14*). NTL9-F5F_{CN}Y25F was characterized by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). The observed average MW was 6228.70 Da and the calculated average MW was 6228.30 Da.

3.2.3. Equilibrium Denaturations

CD monitored denaturation experiments were performed as previously described for NTL9 (*14*). Samples were in 20 mM sodium acetate, 100 mM NaCl. Measurements were conducted at pH 5.4 and 25°C. F_{CN} fluorescence monitored denaturation experiments were conducted in 20 mM sodium acetate, 100 mM NaCl, pH 5.4 and 25°C. F_{CN} fluorescence was excited at 240 nm and emission was monitored at 297 nm. The concentration of urea solutions were determined by measuring the refractive index. Denaturation curves were fit to the equation:

$$f = \frac{a_n + b_n[\text{denaturant}] + (a_d + b_d[\text{denaturant}])\exp^{-[\Delta G^\circ_U([\text{denat.}]/RT)]}}{1 + e^{-[\Delta G^\circ_U([\text{denaturant}]/RT)]}} \qquad (3.1)$$

where:

$$\Delta G^o_U([\text{denaturant}]) = \Delta G^o_U(H_2O) - m[\text{denaturant}] \qquad (3.2)$$

f is the measured signal, ellipticity in CD and fluorescence signal in fluorescence monitored denaturation experiments, ΔG^o_U is the change in free energy for unfolding reaction, T is the temperature and R is the gas constant. a_n is the intercept and b_n is the

slope of the curve extrapolated in the pre-transition region. Similarly, a_d is the intercept and b_d is the slope of the curve in the post-transition region.

3.2.4.Stopped-flow Fluorescence Measurements

Stopped flow studies were performed as previously described for NTL9 with the exception that the excitation wavelength was 240 nm (*14*). A 305 nm cutoff filter was used to collect the fluorescence. NTL9-F5F_{CN} sample was in 20 mM sodium acetate, 100 mM NaCl and NTL9-F5F_{CN}Y25F sample was in 120mM sodium acetate. Measurements were conducted at pH 5.4 and 25°C. Stopped flow experiments used 11 fold dilution. The protein concentration before the jump was around 400 μM. After averaging four to five fluorescence traces at each denaturant concentration, the average trace was fit to a single-exponential function in order to obtain the observed rate constant (k_{obs}) at that denaturant concentration. The plot of natural logarithm of k_{obs} vs denaturant concentration was fit to the following equation:

$$\ln(k_{obs}) = \ln (k_f^{H_2O} \exp(-m_f [\text{denaturant}]/RT) + k_u^{H_2O} \exp(m_u [\text{denaturant}]/RT) \quad (3.3)$$

where $k_f^{H_2O}$ and $k_u^{H_2O}$ are the folding and unfolding rate in the absence of denaturant, respectively.

3.3. Results and Discussion

3.3.1. Stability and Fluorescence Emission Data of the Mutants

An F_{CN} for Phe replacement is well tolerated at position 5 of NTL9 in both $F5F_{CN}$ and $F5F_{CN}$Y25F mutants. CD studies demonstrate that the global fold is not perturbed. The location of Phe-5 is shown on the ribbon diagram of Figure 3.1. Both urea and guanidine HCl induced denaturation experiments were used to measure the stability of NTL9-$F5F_{CN}$ mutant (Figure 3.2). The urea dependent CD and fluorescence denaturation curve of NTL9-$F5F_{CN}$Y25F mutant is discussed in Chapter 2. The use of denaturant induced unfolding to measure $\Delta G°$ depends upon the validity of the linear extrapolation method (*17, 18*). The linear extrapolation method has been shown to be valid for NTL9 and for a large number of NTL9 mutants (*14, 19*). The stabilities measured for the mutants were similar to that of wild-type. The m-values for urea and guanidine induced unfolding report on the change in solvent accessible surface area between the native and unfolded states (*20*). The measured values are essentially identical to those previously measured for wild-type. The thermodynamic parameters are summarized in Table 3.1. The fact that NTL9 tolerates the replacement of a buried phenylalanine without loss of stability or perturbation of the structure illustrates the conservative nature of the F_{CN} substitution.

In Figure 3.3, fluorescence emission spectra of NTL9-$F5F_{CN}$ protein are shown in the folded and in the urea-induced unfolded state. The blue shifted absorbance of F_{CN} compared to Trp and Tyr allows F_{CN} fluorescence to be excited at 240 nm. The fluorescence is low in the folded state due to quenching of Tyr25, which is 7.4 Å away

from $F5F_{CN}$ (Chapter 2) (*21*). Fluorescence intensity increases as the protein unfolds because Tyr is further from F_{CN} in the unfolded state.

Figure 3.4 shows fluorescence emission spectra of NTL9-$F5F_{CN}$Y25F in the folded and in the unfolded state. The fluorescence is high in the folded state as expected after Tyr25 is replaced with a Phe. Suprisingly, the fluorescence intensity is low in the urea induced unfolded state which is most probably due to the Met1 residue. In the unfolded state ensemble, in some hydrophobic collapsed states Met1 might be closer to $F5F_{CN}$ than it is in the folded state because of the flexibility of the polypeptide chain. Possible amino acid side chain quenchers of F_{CN} fluorescence is investigated in Chapter 4. Other than Met1 there is no quenching amino acid side-chain in the protein.

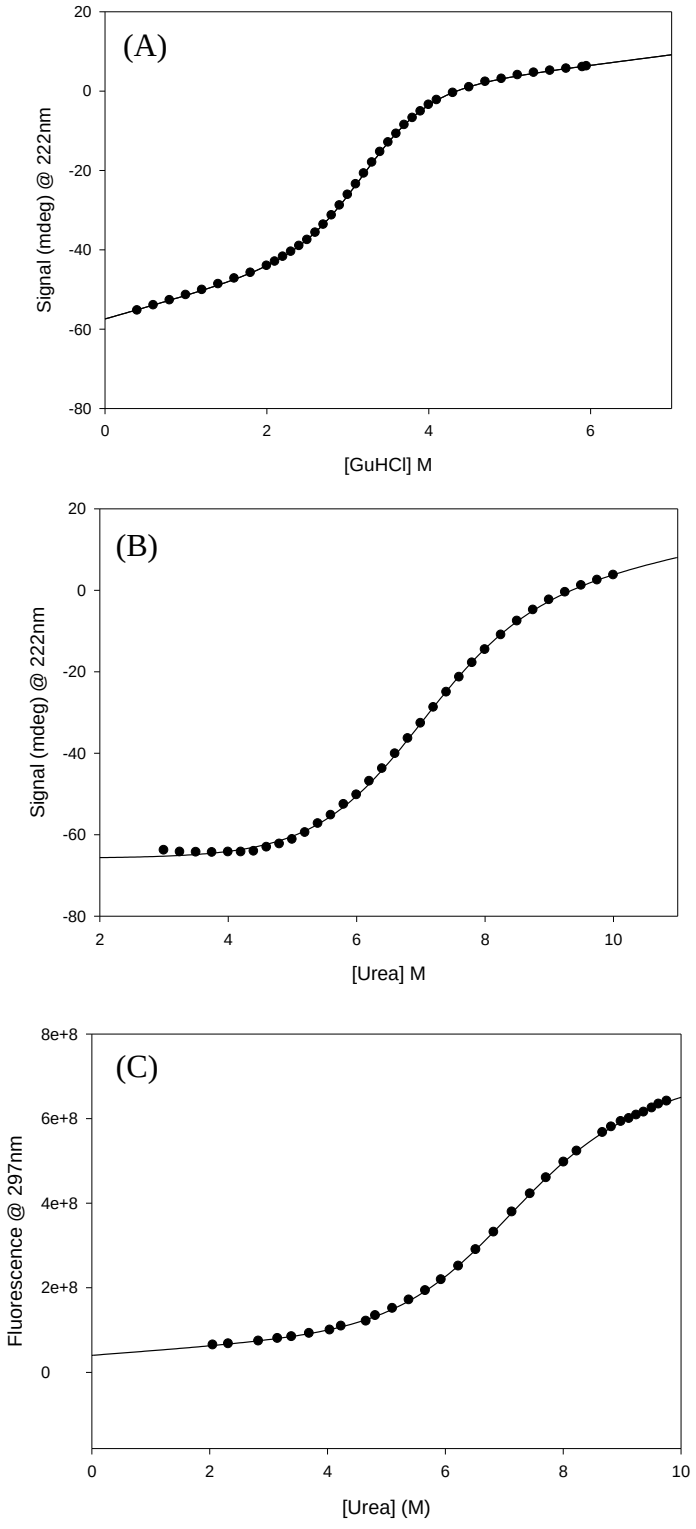


Figure-3.2. Equilibrium unfolding studies of NTL9-F5F$_{CN}$. (A) Guanidine HCl induced unfolding. (B) Urea induced unfolding. The CD signal at 222 nm was monitored. (C) Urea induced unfolding monitored by F_{CN} fluorescence at 297 nm.

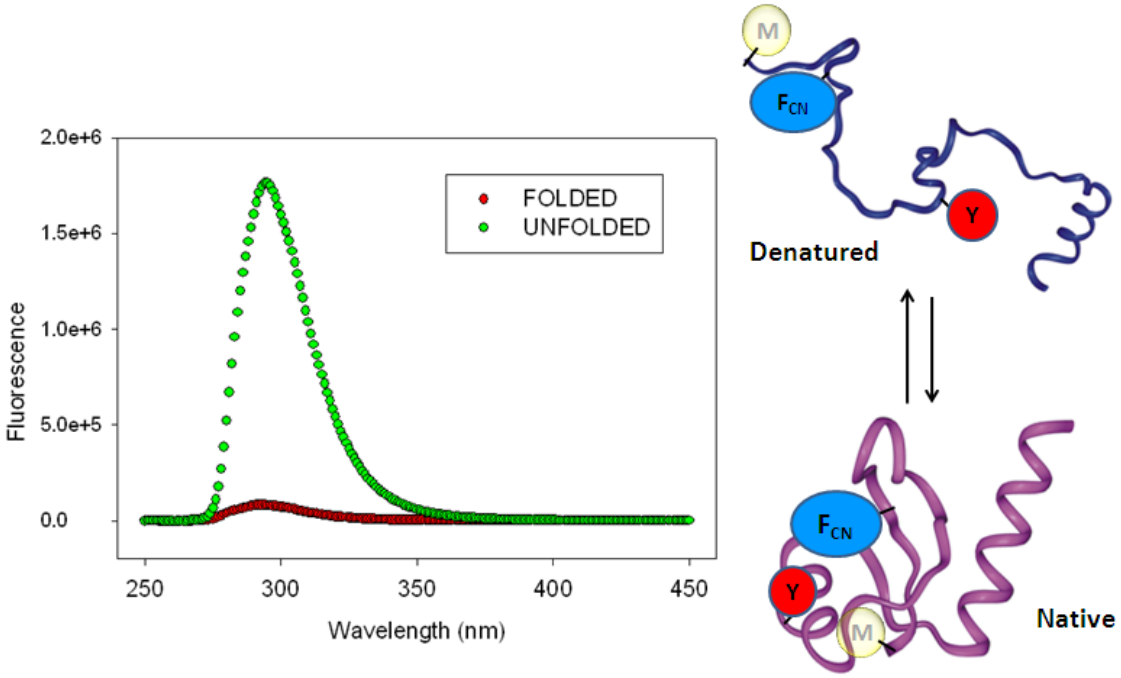


Figure-3.3. Fluorescence emission spectra of NTL9-F5F$_{CN}$ in the folded (red) and unfolded state (green). Experiments were conducted in 20mM sodium acetate, 100mM NaCl, pH 5.4, 25°C. The diagram displays the positions of M1, F5F$_{CN}$, and Y25 in the protein in the folded and in the unfolded state. The distance between the center of the Tyr ring and the F$_{CN}$ ring is 7.4 Å in the native state and the distance between the S on Met1 and the F$_{CN}$ ring is 11.0 Å in the folded state.

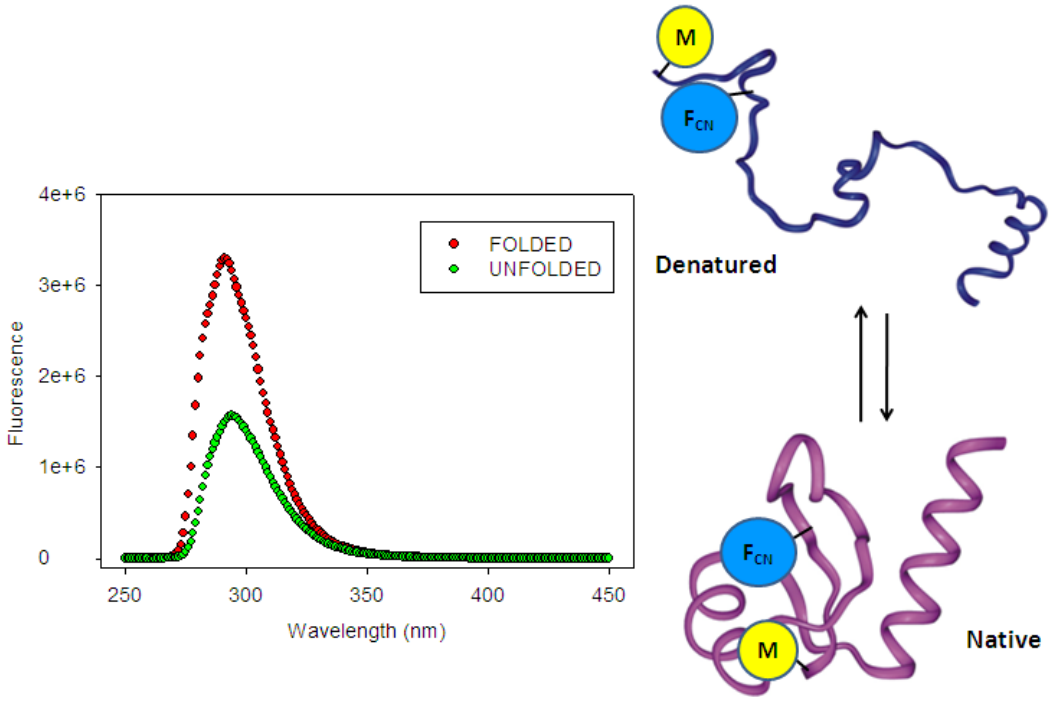


Figure 3.4. Fluorescence emission spectra of NTL9-F5F$_{CN}$Y25F in the folded (red) and in the unfolded state (green). Experiments were conducted in 20mM sodium acetate, 100mM NaCl, pH 5.4, 25°C. The diagram displays the positions of M1 and F5F$_{CN}$ in the protein in the folded and in the unfolded state.

3.3.2. Analysis of the Kinetics of NTL9-F5F$_{CN}$

Both mutants experience a change in the fluorescence emission spectra going from the folded to the unfolded state. In F5F$_{CN}$ the change is due to Tyr25 quenching, and Met1 is the reason for the change in fluorescence intensity of F5F$_{CN}$Y25F.

In NTL9-F5F$_{CN}$, the quenching and unquenching kinetics of F$_{CN}$ fluorescence by Tyr25 as the protein folds and unfolds was measured by stopped flow methods. A plot of the natural log of the observed first order rate constant vs. denaturant concentration, a chevron plot, is shown in Figure 3.5. The first panel displays the results of the guanidine jump experiments (Figure 3.5A) while the second shows the urea jump experiments (Figure 3.5B). The urea-induced unfolding transitions of both wild-type and NTL9-F5F$_{CN}$ are rather broad, as expected for a protein the size of NTL9. Thus the guanidine HCl jump studies were conducted in order to obtain the complete chevron plot. The plot displays the classic V shaped profile expected for cooperative two-state folding. In particular there is no hint of "roll-over", i.e. a deviation from linearity, at low concentration of denaturant. The quenching rate obtained from analysis of the data is slightly faster than folding rate of wild-type protein under the same conditions while the unfolding rate is slightly slower. However, given the lengthy extrapolation required to estimate the unfolding rate, especially for the urea data, the differences in k_u are unlikely to be significant. The unfolding branch of the urea jump experiment is poorly constructed since it is necessarily defined by a limited number of points, thus we feel that the estimated k_u is not reliable even though the standard error to the fit is modest. The slightly faster folding is consistent with previous studies that have shown that increasing the size and hydrophobicity of core residues increases the folding rate of NTL9 through

stabilization of the transition state via increased hydrophobic interactions (*14*). The quenching and unquenching rates measured for NTL9-F5F_{CN} are in good agreement with the folding and unfolding rate of the wild-type protein (Table 3.1). Thus, the rates measured by following F_{CN} fluorescence can be considered as folding and unfolding rates of the protein. In addition, the stability of the mutant calculated from the measured folding and unfolding rates is in good agreement with the value determined from equilibrium measurements.

The chevron plot also yields the dependence of the log of the folding and unfolding rates upon denaturant concentration. These parameters, traditionally denoted as m_f and m_u respectively, are believed to report on the relative change in accessible surface area between the unfolded state and the transition state (m_f) and between the folded state and the transition state (m_u) (*22*). The values are very similar to those obtained for wild-type NTL9 (*13*, *23*). m_f and m_u can be combined to yield the equilibrium m-value for unfolding. The value calculated from the kinetic parameters is in excellent agreement with the equilibrium values, consistent with two-state folding. m_f and m_u can also be used to calculate the dimensionless order parameter β_T, also sometimes denoted θ_m, $m_f/(m_f - m_u)$. β_T reports on the relative solvent exposure of the transition state (*22*). Values near unity reflect a native like transition state while values near zero indicate a denatured like transition state. Most single domain globular proteins have values between 0.5 and 0.8. For wild-type NTL9 the value of β_T is 0.60 to 0.65. NTL9-F5F_{CN} mutant has a value of 0.63 indicating that, by this measure, the relative position of the transition state has not shifted. The kinetic parameters are included in Table 3.1.

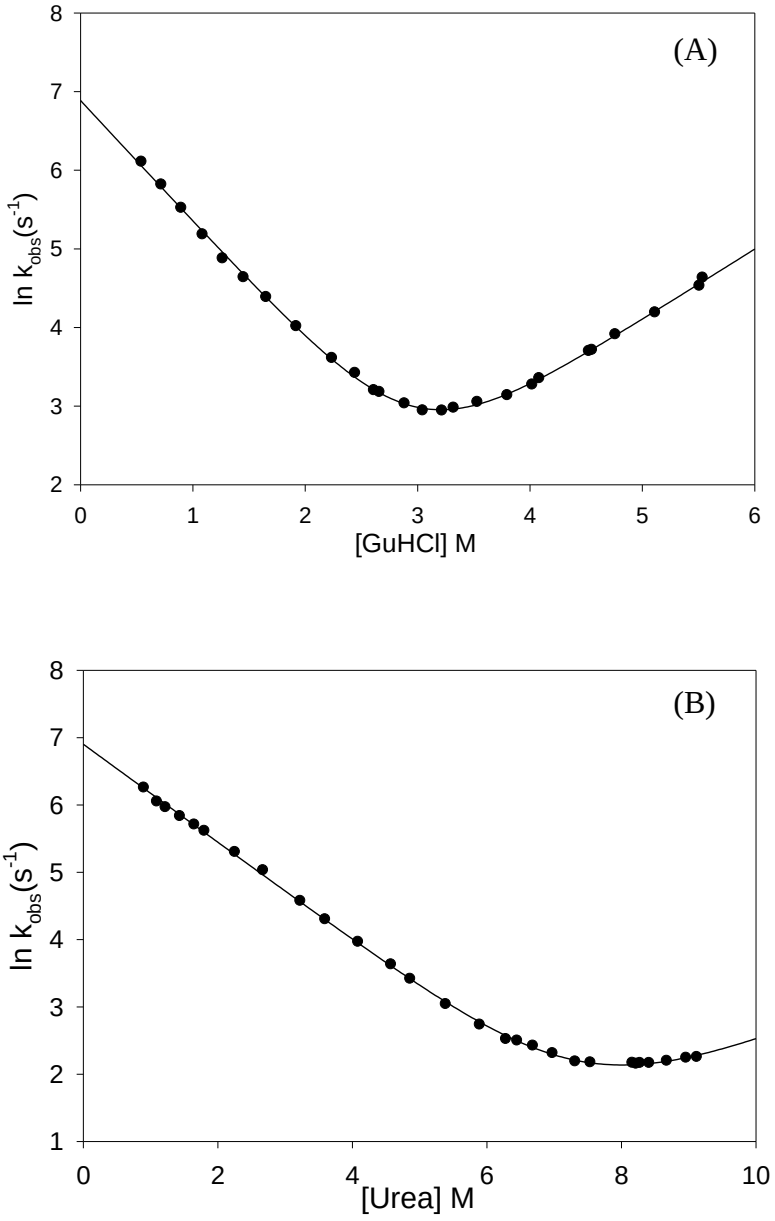


Figure 3.5. Fluorescence detected stopped flow folding studies of the NTL9-F5F$_{CN}$. (A) Guanidine HCl jump studies. (B) Urea jump studies. Experiments were conducted at pH 5.4, 25°C, in 20 mM sodium acetate, 100 mM NaCl.

Table 3.1. **Comparison of the measures of thermodynamic and kinetic parameters for WT, NTL9- $F5F_{CN}$, NTL9- $F5F_{CN}$Y25F as determined by urea and GuHCl denaturation**

	Protein	$\Delta G^{\circ}_{eqlib.}$ *kcal mol^{-1}*	m_{total} *kcal mol^{-1} M^{-1}*	k_f *s^{-1}*	k_u *s^{-1}*	m_f *kcal mol^{-1} M^{-1}*	m_u *kcal mol^{-1} M^{-1}*	β_T
Urea	WT	4.30 ± 0.36[a]	0.66 ± 0.07[a]	870 ± 120[a]	0.90 ± 0.4[a]	0.45 ± 0.06[a]	0.20 ± 0.09[a]	0.69
	$F5F_{CN}$	4.67 ± 0.36[b]	0.65 ± 0.06[b]	-	-	-	-	-
	$F5F_{CN}$	4.34 ± 0.10[c]	0.62 ± 0.02[c]	990 ± 20	0.27 ± 0.08[d]	0.43 ± 0.01	0.22[e]	0.66
	$F5F_{CN}$Y25F	4.54 ± 0.08[b]	0.65 ± 0.01[b]	-	-	-	-	-
	$F5F_{CN}$Y25F	4.32 ± 0.19[f]	0.63 ± 0.03[f]	1680 ± 50	1.2 ± 0.4	0.48± 0.01	0.15 ± 0.02	0.76[g]
GuHCl	WT	4.17 ± 0.07[h]	1.35 ± 0.02[h]	820 ± 40[h]	0.89 ± 0.08[h]	0.86 ± 0.02[h]	0.57 ± 0.02[h]	0.60
	$F5F_{CN}$	4.71 ± 0.06	1.45± 0.02	980 ± 30	0.69 ± 0.05	0.91 ± 0.01	0.53 ± 0.01	0.63

The numbers after the ± sign represent the standard errors to the fit. [a]Values taken from Cho et al. (*13*). [b]Determined by CD monitored equilibrium unfolding. [c]Determined by F_{CN} fluorescence monitored equilibrium unfolding. [d]The value of k_u for the urea jump experiments with $F5F_{CN}$ is believed to be unreliable for the reasons described in the text. [e]Calculated using m_{total} and m_f. [f]Determined from the Chevron plot of stopped-flow experiment. [g]Since the change in fluorescence intensity in $F5F_{CN}$Y25F mutant does not report on folding reaction, β_T value for this mutant is not meaningful. [h]Values taken from Horng et al. (*23*).

3.3.3. Analysis of the Kinetics of NTL9-F5F$_{CN}$Y25F

The kinetic analysis of NTL9-F5F$_{CN}$Y25F mutant is shown in Figure 3.6. The rate of change in the fluorescence signal as the protein folds is 1680 s^{-1}, which is much faster than the folding rate of the wild-type measured by Tyr fluorescence (Table 3.1). The protein does have a Met residue at the first position. This is the only expected quencher of F_{CN} fluorescence in NTL9-F5F$_{CN}$Y25F in these solution conditions. The fluorescence intensity increases as the protein folds. The observed high rate of change in fluorescence as the protein folds might be because of the presence of a non-native hydrophobic cluster in the unfolded state that brings Met1 and F5F$_{CN}$ in close proximity, quenches the F_{CN} fluorescence in the unfolded state. F_{CN} fluorescence gets unquenched and this non-native collapsed state dissociates as the protein gets more ordered. If the reason for the change in F_{CN} fluorescence is Met1 quenching, we cannot argue that the measured decay rate is the true folding rate of the protein since M1 and F5F$_{CN}$ are so close in sequence. It is possible that M1 and F5F$_{CN}$ interact in the unfolded state which causes a change in fluorescence signal when the rest of the protein is still unfolded. On the other hand, the observed decay rate for NTL9-F5F$_{CN}$Y25F unfolding is very similar to the unfolding rate of the wild-type monitored by Tyr fluorescence (Table 3.1). The first β-strand, where Met1 and F5F$_{CN}$ are located, is surrounded by the other secondary structure elements of the protein (Figure 3.1). The first β-strand does not gain flexibility until the whole protein unfolds. As first β-strand becomes more flexible, the probability of Met1 and F5F$_{CN}$ to come closer increases which causes the change in F_{CN} fluorescence by Met quenching during unfolding. The β_T parameter calculated for NTL9-F5F$_{CN}$Y25F is quite different than that of wild-type or NTL9-F5F$_{CN}$ mutant (Table 3.1). Since we hypothesize that the

change in fluorescence intensity in NTL9-F5F$_{CN}$Y25F mutant is due to Met1 quenching, and Met1 and F5F$_{CN}$ are so close in sequence, that the fluorescence signal change may not report on the complete folding reaction. If this is true, then the β_T parameter for this mutant does not report on the position of the transition state.

