Role of Symmetry in Hisactophilin Folding

by

Fadila Ghashut

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Chemistry

Waterloo, Ontario, Canada, 2009

Author's declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Abstract

Hisactophilin is a histidine-rich actin binding protein from *Dictyostelium discoideum*; it is a member of the β-trefoil superfamily and exhibits a characteristic tertiary structure with three-fold symmetry. The effect of various mutations on stability and folding of other β-trefoil proteins has been studied and certain mutations that increase sequence and structural symmetry were found to increase protein stability. This study focuses on the effects of point mutations that increase symmetry on hisactophilin stability and folding designed within the core, minicore and turns of conserved and unconserved residues. The stability and folding of hisactophilin mutants were analyzed both thermodynamically and kinetically using fluorescence spectroscopy and circular dichroism (CD) at pH 6.7, 7.7, and 8.7. The stability of each mutant was measured and compared with the wild type protein (WT), and with data obtained for another β-trefoil protein, Fibroblast Growth Factor (FGF). The folding of WT and mutant hisactophilin was found to be reversible by chemically induced denaturation. Both the equilibrium and kinetic data in urea for WT and mutant hisactophilin can be analyzed using a 2-state model for the transition between the folded and unfolded state. The results indicated that H90G and I85L are more stable than the WT due to increase in the folding rate, while I93L is slightly less stable than the WT due to an increase in the unfolding rate. F6L is more markedly less stable than the WT as indicated by urea denaturation; due to protein precipitation the kinetic analysis was not performed for this mutation. F13Ymutation is less stable than the WT from CD result. and it needs investigation with DSC to confirm the stability as indicated as a future work. The results of the mutations support the notion of increased symmetry resulting in increased stability.

Acknowledgements

I would like to express my gratitude to my supervisor, Professor Elizabeth Meiering. for her invaluable support and guidance. Her guiding advices lighted up my way to go and left me up in my difficult times. Her caring attitude and kind approaches eased every burden in my way, inspired me and imparted me the potential and courage to peruse and finish my research path.

I would like to thank all my lab members for their help and support. I would also like to thank my family and my friends here and back home for their love, I appreciate your involvement and support.

I also wish to acknowledge the Libyan Educational Program funded by the Libya Government for financial support.

Dedications

This thesis is dedicated to my dear husband, Esam, whose support and encouragement has been an invaluable resource for me.

My lovely son, Taha, whose compromise with a busy mom let me finish this work.

My beloved parents, Taher and Galia, and my dear sister, Farida, whose belief has given me the motivation, confidence and determination to start.

Table of Contents

List of Tables

Chapter 1

Chapter 2

Chapter 3

List of Figures

Chapter 1

Chapter 2

Chapter 3

List of Abbreviations

Amp: Ampicillin

A_{280}: Absorbance at 280 nm

β_T: the Tanford value

CD: Circular dichrosim

D.discoideum: Dictyostelium discoideum

3D: Three Dimension

ddH_2O: distilled deionized water

DNA: Deoxyribonucleic acid

DEAE: diethylaminoethyl

DTT: Dithiothreitol

EDTA: Ethylene Diamine Tetraacetic Acid

dNTP: deoxynucleotide triphosphate

DSC: Differential scanning calorimetry

E.coli: Escherichia coli

ETI: Erythrina caffra inhibitor

F: Folded state

F: Phenylalanine

FGF: Fibroblast growth factor

f_u: Fraction of unfolded protein

ΔG_U: Free energy of unfolding in water

$\Delta\Delta G$: difference in ΔG between mutant and pseudo WT

ΔG: difference in Gibb's free energy between two states

G: Glycine

G_U: Gibbs free energy of the unfolded state

G_N: Gibbs free energy of the native state

H: Histidine

His: Hisactophilin

HisI&HisII: Hisactophilin I and II

IPTG: Isopropyl-β-D-thiogalactopyranoside

I: Isoleucine

kDa: Kilo Dalton

kb: Kilobase Pair(s)

K_u: Equilibrium constant of unfolding

k_f: Folding rate constant

k_u: Unfolding rate constant

kcal: kilocalorie

1L-1β: Interleukin-1β

L: Liter

LB: Luria Broth media

L: Leucine

M: Molar

mg: milligram

MW: Molecular Weight

m: Constant describing the denaturant dependence of ΔG

m_{eq}: m determined from equilibrium denaturation

MS: Mass spectroscopy

NMR: nuclear magnetic resonance

NOE: Nuclear Overhauser Effect

OD: Optical Density

PCR: Polymerase Chain Reaction

pI: Isoelectric Point

pHisI: Hisactophilin plasmid

PDB: Protein Data Bank

PWT: Wild Type hisactophilin with 4 extra amino acid

RTB: Ricin B

SDS-PAGE: Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis

SCOP: Structural Classification of Proteins

TBE: Tris-Borate-EDTA-Buffer

T_m: melting Temperature

Tris-HCl: Tris (hydroxymethyl) aminomethane hydrogen chloride

TS‡: Transition state

U: Unfolded state

WT: Wild type hisactophilin

Y: tyrosine

Chapter 1: Introduction

1.1 Protein folding

Protein folding studies not only provide an intresting intellectual challenge, but also have many practical applications. The information required for a protein to fold and adopt its native structure is stored within the primary sequence (Anfinsen, 1973); however, we are still unable to accurately and universally predict the final fold based solely on primary sequence. During the last several years significant advances have been made in this field. As the number of known structures solved by X-ray crystallography and NMR techniques has increased, additional information has been gathered relating sequence to structure. From this data, it has become clear that protein structure is much more highly conserved throughout evolution than primary sequence (Thornton, 2005). For the pharmaceutical industry, the ability to predict tertiary structure holds the prospects of greatly reducing the cost of developing new therapeutic drugs. Furthermore, recent research suggests that certain diseases such as mad cow disease, cancers, and possibly both Alzheimer's and type II diabetes may be caused by misfolding and aggregation of otherwise normally functional proteins (Kuwajima, 1999).

Finding answers to the protein folding problem is a most worthwhile endeavour, ultimately contributing to the prediction of protein structure based only on primary sequence, to the design of recombinant proteins with a stable fold and to the understanding of diseases involving misfolded proteins. Mutagenesis experiments show that limited changes in sequence can have effects on stability and activity but generally do not lead to large shifts in structure (Cordes et al., 1999).

1.1.1 Folding of small model proteins

Many protein folding investigations have focused on studying small single-domain proteins (up to ~100 amino acid residues) that fold and unfold in a two-state transition between native/folded (F) and denatured/unfolded (U) states without intermediate states significantly populated on the folding pathway (Jackson, 1998). Two state folding can be described using the following scheme:

Scheme 1:

$$F \underset{k_f}{\overset{k_u}{\rightleftharpoons}} U \qquad (1.1)$$

$$K_{U\text{-}F} = [U] / [F] = k_u / k_f \qquad (1.2)$$

$$f_F + f_U = 1 \qquad (1.3)$$

where $K_{U\text{-}F}$ is the equilibrium constant for unfolding, which can be determined by urea denaturation, k_u and k_f are the rate constants for unfolding and refolding, respectively, and f_F and f_U are the fraction of folded and unfolded protein, respectively (Pace, 1986). For a globular single domain protein, hydrophilic residues tend to be on the surface of the molecule. Surface residues are relatively flexible and can adopt alternative conformations. Hydrophobic side chains of residues are generally packed together inside the molecule, excluding water to form anhydrous cores as shown in (Figure 1.1).

.