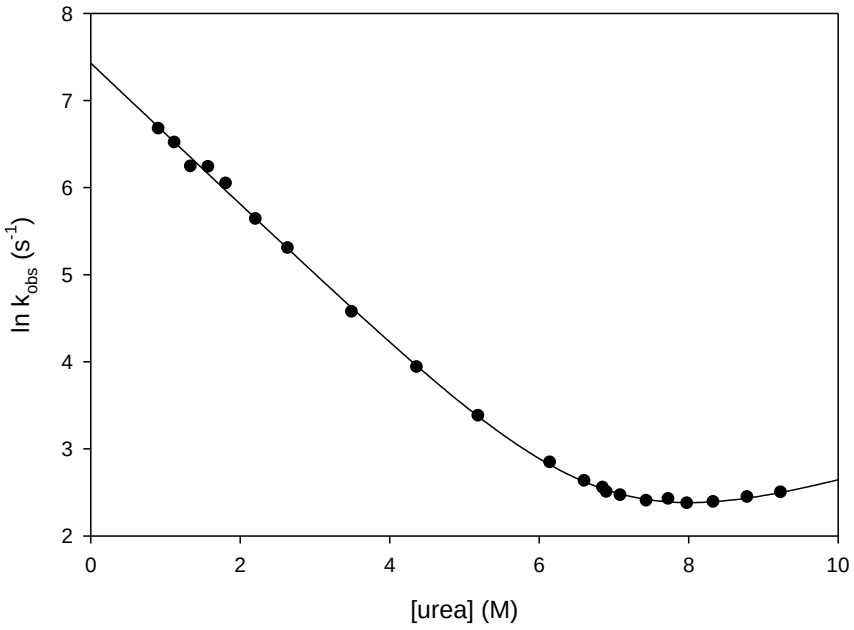


Figure 3.6. Fluorescence detected urea jump stopped flow fluorescence studies of the NTL9-F5F$_{CN}$Y25F.

The NTL9-F5F$_{CN}$ mutant also contains the Met residue at the first position. The reason why Met quenching rate is not observed in F5F$_{CN}$ mutant, but is observed in F5F$_{CN}$Y25F mutant maybe the dead time of the instrument. The dead time is the minimum time after the reactant have mixed that the observation starts. It is 3 millisecond (ms) in the employed stopped-flow instrument. In NTL9-F5F$_{CN}$Y25F, the rate of Met quenching is around 1680 s^{-1} or 0.6 ms. Tyr quenching rate in the F5F$_{CN}$ mutant is 990 s^{-1}

or 1ms. Both of these events occur in the dead time of the instrument. What is observed is the decay curve which is fit to an exponential decay. The fit extrapolates the signal back to time equals to zero. However, it is possible that there is a faster event in the initial 3ms that cannot be observed. This idea is supported by the fluorescence emission intensities of both mutants in the unfolded state which are similar (Figure 3.3 and 3.4). It is possible that Met unquenching in NTL9-F5F_{CN} mutant is the faster event which is over in the dead time, so it is not observed in the stopped-flow kinetics. Tyr quenching follows Met quenching which also starts in the dead time of the instrument, and the observed signal is the decay from Tyr quenching. However, the kinetics of the NTL9-F5F_{CN}Y25F mutant might only report on the decay from Met unquenching which might explain why a faster rate is observed.

3.3.4. Burst Phase Analysis of NTL9-F5F_{CN}

Chevron plots are based upon the analysis of rates. However, the amplitudes of the stopped flow traces also contain important information. The initial amplitudes, calculated by extrapolating the traces backward to time equal zero, can be used in a so-called "burst phase" analysis. For two-state folders the initial amplitudes of refolding studies should fall on the extrapolated baseline for the final intensity at high denaturant concentration. Stated differently, the initial amplitude in a refolding experiment reports on the fluorescence of an unfolded protein under native conditions. If the fluorescence of the unfolded state is a linear function of denaturant concentration then one can extrapolate backwards to low denaturant concentration to calculate the expected amplitude for two-state folding. Unfortunately it is very difficult to conduct a burst phase

analysis of NTL9-F5F_{CN}. Guanidine HCl quenches the fluorescence of F_{CN}, thus the initial amplitudes are affected by factors other than folding and the magnitude of this effect varies as a function of guanidine HCl concentration, note this does not affect the measured rate constants. This is not a problem with urea induced unfolding. However the stability and size of NTL9 conspire to prevent the generation of a complete chevron plot. The incomplete unfolding branch prevents a reliable burst phase analysis. The initial and final amplitudes of the urea jump stopped flow studies are shown in Figure 3.7. It is not possible to accurately define the unfolded baseline for the reasons outlined above but the initial amplitudes are a linear function of denaturant concentration. A plot of the initial amplitudes shows no hint of sigmoidal curvature as would be expected for a burst phase intermediate.

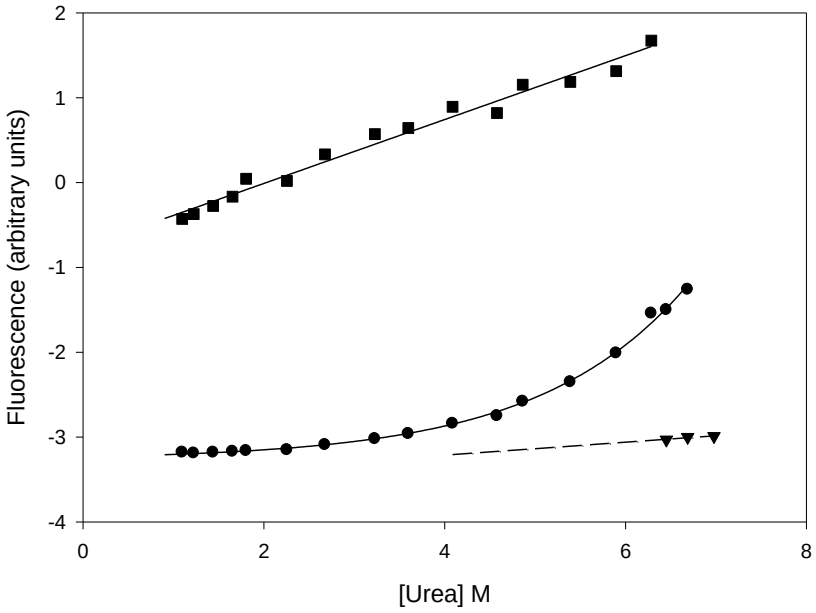


Figure 3.7. Burst-phase analysis of NTL9-F5F_{CN}. Initial (■, ▼) and final amplitudes (●) from the urea jump stopped flow experiments plotted vs. [urea], (■) denotes refolding studies, (▼) denotes unfolding studies. Experimental conditions were as described in the caption to Figure-3. The initial refolding amplitudes show no deviation from linearity, consistent with the lack of a burst phase intermediate. The solid line is drawn through the

initial amplitudes and has no theoretical significance. The dashed line represents extrapolation of the folded baseline.

This study demonstrates the utility of F_{CN} as a fluorescence probe for protein folding studies. The relatively conservative nature of the cyano substitution should allow the replacement of Phe or Tyr residues. In this study, introduction of the non-coded aminoacid at a completely buried position did not alter the structure or stability of the protein.

The use of non-coded amino acids in protein folding studies has been limited to date (*14*, *23-26*). We expect this to change rapidly given recent technical advances. Advances in peptide synthesis have made the assembly of small single domain proteins relatively straightforward (*27*). Larger proteins can be produced using native ligation and expressed protein ligation methodologies (*28*, *29*). In addition the 21st pair methodology developed in Schultz Lab, utilized in this study, is making recombinant production of F_{CN} containing proteins possible (*30*). Thus, incorporation of this useful analog is relatively straightforward and should become even easier as technology continues to mature.

In the present case, use of F_{CN} allowed investigation of the folding reaction of NTL9. All previous studies relied upon the fluorescence of the single tyrosine at position 25. This residue is not part of the hydrophobic core and is partially exposed to solvent in the native state. Thus it does not directly report on the development of the hydrophobic core. In contrast, the kinetic refolding and unfolding experiments conducted with NTL9-F5F_{CN} report folding of the domain, as the polypeptide chain becomes compact and fold F5F_{CN} and Tyr25 come closer leading to quenching of F_{CN} fluorescence. The probe is useful, however, one needs to be careful about interpreting the signal change. The data

presented here represents how the signal change might be due to folding reaction as in the case of NTL9-F5F_{CN}, but it might also be due to the flexibility of the polypeptide chain, as in the case of NTL9-F5F_{CN}Y25F. The data presented here is consistent with cooperative two-state folding of the split β-α-β motif of NTL9.

3.4. References

1. Levinthal, C. (1968) Are there pathways for protein folding?, *J. Chim. Phys. Phys-Chim. Biol. 65*, 44-45.
2. Anfinsen, C. B. (1973) Principles that govern folding of protein chains, *Science 181*, 223-230.
3. Fersht, A. R. (2000) *Structure and Mechanism in Protein Science - A Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman & Company, New York.
4. Onuchic, J. N., Nymeyer, H., Garcia, A. E., Chahine, J., and Socci, N. D. (2000) The energy landscape theory of protein folding: insights into folding mechanisms and scenarios, *Advances in Protein Chemistry, Vol 53 53*, 87-152.
5. Dobson, C. M. (2003) Protein folding and misfolding, *Nature 426*, 884-890.
6. Selkoe, D. J. (2003) Folding proteins in fatal ways, *Nature 426*, 900-904.
7. Rose, G. D., Fleming, P. J., Banavar, J. R., and Maritan, A. (2006) A backbone-based theory of protein folding, *Proc. Nat. Ac. Sci. U.S.A. 103*, 16623-16633.
8. Tucker, M. J., Oyola, R., and Gai, F. (2006) A novel fluorescent probe for protein binding and folding studies: p-cyano-phenylalanine, *Biopolymers 83*, 571-576.
9. Efimov, A. V. (1994) Common structural motifs in small proteins and domains, *Febs Lett. 355*, 213-219.
10. Hoffman, D. W., Cameron, C. S., Davies, C., White, S. W., and Ramakrishnan, V. (1996) Ribosomal protein L9: a structure determination by the combined use of X-ray crystallography and NMR spectroscopy, *J. Mol. Biol. 264*, 1058-1071.

11. Kuhlman, B., Boice, J. A., Fairman, R., and Raleigh, D. P. (1998) Structure and stability of the N-terminal domain of the ribosomal protein L9: evidence for rapid two-state folding, *Biochemistry 37*, 1025-1032.

12. Kuhlman, B., Luisi, D. L., Evans, P. A., and Raleigh, D. P. (1998) Global analysis of the effects of temperature and denaturant on the folding and unfolding kinetics of the N-terminal domain of the protein L9, *J. Mol. Biol. 284*, 1661-1670.

13. Cho, J. H., Sato, S., and Raleigh, D. P. (2004) Thermodynamics and kinetics of non-native interactions in protein folding: a single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state, *J. Mol. Biol. 338*, 827-837.

14. Anil, B., Sato, S., Cho, J. H., and Raleigh, D. P. (2005) Fine structure analysis of a protein folding transition state; distinguishing between hydrophobic stabilization and specific packing, *J. Mol. Biol. 354*, 693-705.

15. Cho, J. H., and Raleigh, D. P. (2005) Mutational analysis demonstrates that specific electrostatic interactions can play a key role in the denatured state ensemble of proteins, *J. Mol. Biol. 353*, 174-185.

16. Luisi, D. L., and Raleigh, D. P. (2000) pH-dependent interactions and the stability and folding kinetics of the n-terminal domain of L9. Electrostatic interactions are only weakly formed in the transition state for folding, *J. Mol. Biol. 299*, 1091-1100.

17. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves, *Met. Enzymol. 131*, 266-280.

18. Santoro, M. M., and Bolen, D. W. (1988) Unfolding free-energy changes determined by the linear extrapolation method .1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants, *Biochemistry 27*, 8063-8068.

19. Vugmeyster, L., Kuhlman, B., and Raleigh, D. P. (1998) Amide proton exchange measurements as a probe of the stability and dynamics of the N-terminal domain of the ribosomal protein L9: comparison with the intact protein, *Protein Sci. 7*, 1994-1997.

20. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding, *Protein Sci. 4*, 2138-2148.

21. Taskent-Sezgin, H., Chung, J., Patsalo, V., Miyake-Stoner, S. J., Miller, A. M., Brewer, S. H., Mehl, R. A., Green, D. F., Raleigh, D. P., and Carrico, I. (2009) Interpretation of p-cyanophenylalanine fluorescence in proteins in terms of solvent exposure and contribution of side-chain quenchers: A combined fluorescence, IR and molecular dynamics study, *Biochemistry 48*, 9040-9046.

22. Tanford, C. (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation, *Advan. Protein Chem. 24*, 1-95.

23. Horng, J. C., and Raleigh, D. P. (2003) Phi-Values beyond the ribosomally encoded amino acids: kinetic and thermodynamic consequences of incorporating trifluoromethyl amino acids in a globular protein, *J. Am. Chem. Soc. 125*, 9286-9287.

24. Anil, B., Song, B. B., Tang, Y. F., and Raleigh, D. P. (2004) Exploiting the right side of the ramachandran plot: substitution of glycines by D-alanine can significantly increase protein stability, *J. Am. Chem. Soc. 126*, 13194-13195.

25. Bann, J. G., and Frieden, C. (2004) Folding and domain-domain interactions of the chaperone PapD measured by F-19 NMR, *Biochemistry 43*, 13775-13786.

26. Muralidharan, V., Cho, J. H., Trester-Zedlitz, M., Kowalik, L., Chait, B. T., Raleigh, D. P., and Muir, T. W. (2004) Domain-specific incorporation of noninvasive optical probes into recombinant proteins, *J. Am. Chem. Soc. 126*, 14004-14012.

27. Schnolzer, M., Alewood, P., Jones, A., Alewood, D., and Kent, S. B. H. (1992) In situ neutralization in Boc-chemistry solid-phase peptide-synthesis - Rapid, high-yield assembly of difficult sequences, *Int. J. Pep. Prot. Res. 40*, 180-193.

28. Dawson, P. E., and Kent, S. B. H. (2000) Synthesis of native proteins by chemical ligation, *Annu. Rev. Biochem. 69*, 923-960.

29. Muralidharan, V., and Muir, T. W. (2006) Protein ligation: an enabling technology for the biophysical analysis of proteins, *Nat. Methods* 3, 429-438.

30. Schultz, K. C., Supekova, L., Ryu, Y. H., Xie, J. M., Perera, R., and Schultz, P. G. (2006) A genetically encoded infrared probe, J. Am. Chem. Soc. 128, 13984-13985.

4. Modulation of p-Cyanophenylalanine Fluorescence by Amino Acid Side-chains and Rational Design of Fluorescence Probes of α-Helix Formation

Abstract

p-Cyanophenylalanine is an extremely useful fluorescence probe of protein structure which can be recombinantly and chemically incorporated into proteins. The probe has been used to study protein folding, protein-membrane interactions, protein-peptide interactions and amyloid formation, however the factors that control its fluorescence are not fully understood. Hydrogen bonding to the cyano group is known to play a major role in modulating the fluorescence quantum yield, but the role of potential side-chain quenchers has not yet been elucidated. A systematic study on the effects of different side-chains on *p*-cyanophenylalanine fluorescence is reported. Tyr is found to have the largest effect followed by deprotonated His, Met, Cys, protonated His, Asn, Arg, and protonated Lys. Deprotonated amino groups are much more effective fluorescence quenchers than protonated amino groups. Free neutral imidazole and hydroxide ion are also effective quenchers of *p*-cyanophenylalanine fluorescence with Stern-Volmer constants of 39.8 M^{-1} and 22.1 M^{-1}, respectively. The quenching of *p*-cyanophenylalanine fluorescence by specific side-chains is exploited to develop specific, high sensitivity, fluorescence probes of helix formation. The approach is demonstrated with Ala based

peptides that contain a *p*-cyanophenylalanine-His or a *p*-cyanophenylalanine-Tyr pair located at positions i and i+4. The *p*-cyanophenylalanine-His pair is most useful when the His side-chain is deprotonated and is, thus, complimentary to Trp-His pair which is most sensitive when the His side-chain is protonated.

Note: The material presented in this chapter has been published (Humeyra Taskent-Sezgin, Peter Marek, Rosanne Thomas, Daniel Goldberg, Juah Chung, Isaac Carrico, and Daniel P. Raleigh. *Biochemistry* **2010,** *49,* 6290-6295). This chapter contains direct excerpts from the manuscript that was written by me with suggestions and revisions by Professor Raleigh. The work on the design and testing of the probes of helix formation was conducted by Ms. R. Thomas, Mr. D. Goldberg under the supervision of Mr. P. Marek. It is included here for completeness.

4.1. Introduction

Fluorescence spectroscopy is widely used to study proteins and peptides, with the vast majority of studies making use of the naturally occurring fluorophores, Tyr and Trp (*1, 2*). However, there is considerable interest in the incorporation of novel fluorophores, because they can offer site specific probes and, in many cases, they can be chosen to allow selective excitation and detection (*3-7*). Recently *p*-cyanophenylalanine (F_{CN}) fluorescence has been developed as a robust probe of protein folding, protein-peptide interactions, protein-membrane interactions, ligand binding and amyloid formation (Figure 4.1) (*8-15*).

The utility of *p*-cyanophenylalanine is due to three features. First, it can be incorporated into proteins recombinantly, using the so called 21st pair methodology, or by solid-phase peptide synthesis (*11, 16*). Secondly, it represents a relatively conservative substitution for the aromatic amino acids since it is much more similar in shape and size than many other fluorophores and the polarity of the cyano group is between that of a methylene and an amide group. This intermediate polarity allows F_{CN} to be accepted in both hydrophobic and hydrophilic environments in a protein. Third, the photophysical properties of F_{CN} nicely complement existing fluorophores. It can be selectively excited in the presence of Trp and Tyr and it forms a useful resonance energy transfer (RET) probe with both Tyr and Trp (*7, 14, 17, 18*). However, the factors that control its quantum yield are not completely understood. It is known that the fluorescence is high when the cyano group is hydrogen bonded and fluorescence can be quenched by RET to Tyr or Trp, but unfortunately the effect of other amino acid side-chains is not known (*14, 19*). A detailed understanding of the factors that control F_{CN} fluorescence is required to fully

exploit this very promising probe and to avoid misinterpretation. A striking example is provided by our recent application of F_{CN} to study the folding of NTL9, a small globular protein. The fluorescence of F_{CN} in NTL9 was very low in the folded state suggesting that the cyano group was sequestered from solvent, however IR measurements showed that it was exposed, and further investigation revealed that the F_{CN} fluorescence was quenched by interactions with a Tyr side-chain in the native state and showed that the cyano group was, in fact, exposed to solvent (*14*).

Here I systematically examine the ability of other amino acids and the termini of polypeptides to modulate F_{CN} fluorescence by examining the fluorescence of F_{CN} in a set of peptides of general sequence $GGF_{CN}XA$ where X represents Ala, Cys, His, Lys, Met, Asn, Arg, and Tyr (Figure 4.1). Based on these results we demonstrate how such interactions can be used as a highly sensitive probe of helix formation. We also show that free imidazole and hydroxide ion are effective quenchers of F_{CN} fluorescence.

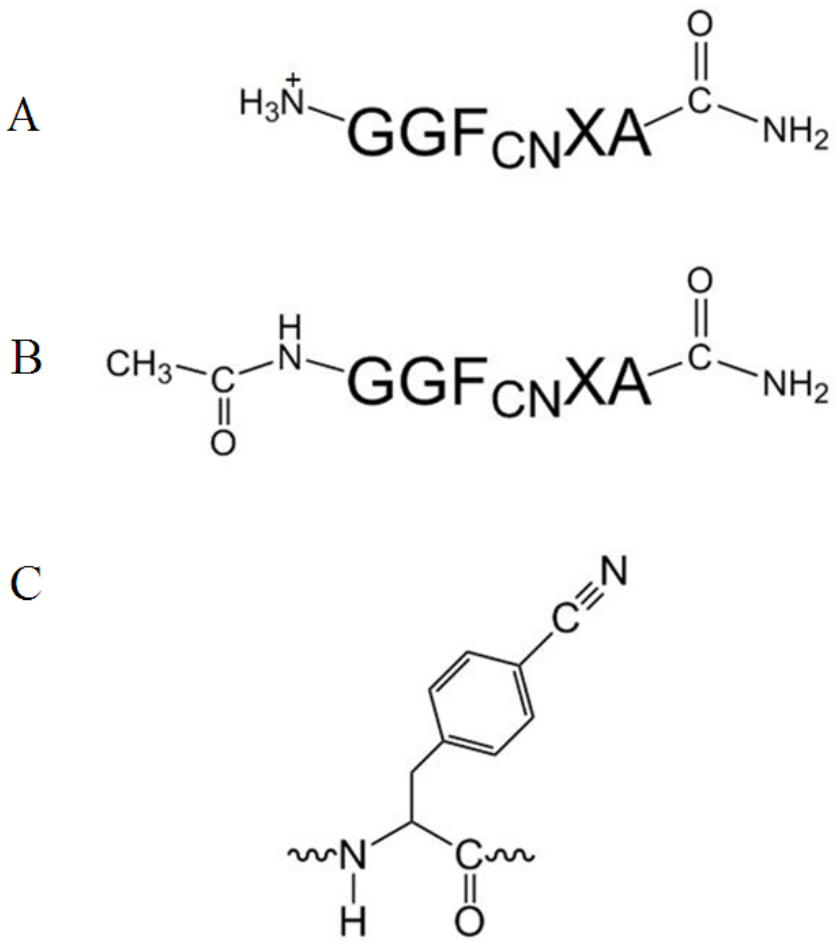


Figure 4.1. General sequence of the peptides studied. (A) Several variants have an amidated C-terminus and a free N-terminus. (B) An additional three peptides $GGF_{CN}AA$, $GGF_{CN}HA$ and $GGF_{CN}KA$ were synthesized with an acetylated N-terminus. (C) The structure of the *p*-cyanophenylalanine side-chain.

4.2. Materials and Methods

4.2.1. Peptide Synthesis and Purification

A set of peptides with a general sequence $GGF_{CN}XA$ (X = A, C, H, K, M, N, R, or Y) were synthesized using standard Fmoc solid-phase methods on an Applied Biosystems 433A peptide synthesizer. PAL-PEG-PS resin was used which leads to an amidated C-

terminus. Unless noted, the peptides have a free N-terminus. Additional $GGF_{CN}XA$ peptides (X = A, H or K) were synthesized with an acetylated N-terminus and an amidated C-terminus. Two 21 residue-long Ala based helical peptides were prepared with the sequence D-P-A-A-K-A-A-A-K-A-A-**X**-A-A-A-$\mathbf{F_{CN}}$-A-A-A-K-A where X is His or Tyr. These peptides were amidated at their C termini and acetylated at their N termini. The peptides were cleaved from the resin using 91% (v/v) trifluoroacetic acid (TFA), 3% (v/v), anisole, 3% (v/v) thioanisole, 3% (v/v) 1,2-ethanedithiol, precipitated using cold ethyl ether, and scavengers were removed under vacuum. The crude peptides were purified with reverse-phase HPLC using a Vydac C18 semi-preparative column. An A-B gradient system was used, with buffer A composed of a 0.1% (v/v) solution of TFA and buffer B composed of 90% (v/v) acetonitrile, 9.9% (v/v) water and 0.1% (v/v) TFA. The gradient was 0-90% B in 90 min. The identities of the purified peptides were confirmed by MALDI-TOF mass spectroscopy. The expected and observed molecular weights for the $GGF_{CN}XA$ peptides are; $GGF_{CN}AA$ 446.5 Da and 446.8 Da; $GGF_{CN}YA$ 538.6 Da and 538.6 Da; $GGF_{CN}MA$ 506.6 Da and 507.8 Da; $GGF_{CN}CA$ 478.5 Da and 478.5 Da; $GGF_{CN}HA$ 512.5 Da and 512.8 Da; $GGF_{CN}KA$ 503.6 Da and 504.0 Da; $GGF_{CN}NA$ 489.5 Da and 489.8 Da; and $GGF_{CN}RA$ 531.5 Da and 531.3 Da, respectively. The expected and observed molecular weights for the Ac-$GGF_{CN}XA$ peptides are; Ac-$GGF_{CN}AA$ 488.5 Da and 488.1 Da; Ac-$GGF_{CN}HA$ 554.5 Da and 554.2 Da, and Ac-$GGF_{CN}KA$ 545.6 Da and 545.3 Da, respectively. The observed masses of the Ala based peptides were 1961.0 Da for the H-F_{CN} peptide and 1987.3 Da for the Y-F_{CN} peptide. The expected masses were 1961.1 Da for the H-F_{CN} peptide and 1987.2 Da for the Y-F_{CN} peptide.