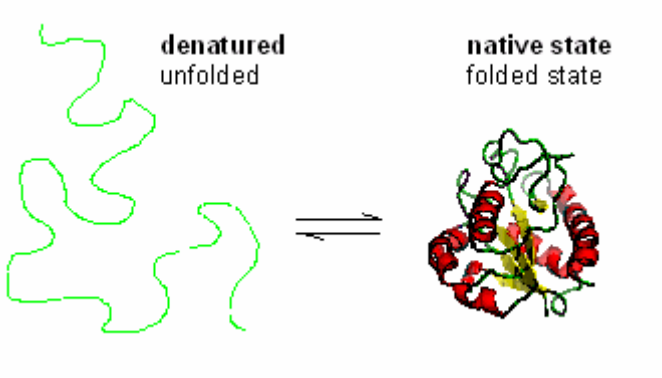


Figure 1. 1 Native and denatured states.
The native state has less hydrophobic surface area exposed to solvent, and in the denatures state more of hydrophobic surface area is exposed to solvent.

Reversible folding of two states is a cooperative reaction, which means that the transition between folding and unfolding occurs over a narrow range of changing environmental conditions.

For larger proteins, three steps are more frequently observed due to formation of an intermediate. The intermediate is either a so-called molten globule, which is formed by a process called hydrophobic collapse (in which all hydrophobic side-chains suddenly fall inside the protein or cluster together) or a structure in which the secondary structures of the protein are already fully formed (Fersht, 1999). However, there is controversy concerning about whether these intermediates are formed en route to the correct folding pattern, or whether they represent structural dead end.

For small proteins, intermediates may tend to act as kinetic traps and slow the folding process. However, intermediates may also play a role in helping proteins fold properly, especially large proteins, or facilitate aggregation. Some proteins related to misfolding diseases involving prion, amyloid formation and protein aggregation are proposed to fold via formation of intermediates which favour protein misfolding (Galani et al., 2002; Zhu et al., 2003). The role of folding intermediates is not well defined and requires further study.

Hisactophilin, the protein which has been studied in this thesis, displays predominantly two-state folding with intermediates observed only under highly stabilizing solution conditions (Liu et al., 2002).

1.2 Protein stability

Protein folding studies rely on measuring thermodynamic and kinetic parameters for folding. Thermodynamics measure the energy difference between the folded and unfolded states at equilibrium, while kinetics measure free energy differences between transition states and the initial ground state (native or unfolded).

Various methods such as fluorescence, circular diochroism and DSC can be used to monitor the structure of these different states. The folded functional state of a protein is generally believed to be energetically favored over other conformations (Pace, 1986). The stability of a small monomeric protein is typically about 5-10 kcal/mol, which is very small.

1.3 Methods to study protein folding

1.3.1 Equilibrium denaturation

One approach to study protein folding is to determine the properties of proteins under equilibrium conditions, and the properties of the transition between different states that are populated at equilibrium.

Proteins can denature by changes of various solution conditions: increasing temperature, pressure, decreasing or increasing pH, and denaturant concentration. Urea and guanidine hydrochloride (GdnHCl) are the most commonly used reagents to unfold proteins. These compounds denature proteins by interfering with the hydrophobic effect and so promote protein unfolding and exposure of hydrophobic groups. The unfolded state of

proteins is more favored than the native state in high denaturants concentrations because the denaturants interact more favorably with hydrophobic groups. Urea is a neutral denaturant, while guanidine is a much stronger denaturant than urea and may be necessary in some systems to completely denature a protein (Jaenicke, 1989).

The unfolded state refers to the highly solvent exposed states that proteins assume under strongly denaturing conditions. The free energy of transfer of the side chains and polypeptide backbone from water to solution of denaturant is linearly proportional to the concentration of denaturant (Fresht, 1997).

Proteins that denature reversibly *in vitro* will reach thermodynamic equilibrium after unfolding/refolding for a sufficient period of time in denaturant. For the 2-state model in which the protein exists in either the unfolded state (U) or the native state (N), the stability is measured by the difference in Gibbs free energy between the N and U (ΔG). This value can be determined from the equilibrium constant as:

$$\Delta G_U = G_U - G_N = -RT\ln K_U \qquad (1.4)$$

where G_U is the Gibbs free energy of the unfolded state, G_N is the Gibbs free energy of the native state, R is the gas constant and T is absolute temperature. Proteins that unfold by a two-state mechanism exhibit a symmetrical sigmoidal transition in denaturant shown in (Figure 1.2).

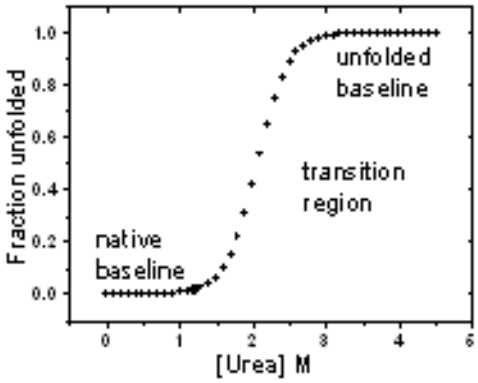


Figure 1. 2 A typical urea denaturation curve for a two state unfolding transition.
The fractional of unfolded protein is plotted as a function of denaturant concentration.

Equilibrium denaturation curves are commonly used to characterize the unfolding transitions of proteins (Pace, 1986). Protein samples are diluted into a series of denaturant concentrations and once the protein sample reaches equilibrium the properties of the protein are measured using optical probes such as CD and fluorescence. Over a range of denaturant concentrations there is a transition region (Figure 1.2), where the protein undergoes a co-operative transition between the native and the unfolded state. The fraction of the unfolded protein, f_u can be calculated using the following equation:

$$f_u = \frac{Y_{obs} - Y_U}{Y_N - Y_U} \tag{1.5}$$

where Y_{obs} is the observed signal, Y_U is the signal of the unfolded state, and Y_N is the signal of the native state. As the denatured state is more exposed to solvent than the native state, the denatured state is preferentially stabilized by denaturant. The free energy of unfolding at any urea concentration, ΔG_U[urea], is expressed by the following equation:

$$\Delta G_U\,[\text{urea}] = \Delta G^\circ_U + m_1\,[\text{urea}] - m_2\,[\text{urea}]^2 \tag{1.6}$$

where ΔG_U[urea] is the free energy of unfolding at any urea concentration, $\Delta G°_U$ is the free energy of unfolding in 0 M urea, and m_1 and m_2 are constants describing the denaturant dependence of the Gibbs free energy of unfolding.

The difference in the free energy of activation ($\Delta\Delta G_u$) as described by Pace and Scholtz between the WT and mutant can be determined with more certainty than the absolute ΔG_u:

$$\Delta\Delta G_u = \Delta G_{WT} - \Delta G_{mut} \quad (1.7)$$

$$= (C_{mWT} - C_{m\ mutant})\ (m_{WT} + m_{mutant})/\ 2 \quad (1.7.1)$$

A positive value of $\Delta\Delta G_u$ indictes that a mutant is less stable than the WT and a negative value of $\Delta\Delta G_u$ indicates that a mutant is more stable than the WT.

1.3.2 Kinetic measurements of unfolding and folding

1.3.2.1 Introduction

Proteins that appear half-denatured in the test tube are, in fact, mixtures of molecules that flip forwards and backwards between the unfolded and fully native states. Individual molecules pass continuously the top of the free energy barrier separating the N and U states. By analogy to chemical reactions, the state corresponding to the free energy barrier between two states is called the transition state (TS). The rate constant for refolding (k_f) and unfolding (k_U) is related to the energy difference between U and TS and between N and TS, respectively, as illustrated in (Figure 1.3).

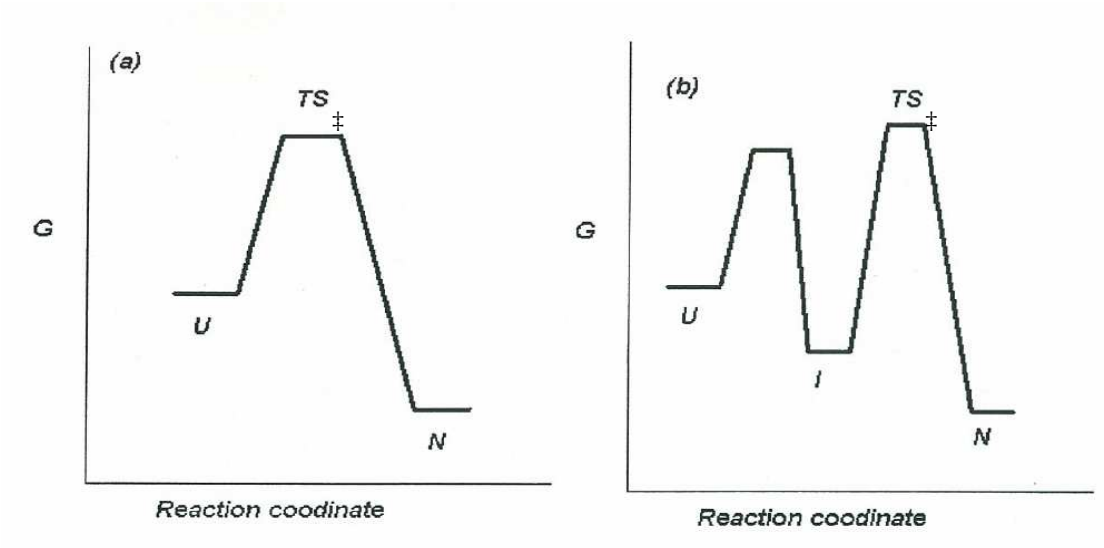


Figure 1. 3 Free energy diagram for folding of protein.
(a). Two state transition. N, U and $TS^{‡}$ represent the native state, unfolded state and the transition state, respectively. (b). Three state folding of protein. I represents the intermediate.

To study refolding, a protein is subjected to denaturing conditions, which then are very rapidly changed back to folding conditions, e.g. by dilution of a high denaturant solution. The reverse is done to study unfolding.

1.3.2.2 One-step folding

The equilibrium constant for unfolding (k_U) is given by the ratio of unfolding and refolding rate constant:

$$K_U = [U]/[N] = k_u/k_f \quad (1.8)$$

After a jump to folding or unfolding solution conditions, relaxation to equilibrium can be monitored spectroscopically and described by a single exponential with a drift:

$$A(t) = a*t + b + c1*\exp(-k_1*t) \quad (1.9)$$

where $A(t)$ is the signal at time t, a is the drift, b is the offset (signal at $A(\infty)$), $c_1$is the amplitude of the signal, and k_1 is the rate constant.

Raw kinetic data that are ill fit to a single exponential may instead fit to a double exponential with drift:

$$A(t)= a*t +b+c1*exp (-k_1*t) + c2*exp ((-k_2*t) \quad (1.10)$$

where *A(t)* is the signal at time *t*, a is the drift, *b* is the offset (signal at *A* (∞)), c_1 is the amplitude of the fast phase with corresponding k_1 rate constant and *c2* is the amplitude of the slow phase with corresponding k_2 rate constant.

$$k_{obs} = k_f + k_u \quad (1.11)$$

At lower concentrations of the denaturant, the folding rate is much larger than the unfolding rate, so that k_{obs} approximates k_f. Conversely, at higher denaturant concentrations, the unfolding rate is much higher than that of folding, and k_{obs} should approximate k_u. The half-life of the reaction is determined as:

$$t_{1/2} = \ln 2/k_{obs} \quad (1.12)$$

1.3.2.3 Chevron plot

The criteria by which a protein can be shown to fold according to two state model has been well studied (Johnson and Fersht, 1995). For such proteins plots of the natural logarithm of the rate constants for unfolding, $\ln k_u$ and refolding, $\ln k_f$, versus denaturent concentration results in a V-shaped profile often referred to as a chevron polt see (Figure 1.4). For a 2-state model the unfolding and refolding part of the logarithm curve are linear. The observed or measured rate constan k_{obs} is the sum of the folding and unfolding rates constants as shown in Equation 1.11 (Fersht, 1999).

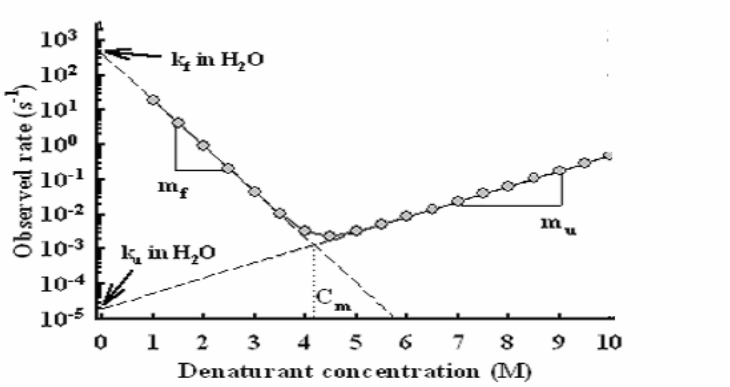


Figure 1. 4 A typical chevron plot observed in two-state model.
Where K_f in H_2O is the rate constant of folding in water, k_u in H_2O is the rate constant of unfolding in water, m_u and m_f are the dependence of $\ln k_u$ and $\ln k_f$ on [urea], and C_m is the denaturation midpoint.

More importantly, equlibrium parameters for the change in free energy of unfolding in the absence of denaturent, ΔG^o_u, and m-values, $m_f^{\ddagger}$ and $m_u^{\ddagger}$, reflect the change in solvent accessibility associated with the formation of the transition state in the refolding and unfolding reaction, respectively. For a 2-state reaction, the addition of $m_f^{\ddagger}$ and $m_u^{\ddagger}$ will equal the m-value measured using equilibrium denaturation experiments:

$$m_{eq} = m_f^{\ddagger} + m_u^{\ddagger} \quad (1.13)$$

Also ΔG can be calculated from the kinetic data and should agree with that obtained directly from equilibrium data (Myers et al., 1995).

$$\Delta G_u = G_U - G_N = -RT \ln(k_{unf}/k_{ref}) \quad (1.14)$$

1.3.2.4 Transition state

The protein must pass through a higher-free-energy transition state between the unfolded and native low-free-energy states. The transition state is intermediate between the folded and unfolded states in the degree of exposure to solvent as shown in (Figure 1.3) (Matouschek et al., 1994). The transition state is characterized by two properties which are

the Tanford value, β_T, and the Phi value, Φ (Fersht, 1999). The Tanford value, β_T, is a measure of the average degree of exposure in the transition state relative to the state that of unfolding from folding and, as such, an indicator of the compactness of the transition state ensemble. (β_T) can be calculated from the equilibrium and kinetic m-values (Matouschek and Fersht, 1993).

$$\beta_T = 1- (m_{unf} / (m_{unf} + m_{ref})) \quad (1.15)$$

When β_T is close to 1, the transition state solvent-accessible surface area is close to that of native protein,where as β_T is close to 0, the transition state is close to the unfolded state (Matouschek et al., 1994).

The Φ values are another measure of structure in protein folding transitions states, which give site specific information. Φ is defined as the ratio of the energetic destabilization introduced by the mutation to the transition state versus that introduced to the native folded state. The phi values have been used extensively to characterize the structure of protein folding transition states at atomic resolution (Fersht., 1999). In Equation (1.16), Φ_u is the ratio of the difference in free energy determined in the unfolding direction between mutant and wild-type protein to the difference in the overall free energy:

$$\Phi_u = \Delta\Delta G_u^{\ddagger} / \Delta\Delta G_{eq} \quad (1.16)$$

where $\Delta\Delta G_u^{\ddagger}$ represent the energy difference between the transition state and the folded state for the WT protein, $\Delta\Delta G_{eq}$ represent the energy difference between the native state and the denatured state.