4.2.2. Fluorescence Spectroscopy

Fluorescence emission spectra of the peptides were collected using an Applied Photon Technologies fluorimeter. The peptide concentration was determined by measuring the absorbance of the sample at 280 nm. The extinction coefficient for F_{CN} is 850 $M^{-1}cm^{-1}$ and for Tyr is 1490 $M^{-1}cm^{-1}$ (*9*). Fluorescence was excited at 240 nm and the emission signal was recorded from 250-350 nm. For the pH dependent fluorescence experiments, the excitation wavelength for the F_{CN} group was 240 nm, and the emission signal was followed at 291 nm. The peptide concentration for both the fluorescence emission spectra and pH dependent fluorescence experiments was 20 uM.

Emission spectra of the Ala based peptides were collected in buffer and 8 M urea over the range of 250 to 350 nm with excitation at 240 nm. The peptide concentration was 25 uM. The Tyr-F_{CN} Ala rich peptide was prepared at pH 5.5 in 10 mM sodium acetate. Fluorescence emission spectra of the His-F_{CN} peptide were taken as a function of pH. The His-F_{CN} Ala rich peptide was examined in 10 mM sodium acetate at pH 5.5 and 10 mM Tris at pH 8.3.

4.2.3. Förster Distance Calculation

The Förster distance (R_o) of the F_{CN}-Tyr pair was calculated according to the following equation:

$$R_o^6 = \frac{9000(\ln 10)\kappa^2\theta_D}{128\pi^5\eta^4 N}\int f_D(\lambda)\varepsilon_A(\lambda)\lambda^4 d\lambda \qquad (4.1)$$

where κ^2 is the dipole orientation factor, taken as 2/3 which is the value for isotropic averaging, θ_D is the quantum yield of the donor, which is 0.11 (*9*), η is the refractive index of the medium which is 1.33 for water, N is Avagadro's number, and the integral is the spectral overlap integral of the peak-normalized donor emission (Figure 4.2A) and molar absorption coefficient of the acceptor (Figure 4.2B). The calculated R_o for the F_{CN}-Tyr pair was 12 Å.

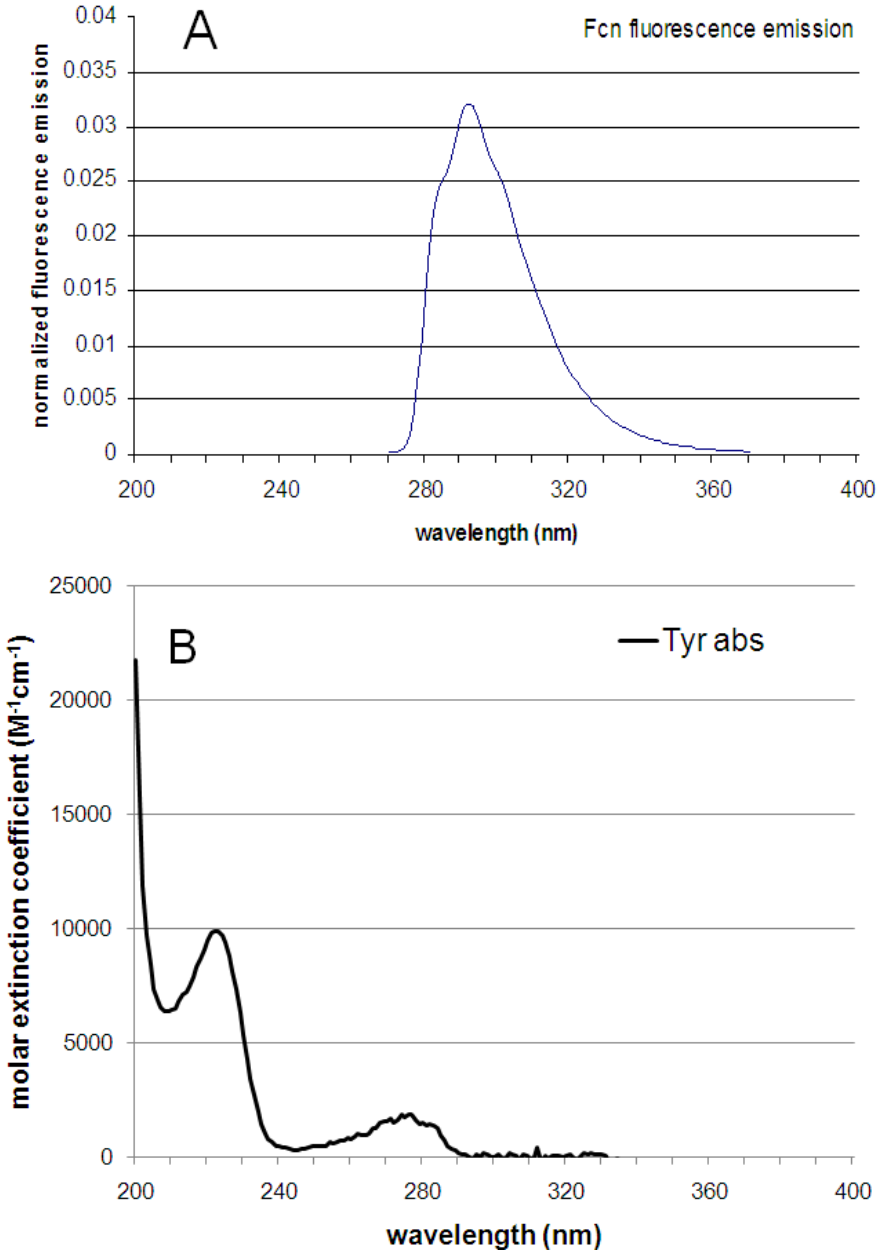


Figure 4.2. (A) Fluorescence emission spectrum of donor, F_{CN} amino acid following excitation at 240 nm in water in the absence of acceptor. The spectrum was normalized so the integrated intensity is 1.00. (B) Normalized absorption spectrum of acceptor, Tyr in water.

4.2.4. Circular Dichroism

Circular dichroism spectra of the 21-residue Ala based peptides were collected using a Chirascan Applied Photophysics CD spectrophotometer. The peptides were in 10mM sodium acetate at pH 5.5 or 10 mM tris at pH 8.3 for studies of the folded form. The unfolded peptides were in 8 M urea and either 10 mM sodium acetate at pH 5.5 or 10 mM Tris at pH 8.3. The urea concentration was determined using refractive index measurements. Wavelength scans were recorded at 25 ºC from 190-260 nm for the folded peptide and from 210-260 nm for the urea unfolded peptide.

4.3. Results and Discussion

4.3.1. Effect of Amino Acid Side-chains to F_{CN} Fluorescence

A set of peptides with the general sequence GGF_{CN} XA (X= A, C, H, K, M, N, R, or Y) were synthesized to probe the ability of different side-chains to modulate F_{CN} fluorescence. These side-chains include all of the potential quenching groups in proteins with the exception of Trp which has already been analyzed. Ser was not examined because it is known that hydrogen bonding to the cyano group leads to high fluorescence. The potential effect of protonated and deprotonated carboxylic acid were examined by a Stern-Volmer analysis of the quenching by acetic acid and sodium acetate. All of the peptides have an amidated C-terminus. Two GGF_{CN}AA peptides were prepared, one with a free N-terminus and the other with an acetylated N-terminus (Figure 4.1). The structure of F_{CN} is shown in Figure 4.1. The GGF_{CN}AA peptide sequence is the parent (reference) peptide, and was chosen as such since the Ala side-chain will not quench the fluorescence.

In order to determine the potential effects of the N-terminus we examined the fluorescence of the $GGF_{CN}AA$ variant with a free N-terminus as a function of pH (Figure 4.3). The fluorescence intensity titrates with a pKa of 9.0, consistent with that of the amino group, decreasing as the amino group is deprotonated. This observation is consistent with reports on benzonitrile fluorescence quenching with diethylamine (*20*). As a control, a parent peptide with an acetylated N-terminus was also examined and showed no change in fluorescence up to pH 12. The decrease at pH 12 is reproducible and is due to quenching by hydroxide ion (Figure 4.4).

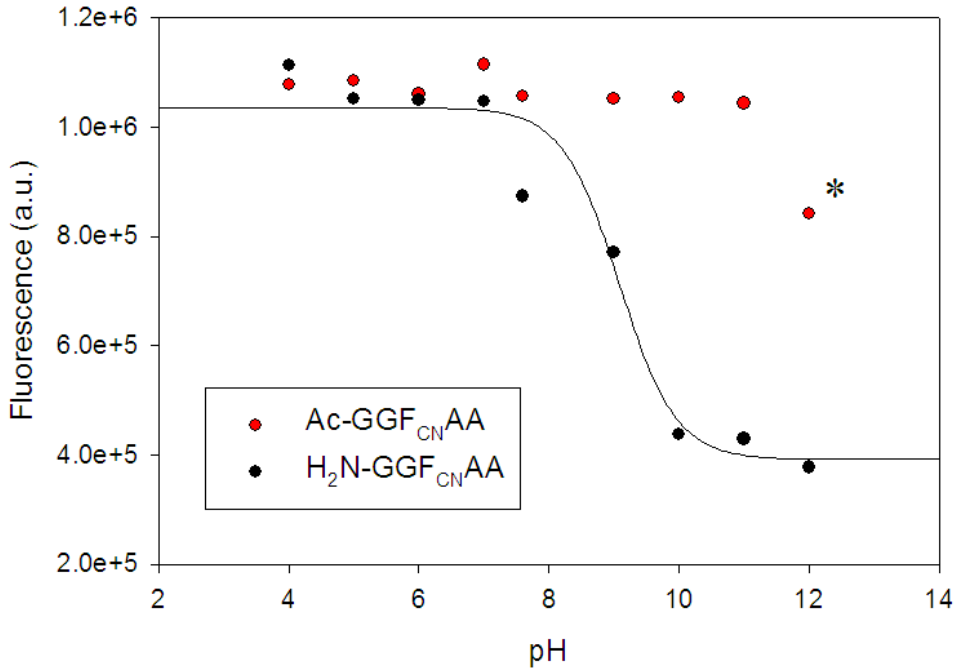


Figure 4.3. Comparison of the pH dependence of the fluorescence emission intensity for $GGF_{CN}AA$ peptide with a free N-terminus (black) and with an acetylated N-terminus (Ac-$GGF_{CN}AA$, red). The solid black line is a fit of the data to the Henderson-Hasselbalch equation with an apparent pKa of 9.0. The decrease in intensity for the Ac-$GGF_{CN}AA$ peptide above pH 11 is due to quenching by hydroxide ion (labeled with a star). Fluorescence was excited at 240 nm and the emission was collected at 291 nm. The peptide concentration was 20 uM. The experiments were conducted in water at the indicated pH values.

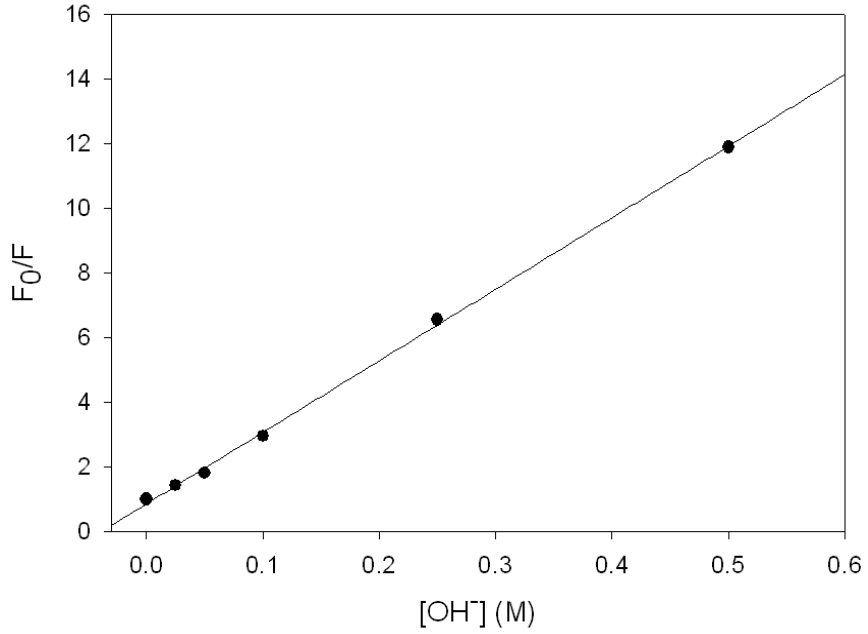


Figure 4.4. Stern-Volmer plot of the quenching of F_{CN} fluorescence of the acetylated $GGF_{CN}AA$ (Ac-$GGF_{CN}AA$) peptide by hydroxide ion. The Stern-Volmer constant is 22.1 M^{-1}.

The ability of individual side-chains to modulate F_{CN} fluorescence was examined in the context of the $GGF_{CN}AA$ peptides at pH 5. The peptides have a free N-terminus, but the pH dependent studies of the parent peptide with the acetylated and unblocked N-terminus demonstrate that a protonated amino group has no effect, thus alleviating any potential concerns about the interpretation of the spectra, since quenching is dominated by the side-chain. Figure 4.5 displays fluorescence emission spectra of various peptides. Each panel of the figure includes the spectrum of the parent peptide, $GGF_{CN}AA$ for comparison. At pH 5.0 the rank order of effectiveness in reducing the F_{CN} fluorescence is Y > M > C > N > R = K^+, where K^+ a positively charged Lys side-chain. Arg and positively charged Lys had no effect on F_{CN} fluorescence.

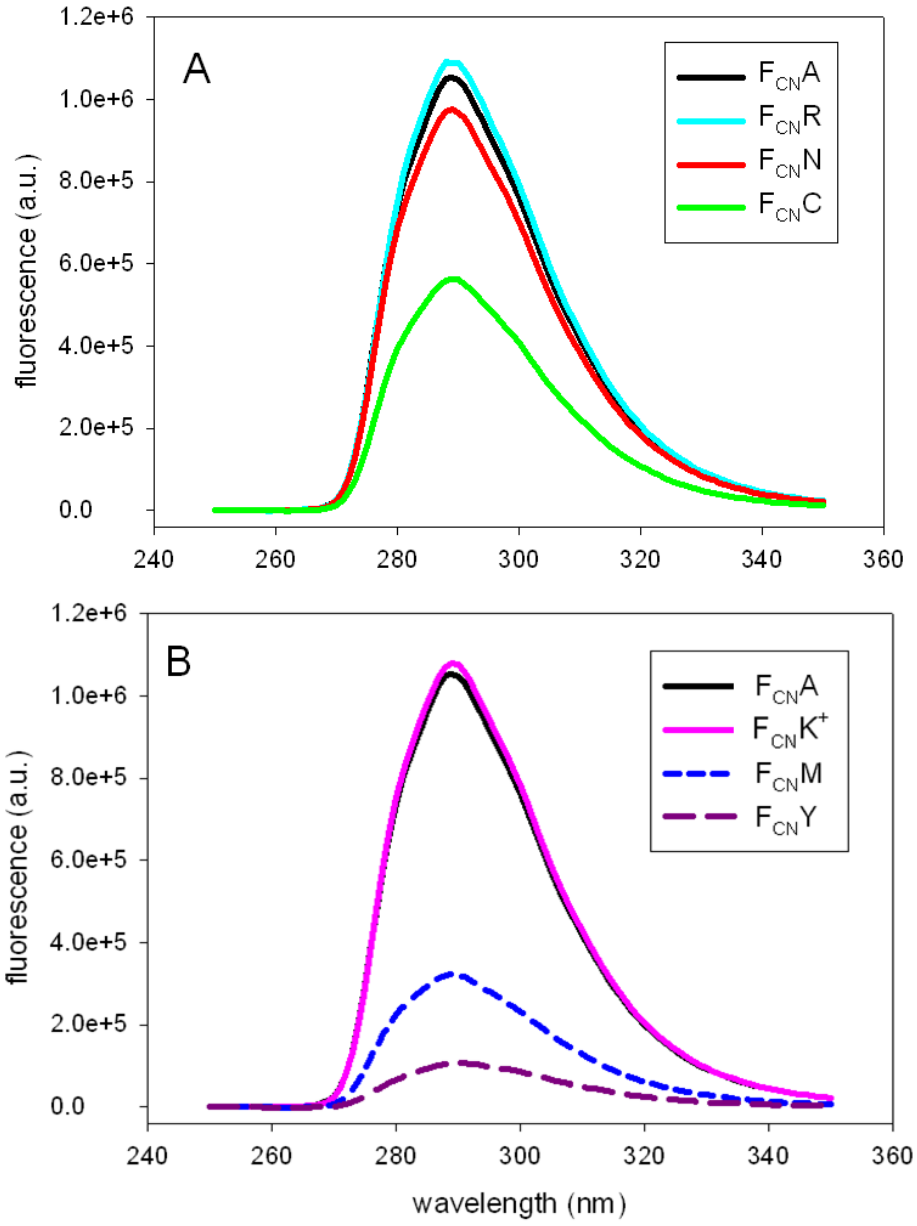


Figure 4.5. Fluorescence emission spectra of $GGF_{CN}XA$ peptides with a free N-terminus at pH 5. X being (A) Ala, Arg, Asn, and Cys; (B) Ala, Met, Lys, and Tyr. Fluorescence was excited at 240 nm. The experiments were conducted in water.

The pH dependent data collected for the $GGF_{CN}AA$ peptide suggests that it is worthwhile to examine the pH dependence of the quenching by His and Lys. In order to examine the effect of the protonation state of the His and Lys side-chains, we recorded fluorescence emission spectra of N-terminally acetylated variants of the $GGF_{CN}AA$, $GGF_{CN}HA$, and $GGF_{CN}KA$ peptides. Acetylation blocks the N-terminus and the resulting amide linkage is not an effective quencher. This allows the effect of the side-chain to be probed without complications from quenching by a deprotonated N-terminal amino group at high pH. The data indicates that a deprotonated His side-chain is a much more effective quencher than a protonated one (Figure 4.6).

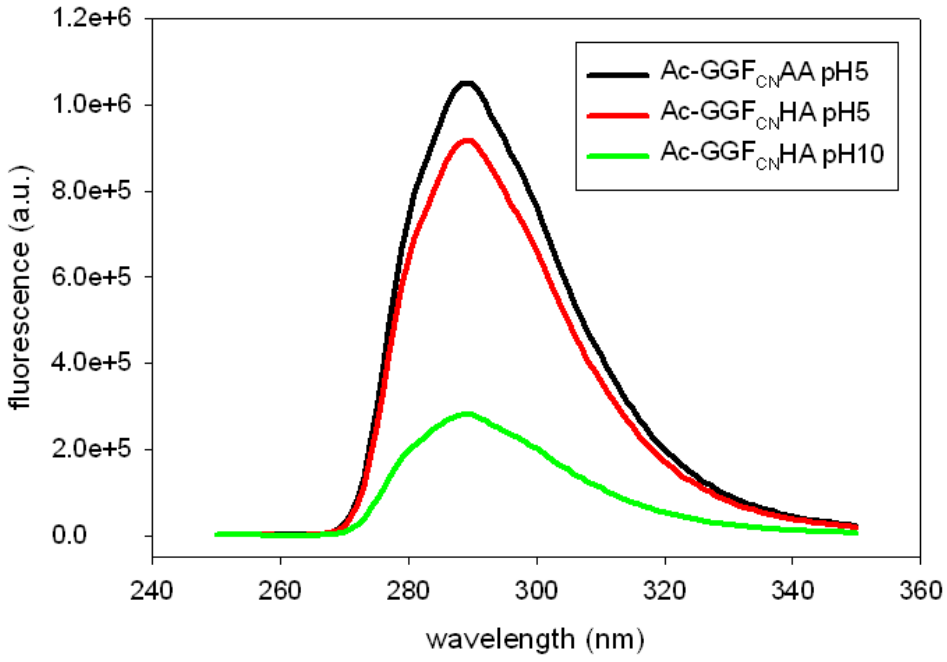


Figure 4.6. Comparison of the effects of a protonated and deprotonated His side-chain on F_{CN} fluorescence measured for the $GGF_{CN}XA$ peptides with an acetylated (blocked) N-terminus. Black, X=Ala at pH 5; red X=His at pH 5; blue, X=His at pH 10. All spectra were recorded in water at the indicated pH. Sample concentration was 20 uM and the excitation was at 240 nm.

Figure 4.7 displays the fluorescence emission of His and Lys peptides with an acetylated N-terminus as a function of pH. In the His peptide, the F_{CN} fluorescence decreases as the His side-chain is deprotonated at higher pH and the intensity vs. pH profile is well fit to an apparent pKa of 7.3. It is interesting to note that protonated His is a much more effective quencher of Trp fluorescence than neutral His, but the situation is reversed here (*21, 22*). Likewise, for the Lys peptide, the F_{CN} fluorescence intensity does not change until the pH approaches the pKa of the Lys side-chain and then decreases. It was not possible to accurately monitor F_{CN} fluorescence of the Lys peptide over the entire pH range, because quenching by hydroxide is significant at pH 12 and above. However, our observation that a deprotonated N-terminus is an effective quencher indicates that a deprotonated Lys side-chain will be a better quencher than a protonated Lys side-chain. Of course, the effect is likely to be of little practical significance in most studies of proteins given that the intrinsic pKa of Lys is above 10. We examined the ability of protonated and deprotonated carboxylic acids to quench F_{CN} fluorescence by conducting Stern-Volmer plots of the quenching by acetic acid and sodium acetate. Both are inefficient quenchers with Stern-Volmer constants of 0.31 M^{-1} and 0.25 M^{-1} (Figure 4.8). These results confirm that caboxylate does not have a significant quenching effect on F_{CN} fluorescence. Thus, Asp, Glu amino acid side-chains and C-terminus of a peptide are not expected to significantly affect F_{CN} fluorescence.

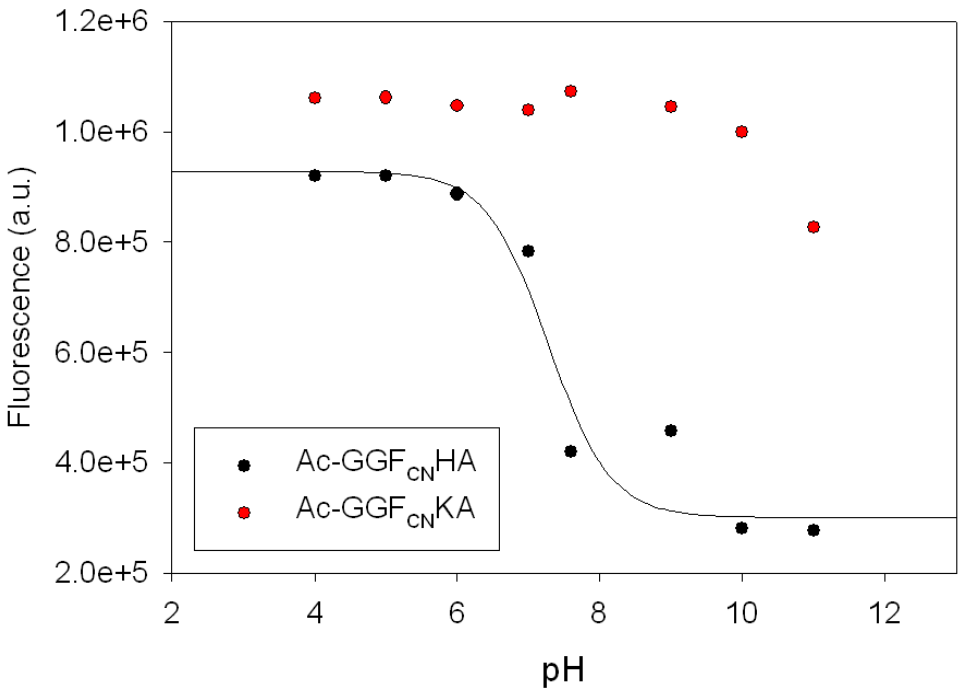


Figure 4.7. Fluorescence emission intensity of the Ac-GGF$_{CN}$HA (black) and Ac-GGF$_{CN}$KA (red) N-teminal acetylated peptides as a function of pH. The F_{CN} fluorescence is quenched as the pH approaches the pKa of His, indicating that deprotonated His is a better quencher of F_{CN} fluorescence than protonated His. A decrease in fluorescence intensity of the Ac-GGF$_{CN}$KA peptide is observed as the pH approaches the pKa of Lys. Data is not reported above pH 11 because quenching by OH^- is significant. The solid curve is a fit of the His data to the Henderson Hasselbalch equation and yields an apparent pKa of 7.3.

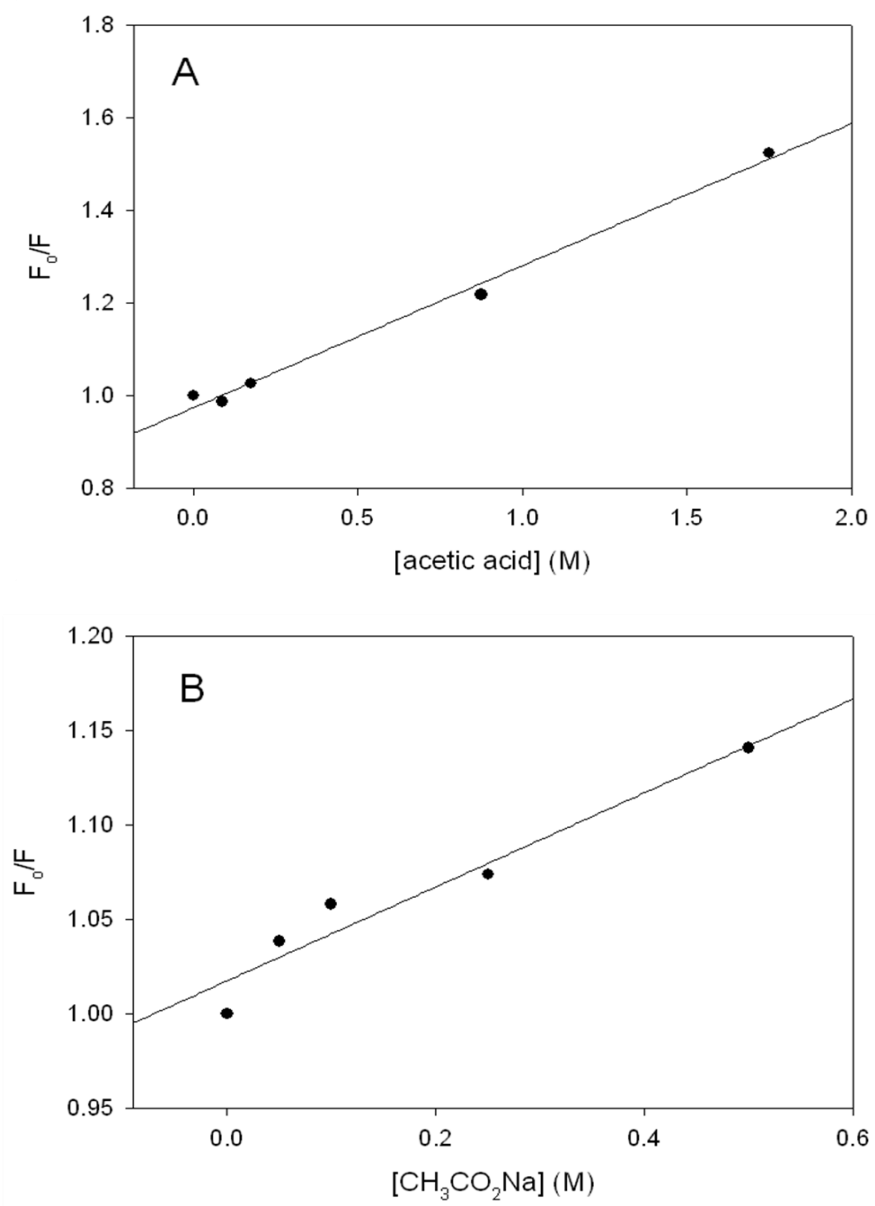


Figure 4.8. Stern-Volmer plot of the quenching of F_{CN} fluorescence of the acetylated $GGF_{CN}AA$ (Ac-$GGF_{CN}AA$) peptide by acetic acid (A) and sodium acetate (B). The Stern-Volmer constants are 0.31 M^{-1} and 0.25 M^{-1}, respectively.

Combining the peptide data with the pH dependent studies shows that the rank order of quenching is Y > H° > M > deprotonated N-terminus > C > H^+ > N > R = K^+ where H° denotes a neutral His side-chain and H^+ denotes a positively charged His side-

chain (Table 4.1). As noted, the quenching by a neutral Lys side-chain could not be measured, but it is expected to be comparable to that of the deprotonated N-terminus.

Side-chain	Intensity ratio relative to control: $GGF_{CN}XA$ / $GGF_{CN}AA$
Tyr	0.10
His (pH:10)	0.27
Met	0.31
Cys	0.54
His^+ (pH 5)	0.85
Asn	0.93
Lys^+ (pH 5)	1.00
Arg	1.00

Table 4.1. Effect of various side-chains on *p*-cyanophenylalanine fluorescence. Measurements were performed in water at pH 5 if not otherwise indicated. Acetylated peptides were used to obtain the data for both of the His peptides.

The results with the His peptide indicate that free imidazole should be a quencher of F_{CN} fluorescence, and suggest that the neutral form should be more efficient than the protonated form. This was confirmed by Stern-Volmer analysis of quenching data for the acetylated $GGF_{CN}AA$ peptide (Figure 4.9). A Stern-Volmer plot of the ratio of the F_{CN} fluorescence in the absence and presence of the quencher, F_0/F vs. quencher concentration is linear. The slope of the curve, the Stern-Volmer constant (K_{SV}), is a measure of the efficiency of a quencher. K_{SV} for imidazole at pH 5 is 18.3 M^{-1} and at pH

9, it is 39.8 M^{-1}, confirming that neutral imidazole is a significantly better quencher of F_{CN} fluorescence than protonated imidazole. For comparison the K_{sv} value for hydroxide ion is 22.1 M^{-1} (Figure 4.4). Chloride ion is the only other reported quencher of F_{CN} fluorescence and its K_{SV} value is 9.3 M^{-1} (*9*, *23*).

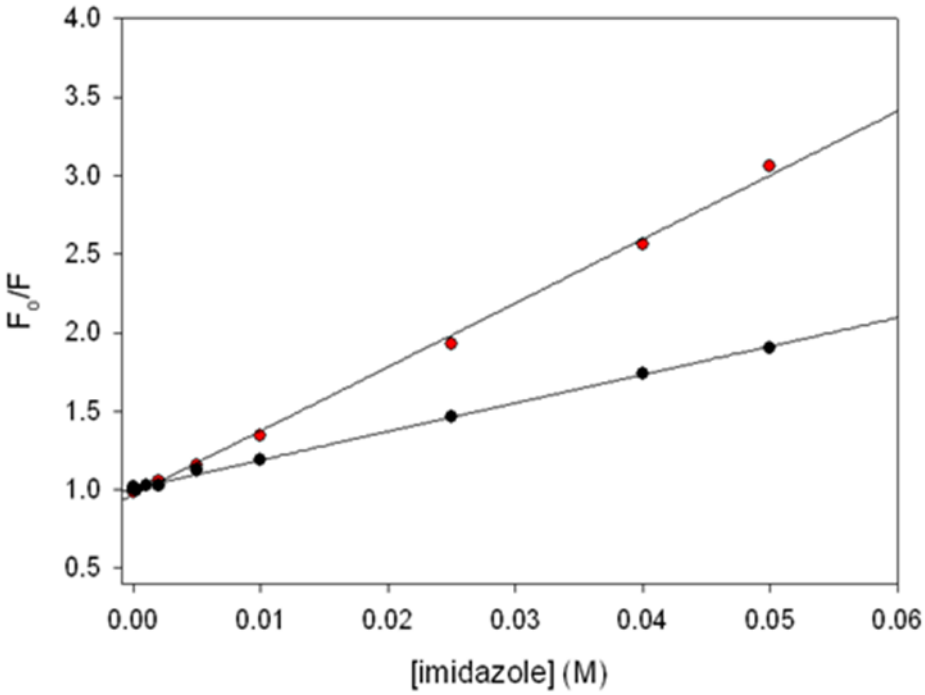


Figure 4.9. Stern-Volmer plots of the quenching of F_{CN} fluorescence of the acetylated GGF_{CN}AA (Ac-GGF_{CN}AA) peptide by imidazole at pH 5 (black) and pH 9 (red). The Stern-Volmer constants are 18.3 M^{-1} at pH 5 and 39.8 M^{-1} at pH 9.

4.3.2. F_{CN} Based Site Specific Probes of Helix Formation

The short range nature of the quenching suggests that suitable X-F_{CN} pairs could be used to probe local structure formation. We tested this hypothesis by examining *de novo* designed helical peptides which contained either a Tyr-F_{CN} pair or a His-F_{CN} pair

located one helical turn apart. Two 21 residue Ala based peptides were designed to test the ability of F_{CN} fluorescence quenching to probe helix formation. The designed peptides each contain an Asp residue at position one since it is an excellent capping residue and a Pro at position two since prolines are favorable at this position in a helix (*24*). Several Lys residues were included to ensure solubility and the termini were capped to avoid unfavorable interactions with the helix dipole. The method will be easiest to interpret quantitatively in terms of helix formation when the F_{CN} group is exposed to solvent in both states since the fluorescence quantum yield is influenced by hydrogen bonding to the cyano group as well as by quenching from other protein side-chains. If the F_{CN} group were buried in one conformation, but not in the other, then the observed intensity change would be a convolution of the effects due to changes in side-chain quencher interactions and the effects of changes in solvation (*18*). However, note that both quenching and side-chain burial reduce F_{CN} fluorescence. Thus, the pair can still be used as a probe of global folding even if the side-chain is buried.

We first examined the use of a Tyr-F_{CN} pair. In the present case, Tyr was incorporated at residue-12 and F_{CN} at residue-16. The distance between the Tyr and F_{CN} residues, when folded into a helical conformation is well within the Förster distance of 12 Å (calculation in materials and methods) and the cyano group is exposed to solvent. The CD spectrum indicates that the Tyr-F_{CN} peptide is partially helical in buffer. The spectrum shows the characteristic helical trend in ellipticity at 208 nm and 222 nm. The mean residue ellipticity is -19,300 deg cm^2 $dmol^{-1}$ at 222 nm, which corresponds to a helical content of 55%, as estimated using the method of Baldwin (Figure 4.10A) (*25*). CD reveals, as expected, that the peptide is unstructured in 8 M urea. The change in the

conformation of the peptide is reflected by significant changes in its fluorescence properties. The fluorescence intensity is high in the urea induced unfolded state, but is significantly reduced in the partially folded state in buffer, owing to helix formation which brings the Tyr-F_{CN} pair into closer proximity (Figure 4.10B). The ratio of the fluorescence of the partially folded form to that of the unfolded form is 0.50. The data demonstrates that a solvent exposed Tyr-F_{CN} pair can be used as a local fluorescence probe of helix formation.

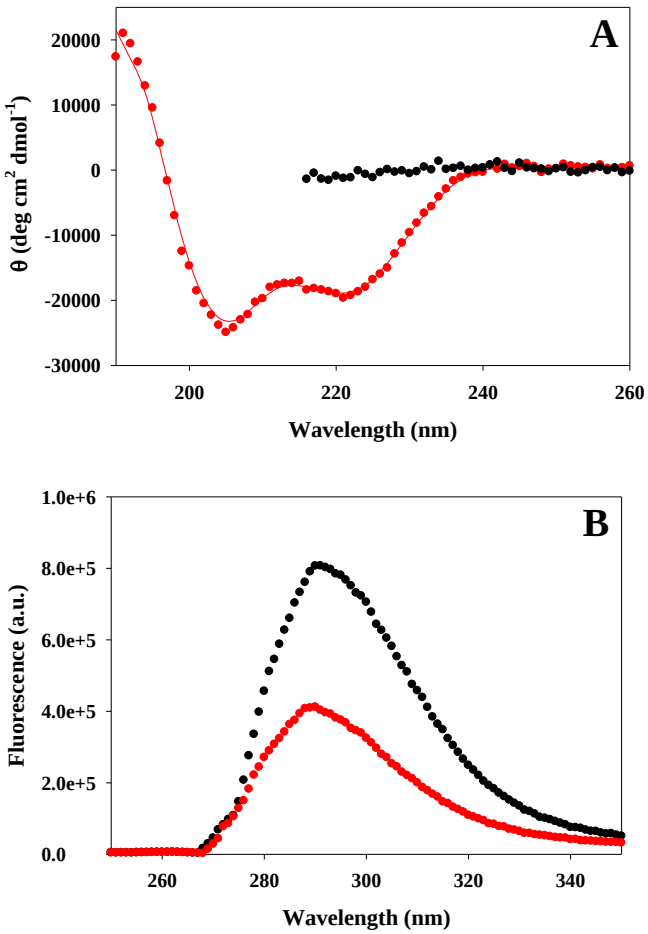
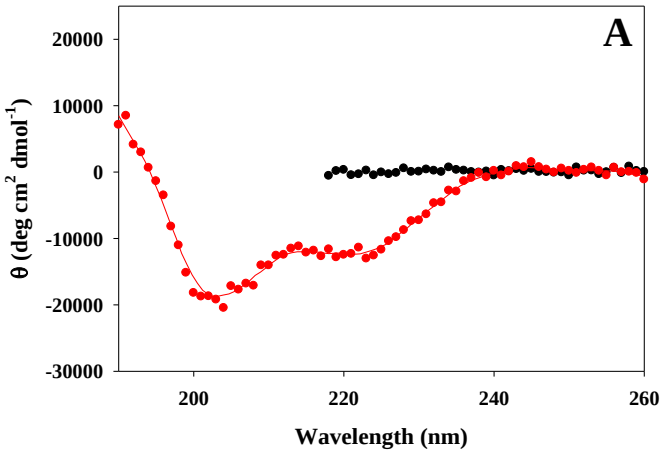
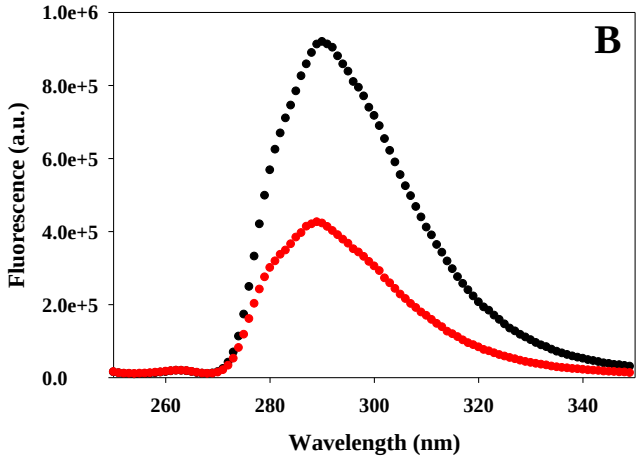


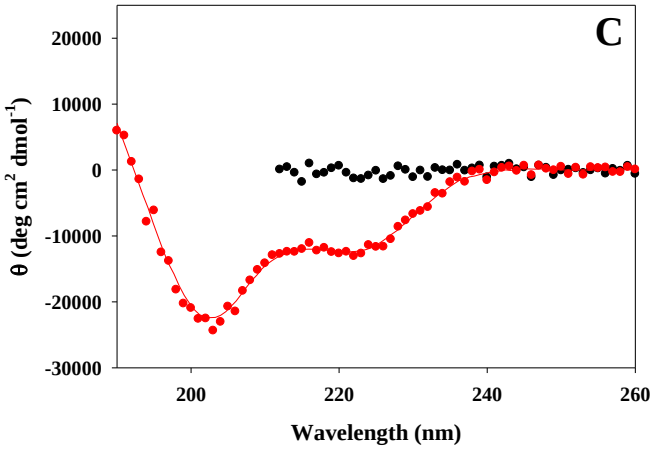
Figure 4.10. (A) CD spectra of the Tyr-F_{CN} peptide indicates that it is partially helical in buffer (red), but is unfolded in 8 M urea (black). Spectra were recorded at 25ºC at pH 5.5. (B) Fluorescence emission spectra confirm that the Tyr-F_{CN} pair provides a probe of helix formation. Fluorescence emission spectra of the Tyr-F_{CN} peptide in the partially helical state (red) and in the urea unfolded state (black) are shown.

The data collected on the GGF_{CN}HA peptide suggests that a solvent exposed His-F_{CN} pair could also be exploited to provide a fluorescence based probe of helix formation. Our analysis of the small peptides indicates that the method should be more sensitive at pH values where the His sidechain is neutral. The quenching of Trp by protonated His has been used to probe protein conformation changes, but is limited to pHs where His is protonated (*26*). Thus, a His-F_{CN} pair should be complimentary since it will be most

effective when the His side-chain is neutral. We examined the fluorescence properties of the His-F_{CN} containing Ala based peptide at pH 5.5, where the His side-chain is predominantly protonated, and at pH 8.3, where it is largely deprotonated. The peptide is partially helical in buffer at both pH 5.5 and pH 8.3 as judged by CD (Figure 4.11). The helical content is less than observed for the Tyr containing peptide, but it is still significant. The shape of the spectra suggests that the His containing peptide is slightly more helical at the higher pH, although the helical content is on the order of 37 to 42% for both pHs. The addition of 8 M molar urea disrupts the helical structure (Figure 4.11). The fluorescence is lower in the partially folded state than in the unfolded state at both pHs, however the change is much larger at pH 8.3, where the ratio of the fluorescence of the partially folded form to that of the unfolded form is 0.46 (Figure 4.11B). At pH 5.5 the ratio is 0.74 (Figure 4.11B). The data confirms that solvent exposed F_{CN}-His pairs can be effectively exploited to probe helical structure and demonstrates that the method is most sensitive when the His side-chain is neutral.


A
20000
10000
0
-10000
-20000
-30000
θ (deg cm^2 dmol^{-1})
200
220
240
260
Wavelength (nm)


B
1.0e+6
8.0e+5
6.0e+5
4.0e+5
2.0e+5
0.0
Fluorescence (a.u.)
260
280
300
320
340
Wavelength (nm)


C
20000
10000
0
-10000
-20000
-30000
θ (deg cm^2 dmol^{-1})
200
220
240
260
Wavelength (nm)

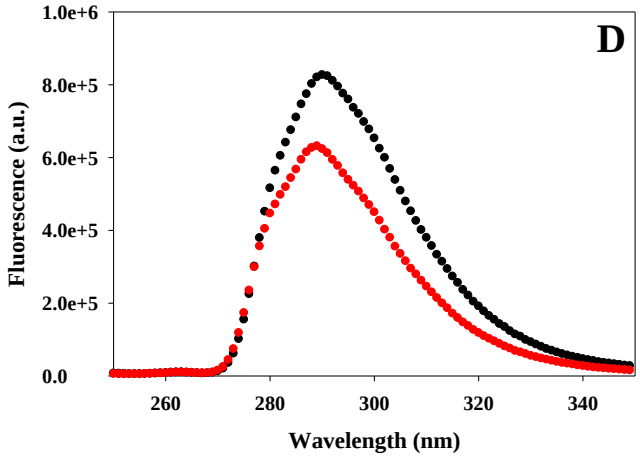


Figure 4.11 (A) CD spectra of the His-F_{CN} peptide at pH 8.3 indicates that it is partially helical in buffer (red), but is unfolded in 8 M urea (black). The red curve in the CD spectra is included as an aid to visualize. (B) Fluorescence emission spectra confirm that the His-F_{CN} pair provides a sensitive probe of helix formation when the His side-chain is neutral. Fluorescence emission spectra of the His-F_{CN} peptide in the partially helical state (red) and in the urea-unfolded state (black) are shown at pH 8.3. (C) CD spectra of the His-F_{CN} peptide at pH 5.5 indicates that it is partially helical in buffer (red), but is unfolded in 8 M urea (black). (D) Fluorescence emission spectra confirm that the His-F_{CN} pair is a less sensitive probe of helix formation when the His side-chain is charged. Fluorescence emission spectra of the His-F_{CN} peptide in the partially helical state (red) and in the urea unfolded state (black) are shown at pH 5.5.

The data presented here provides a comprehensive catalog of the effect of protein side-chains on F_{CN} fluorescence. The short range of the quenching by His and Tyr can be exploited to design fluorescence-based sensors of specific elements of secondary structure. The strategy was demonstrated here using α-helical peptides, but can obviously be applied to globular proteins as well. In this case the F_{CN}-Tyr/His pair should be located in solvent exposed sites so that the change in fluorescence intensity upon unfolding is dominated by changes in the quencher-flurophore distance and not by changes in the solvation of the cyano group, which may arise if the probe were incorporated into a site

that is buried in the native state but exposed in the unfolded state. The change in fluorescence should be even larger than observed for the Ala-rich peptides, since the helical structure will be fully formed in a folded globular protein, but is only partially formed in the peptides. The approach is clearly not limited to helices and could be used to follow β-hairpin or β-sheet formation. The quenching studies of the pentapeptides indicate that a F_{CN}-Met pair could also be exploited to follow local structure formation.

4.4. References:

1. Lakowicz, J. R. (2006) *Principles of Fluorescence Spectroscopy*, Springer Science+Business Media, LLC New York, NY.
2. Beechem, J. M., and Brand, L. (1985) Time-resolved fluorescence of proteins, *Annu. Rev. Biochem. 54*, 43-71.
3. Ross, J. B. A., Szabo, A. G., and Hogue, C. W. V. (1997) Enhancement of protein spectra with tryptophan analogs: Fluorescence spectroscopy of protein-protein and protein-nucleic acid interactions, *Fluorescence Spectroscopy 278*, 151-190.
4. Wong, C. Y., and Eftink, M. R. (1998) Incorporation of tryptophan analogues into staphylococcal nuclease, its V66W mutant, and Delta 137-149 fragment: Spectroscopic studies, *Biochemistry 37*, 8938-8946.
5. Twine, S. M., and Szabo, A. G. (2003) Fluorescent amino acid analogs, *Method Enzymol 360*, 104-127.
6. Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code, *Annu. Rev. Biophys.Biomol. Struc. 35*, 225-249.
7. Rogers, J. M. G., Lippert, L. G., and Gai, F. (2010) Non-natural amino acid fluorophores for one- and two-step fluorescence resonance energy transfer applications, *Anal. Biochem. 399*, 182-189.
8. Tucker, M. J., Getahun, Z., Nanda, V., DeGrado, W. F., and Gai, F. (2004) A new method for determining the local environment and orientation of individual side chains of membrane-binding peptides, *J. Am. Chem. Soc. 126*, 5078-5079.
9. Tucker, M. J., Oyola, R., and Gai, F. (2006) A novel fluorescent probe for protein binding and folding studies: p-cyano-phenylalanine, *Biopolymers 83*, 571-576.

10. Tucker, M. J., Tang, J., and Gai, F. (2006) Probing the kinetics of membrane-mediated helix folding, *J. Phys. Chem. B 110*, 8105-8109.

11. Aprilakis, K. N., Taskent, H., and Raleigh, D. P. (2007) Use of the novel fluorescent amino acid p-cyanophenylalanine offers a direct probe of hydrophobic core formation during the folding of the N-terminal domain of the ribosomal protein L9 and provides evidence for two-state folding, *Biochemistry 46*, 12308-12313.

12. Marek, P., Gupta, R., and Raleigh, D. P. (2008) The fluorescent amino acid p-cyanophenylalanine provides an intrinsic probe of amyloid formation, *Chembiochem 9*, 1372-1374.

13. Miyake-Stoner, S. J., Miller, A. M., Hammill, J. T., Peeler, J. C., Hess, K. R., Mehl, R. A., and Brewer, S. H. (2009) Probing protein folding using site-specifically encoded unnatural amino acids as FRET donors with tryptophan, *Biochemistry 48*, 5953-5962.

14. Taskent-Sezgin, H., Chung, J., Patsalo, V., Miyake-Stoner, S. J., Miller, A. M., Brewer, S. H., Mehl, R. A., Green, D. F., Raleigh, D. P., and Carrico, I. (2009) Interpretation of p-cyanophenylalanine fluorescence in proteins in terms of solvent exposure and contribution of side-chain quenchers: A combined fluorescence, IR and molecular dynamics study, *Biochemistry 48*, 9040-9046.

15. Liu, J., Strzalka, J., Tronin, A., Johansson, J. S., and Blasie, J. K. (2009) Mechanism of interaction between the general anesthetic halothane and a model ion channel protein, II: fluorescence and vibrational spectroscopy using a cyanophenylalanine probe, *Biophys. J. 96*, 4176-4187.

16. Schultz, K. C., Supekova, L., Ryu, Y. H., Xie, J. M., Perera, R., and Schultz, P. G. (2006) A genetically encoded infrared probe, *J. Am. Chem. Soc. 128*, 13984-13985.