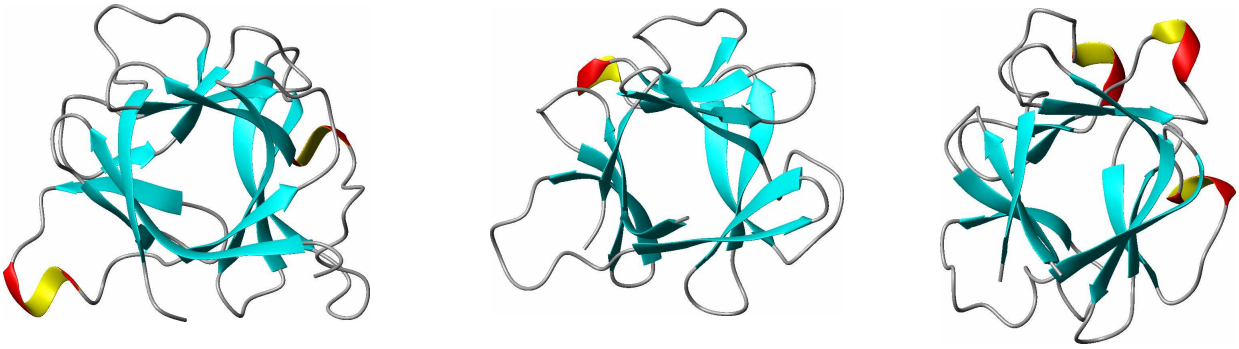
$$\Delta\Delta G_u^{\ddagger} = -RT\ln (k_{f'}/k_f) \quad (1.17)$$

Φ-value near 0 indicates that the region surrounding the mutation is relatively unfolded in the transition state, while a Φ=1 indicates that the local structure around the mutation site in the transition state closely resembles the structure in the native state.

β-trefoil proteins are of interest for folding studies because this is an abundant all β-fold, that is relatively large and may exhibit interesting complexities relative to smaller proteins. The structure and folding of β-trefoil proteins will discussed in more detail in the next section.

1.4 β-Trefoil proteins

Beta-trefoil proteins possess distinct functions and occur in different subcellular locations. Proteins with this structural fold include many superfamilies: Cytokines (including basic and acidic fibroblast growth factors (FGF), Interleukin-1β (IL-1β)); MIR domain; Ricin B-like; Agglutinin; STI-like (Kunitz (STI) inhibitors; Clostridium neurotoxins); DNA-binding protein LAG-1; AbfB domain; and Actin-crosslinking proteins (Fascin, Hisactophilin) (Andreeva et al., 2004). The β-trefoil structure (Murzin, et al., 1992) is formed by a six-stranded β-barrel closed off at one end by three β-hairpin structure, illustrated in (Figure 1.5), which shows ribbon representations of 3 β-trefoil proteins whose folding has been characterized in detail: hisactophilin, interleukin-1β and fibroblast growth factor (FGF).



Interleukin 1-β (2I1B) **Hisactophilin (1HCD)** **FGF (2AFG)**

Figure 1. 5 Ribbon representation of β-trefoil proteins.
β-strand are shown as blue arrows, and loops as the gray tubes, this diagram was prepared using the program MolMol, (Koradi et al., 1996) and PDB accession codes 1HCD, 2I1B, and 2AFG.

β-trefoil proteins fold very slowly despite having a low contact order. It has been proposed that all β-trefoil proteins are homologues, descended from a common ancestor (Ponting and

Russell, 2000). These β-trefoil proteins have less than 10% sequence homology, but nevertheless retain the same overall fold (Murzin et al., 1992). In β-trefoil proteins the arrangement of the three units gives the molecule a pseudo 3-fold axis of symmetry. Hisactophilin and other β-trefoil protein family members are interesting model systems for studying protein folding, and how highly diverse sequences can adopt highly similar structures, due to the wealth of information that has been collected on these systems.

1.4.1 Duplications

The structure of β-trefoil proteins can be equally divided into three subdomains of approximately 40-60 residues, which have structural homology each related by a pseudo-threefold axis of symmetry through the barrel axis. There is a high degree of sequence and structural conservation between the three subdomains which suggests that all β-trefoil proteins have evolved from a series of gene duplication (Figure 1.6 (Dubey et al., 2005)). (Figure 1.6) illustrates how gene duplication events could give rise to a beta-trefoil domain.

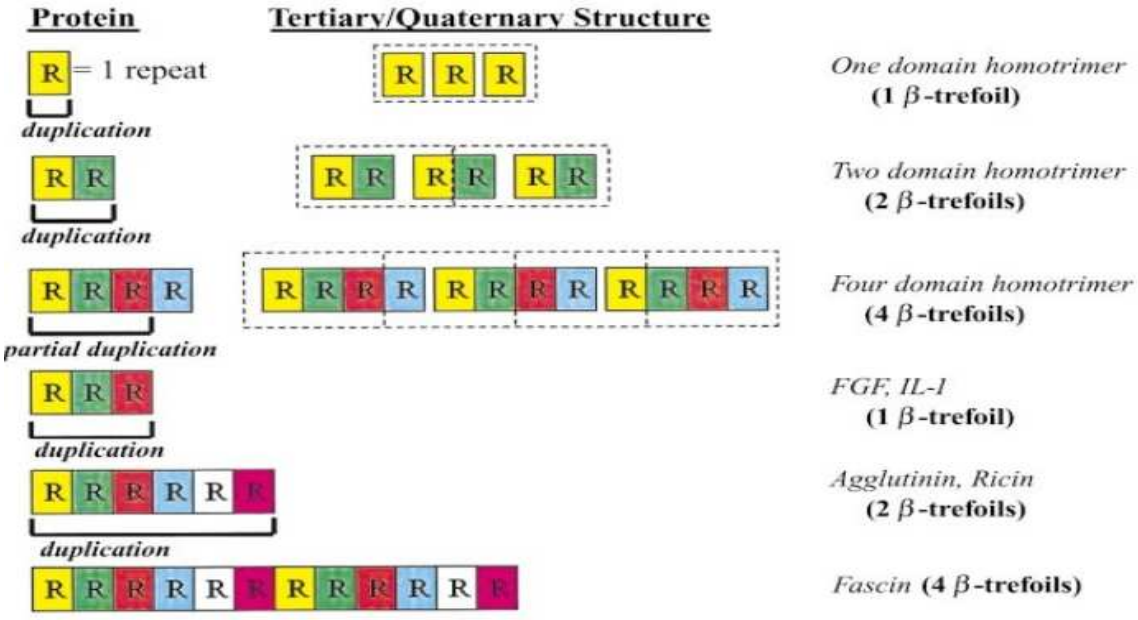


Figure 1. 6 Hypothetical scheme for β-trefoil protein.
Show how modern β-trefoil proteins could have arisen from a progenitor single repeat (R) that formed homotrimers from a series of gene duplication events (Ponting and Russell, 2000).

1.4.2 Common ancestor

Proteins that share similar function often have similar folds; this is thought to be a result of descending from a common ancestral protein. Within a superfamily however, functional divergence is actually very common (Watson et al., 2005). In addition to analyzing structure and sequence databases (Figure 1.7), analysis of the sequence space of all β-trefoil proteins have lead to conclusion that a second group of actin binding proteins (Fascins) and other proteins of unknown function are β-trefoil homologues (Ponting and Russell, 2000).

As structure is better conserved than sequence, hisactophilin likely evolved from a common ancestor. There are however, residues within the proteins that are conserved to allow for the high similarity in folds (Murzin et al., 1992). These minor sequence similarities as shown in (Figure 1.7) have led (Ponting and Russell, 2000) to the conclusion that β-trefoil proteins may have descended from a common protein ancestor which was a homotrimer of trefoil elements. Hisactophilin is a β-trefoil protein is an interesting model system for studying protein folding, which will be discussed further in the following section.