17. Tucker, M. J., Oyola, R., and Gai, F. (2005) Conformational distribution of a 14-residue peptide in solution: A fluorescence resonance energy transfer study, *J. Phys. Chem. B 109*, 4788-4795.

18. Glasscock, J. M., Zhu, Y. J., Chowdhury, P., Tang, J., and Gai, F. (2008) Using an amino acid fluorescence resonance energy transfer pair to probe protein unfolding: Application to the villin headpiece subdomain and the LysM domain, *Biochemistry 47*, 11070-11076.

19. Serrano, A. L., Troxler, T., Tucker, M.J., Gai, F. (2010) Photophysics of a fluorescent non-natural amino acid: p-Cyanophenylalanine, *Chem. Phys. Lett. 487*, 303-306.

20. Yoshida, M., Kaneko, H., Kitamura, A., Ito, T., Oohashi, K., Morikawa, N., Sakuragi, H., and Tokumaru, K. (1976) Mechanism for light-induced hydrogen isotope exchange in benzonitrile, *B. Chem. Soc. Jpn. 49*, 1697-1700.

21. Chen, Y., and Barkley, M. D. (1998) Toward understanding tryptophan fluorescence in proteins, *Biochemistry 37*, 9976-9982.

22. VanGilst, M., and Hudson, B. S. (1996) Histidine-tryptophan interactions in T4 lysozyme: 'Anomalous' pH dependence of fluorescence, *Biophys. Chem. 63*, 17-25.

23. Marek, P., Mukherjee, S., Zanni, M. T., Raleigh, D. P. (2010) Residue specific, real time characterization of lag phase species and fibril growth during amyloid

formation: A combined fluorescence and IR study of p-cyanophenylalanine analogs of islet amyloid polypeptide, *J. Mol. Biol. In press.*

24. Aurora, R., and Rose, G. D. (1998) Helix capping, *Protein Sci. 7*, 21-38.

25. Luo, P. Z., and Baldwin, R. L. (1997) Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water, *Biochemistry 36*, 8413-8421.

26. Kubelka, J., Chiu, T. K., Davies, D. R., Eaton, W. A., and Hofrichter, J. (2006) Sub-microsecond protein folding, *J. Mol. Biol. 359*, 546-553.

5. Azidohomoalanine Provides a Conformationally Sensitive IR Probe of Protein Folding, Protein Structure and Electrostatics Which Can Be Readily Incorporated into Proteins

Abstract

IR active probes are finding increasing use in studies of protein structure and dynamics. The vibrational modes of an azido group are known to be sensitive to solvation and the mode has a relatively high extinction coefficient, making it a potentially attractive probe. We demonstrate that the azido analog of methionine, azidohomoalanine can be readily incorporated into proteins by solid-phase peptide synthesis and recombinantly in high yield using methionine auxotrophic strains. Azidohomoalanine was incorporated into two sites in the α-β protein NTL9. Both of the variants are folded as judged by NMR and CD and adopt the wild-type structure. The frequency of the azido mode is observed to undergo a significant blue shift in the thermally unfolded state, indicating that the group provides a sensitive probe of protein folding and sidechain burial.

Note: Part of the material presented in this chapter has been published (H. Taskent-Sezgin, J. Chung, P. S. Banerjee, S. Nagarajan, R. B. Dyer, I. Carrico, and D. P. Raleigh. *Angewandte Chemie* **2010**). This chapter contains direct excerpts from the manuscript that was written by me with suggestions and revisions by Professor Raleigh. The unnatural amino acid, azidohomoalanine, was synthesized by Mr. P.S. Banerjee in Professor Isaac Carrico's laboratory. The FTIR spectra were taken in Professor R.B. Dyer's laboratory at Emory University.

5.1 Introduction

IR probes are widely used to study protein dynamics, protein-ligand interactions and electric field effects in proteins, but such probes can suffer from poor sensitivity or can be difficult to introduce(*1-12*). The approach requires the introduction of an IR active reporter group into the protein of interest. In this chapter we show that azido bearing unnatural amino acid, azidohomoalanine (Aha), provides a high sensitivity probe of protein structure, protein folding and protein electrostatics. Interest in the azido group was initially driven by its application in bioorthagonal protein labeling and click chemistry (*13*). Use of the azido group as an IR probe has a number of attractive features. Its absorbance falls in an otherwise transparent region of the IR spectrum and the frequency is sensitive to the environment (*14*, *15*). Of particular importance, the extinction coefficient of an azido group is approximately twenty fold larger than that of the commonly employed cyano group (*14*). Azidohomoalanine (Aha) can be viewed as an analog of Met and, thus, one can realize its incorporation using well developed Met auxotrophic strains (*13*, *16*) (Figure 5.1A). The group is also compatible with the conditions of Fmoc solid-phase peptide synthesis.

We used the N-terminal domain of the ribosomal protein L9 (NTL9) as a test case. NTL9 is a 56 residue, with a mixed β-α structure protein that has been studied extensively as a model system for protein folding and stability. It folds cooperatively under a wide range of conditions (*17-19*). We targeted Met 1 by expression (NTL9-M1Aha) and Ile 4 (NTL9*-I4Aha) by chemical synthesis as sites for incorporation of the Aha substitutions (Figure 5.1B). NTL9* refers to a K12A mutant of NTL9 which adopts the same fold, but is more stable than wild type. Both of these positions are in the

hydrophobic core of the protein (Figure 5.1B). Figure 5.2 shows the IR spectra of azidohomoalanine in DMF, DMSO and D_2O. The peak positions are 2098, 2103 and 2115 cm^{-1} in DMF, DMSO and D_2O, respectively. There is a significant band shift observed with the change in the environment.

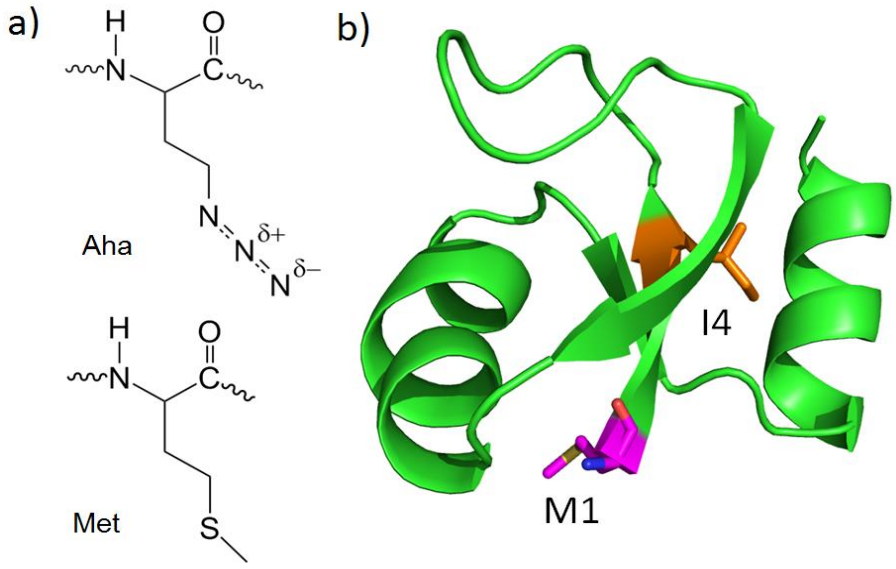


Figure 5.1. **(a) Structure of azidohomoalanine (Aha) and methionine (Met). (b) A ribbon diagram of NTL9 showing the location of Met1 and Ile4, in pink and orange, respectively.** The diagram was generated using the pdb file 2HBA and the program Pymol.

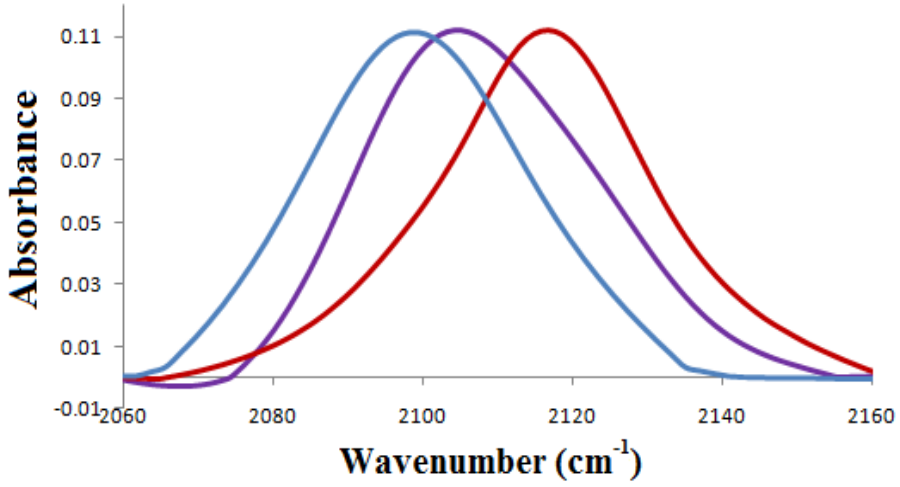


Figure 5.2. FTIR spectra of Aha in DMF (blue), DMSO (violet), and D_2O (red). Azidohomoalanine concentration was 1mg/ml in DMF, DMSO and D_2O (pH* 5.4). Spectra were recorded at 20°C.

5.2. Materials and Methods

5.2.1. Azidohomoalanine Synthesis:

Azidohomoalanine was synthesized in four steps starting from homoserine in 17% overall yield (shown in scheme 5.1). The synthesis was performed by P.S. Banerjee in Professor Isaac Carrico's laboratory.

Compound 3. L-homoserine, compound 1, (1.45 g, 12.7 mmol) was added to a solution of 9-BBN, 9-Borabicyclo(3.3.1)nonane (1.51 g, 12.4 mmol) in methanol (25 mL). The reaction mixture was refluxed for 3 hours under argon. TLC analysis showed the reaction to be complete. The reaction mixture was concentrated and the crude product purified by silica gel chromatography eluting with 50% ethylacetate-hexane to 100% ethyacetate. The 9-BBN protected L-homoserine was obtained as a white solid with 50% yield and purity was assessed as a single spot in TLC. ^{1}H NMR (300 MHz, CD_3OD): δ 1.4-1.5 (2H,

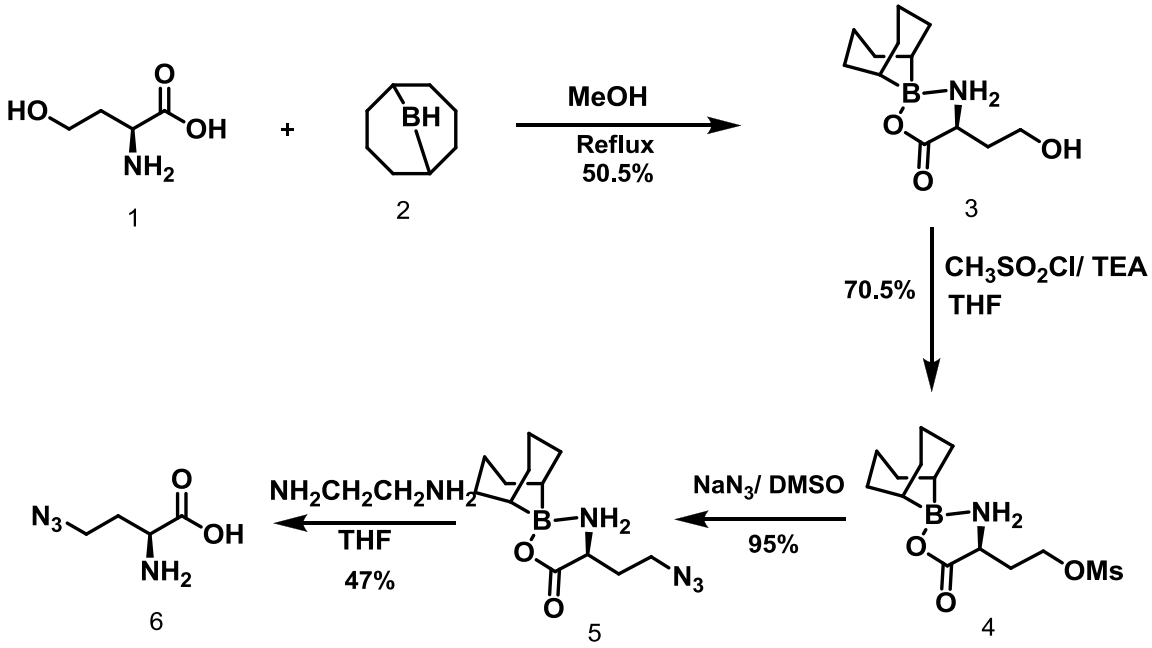
m), 1.6-2.0 (13H, m), 2.2 (1H, m), 3.75-3.85 (3H, m) ppm. High Resolution Mass Spectrum (HRMS) calculated for $C_{12}H_{22}BO_3N$ 238.81 found at 239.1(MH^+).

Compound 4. Methylsulfonyl chloride (0.87 mL, 7.5 mmol) was added to a solution of Compound 3 (1.5g, 6.3 mmol) in tetrahydrofurane (20 ml) followed by triethylamine (1.2 mL, 12 mmol) at 0°C. The reaction mixture was slowly warmed to room temperature and stirred overnight. TLC analysis showed that the reaction was complete. The mixture was concentrated and purified using silica gel chromatography eluting with 50% ethylacetate-hexane to 100% ethylacetate. The product was obtained as a white solid with 70% yield. Purity was ascertained by TLC and ^{1}H NMR. ^{1}H NMR (300MHz, CD_3OD): δ 1.4-1.5 (2H, m), 1.6-2.0 (12H, m), 2.1 (1H, m), 2.4 (1H, m), 3.1 (3H, s), 3.8 (1H, t), 4.5 (2H, m) ppm. HRMS calculated for $C_{13}H_{24}BSO_4$ 317.81 found at 319.1(MH^+).

Compound 5. Sodium azide (0.5 g, 7 mmol) was added to a solution of compound 4 (1.5 g, 5 mmol) in dimethylsulfoxide (10 mL). The mixture was heated at 60°C for 3 hours. The mixture formed a thick solid which was washed with water and extracted into ethylacetate. The product was a light yellow oil obtained at 95% yield. ^{1}H NMR (300MHz, CD_3OD): δ 1.4-1.5 (2H, m), 1.6-2.0 (13H, m), 2.2 (1H, m), 3.5-3.7 (2H, m), 3.75 (1H, t) ppm. HRMS calculated for $C_{12}H_{22}BO_2N_4$ 264.1 found at 264.1(MH^+).

Azidohomoalanine (compound 6). Ethylene diamine (1 mL, 21 mmol) was added to a solution of compound 5 (1.5 g, 5.6 mmol) in tetrahydrofurane (10 ml) and refluxed for 2 hours. The mixture was concentrated and triturated with diethylether, then filtered through a glass column over sand and cotton. The solid residue was dissolved in methanol and purified by silica gel chromatography eluting with 50:30:20 ethylacetate: methanol: water. The pure amino acid was obtained as white solid in 47% yield. ^{1}H and

2D NMR confirmed its purity. ^{1}H NMR (300Mz, D_2O): δ 1.9-2.1 (2H, m), 3.4-3.5 (2H, m), 3.7 (1H, t) ppm. HRMS calculated for $C_4H_8O_2N_4$ 144.1 found at 145.1(MH^+).



Scheme 5.1. Synthesis Scheme for Azidohomoalanine (Aha).

5.2.2. Protein Expression and Purification:

An overnight culture of B834-pET3a-NTL9 was grown in LB rich media with ampicillin. This starter culture was added to 1 liter of M9 minimal media supplemented with 18 amino acids (including methionine, but no tyrosine and cysteine). Cells were grown to an OD_{600} of 0.8-1, harvested and resuspended in M9 media salt solution. After agitation for several minutes at room temperature, cells were harvested again and resuspended in M9 minimal media supplemented with 17 amino acids (no methionine, tyrosine and cysteine) plus 40 mg azidohomoalanine. Protein expression was induced with 1mM IPTG overnight at 25°C. Protein was purified from the supernatant of the cell

lysate by cation exchange chromatography followed by reversed-phase HPLC on a Vydac C4 semi-preparative column. An A-B gradient system was used, with buffer A composed of 0.1% (v/v) solution of TFA and buffer B composed of 90% (v/v) acetonitrile, 9.9% (v/v) water and 0.1% (v/v) TFA. The gradient was 0-90% B in 90 min. The expression yield was, 15 mg which was around 30% of the wild type expression yield obtained in LB media.

5.2.3. Protein Synthesis and Purification:

NTL9 containing azidohomoalanine at position 4 was synthesized using Fmoc chemistry on an Applied Biosystems 433A peptide synthesizer. Standard Fmoc protocols were used as described elsewhere (*20*). The substitution was made in a K12A background, which is denoted as NTL9*. This background was chosen because it is more stable than wild-type and thus we reasoned that it would be able to tolerate the Aha substitution even if it were very destabilizing. In fact the mutation decreased the stability by 1.9 kcal.mol^{-1} and hence could have been tolerated in the wild-type background. The K12A mutation has been previously characterized and does not alter the structure or the cooperativity of folding (*21*). K12 is a surface residue and the mutation increases stability by modulating electrostatic interactions. The Fmoc-azidohomoalanine derivative was synthesized as described. The protein was purified using reverse phase HPLC on a Vydac C4 column as described (*20*).

5.2.4. Mass Spectroscopy:

NTL9-M1Aha was characterized by LTQ-Orbitrap XL Mass Spectrometer (Figure 5.3). NTL9-I4AhaK12A protein was characterized by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). For NTL9*-I4Aha the observed (6171.75 Da) and calculated (6171.38 Da) monoisotopic masses match well.

5.2.5. Equilibrium Denaturations:

CD wavelength spectra, CD monitored urea and temperature denaturation experiments were performed on an Applied Photophysics Chirascan instrument. Samples were, 20 mM sodium acetate, 100 mM NaCl. Measurements were conducted at pH 5.4 and 25°C. The concentration of urea solutions were determined by measuring the refractive index. Denaturation curves were fit to the equation:

$$f = \frac{a_n + b_n[denaturant] + (a_d + b_d[denaturant]) \exp(-\Delta G_u^o([denaturant]/RT))}{1 + \exp(-\Delta G_u^o([denaturant]/RT))} \quad (5.1)$$

where:

$$\Delta G_u^o([denaturant]) = \Delta G_u^o(H_2O) - m[denaturant] \quad (5.2)$$

f is the ellipticity in CD monitored denaturation experiments, ΔG_u^o is the change in free energy for unfolding reaction, T is the temperature and R is the gas constant. a_n is the intercept and b_n is the slope of the curve extrapolated in the pre-transition region. Similarly, a_d is the intercept and b_d is the slope of the curve in the post-transition region. Thermal denaturations were fit with an expression of the same type as equation (1) with

$\Delta G_u^o[denaturant]$ replaced by $\Delta G^o(T).\Delta G^o(T)$ is given by the Gibbs–Helmholtz expression:

$$\Delta G^o(T) = \Delta H^o(T_m)(1-\frac{T}{T_m}) - \Delta C_p^o\left[(T_m - T) + T.\ln(\frac{T}{T_m})\right] \quad (5.3)$$

where T_m is the midpoint temperature, T is the temperature, $\Delta H^o(T_m)$ is the change in enthalpy for the unfolding reaction at T_m, and ΔC_p^o is the change in heat capacity and was set to 0.53 kcal.mol^{-1}.deg^{-1} as previously determined (*17*).

5.2.6. NMR:

The 1D NMR spectra were collected using a Varian Inova 600 MHz spectrometer and the data were processed with MestReNova from Mestrelab Research S. L. A 2D ^{1}H-^{1}H NOESY experiment in D_2O for NTL9*-I4Aha was performed on Bruker 700 MHz spectrometer with mixing time of 150 ms. The spectrum was collected with spectral widths of 8389 for both ^{1}H dimensions in a data matrix of 2048 (direct ^{1}H) × 512 (indirect ^{1}H) complex points. A 2D ^{1}H-^{1}H NOESY experiment in D_2O for NTL9-M1Aha was performed on a Varian Inova 600 Mhz spectrometer with mixing time of 150 ms. The spectrum was collected with spectral widths of 6600 for both ^{1}H dimensions in a data matrix of 2048 (direct ^{1}H) × 1024 (indirect ^{1}H) complex points. Both NOESY spectra were processed with the NMRPipe software package developed by Delaglio and coworkers, and the chemical shifts were analyzed with the NMRViewJ software package developed by Johnson and coworkers (*22*, *23*).

Protein samples for NMR experiments were dissolved in 20 mM deuterated sodium actate, 100 mM NaCl in 100% D_2O at pD 5.6 with concentration about 2mM. The spectrum was internally referenced to sodium 4,4-dimethyl-4-silapentane-1-sulphonate (DSS).

5.2.7. Stopped-flow Fluorescence Measurements:

Stopped-flow studies were performed as previously described for NTL9 using an SX20 stopped-flow spectrometer from Applied Photophysics *(20)*. A 305 nm cutoff filter was used to collect the fluorescence. Samples were 20 mM sodium acetate, 100 mM NaCl, pH 5.4 at 37°C. The protein concentration before the jump was around 400 uM. Stopped flow experiments used 11 fold dilution. Four to five fluorescence traces were averaged at each denaturant concentration, and the average trace was fit to a single-exponential function in order to obtain the observed rate constant (k_{obs}) at that denaturant concentration. The concentration of urea solutions were determined by measuring the refractive index. The plot of natural logarithm of k_{obs} vs denaturant concentration was fit to the following equation:

$$\ln(k_{obs}) = \ln(k_f(H_2O)\exp(-m_f[denaturant]/RT) + k_u(H_2O)\exp(m_u[denaturant]/RT) \quad (5.4)$$

where $k_f(H_2O)$ and $k_u(H_2O)$ are the folding and unfolding rate in the absence of denaturant, respectively.

5.2.8. FTIR Spectroscopy:

Equilibrium FTIR temperature-dependent spectra were recorded on a Varian 3100 FTIR spectrometer equipped with a liquid nitrogen cooled mercury cadmium telluride (MCT) detector. The spectra were the result of 256 scans recorded at a resolution of 2 cm^{-1}. The proteins were dissolved in a buffer containing 10 mM sodium phosphate and 150 mM sodium chloride at a pH* of 5.4, 6 and 8.8 in D_2O (pH* refers to the uncorrected (for D_2O) pH-meter reading at 25°C). The NTL9 protein concentration for IR experiments is ~ 2 mM. A split IR cell composed of CaF_2 windows was utilized with a path length of 100 μm to record the spectrum of both the reference (buffer in D_2O) and the sample (protein in the D_2O buffer) side of the IR transmission cell under identical conditions at each temperature. The temperature of the IR cell was controlled by a water bath, and the sample temperature was measured by a thermocouple attached to the cell. The absorbance spectra of the protein were determined from the negative logarithm of the ratio of the single beam spectra of the sample to the reference side of the IR split cell at each temperature.

5.3. Results and Discussion

Protein expression was performed with the standard methionine auxotrophic *E. coli* strain, B834. Mass spectroscopy was used to test the level of incorporation (Figure 5.3). The experimental and theoretical mass spectra of NTL9-M1Aha match extremely well and a minimum incorporation of 94% was detected by MS-MS analysis. The small amount of wild-type protein is spectroscopically silent in the azido vibration region of the infrared spectrum.

Spectroscopic analysis shows that the substitution does not perturb the structure of the protein. The far-UV CD spectrum of the M1Aha mutant is similar to the spectrum of the wild-type indicating that they have similar secondary structure (Figure 5.4). The 1D NMR spectrum of the mutant displays the characteristic ring current shifted methyl resonances of NTL9 (Figure 5.5), as well as Cα proton chemical shifts downfield of H_2O, indicative of β-sheet structure, and the NOE spectrum shows that the native registry of the β-sheet is maintained (Figure 5.6). The potential effects of the substitution on the stability and cooperativity of folding were also examined. Urea denaturation shows a sigmoidal curve, as expected for a cooperatively folded unit (Figure 5.7). The stability is decreased relative to wild-type by 0.81 $kcal.mol^{-1}$. The m-value, which is the slope of the ΔG° vs urea curve, is very similar to that of wild-type (Table 5.1). M-values are widely believed to be related to the change in accessible surface area between the folded and unfolded states and the good agreement between wild-type and mutant provides additional evidence of the structural integrity of the mutant (*24*). Thermal unfolding is also cooperative and the Tm is decreased by 5 to 6 °C (Figure 5.8). Stopped-flow kinetic refolding studies confirm that folding is two-state (Figure 5.9).