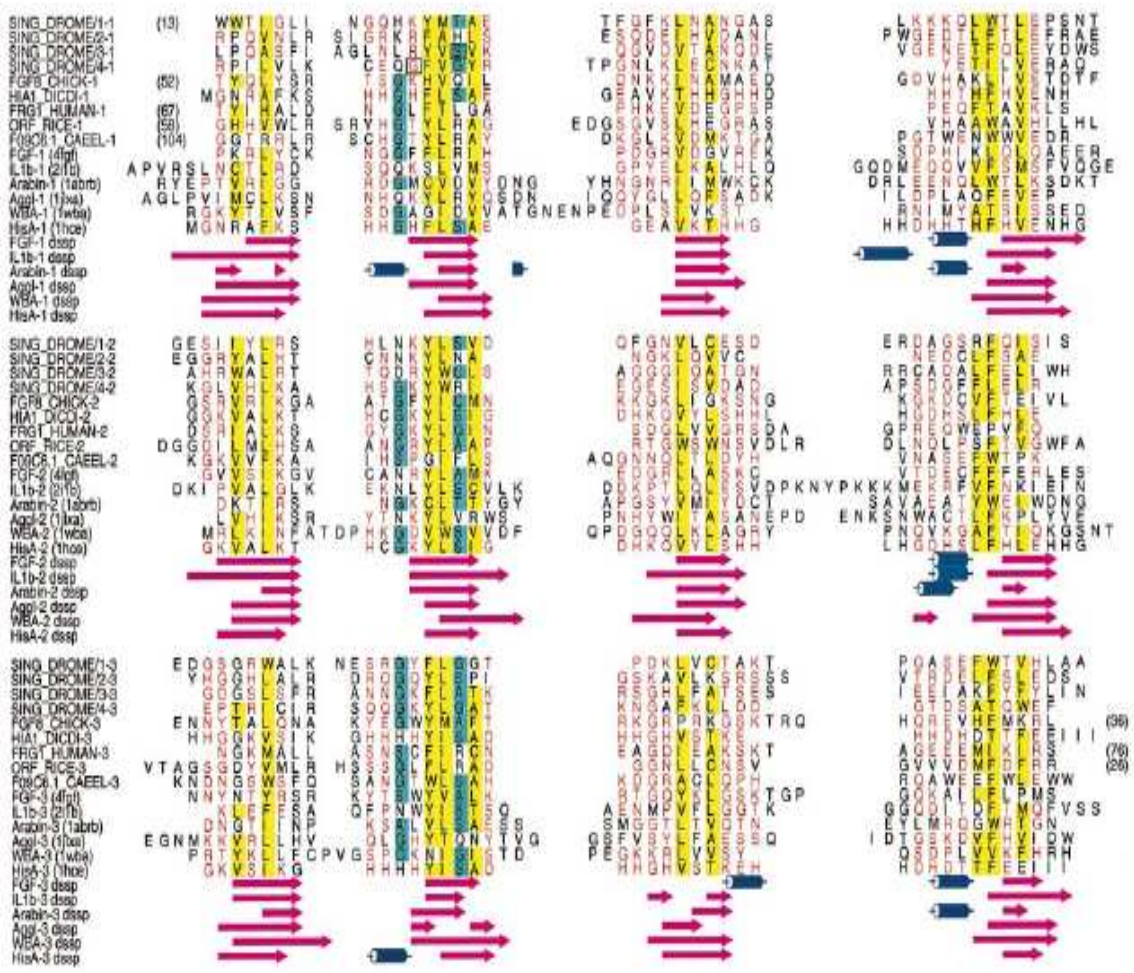


Figure 1. 7 Multiple alignments of representative β-trefoil proteins sequences.
Residues are colored if >60% of sequences at a position share the same amino acid property. Yellow colour is hydrophobic amino acids; blue is polar amino acids; secondary structure is shown beneath the sequences for some proteins (reproduced from ((Ponting and Russell, 2000)).

1.5 Hisactophilin

Hisactophilin is a histidine rich, actin binding protein found in the highly motile amoeboid cells of the slime mold, *Dictyostelium discoideum* (Scheel et al., 1989). In *D. discoideum*, hisactophilin exists free in the cytoplasm as well as membrane bound (Scheel et al., 1989). Also, this protein exists in two isoforms (denoted as HisI (13.5 kDa) and HisII (13.7 kDa)), which are very similar in sequence and function (Hanakam et al., 1996). We will focus here on HisI, which is the form that has been characterized in most studies.

Hisactophilin consists of 118 amino acid residues of which 31 are histidine for HisI .The primary amino acid sequence of hisactophilin is shown in (Figure 1.8) 28 of the histidines are located on the surface of the protein in turns and loops exposed to solvent (Figure 1.9).

GNRAF^{6}KSHHGHF^{13}LSAEGEAVKTHHGHHDHHTHFHVENHG

GKVALKTHCGKYLSIGDHKQVYLSHHLHGDHSLFHLEHHG

GKVSI^{85}KGHHH^{90}HYI^{93}SADHHGHVSTKEHHDHDTTFEEIII

Figure 1. 8 Primary sequence of hisactophilin.
The sequence is shown on 3 lines which correspond to the 3 trefoils. The 3 trefoil units are aligned to show the internal 3-fold symmetry characteristic of the β-trefoil fold. Colored numbers indicated the mutation in the hisactophilin gene that are characterized in this thesis.

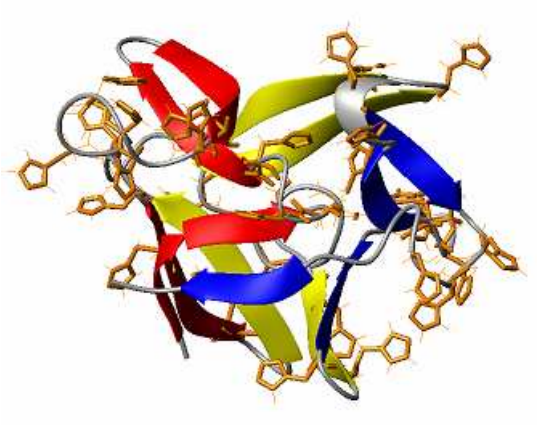


Figure 1. 9 Side view of hisactophilin structure.
The structure is colored to highlight β-hairpin formed by β-strands: 1- 4 (red), 5-8 (yellow), 9-12 (blue). The 31 histidine residues are shown in yellow. The ribbon diagram was prepared using the program MolMol (Koradi et al., 1996).

The histidine residues have pKa values of ~ 6.8 (Hammond et al., 1998) (Figure 1.10) which favours binding of hisactophilin to negatively charged species (like Actin) at pH lower than ~7. The high histidine content results in an isoelectric point of 6.9. Most of the residues are hydrophilic which makes it water soluble. Hisactophilin is folded within the pH range from 4.7~10.7; below 4.7 it exists in an unfolded state.