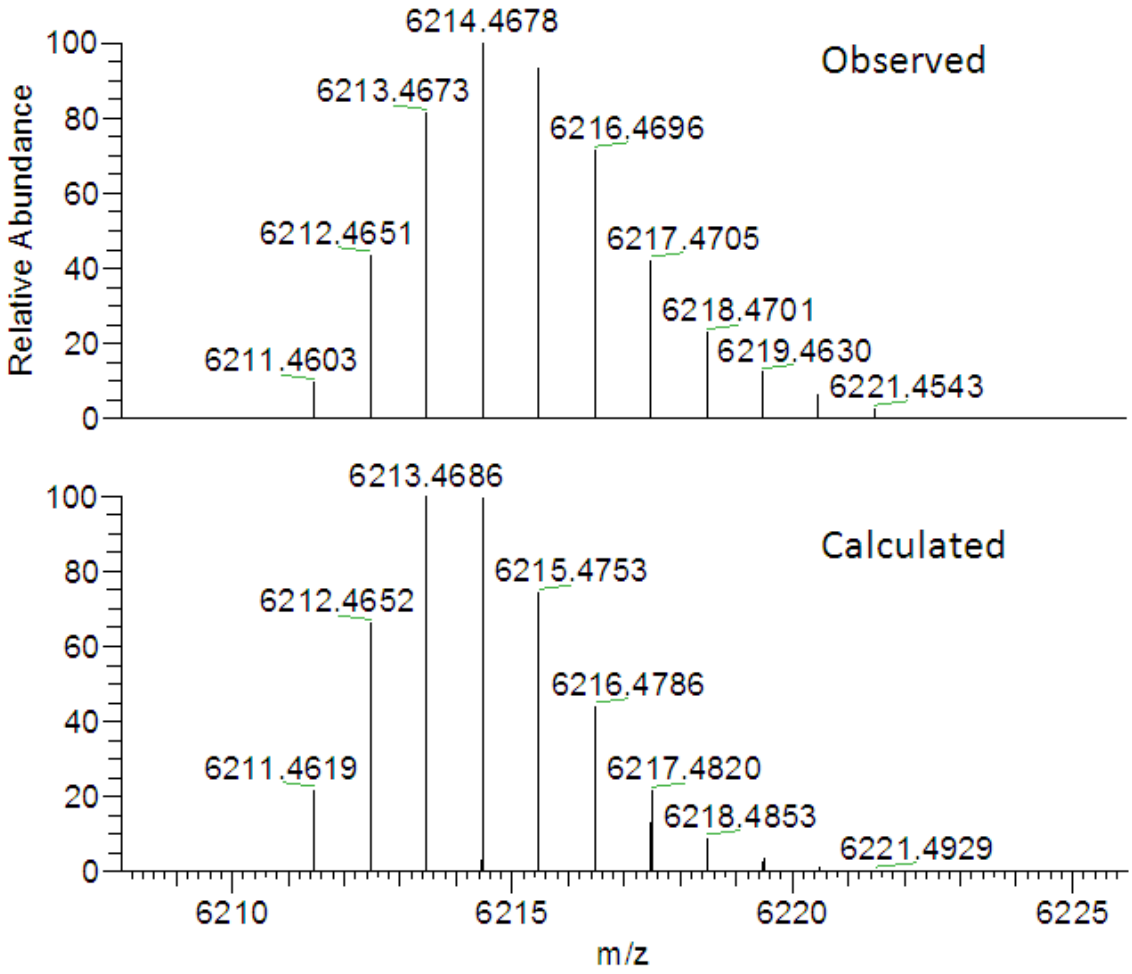


Figure 5.3. Mass Spectrum of NLT9-M1Aha. Observed (6211.460 Da) and calculated (6211.462) monoisotopic masses are in excellent agreement. There is estimated to be 6% wild-type protein contamination as detected by a separate MS/MS analysis.

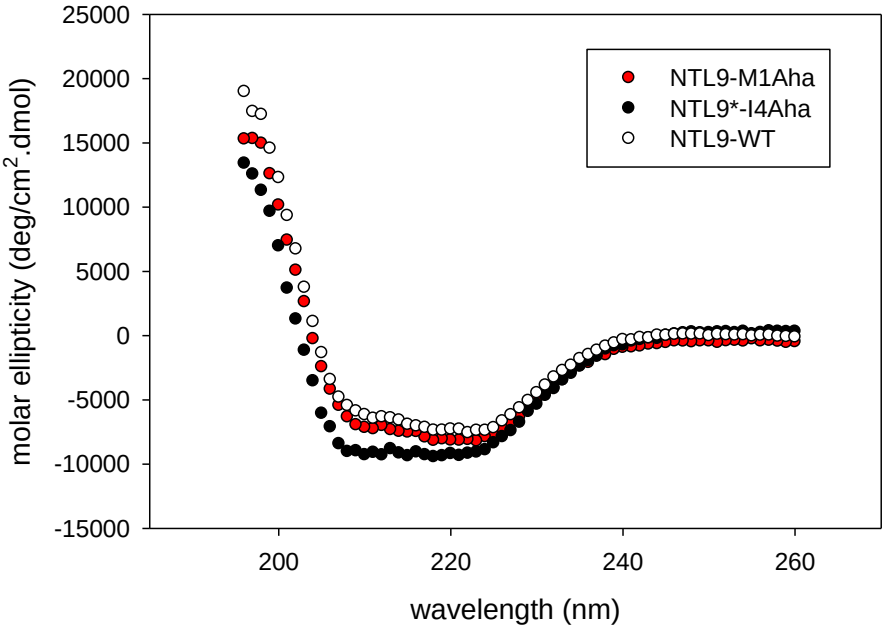


Figure 5.4. Far-UV CD wavelength spectra of NTL9 wild-type, NTL9-M1Aha, and NTL9*-I4Aha mutants. The similarity in the curves indicates that the mutants have similar secondary structure. Samples were in 20 mM sodium acetate, 100 mM NaCl, at pH 5.4. The spectra were collected at 25°C. Protein concentration was 20 uM.

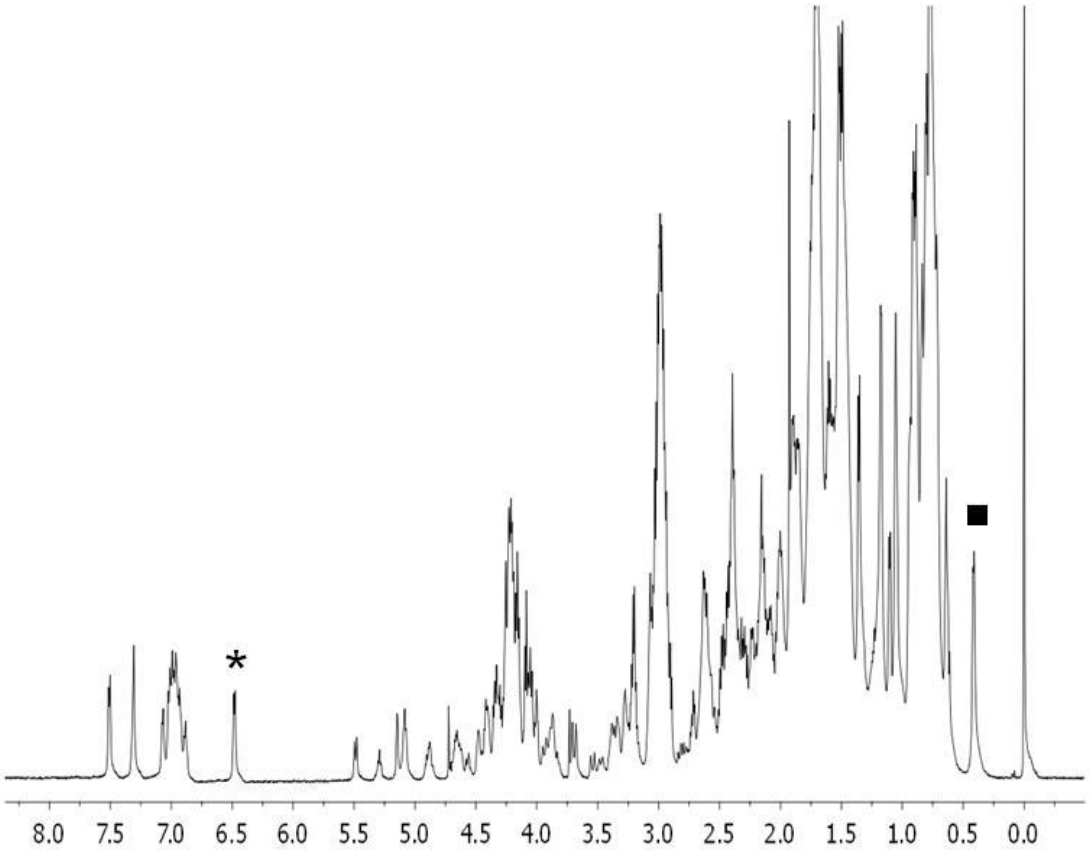


Figure 5.5. 1D NMR spectrum of NTL9-M1Aha. There are signature C^{α} proton chemical shifts in the 1D NMR spectrum of folded wild type NTL9 which are also present in the spectrum of the mutant. These are located downfield of water chemical shift and include the α-protons of Ile4 at 5.08 ppm, Val3 at 5.15 ppm, Asn20 at 5.29 ppm and Ile18 at 5.49 ppm. In addition to these, the doublet from the protons on the phenol ring of Tyr25 at 6.47 ppm is diagnostic of the folded structure of NTL9 (labeled with *). Note, also, that the characteristic ring current shifted methyl resonance of Val3 is observed at 0.42 ppm (labeled with ■). These chemical shifts indicate that the mutation did not perturb the structure of the protein. The experiment was conducted at 25°C, pD 5.6 in 20 mM deutrated sodium adetate, 100 mM NaCl in D_2O. The sharp peak at 0.0 ppm is due to the chemical shift standard, DSS.

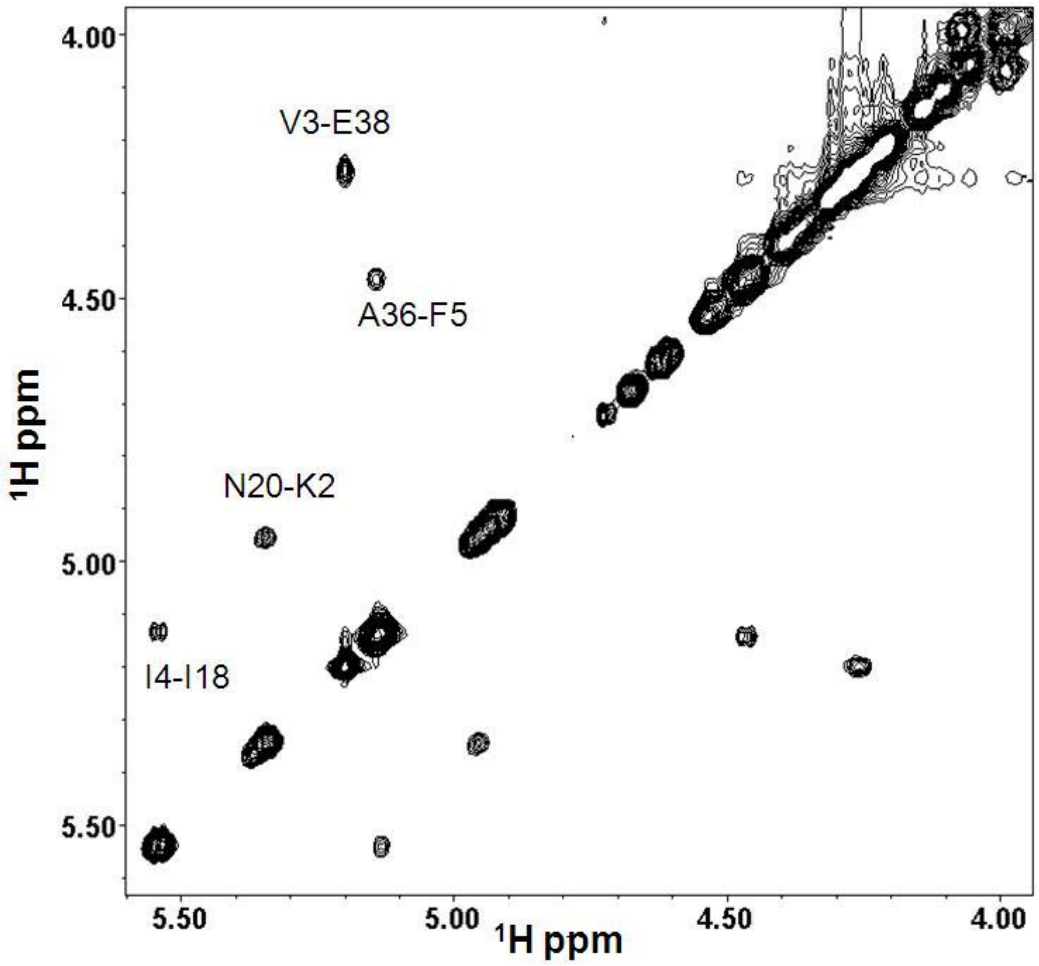


Figure 5.6. α to α region of the NOESY-NMR spectrum of the NTL9-M1Aha mutant in D_2O. The spectrum shows NOEs between α-protons on adjacent β-strands: Ile 18-Ile 4, Asn 20-Lys 2, Ala 36-Phe 5, and Glu 38-Val 3. These cross-strand NOEs show that the β-strands in the mutant are in the correct registry. These are all of the NOEs which are observed for the wild-type protein in this region.(*17*) The experiment was conducted at 25°C, pD 5.6 in 20 mM deutrated sodium acetate, 100 mM NaCl in D_2O.

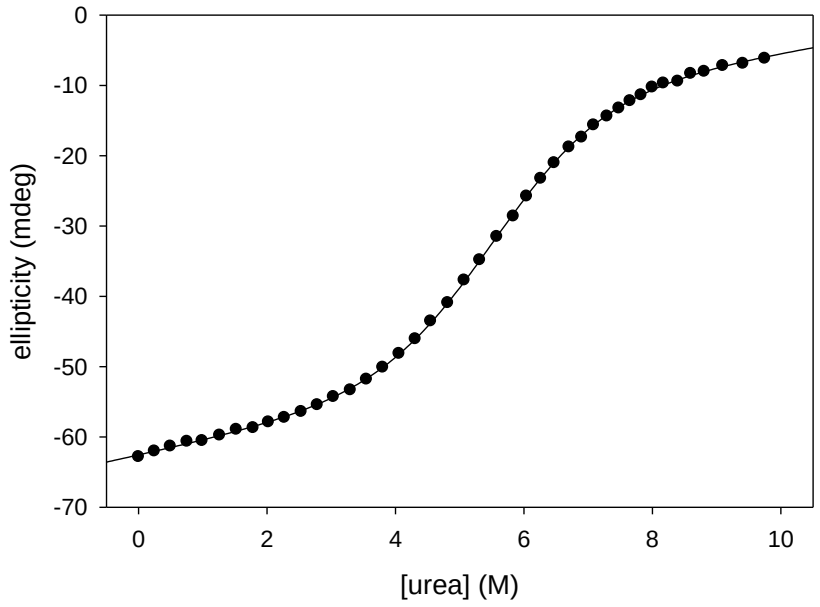


Figure 5.7. Equilibrium unfolding studies of NTL9-M1Aha mutant induced by urea. The CD signal at 222 nm was monitored. Experiments were conducted at pH 5.4, 25°C, in 20 mM NaAcetate, 100 mM NaCl.

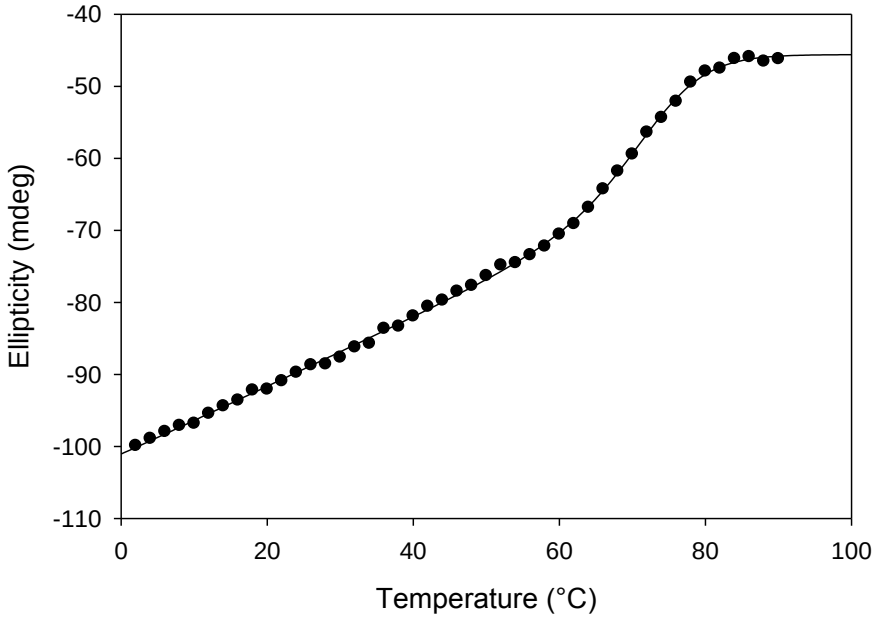


Figure 5.8. Thermal denaturation curve for NTL9-M1Aha mutant at pH 5.4. Experiments were conducted at 25°C, in 20 mM NaAcetate, 100 mM NaCl.

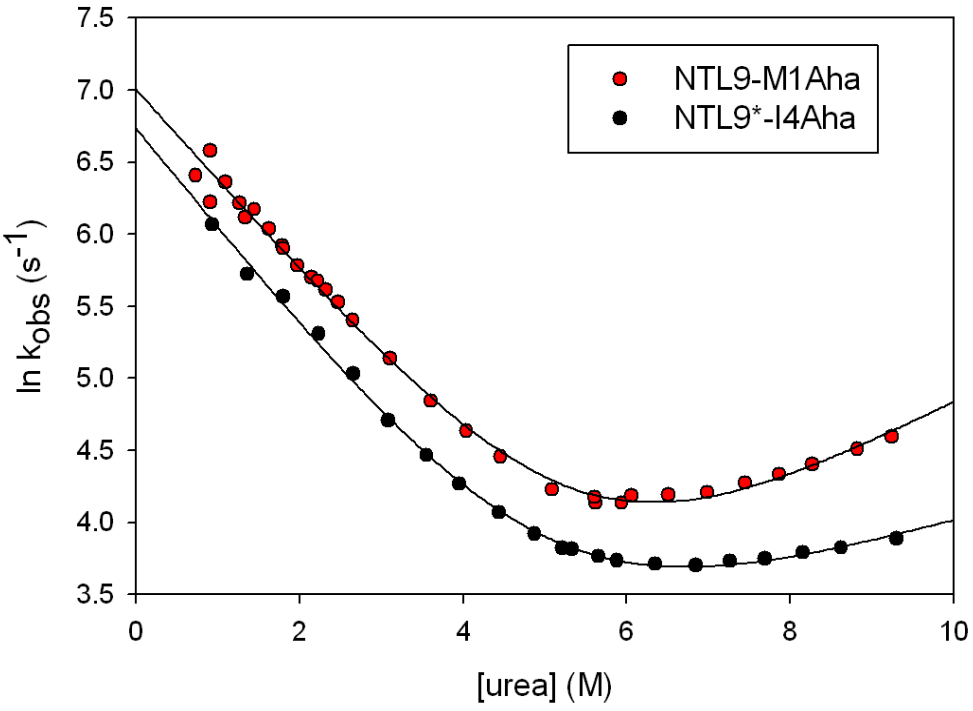


Figure 5.9. Fluorescence detected stopped flow urea jump folding studies of NTL9-M1Aha and NTL9*-I4Aha. Experiments were conducted at pH 5.4, 37°C, in 20 mM NaAcetate, 100 mM NaCl.

protein	ΔG° (kcal/mol)	m (kcal/mol.M)	Tm (°C) (fit / first derivative)[a]
I4Aha,K12A	3.03 ± 0.07	0.53 ± 0.01	72.1 / 69.5
K12A[b]	4.91 ± 0.15[b]	0.66 ± 0.09[b]	_
WT[c]	4.30 ± 0.36[c]	0.66 ± 0.07[c]	77.6[c]
M1Aha	3.49 ± 0.07	0.62 ± 0.01	72.4 / 70.9

Table 5.1. Thermodynamic parameters of NTL9 wild-type and mutants at pH 5.4 in 20 mM sodium acetate and 100 mM NaCl at 25°C in H_2O. [a]Tm values were calculated by directly fitting the curve and by calculating the first derivative. [b]Values taken from Cho et. al. (*21*). [c]Values taken from Cho et. al. (*25*). ΔG° values are given as unfolding free energies and were measured at 25°C.

IR spectra were recorded in the folded state and in the unfolded state induced by temperature. The folded state spectrum was recorded at pD 8.8 to ensure a single protonation state for the N-terminus since a partially protonated N-terminus might lead to a splitting of the folded state peak due to electrostatic effects. NTL9 is fully folded at pD 8.8, 20 °C (Figure 5.10 and Figure 5.11). In the folded state, the azido band is at 2094 cm^{-1}. In the unfolded state at 90ºC, a single broad band is found at 2113 cm^{-1}, indicating exposure of the azido group to water (Figure 5.12). The significant band shift observed between the folded and unfolded states indicates that the azido vibration can be used as a probe of protein folding and hydrophobic burial of side chain.

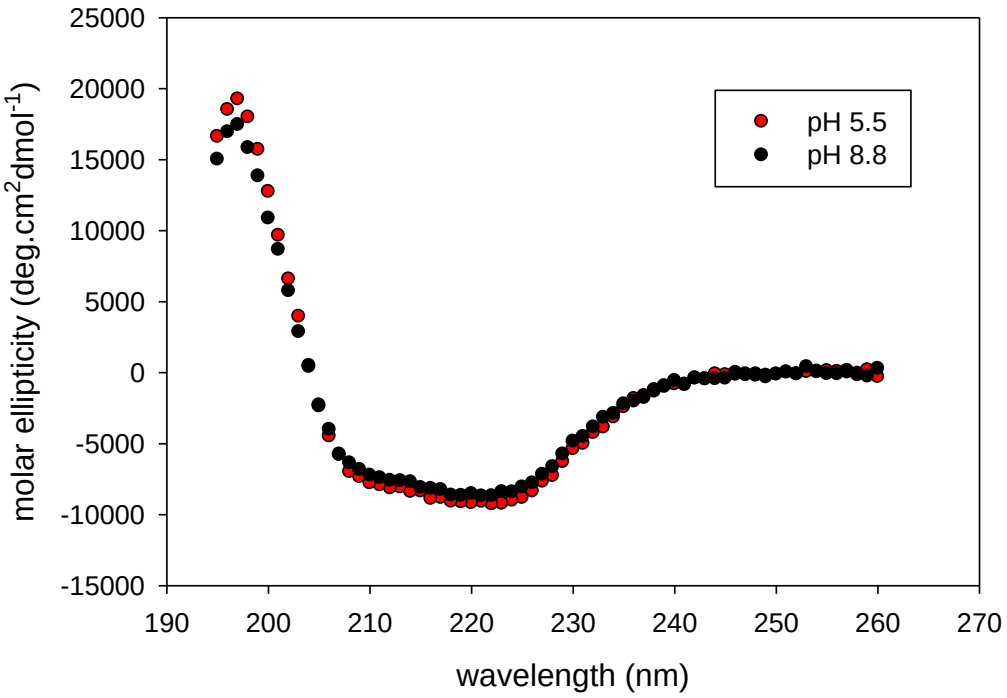


Figure 5.10. Far-UV CD wavelength spectra of NTL9-M1Aha at pH 5.5 (black) and at pH 8.8 (red). Experiment was conducted in 10mM sodium phosphate and 150mM NaCl. The similarity of the curves indicates that the protein has similar secondary structure at both pH values.

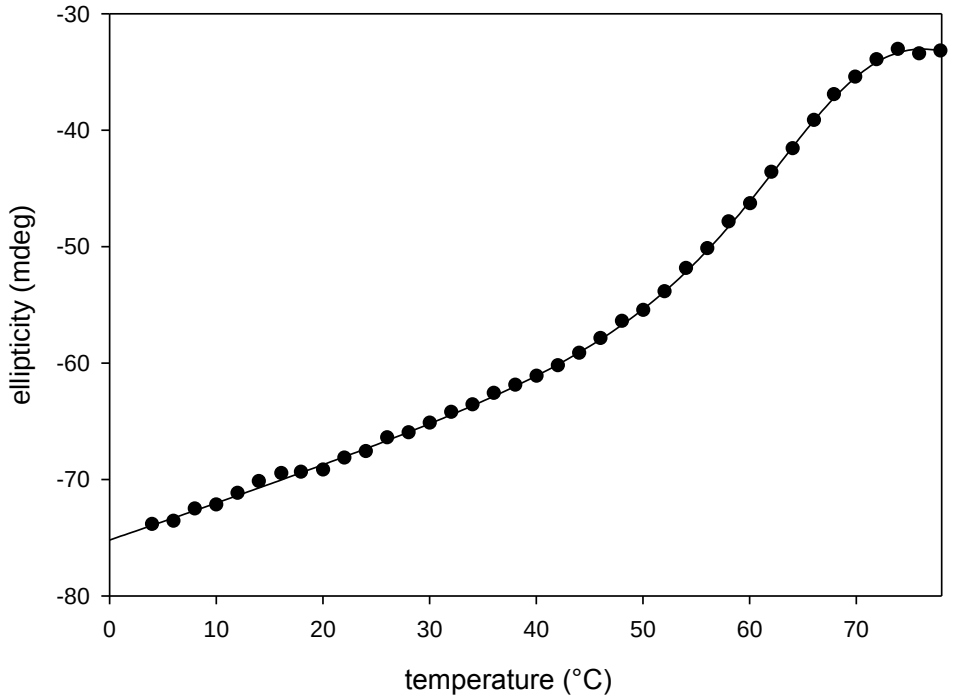


Figure 5.11. CD monitored temperature melt of NTL9-M1Aha at pD 8.8. Experiment was conducted in 10 mM sodium phosphate, 150 mM sodium chloride at indicated pD (100% D_2O). The apparent Tm is 67.6°C. The data indicates that the protein is fully folded at 20°C.

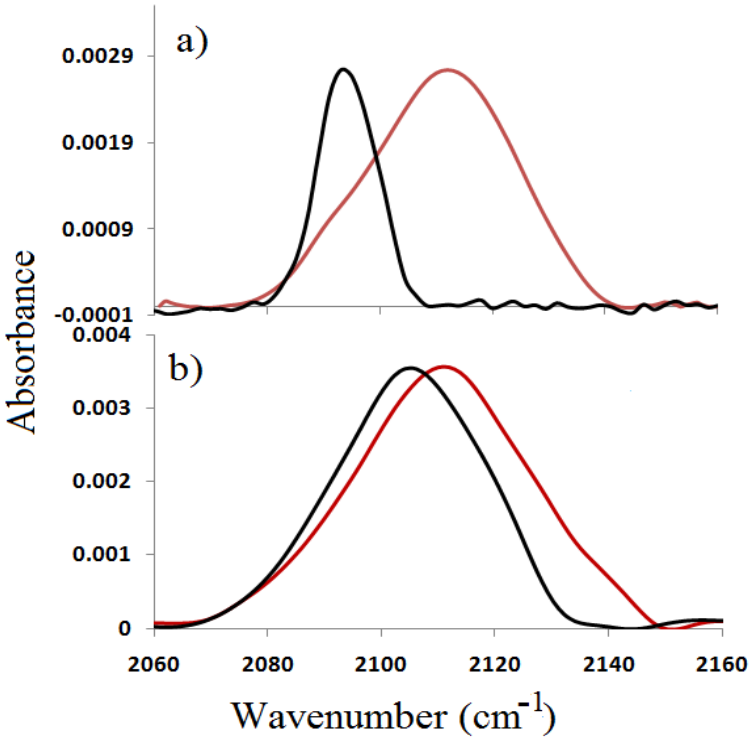


Figure 5.12. FTIR spectra of (a) NTL9-M1Aha and (b) NTL9*-I4Aha in the folded state at 20ºC (black) and in the unfolded state at 90ºC (red).

The methodology is not limited to selective incorporation of a Met at the N-terminus since methionyl aminopeptidase has been shown to cleave azidohomoalanine from proteins expressed in E. coli depending on the solvent exposure of Met 1 and the amino acid following Met 1. Methionyl aminopeptidase cleaves N-terminal Aha if the next amino acid is small (*16*). In NTL9, Met 1 is part of the hydrophobic core and is retained when the protein is expressed in normal cell lines.

We demonstrated that the probe can also be incorporated into proteins by solid-phase peptide synthesis. Ile 4 in NTL9 was replaced by Aha using Fmoc based methods. The I4Aha variant adopted the same fold as wild-type as judged by CD (Figure 5.4), 1D

(Figure 5.13) and 2D-NMR (Figure 5.14). Denaturant induced and thermal unfolding are still cooperative (Figure 5.15 and 5.16) and stopped-flow kinetic studies confirm that folding is two state (Figure 5.9). Substitution of Aha for I4 is more destabilizing than for M1 with a $\Delta\Delta G°$ of 1.9 kcal.mol^{-1} (Table 5.1). The larger effect is likely due to the fact that Aha is approximately isosteric for Met but not for Ile. IR spectra of the I4Aha mutant were taken in the folded state and in the thermally unfolded state. In the folded state the azido vibration was observed at 2105 cm^{-1} and in the unfolded state the vibration blue shifts to 2112 cm^{-1} (Figure 5.12). The observed frequency for Aha in the folded state of the I4Aha mutant at 20°C indicates that azido group is more exposed to solvent than in the M1Aha mutant. I4 packs against the C-terminal helix and this helix is known to fray at its C-terminus in the wild-type. The change in shape and polarity of the substitution may enhance this effect in the mutant and account for the partial solvent exposure. Irrespective of the details, the experiments with I4Aha demonstrate that Aha can be incorporated by solid phase peptide synthesis and reiterate that the IR vibration of the azido group is sensitive to its environment.

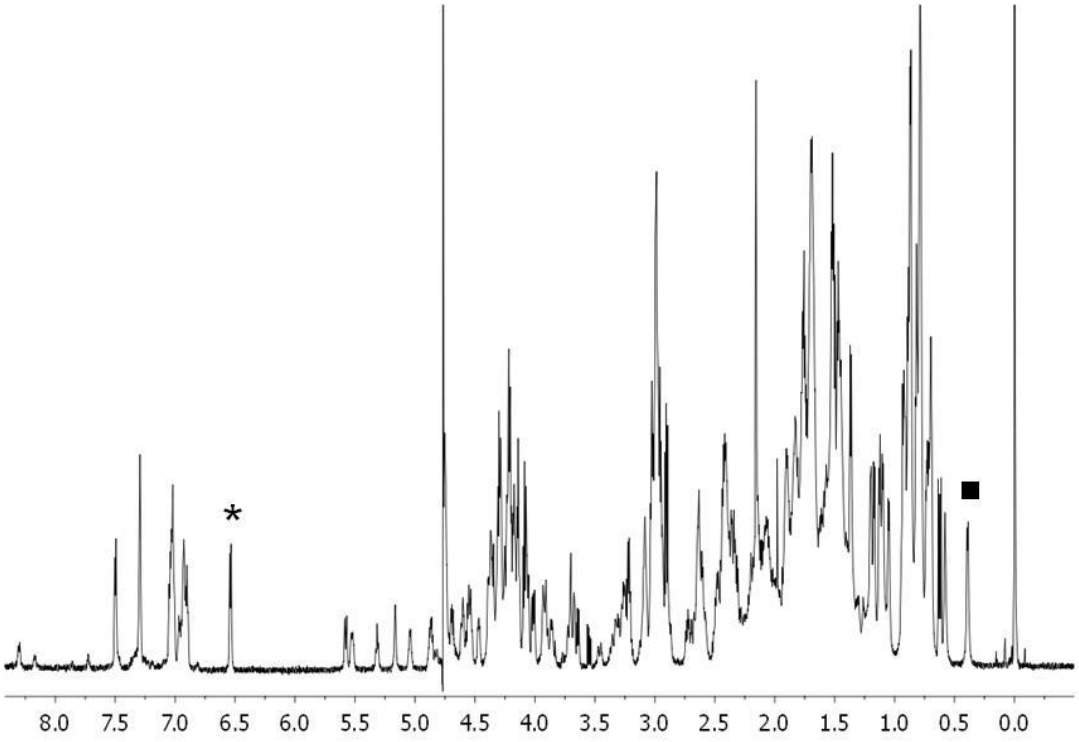


Figure 5.13. 1D NMR spectrum of NTL9*-I4Aha. The signature Cα-proton chemical shifts downfield of water are also present in NTL9*-I4Aha mutant: These are Aha4 at 5.05 ppm, Val3 at 5.18 ppm, Asn20 at 5.30 ppm and Ile18 at 5.60 ppm. The doublet from the protons of the phenol ring of Tyr25 is observed at 6.55 ppm which is typical for folded NTL9 structure (labeled with *). The ring current shifted methyl resonance due to Val3 is observed at 0.40 ppm (labeled with ■). These chemical shifts indicate the similarity between the structure of the mutant and the wild-type protein. The experiment was conducted at 25°C, pD 5.6 in 20 mM deutrated sodium acetate, 100 mM NaCl in D_2O. The sharp peak at 0.0 ppm is due to the chemical shift standard, DSS.

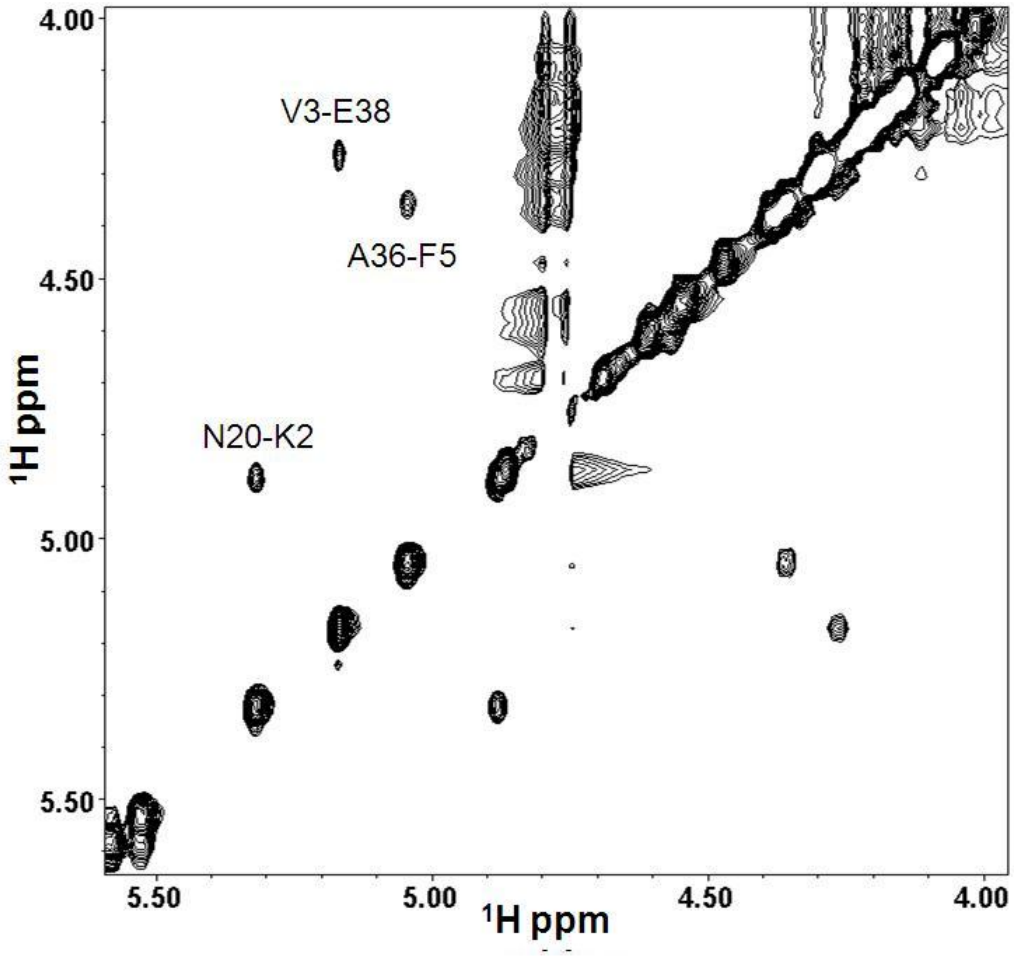


Figure 5.14. α to α region of the NOESY-NMR spectrum of the NTL9*-I4Aha mutant. NOEs are shown between α-protons on adjacent β-strands: Asn20-Lys2, Ala36-Phe5, and Val3-Glu38. These are 3 of the 4 NOEs observed for the wild-type protein. These NOEs confirm the correct registry of the β-strands in the mutant protein. The experiment was conducted at 25°C, pD 5.6 in 20 mM deutrated sodium acetate, 100 mM NaCl in D_2O.

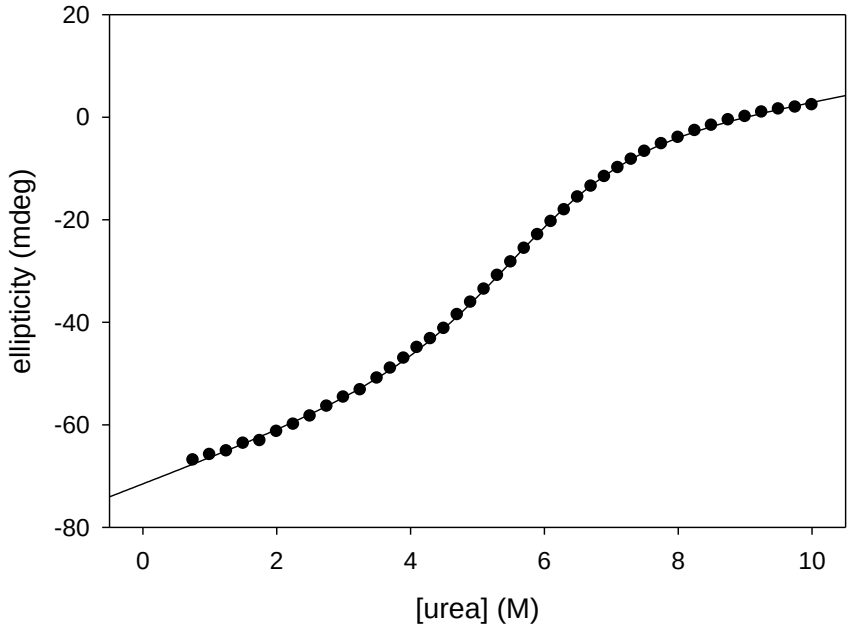


Figure 5.15. Equilibrium unfolding studies of NTL9*-I4Aha mutant induced by urea. The CD signal at 222 nm was monitored. Experiments were conducted at pH 5.4, 25°C, in 20 mM NaAcetate, 100 mM NaCl.

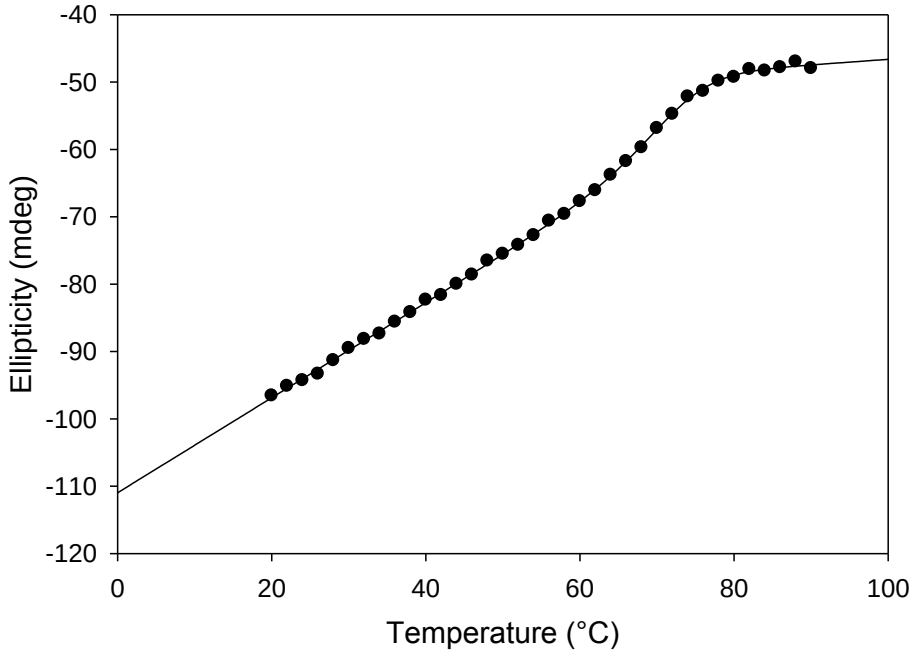


Figure 5.16. Thermal denaturation curve for NTL9*-I4Aha mutant. Experiments were conducted at pH 5.4, 25°C, in 20 mM NaAcetate, 100 mM NaCl.

Spectra were also recorded of the M1Aha variant at pD 6.0 to test the sensitivity of the probe to varying nearby charges. NTL9 is fully folded under these conditions but the N-terminus is partially protonated. Figure 5.17 compares the spectrum recorded at pD 5.4, 6.0, 8.8 in the folded state, and thermally unfolded state. The pD 5.4 and 6.0 spectra display a second partially resolved peak at 2116 cm^{-1} which is due to the protonated form and a major peak at 2094 cm^{-1}. The drastic change observed upon partial protonation of the N-terminus demonstrates the sensitivity of the azido group vibrational mode to changes in electrostatic effects and minor native state structural variations.

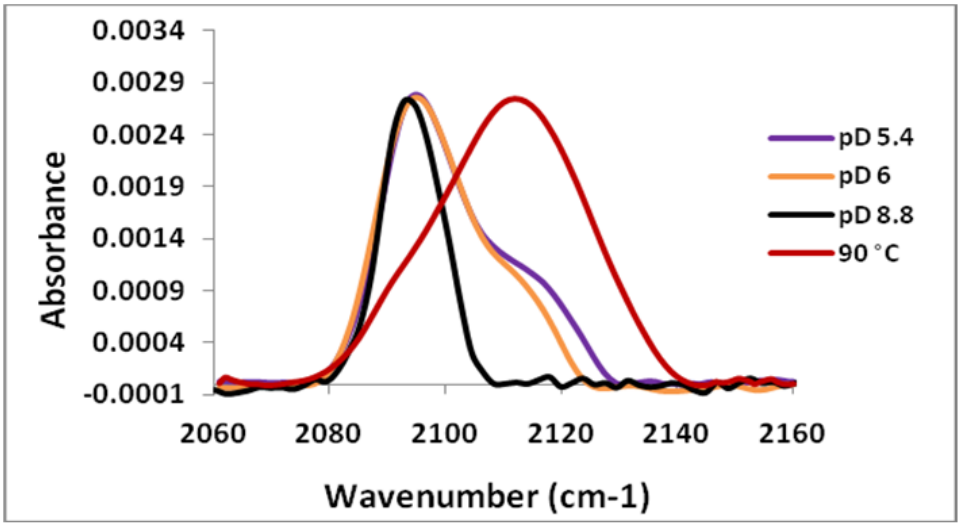


Figure 5.17. pH dependent FTIR spectra of NTL9-M1Aha. Purple pD 5.4; orange pD 6; black pD 8.8 at 20°C in the folded state and pD 6 at 90°C in the unfolded state. The experiments were conducted in 10 mM sodium phosphate and 150 mM NaCl.

We have demonstrated that azidohomoalanine is a sensitive IR probe of protein folding, protein structure and electrostatic effects. The probe can be easily incorporated

into proteins in high yield in a site specific manner using simple, readily available auxotrophic expression systems as well as by solid-phase peptide synthesis. The probe provides a complimentary approach to methods involving orthogonal aminoacyl tRNA synthatases or approaches that involve the attachment of probes to introduced cysteine residues and should be generally accessible (*6*, *7*).

5.4. References:

1. Bublitz, G. U., and Boxer, S. G. (1997) Stark spectroscopy: Applications in chemistry, biology, and materials science, *Annu. Rev. Phys. Chem. 48*, 213-242.
2. Zanni, M. T., and Hochstrasser, R. M. (2001) Two-dimensional infrared spectroscopy: a promising new method for the time resolution of structures, *Curr. Opin. Struct. Biol. 11*, 516-522.
3. Xu, Y., Oyola, R., Gai, F. (2003) Infrared study of the stability and folding kinetics of a 15-residue beta-hairpin., *J. Am. Chem. Soc. 125*, 15388-15394.
4. Suydam, I. T., and Boxer, S. G. (2003) Vibrational Stark effects calibrate the sensitivity of vibrational probes for electric fields in proteins, *Biochemistry 42*, 12050-12055.
5. Suydam, I. T. (2006) Electric fields at the active site of an enzyme: direct comparison of experiment with theory (vol 313, pg 200, 2006), *Science 313*, 1887-1887.
6. Schultz, K. C., Supekova, L., Ryu, Y. H., Xie, J. M., Perera, R., and Schultz, P. G. (2006) A genetically encoded infrared probe, *J. Am. Chem. Soc. 128*, 13984-13985.
7. Fafarman, A. T., Webb, L. J., Chuang, J. I., and Boxer, S. G. (2006) Site-specific conversion of cysteine thiols into thiocyanate creates an IR probe for electric fields in proteins, *J. Am. Chem. Soc. 128*, 13356-13357.
8. Brewer, S. H., Song, B. B., Raleigh, D. P., and Dyer, R. B. (2007) Residue specific resolution of protein folding dynamics using isotope-edited infrared temperature jump spectroscopy, *Biochemistry 46*, 3279-3285.

9. Hauser, K., Krejtschi, C., Huang, R., Wu, L., and Keiderling, T. A. (2008) Site-specific relaxation kinetics of a tryptophan zipper hairpin peptide using temperature-jump IR spectroscopy and isotopic labeling, *J. Am. Chem. Soc. 130*, 2984-2992.

10. Hunt, N. T. (2009) 2D-IR spectroscopy: ultrafast insights into biomolecule structure and function, *Chem. Soc. Rev. 38*, 1837-1848.

11. Ye, S. X., Huber, T., Vogel, R., and Sakmar, T. P. (2009) FTIR analysis of GPCR activation using azido probes, *Nat. Chem. Biol. 5*, 397-399.

12. Kim, Y. S., and Hochstrasser, R. M. (2009) Applications of 2D IR spectroscopy to peptides, proteins, and hydrogen-bond dynamics, *J. Phys. Chem. B 113*, 8231-8251.

13. Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation, *Proc. Natl. Acad. Sci. U.S.A. 99*, 19-24.

14. Oh, K. I., Lee, J. H., Joo, C., Han, H., and Cho, M. (2008) beta-Azidoalanine as an IR probe: Application to amyloid A beta(16-22) aggregation, *J Phys Chem B 112*, 10352-10357.

15. Choi, J. H., Oh, K. I., and Cho, M. H. (2008) Azido-derivatized compounds as IR probes of local electrostatic environment: Theoretical studies, *J. Chem. Phys. 129*, 17.

16. Wang, A., Nairn, N. W., Johnson, R. S., Tirrell, D. A., and Grabstein, K. (2008) Processing of N-terminal unnatural amino acids in recombinant human interferon-beta in Escherichia coli, *Chembiochem 9*, 324-330.

17. Kuhlman, B., Boice, J. A., Fairman, R., and Raleigh, D. P. (1998) Structure and stability of the N-terminal domain of the ribosomal protein L9: Evidence for rapid two-state folding, *Biochemistry 37*, 1025-1032.

18. Kuhlman, B., and Raleigh, D. P. (1998) Global analysis of the thermal and chemical denaturation of the N-terminal domain of the ribosomal protein L9 in H2O and D2O. Determination of the thermodynamic parameters, Delta H degrees, Delta S degrees, and Delta C degrees(p), and evaluation of solvent isotope effects, *Protein. Sci. 7*, 2405-2412.

19. Kuhlman, B., Luisi, D.L., Evans, P.A., Raleigh, D.P. (1998) Global analysis of the effects of temperature and denaturant on the folding and unfolding kinetics of the N-terminal domain of the protein L9., *J. Mol. Biol. 284*, 1661-1670.

20. Anil, B., Sato, S., Cho, J. H., and Raleigh, D. P. (2005) Fine structure analysis of a protein folding transition state; distinguishing between hydrophobic stabilization and specific packing, *J. Mol. Biol. 354*, 693-705.

21. Cho, J. H. (2006) The role of the unfolded state in protein stability and protein folding *Department of Biochemistry and Structural Biology* Ph.D. Thesis State University of New York at Stony Brook., Stony Brook.

22. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) Nmrpipe - A multidimensional spectral processing system based on unix pipes, *J. Biomol. NMR 6*, 277-293.

23. Johnson, B. A. (2004) Using NMRView to visualize and analyze NMR spectra of macromolecules, *Meth. Mol. Biol. 278*, 313-352.

24. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant M-Values and Heat-Capacity Changes - Relation to Changes in Accessible Surface-Areas of Protein Unfolding, *Protein. Sci. 4*, 2138-2148.

25. Cho, J. H., Sato, S., and Raleigh, D. P. (2004) Thermodynamics and kinetics of non-native interactions in protein folding: A single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state, *J. Mol. Biol. 338*, 827-837.

References

1. Anfinsen, C. B. (1973) Principles that govern folding of protein chains, *Science 181*, 223-230.

2. Levinthal, C. (1968) Are there pathways for protein folding?, *J. Chim. Phys. Pcb. 65*, 44-&.

3. Bartlett, A. I., and Radford, S. E. (2009) An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms, *Nat. Struct. Mol. Biol. 16*, 582-588.