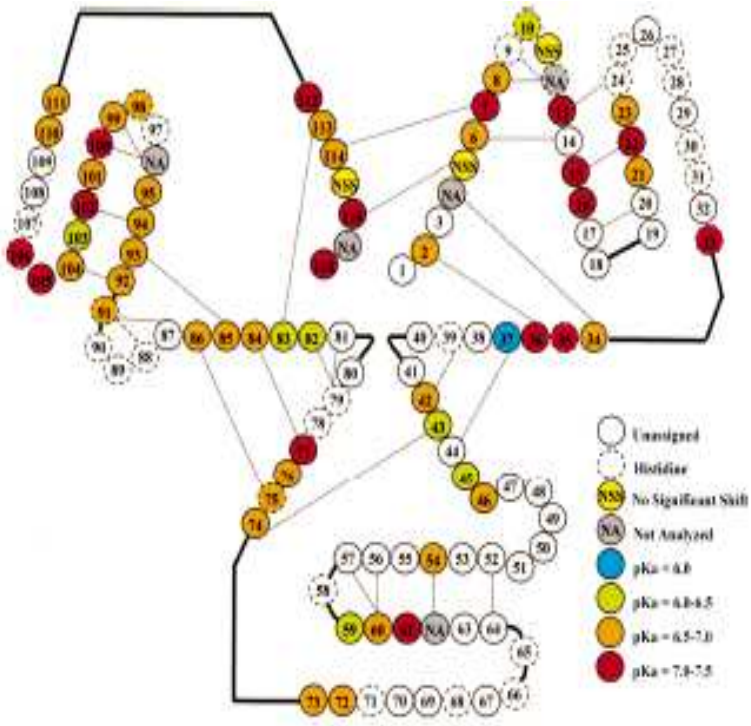


Figure 1. 10 Bubble diagram representation of the structure of hisactophilin.
Each residue in the protein is depicted by a circle, with residue numbers given inside the circles. Colour coding represents the different pKa values for each amino acid. Hydrogen bonds between residues are shown by broken lines. From this figure the β-trefoil structure is clearly evident. The six β-strands in the centre make up the β-barrel, while the three β-hairpins are shown around the edges (Hammond et al., 1998).

1.5.1 Hisactophilin structure

The three-dimensional structure of hisactophilin has been solved by NMR spectroscopy (Habazettl et al., 1992). The structure of β-trefoil fold in hisactophilin consists of 12 antiparallel β-strands arranged around a 3-fold axis of symmetry.

Each trefoil unit consists of a β-β-β-loop-β motif, where the first and the fourth β-strand of each unit make up an anti-parallel β-sheet. Three of these β-sheets (one from each trefoil unit) form a β-barrel (strands β1, β4, β5, β8, β9, β12) (Figure 1.10). The three remaining pairs of antiparallel β-hairpins (the strands β2-β3, β6-β7, and β10-β11) pack together to form a triangular hairpin triplet (Figure 1.10). The N- and the C-terminus are located close to the central axis (Habazettl et al., 1992). There are 18 structurally conserved

hydrophobic residues in β-trefoil proteins that can be subdivided into four layers (Figure 1.12), middle β-barrel layer (M), bottom β-barrel layer (B), upper β-hairpin layer (U), and lower β-hairpin layer (L) (Figures 1.11 and 1.12).

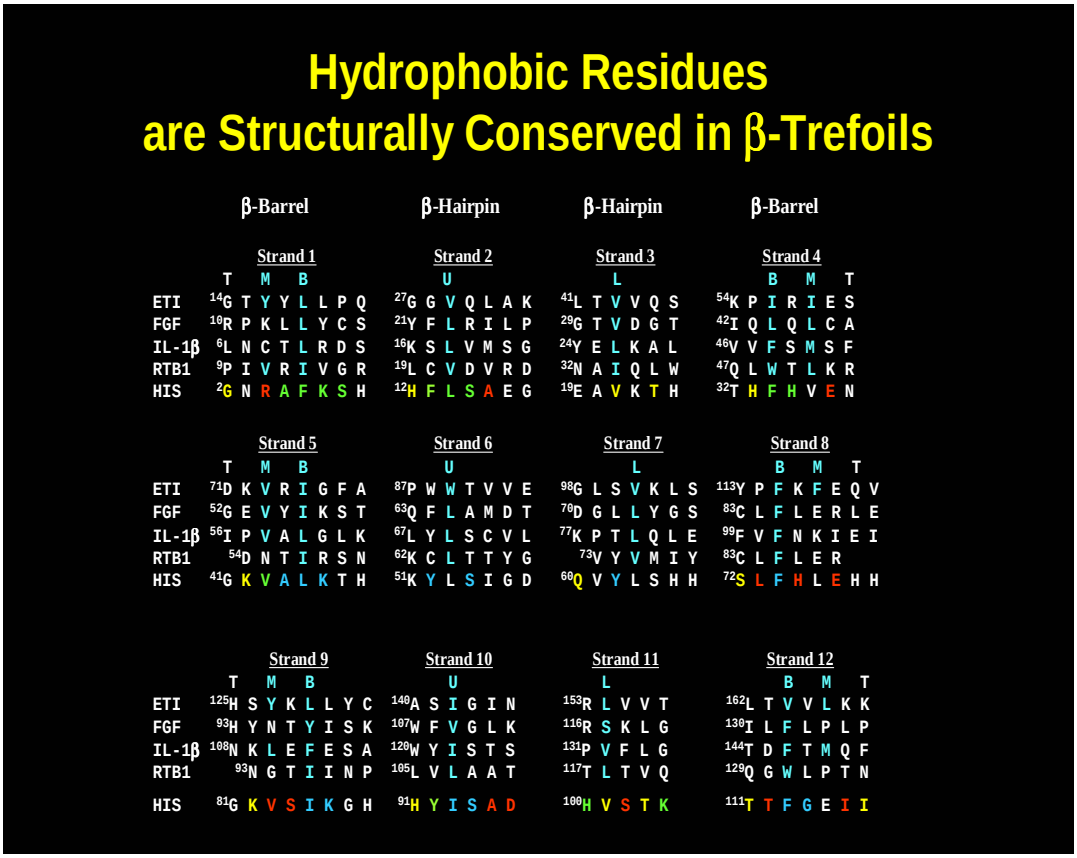


Figure 1. 11 Sequence alignment of the 18 conserved hydrophobic residues in β-trefoil proteins.
The hydrophobic conserved residues are shown in blue color, Erythrina caffra inhibitor (ETI), Fibroblast growth factor (FGF), Interleukin 1-Beta (IL-1β), Ricin B (RTB1), and Hisactophilin (HIS) (Murzin et al., 1992). Above the aligned sequences, T, M and B indicate residues in the top, middle and bottom layers, respectively, of the □-barrel, while U and L indicate residues in the upper and lower layers, respectively, of the hairpin triplet. The residues in the layers are shown within the hisactophilin 3D structure in Figure 1.12.

The 18 conserved hydrophobic residues appear in the same symmetry related positions in each symmetric trefoil element. Twelve out of 18 hydrophobic residues in hisactophilin are located inside the barrel and stabilize this structure. The top layer is closer to the open end of the barrel, is partially exposed to solvent and displays no significant sequence similarity with the

known β-trefoil structure. The remaining conserved hydrophobic residues are three residues from the upper and three residues from the lower β-hairpin layers (Figure 1.11).

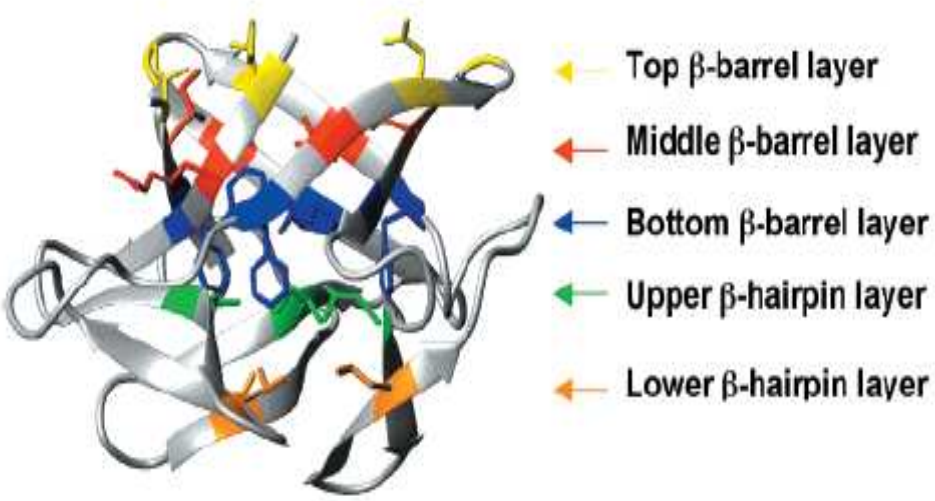


Figure 1. 12 Ribbon representation of hisactophilin illustrating conserved layers.
Side chains of residues in various layers are shown in stick representation. The ribbon diagram was prepared using the program MolMol, yellow color is the top β-barrel layer, red color is the middle β-barrel layer, blue color is the bottom β-barrel layer, green color is the upper β-hairpin layer, and orange color is the lower β-hairpin layer (Koradi et al., 1996).