4. Dill, K. A., Bromberg, S., Yue, K. Z., Fiebig, K. M., Yee, D. P., Thomas, P. D., and Chan, H. S. (1995) Principles of protein-folding - A perspective from simple exact models, *Protein Sci. 4*, 561-602.

5. Magg, C., and Schmid, F. X. (2004) Rapid collapse precedes the fast two-state folding of the cold shock protein, *J. Mol. Biol. 335*, 1309-1323.

6. Ratner, V., Amir, D., Kahana, E., and Haas, E. (2005) Fast collapse but slow formation of secondary structure elements in the refolding transition of E. coli adenylate kinase, *J. Mol. Biol. 352*, 683-699.

7. Magg, C., Kubelka, J., Holtermann, G., Haas, E., and Schmid, F. X. (2006) Specificity of the initial collapse in the folding of the cold shock protein, *J. Mol. Biol. 360*, 1067-1080.

8. Arai, M., Kondrashkina, E., Kayatekin, C., Matthews, C. R., Iwakura, M., and Bilsel, O. (2007) Microsecond hydrophobic collapse in the folding of Escherichia coli dihydrofolate reductase, an alpha/beta-type protein, *J. Mol. Biol. 368*, 219-229.

9. Daggett, V., and Fersht, A. R. (2003) Is there a unifying mechanism for protein folding?, *Trends Biochem. Sci. 28*, 18-25.

10. Snow, C. D., Nguyen, N., Pande, V. S., and Gruebele, M. (2002) Absolute comparison of simulated and experimental protein-folding dynamics, *Nature 420*, 102-106.

11. Santoro, M. M., and Bolen, D. W. (1988) Unfolding free-energy changes determined by the linear extrapolation method .1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants, *Biochemistry 27*, 8063-8068.

12. Tanford, C. (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation., *Advan. Protein Chem. 24*, 1-95.

13. Chan, C. K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A., and Hofrichter, J. (1997) Submillisecond protein folding kinetics studied by ultrarapid mixing, *Proc. Natl. Acad. Sci. U.S.A. 94*, 1779-1784.

14. Hagen, S. J., and Eaton, W. A. (2000) Two-state expansion and collapse of a polypeptide, *J. Mol. Biol. 301*, 1019-1027.

15. Jones, C. M., Henry, E. R., Hu, Y., Chan, C. K., Luck, S. D., Bhuyan, A., Roder, H., Hofrichter, J., and Eaton, W. A. (1993) Fast events in protein-folding initiated by nanosecond laser photolysis, *Proc. Natl. Acad. Sci. U.S.A. 90*, 11860-11864.

16. Yang, W. Y., and Gruebele, M. (2003) Folding at the speed limit, *Nature 423*, 193-197.

17. Wu, Y., Kondrashkina, E., Kayatekin, C., Matthews, C. R., and Bilsel, O. (2008) Microsecond acquisition of heterogeneous structure in the folding of a TIM barrel protein, *Proc. Natl. Acad. Sci. U.S.A. 105*, 13367-13372.

18. Vu, D. M., Myers, J. K., Oas, T. G., and Dyer, R. B. (2004) Probing the folding and unfolding dynamics of secondary and tertiary structures in a three-helix bundle protein, *Biochemistry 43*, 3582-3589.

19. Fafarman, A. T., Webb, L. J., Chuang, J. I., and Boxer, S. G. (2006) Site-specific conversion of cysteine thiols into thiocyanate creates an IR probe for electric fields in proteins, *J. Am. Chem. Soc. 128*, 13356-13357.

20. Mukherjee, S., Chowdhury, P., DeGrado, W. F., and Gai, F. (2007) Site-specific hydration status of an amphipathic peptide in AOT reverse micelles, *Langmuir 23*, 11174-11179.

21. Muralidharan, V., Cho, J. H., Trester-Zedlitz, M., Kowalik, L., Chait, B. T., Raleigh, D. P., and Muir, T. W. (2004) Domain-specific incorporation of noninvasive optical probes into recombinant proteins, *J. Am. Chem. Soc. 126*, 14004-14012.

22. Miyake-Stoner, S. J., Miller, A. M., Hammill, J. T., Peeler, J. C., Hess, K. R., Mehl, R. A., and Brewer, S. H. (2009) Probing Protein Folding Using Site-Specifically Encoded Unnatural Amino Acids as FRET Donors with Tryptophan, *Biochemistry 48*, 5953-5962.

23. Ye, S. X., Zaitseva, E., Caltabiano, G., Schertler, G. F. X., Sakmar, T. P., Deupi, X., and Vogel, R. (2010) Tracking G-protein-coupled receptor activation using genetically encoded infrared probes, *Nature 464*, 1386-U1314.

24. Kent, S. B. H. (1988) Chemical synthesis of peptides and proteins, *Annu. Rev. Biochem. 57*, 957-989.

25. Dawson, P. E., and Kent, S. B. H. (2000) Synthesis of native proteins by chemical ligation, *Annu. Rev. Biochem. 69*, 923-960.

26. Muralidharan, V., and Muir, T. W. (2006) Protein ligation: an enabling technology for the biophysical analysis of proteins, *Nat. Methods 3*, 429-438.

27. Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code, *Annu. Rev. Biophys. Biomol. Struc. 35*, 225-249.

28. Hendrickson, W. A., Horton, J. R., and Lemaster, D. M. (1990) Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD) - A vehicle for direct determination of 3-dimensional structure, *EMBO J. 9*, 1665-1672.

29. Kiick, K. L., and Tirrell, D. A. (2000) Protein engineering by in vivo incorporation of non-natural amino acids: Control of incorporation of methionine analogues by methionyl-tRNA synthetase, *Tetrahedron 56*, 9487-9493.

30. Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation, *Proc. Natl. Acad. Sci. U.S.A. 99*, 19-24.

31. Tucker, M. J., Oyola, R., and Gai, F. (2006) A novel fluorescent probe for protein binding and folding studies: p-cyano-phenylalanine, *Biopolymers 83*, 571-576.

32. Tucker, M. J., Oyola, R., and Gai, F. (2005) Conformational distribution of a 14-residue peptide in solution: A fluorescence resonance energy transfer study, *J. Phys. Chem. B 109*, 4788-4795.

33. Tang, J., Yin, H., Qiu, J. D., Tucker, M. J., DeGrado, W. F., and Gai, F. (2009) Using two fluorescent probes to dissect the binding, insertion, and dimerization kinetics of a model membrane peptide, *J. Am. Chem. Soc. 131*, 3816.

34. Tucker, M. J., Tang, J., and Gai, F. (2006) Probing the kinetics of membrane-mediated helix folding, *J. Phys. Chem. B 110*, 8105-8109.

35. Marek, P., Gupta, R., and Raleigh, D. P. (2008) The fluorescent amino acid p-cyanophenylalanine provides an intrinsic probe of amyloid formation, *Chembiochem 9*, 1372-1374.

36. Schultz, K. C., Supekova, L., Ryu, Y. H., Xie, J. M., Perera, R., and Schultz, P. G. (2006) A genetically encoded infrared probe, *J. Am. Chem. Soc. 128*, 13984-13985.

37. Oh, K. I., Lee, J. H., Joo, C., Han, H., and Cho, M. (2008) beta-Azidoalanine as an IR probe: Application to amyloid A beta(16-22) aggregation, *J Phys Chem B 112*, 10352-10357.

38. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., and Cate, J. H. D. (2005) Structures of the bacterial ribosome at 3.5 angstrom resolution, *Science 310*, 827-834.

39. Hoffman, D. W., Cameron, C. S., Davies, C., White, S. W., and Ramakrishnan, V. (1996) Ribosomal protein L9: a structure determination by the combined use of X-ray crystallography and NMR spectroscopy, *J. Mol. Biol. 264*, 1058-1071.

40. Herr, A. J., Nelson, C. C., Wills, N. M., Gesteland, R. F., and Atkins, J. F. (2001) Analysis of the roles of tRNA structure, ribosomal protein L9, and the bacteriophage T4 gene 60 bypassing signals during ribosome slippage on mRNA, *J. Mol. Biol. 309*, 1029-1048.

41. Kuhlman, B., Boice, J. A., Fairman, R., and Raleigh, D. P. (1998) Structure and stability of the N-terminal domain of the ribosomal protein L9: evidence for rapid two-state folding, *Biochemistry 37*, 1025-1032.

42. Kuhlman, B., and Raleigh, D. P. (1998) Global analysis of the thermal and chemical denaturation of the N-terminal domain of the ribosomal protein L9 in H2O and D2O. Determination of the thermodynamic parameters, Delta H degrees, Delta S degrees, and Delta C degrees(p), and evaluation of solvent isotope effects, *Protein Sci. 7*, 2405-2412.

43. Cho, J. H., Sato, S., and Raleigh, D. P. (2004) Thermodynamics and kinetics of non-native interactions in protein folding: A single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state, *J. Mol. Biol. 338*, 827-837.

44. Cho, J. H., and Raleigh, D. P. (2005) Mutational analysis demonstrates that specific electrostatic interactions can play a key role in the denatured state ensemble of proteins, *J. Mol. Biol. 353*, 174-185.

45. Anil, B., Li, Y., Cho, J. H., and Raleigh, D. P. (2006) The unfolded state of NTL9 is compact in the absence of denaturant, *Biochemistry 45*, 10110-10116.

46. Selkoe, D. J. (2003) Folding proteins in fatal ways, *Nature 426*, 900-904.

47. Dyer, R. B. (2007) Ultrafast and downhill protein folding, *Curr. Op. Struc. Biol. 17*, 38-47.

48. Kubelka, J., Hofrichter, J., and Eaton, W. A. (2004) The protein folding 'speed limit', *Curr. Op. Struc. Biol. 14*, 76-88.

49. Fersht, A. R. (2000) *Structure and Mechanism in Protein Science - A Guide to Enzyme Catalysis and Protein Folding*, W. H. Freeman & Company, New York.

50. Royer, C. A. (2006) Probing protein folding and conformational transitions with fluorescence, *Chem. Rev. 106*, 1769-1784.

51. Tang, J., Signarvic, R. S., DeGrado, W.F., and Gai, F. (2007) Role of helix nucleation in the kinetics of binding of Mastoparan X to phospholipid bilayer, *Biochemistry 46*, 13856-12863.

52. Tucker, M. J., Tang J., and Gai F. (2006) Probing the kinetics of membrane-mediated helix folding, *J. Phys. Chem. B 110*, 8105-8109.

53. Marek, P., Gupta, R., and Raleigh, D.P. (2008) The fluorescent amino acid p-cyanophenylalanine provides an intrinsic probe of amyloid formation, *ChemBioChem 9*, 1372-1374.

54. Aprilakis, K. N., Taskent, H., and Raleigh, D. P. (2007) Use of the novel fluorescent amino acid p-cyanophenylalanine offers a direct probe of hydrophobic core formation during the folding of the N-terminal domain of the ribosomal protein L9 and provides evidence for two-state folding, *Biochemistry 46*, 12308-12313.

55. Getahun, Z., Huang, C. Y., Wang, T., De Leon, B., DeGrado, W. F., and Gai, F. (2003) Using nitrile-derivatized amino acids as infrared probes of local environment, *J. Am. Chem. Soc. 125*, 405-411.

56. Suydam, I. T., Snow, C. D., Pande, V. S., and Boxer, S. G. (2006) Electric fields at the active site of an enzyme: direct comparison of experiment with theory, *Science 313*, 200-204.

57. Anil, B., Sato, S., Cho, J. H., and Raleigh, D. P. (2005) Fine structure analysis of a protein folding transition state; distinguishing between hydrophobic stabilization and specific packing, *J. Mol. Biol. 354*, 693-705.

58. Horng, J. C., and Raleigh, D. P. (2003) Phi-Values beyond the ribosomally encoded amino acids: kinetic and thermodynamic consequences of incorporating trifluoromethyl amino acids in a globular protein, *J. Am. Chem. Soc. 125*, 9286-9287.

59. Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code, *Annu. Rev. of Biophys. & Biomol. Struc. 35*, 225-249.

60. Jackson, J. C., Hammill, J. T., and Mehl, R. A. (2007) Site-specific incorporation of a (19)F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity, *J. Am. Chem. Soc. 129*, 1160-1166.

61. Hammill, J. T., Miyake-Stoner, S., Hazen, J.L., Jackson, J.C. and Mehl, R.A. (2007) Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization, *Nature Prot. 2*, 2601-2607.

62. Brunger, A. T., and Karplus, M. (1988) Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison, *Proteins 4*, 148-156.

63. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) Charmm - a Program for Macromolecular Energy, Minimization, and Dynamics Calculations, *J. Comp. Chem. 4*, 187-217.

64. MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B 102*, 3586-3616.

65. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD, *J. Comp. Chem. 26*, 1781-1802.

66. Wang, L., Brock, A., Herberich, B., and Schultz, P. G. (2001) Expanding the genetic code of Escherichia coli, *Science 292*, 498-500.

67. Yin, D. (1997) Ph.D. Thesis, "Parameterization for Empirical Force Field Calculations and A Theoretical Study of Membrane Permeability of Pyridine Derivative", *Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland*.

68. Cho, J. H., Sato, S., and Raleigh, D. P. (2004) Thermodynamics and kinetics of non-native interactions in protein folding: a single point mutant significantly stabilizes the N-terminal domain of L9 by modulating non-native interactions in the denatured state, *J. Mol. Biol. 338*, 827-837.

69. Link, A. J., Mock, M. L., and Tirrell, D. A. (2003) Non-canonical amino acids in protein engineering, *Curr. Op. Biotech. 14*, 603-609.

70. Levinthal, C. (1968) Are there pathways for protein folding?, *J. Chim. Phys. Phys-Chim. Biol. 65*, 44-45.

71. Onuchic, J. N., Nymeyer, H., Garcia, A. E., Chahine, J., and Socci, N. D. (2000) The energy landscape theory of protein folding: insights into folding mechanisms and scenarios, *Advances in Protein Chemistry, Vol 53 53*, 87-152.

72. Dobson, C. M. (2003) Protein folding and misfolding, *Nature 426*, 884-890.

73. Rose, G. D., Fleming, P. J., Banavar, J. R., and Maritan, A. (2006) A backbone-based theory of protein folding, *Proc. Nat. Ac. Sci. U.S.A. 103*, 16623-16633.

74. Efimov, A. V. (1994) Common structural motifs in small proteins and domains, *Febs Lett. 355*, 213-219.

75. Kuhlman, B., Luisi, D. L., Evans, P. A., and Raleigh, D. P. (1998) Global analysis of the effects of temperature and denaturant on the folding and unfolding kinetics of the N-terminal domain of the protein L9, *J. Mol. Biol. 284*, 1661-1670.

76. Anil, B., Sato, S., Cho, J. H., and Raleigh, D. P. (2005) Fine structure analysis of a protein folding transition state; distinguishing between hydrophobic stabilization and specific packing, *J. Mol. Biol. 354*, 693-705.

77. Luisi, D. L., and Raleigh, D. P. (2000) pH-dependent interactions and the stability and folding kinetics of the n-terminal domain of L9. Electrostatic interactions are only weakly formed in the transition state for folding, *J. Mol. Biol. 299*, 1091-1100.

78. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves, *Met. Enzymol. 131*, 266-280.

79. Vugmeyster, L., Kuhlman, B., and Raleigh, D. P. (1998) Amide proton exchange measurements as a probe of the stability and dynamics of the N-terminal domain of the ribosomal protein L9: comparison with the intact protein, *Protein Sci. 7*, 1994-1997.

80. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding, *Protein Sci. 4*, 2138-2148.

81. Taskent-Sezgin, H., Chung, J., Patsalo, V., Miyake-Stoner, S. J., Miller, A. M., Brewer, S. H., Mehl, R. A., Green, D. F., Raleigh, D. P., and Carrico, I. (2009) Interpretation of p-cyanophenylalanine fluorescence in proteins in terms of solvent exposure and contribution of side-chain quenchers: A combined fluorescence, IR and molecular dynamics study, *Biochemistry 48*, 9040-9046.

82. Tanford, C. (1970) Protein denaturation. C. Theoretical models for the mechanism of denaturation, *Advan. Protein Chem. 24*, 1-95.

83. Horng, J. C., and Raleigh, D. P. (2003) Phi-Values beyond the ribosomally encoded amino acids: kinetic and thermodynamic consequences of incorporating trifluoromethyl amino acids in a globular protein, *J. Am. Chem. Soc. 125*, 9286-9287.

84. Anil, B., Song, B. B., Tang, Y. F., and Raleigh, D. P. (2004) Exploiting the right side of the ramachandran plot: substitution of glycines by D-alanine can significantly increase protein stability, *J. Am. Chem. Soc. 126*, 13194-13195.

85. Bann, J. G., and Frieden, C. (2004) Folding and domain-domain interactions of the chaperone PapD measured by F-19 NMR, *Biochemistry 43*, 13775-13786.

86. Schnolzer, M., Alewood, P., Jones, A., Alewood, D., and Kent, S. B. H. (1992) In situ neutralization in Boc-chemistry solid-phase peptide-synthesis - Rapid, high-yield assembly of difficult sequences, *Int. J. Pep. Prot. Res. 40*, 180-193.

87. Lakowicz, J. R. (2006) *Principles of Fluorescence Spectroscopy*, Springer Science+Business Media, LLC New York, NY.

88. Beechem, J. M., and Brand, L. (1985) Time-resolved fluorescence of proteins, *Annu. Rev. Biochem. 54*, 43-71.

89. Ross, J. B. A., Szabo, A. G., and Hogue, C. W. V. (1997) Enhancement of protein spectra with tryptophan analogs: Fluorescence spectroscopy of protein-protein and protein-nucleic acid interactions, *Fluorescence Spectroscopy 278*, 151-190.

90. Wong, C. Y., and Eftink, M. R. (1998) Incorporation of tryptophan analogues into staphylococcal nuclease, its V66W mutant, and Delta 137-149 fragment: Spectroscopic studies, *Biochemistry 37*, 8938-8946.

91. Twine, S. M., and Szabo, A. G. (2003) Fluorescent amino acid analogs, *Method Enzymol 360*, 104-127.

92. Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code, *Annu. Rev. Biophys.Biomol. Struc. 35*, 225-249.

93. Rogers, J. M. G., Lippert, L. G., and Gai, F. (2010) Non-natural amino acid fluorophores for one- and two-step fluorescence resonance energy transfer applications, *Anal. Biochem. 399*, 182-189.

94. Tucker, M. J., Getahun, Z., Nanda, V., DeGrado, W. F., and Gai, F. (2004) A new method for determining the local environment and orientation of individual side chains of membrane-binding peptides, *J. Am. Chem. Soc. 126*, 5078-5079.

95. Tucker, M. J., Tang, J., and Gai, F. (2006) Probing the kinetics of membrane-mediated helix folding, *J. Phys. Chem. B 110*, 8105-8109.

96. Liu, J., Strzalka, J., Tronin, A., Johansson, J. S., and Blasie, J. K. (2009) Mechanism of interaction between the general anesthetic halothane and a model ion channel protein, II: fluorescence and vibrational spectroscopy using a cyanophenylalanine probe, *Biophys. J. 96*, 4176-4187.

97. Schultz, K. C., Supekova, L., Ryu, Y. H., Xie, J. M., Perera, R., and Schultz, P. G. (2006) A genetically encoded infrared probe, *J. Am. Chem. Soc. 128*, 13984-13985.

98. Tucker, M. J., Oyola, R., and Gai, F. (2005) Conformational distribution of a 14-residue peptide in solution: A fluorescence resonance energy transfer study, *J. Phys. Chem. B 109*, 4788-4795.

99. Glasscock, J. M., Zhu, Y. J., Chowdhury, P., Tang, J., and Gai, F. (2008) Using an amino acid fluorescence resonance energy transfer pair to probe protein unfolding: Application to the villin headpiece subdomain and the LysM domain, *Biochemistry 47*, 11070-11076.

100. Serrano, A. L., Troxler, T., Tucker, M.J., Gai, F. (2010) Photophysics of a fluorescent non-natural amino acid: p-Cyanophenylalanine, *Chem. Phys. Lett. 487*, 303-306.

101. Yoshida, M., Kaneko, H., Kitamura, A., Ito, T., Oohashi, K., Morikawa, N., Sakuragi, H., and Tokumaru, K. (1976) Mechanism for light-induced hydrogen isotope exchange in benzonitrile, *B. Chem. Soc. Jpn. 49*, 1697-1700.

102. Chen, Y., and Barkley, M. D. (1998) Toward understanding tryptophan fluorescence in proteins, *Biochemistry 37*, 9976-9982.

103. VanGilst, M., and Hudson, B. S. (1996) Histidine-tryptophan interactions in T4 lysozyme: 'Anomalous' pH dependence of fluorescence, *Biophys. Chem. 63*, 17-25.

104. Marek, P., Mukherjee, S., Zanni, M. T., Raleigh, D. P. (2010) Residue specific, real time characterization of lag phase species and fibril growth during amyloid formation: A combined fluorescence and IR study of p-cyanophenylalanine analogs of islet amyloid polypeptide, *J. Mol. Biol. In press.*

105. Aurora, R., and Rose, G. D. (1998) Helix capping, *Protein Sci. 7*, 21-38.

106. Luo, P. Z., and Baldwin, R. L. (1997) Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of

peptides from trifluoroethanol/water mixtures back to water, *Biochemistry 36*, 8413-8421.

107. Kubelka, J., Chiu, T. K., Davies, D. R., Eaton, W. A., and Hofrichter, J. (2006) Sub-microsecond protein folding, *J. Mol. Biol. 359*, 546-553.

108. Bublitz, G. U., and Boxer, S. G. (1997) Stark spectroscopy: Applications in chemistry, biology, and materials science, *Annu. Rev. Phys. Chem. 48*, 213-242.

109. Zanni, M. T., and Hochstrasser, R. M. (2001) Two-dimensional infrared spectroscopy: a promising new method for the time resolution of structures, *Curr. Opin. Struct. Biol. 11*, 516-522.

110. Xu, Y., Oyola, R., Gai, F. (2003) Infrared study of the stability and folding kinetics of a 15-residue beta-hairpin., *J. Am. Chem. Soc. 125*, 15388-15394.

111. Suydam, I. T., and Boxer, S. G. (2003) Vibrational Stark effects calibrate the sensitivity of vibrational probes for electric fields in proteins, *Biochemistry 42*, 12050-12055.

112. Suydam, I. T. (2006) Electric fields at the active site of an enzyme: direct comparison of experiment with theory (vol 313, pg 200, 2006), *Science 313*, 1887-1887.

113. Brewer, S. H., Song, B. B., Raleigh, D. P., and Dyer, R. B. (2007) Residue specific resolution of protein folding dynamics using isotope-edited infrared temperature jump spectroscopy, *Biochemistry 46*, 3279-3285.

114. Hauser, K., Krejtschi, C., Huang, R., Wu, L., and Keiderling, T. A. (2008) Site-specific relaxation kinetics of a tryptophan zipper hairpin peptide using temperature-jump IR spectroscopy and isotopic labeling, *J. Am. Chem. Soc. 130*, 2984-2992.

115. Hunt, N. T. (2009) 2D-IR spectroscopy: ultrafast insights into biomolecule structure and function, *Chem. Soc. Rev. 38*, 1837-1848.

116. Ye, S. X., Huber, T., Vogel, R., and Sakmar, T. P. (2009) FTIR analysis of GPCR activation using azido probes, *Nat. Chem. Biol. 5*, 397-399.

117. Kim, Y. S., and Hochstrasser, R. M. (2009) Applications of 2D IR spectroscopy to peptides, proteins, and hydrogen-bond dynamics, *J. Phys. Chem. B 113*, 8231-8251.

118. Choi, J. H., Oh, K. I., and Cho, M. H. (2008) Azido-derivatized compounds as IR probes of local electrostatic environment: Theoretical studies, *J. Chem. Phys. 129*, 17.

119. Wang, A., Nairn, N. W., Johnson, R. S., Tirrell, D. A., and Grabstein, K. (2008) Processing of N-terminal unnatural amino acids in recombinant human interferon-beta in Escherichia coli, *Chembiochem 9*, 324-330.

120. Kuhlman, B., Luisi, D.L., Evans, P.A., Raleigh, D.P. (1998) Global analysis of the effects of temperature and denaturant on the folding and unfolding kinetics of the N-terminal domain of the protein L9., *J. Mol. Biol. 284*, 1661-1670.

121. Cho, J. H. (2006) The role of the unfolded state in protein stability and protein folding
Department of Biochemistry and Structural Biology Ph.D. Thesis State University of New York at Stony Brook., Stony Brook.

122. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) Nmrpipe - A multidimensional spectral processing system based on unix pipes, *J. Biomol. NMR 6*, 277-293.

123. Johnson, B. A. (2004) Using NMRView to visualize and analyze NMR spectra of macromolecules, *Meth. Mol. Biol. 278*, 313-352.

124. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant M-Values and Heat-Capacity Changes - Relation to Changes in Accessible Surface-Areas of Protein Unfolding, *Protein. Sci. 4*, 2138-2148.