1.5.2 Stability and folding of hisactophilin

It was determined that unfolding of PWT (with 4 extra amino acid) hisactophilin is greater than 90% reversible for urea-induced unfolding and at least 80% reversible for thermal denaturation (Liu et al., 2001). For stability measurements of WT and mutant hisactophilin, chemical denaturation by urea was used.

1.5.2.1 Previous equilibrium studies on hisactophilin

The stability and folding of PWT and WT hisactophilin has been characterized in detail previously in our group. Fluorescence and CD measurements were used to monitor equilibrium urea and pH denaturation curves (Liu et al., 2001). PWT and WT denaturation by urea is reversible from pH 5.7 - 9.7 and data fit well to a 2-state transition between native and denatured states (Liu et al., 2001) (Figure 1.12) PWT hisactophilin has moderate and

approximately constant stability from pH 7.7-9.7; however below pH 7.7 stability decreases dramatically, most likely due to the electrostatic repulsion of positively charged histidine residues.

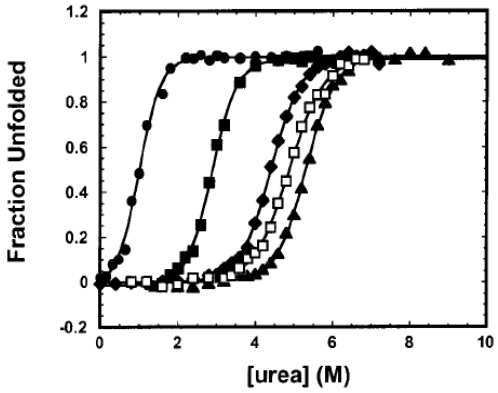


Figure 1. 13 Fluorescence-monitored urea denaturation curve of PWT.
As a function of pH conditions: pH 5.7 (●), pH 6.7 (■), pH 7.7 (♦), pH 8.7 (▲), and pH 9.7 (□), displayed in terms of the fraction of unfolded protein. PWT stability increases markedly from pH 5.7 to 7.7 and then remains relatively constant from pH 7.7 to 9.7 (Liu et al., 2001).

PWT has a characteristic CD spectrum similar to what is expected of an all β-protein. The native spectrum contains broad minima at 209 nm and 200 nm and a maximum at 227 nm (Liu et al., 2001) (Figure 1.13).

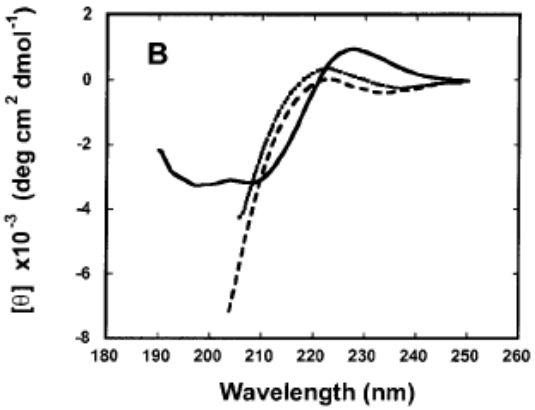


Figure 1. 14 CD spectra of PWT hisactophilin.
Spectra obtained of protein at 20°c, pH 5.7 (50 mM KH_2PO_4 /K_2HPO_4 buffer with 1mM EDTA and 1mM DTT, protein concentration 2.3 mg ml^{-1}) in 0 M urea (solid line), 3 M urea (dashed line) and 8 M urea (dotted line) (Liu et al., 2001).

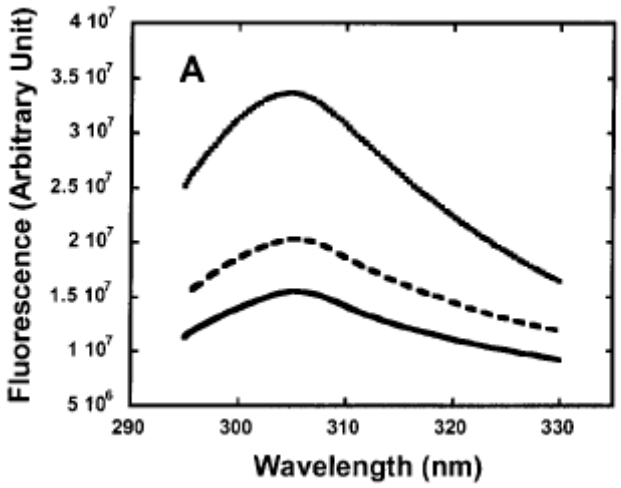


Figure 1. 15 Fluorescence spectra of PWT hisactophilin.
Spectra obtained of protein at 20 °c, pH 5.7 (50 mM $KH_2 po_4 / K_2Hpo_4$ buffer with 1mM EDTA and 1mM DTT, protein concentration 0.24 mg ml^{-1}) in 0 M urea (solid lower line), 3 M urea (dashed line) and 8 M urea (top solid line) (Liu et al., 2001).

Even though the fluorescence signal of WT hisactophilin is low, it is still valid as a probe for structure. Fluorescence studies of PWT and WT hisactophilin show a spectrum that contains a maximum at 306 nm both in the absence and presence of urea and the intensity increases with increased urea concentration as shown in (Figure 1.14) (Liu et al., 2001).

1.5.2.2 Previous kinetic studies on hisactophilin

The PWT and WT hisactophilin, have been studied by using optical and quench-flow NMR methods and have been compared with the folding/unfolding process of other β-trefoil proteins, 1L-1β and FGF (Liu et al., 2001). Most β-trefoil proteins, fold slowly through a population of intermediates. In the case of PWT and WT hisactophilin folding, folding is relating fast and apparently occurs via a two step process except under the most native stabilizing solution conditions (Figure 1.15). Hisactophilin dose not contain any prolines, so the slow phase cannot arise from proline isomerization (Liu et al., 2001). Furthermore, the folding kinetics were measured at low protein concentration, so the slow phase dose not have property chractristics of a phase caused by protein aggregation. The folding intermediate for

WT hisactophilin was found to resemble that found in 1L-1β but differs from that of FGF. These differences are thought to be due to the differences in the non-conserved loops.

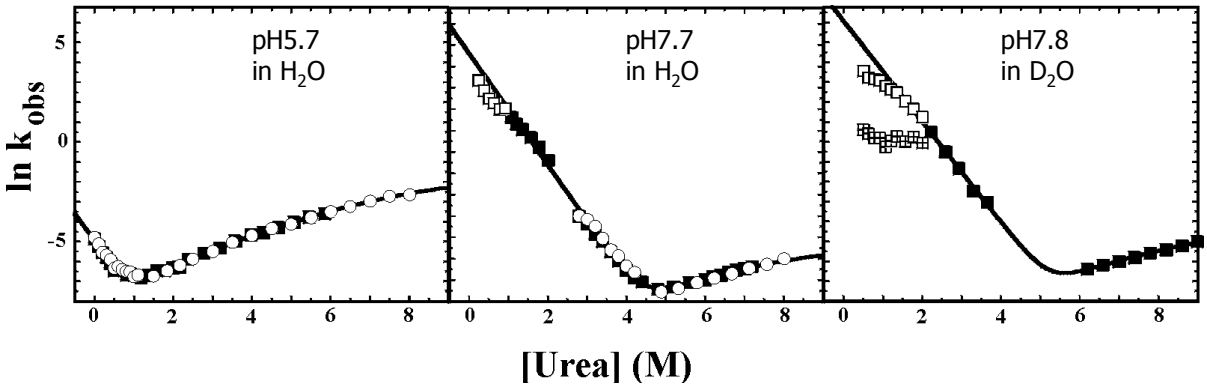


Figure 1. 16 Observed unfolding and refolding rates for hisactophilin at 20°C.
Measured by CD (open circles) and fluorescence (solid, open, quartered squares).Conditions: (A) pH 5.7 in H_2O, (B) pH 7.7 in H_2O, (C) pH 7.8 in D_2O. ○ observed CD rates used for the fit, ■ observed fluorescence rates used for the fit, □ observed fluorescence rates not used for the fit, ⊞ observed rates of the slower refolding phase in D_2O at pH 7.8 (Liu et al., 2001).

1.6 Folding of FGF

Several mutational studies on Fibroblast Growth Factor (FGF) have been performed in order to assess the stability and folding of this β-trefoil structure (Brych et al., 2001). These mutations are located within the hydrophobic core, either structurally conserved or non-conserved, and in various turns and loop regions. In order to increase the overall stability of the structure, the hydrophobic core must be properly packed, the energy associated with bond strains minimized and the relative rate of folding and unfolding should be increased and decreased, respectively. The folding rate can be increased by the efficient packing of the hydrophobic core or proper positioning of secondary structure elements so that they can readily adopt their native state. Based on the hypothesis that an increase in trefoil symmetry would result in increased protein stability, mutational targets were identified in the structurally conserved residues within the bottom β-barrel, lower β-hairpin and upper β-hairpin (Figure 1.17) (Table 1.1).

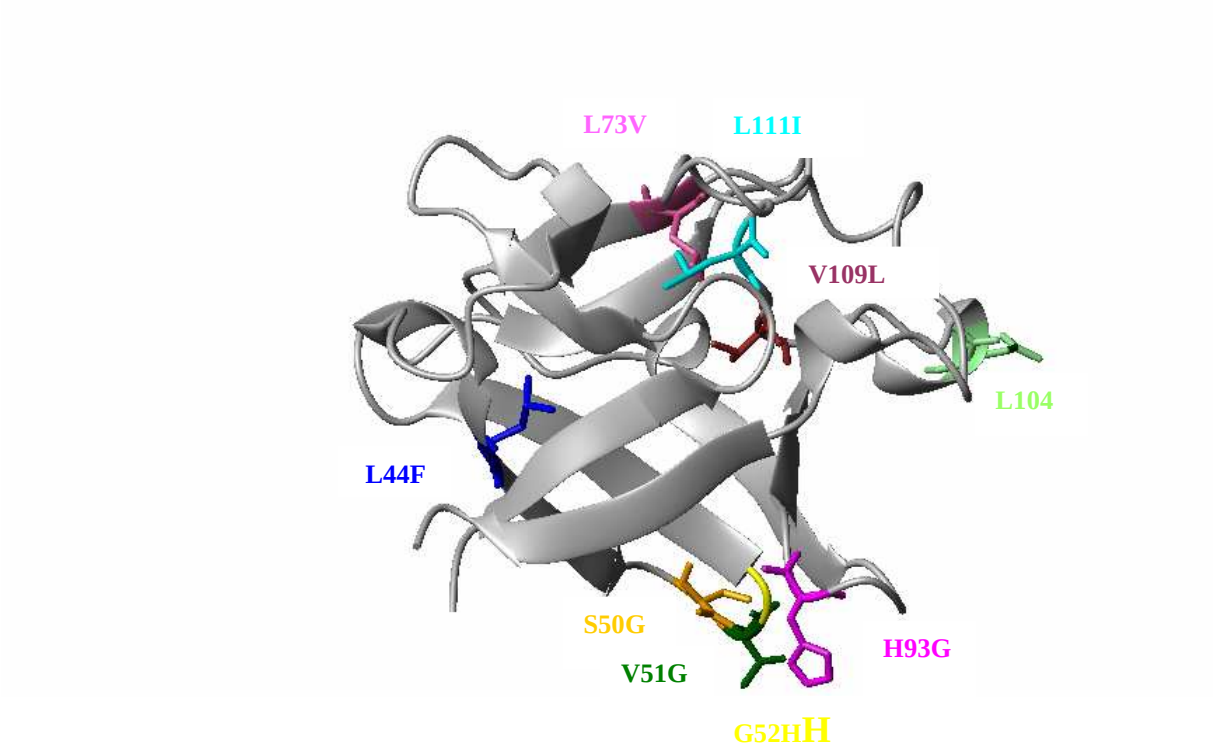


Figure 1. 17 FGF Mutations.
The ribbon diagram was prepared using the program MolMol green color indicates to V51G, orange color S50G, magenta color H93G, blue color L44F, pink color L73V, cyan color L111I, brown color V109L, and Pale green color ΔL104 (Koradi et al., 1996).

The stability measurements of mutants relative to the WT protein are summarized in (Table 1.1). Of the 4 structurally conserved single point mutants only L44F, located in the bottom β-barrel layer, resulted in an increased thermostability, which was attributed to the partial occupation of a microcavity by the bulkier side chain. In other β-trefoil family members, interleukin-1α and β and WT hisacophilin have a F residue at the position corresponding to 44 in FGF and F is most frequent in all 3 paterns in repeats 2^{nd} and 3^{rd}, so this supports the hypothesis that an increase in trefoil symmetry would result in increased protein stability. In addition, spatially proximal I 25 partially shifts from the WT conformation filling an additional microcavity, and contributing to the overall mutant stability (Brych et al., 2001). The destabilizing effect of the other single point mutations was attributed to an increase in overall cavity space within the core due to less suitable packing of the core residues. Also the

observed decrease in folding rates (Brych et al., 2001; Kim et al., 2003) were attributed to a reduction in the total buried area.

Table 1.1: Summary of difference in Gibbs free energy between WT and mutant Fibroblast Growth Factor

Mutation/Deletion	Location	ΔΔG (kJ/mol)
L44F[1]	Structural Conserved Core	-2.9
L73V[1]	Structural Conserved Core	+6.1
V109L[1]	Structural Conserved Core	+2.4
C117V[2]	Structural Conserved Core	+1.2
L111I[2]	Structural Conserved Core	+1.0
L73V/V109L[1]	Structural Conserved Core	+3.7
L44F/L73V/V109L[1]	Structural Conserved Core	+0.6
M67I[2]	Core	+9.4
F22Y[5]	Mini-core	-0.6
Δ104-106[3]	Loop Insertion	-1.6
Δ120-122[3]	Loop Insertion	NP
2nd [2]	Structural Conserved Core	5.3
3rd [2]	Structural Conserved Core	3.2
4th [2]	Structural Conserved Core	1.9
5th [3]	Structural Conserved Core	+0.2
5th/Δ104-106[3]	Structural Conserved Core+ Loop Deletion	-7.6
5th/Δ120-122[3]	Structural Conserved Core + Loop Deletion	+1.1
5th/ΔΔ[3]	Structura lConserved Core + Loop Deletion	-15.9
6th [3]	Conserved and non-conserved core	+10.2
6th/Δ104-106 [3]	Conserved and non-conserved core + Loop Insertion	-2.8
6th/Δ120-122[3]	Conserved and non-conserved core + Loop Insertion	+5.4
6th/ΔΔ[3]	Conserved and non-conserved core + Loop Insertion	-16.1
S50G/V51G[4]	Turn	-1.3
E49G/S50G/V51G[4]	Turn	- 0.4
H93G[4]	Turn	-8.3
G52H[4]	Turn	5.5
S50E/V51N/G52H[4]	Turn	3.6
E91S/N92V[4]	Turn	-5.7
E91S/N92V/H93G[4]	Turn	-2.8

2nd mutation (L73V /V109L)
3rd mutation (L44F/L73V /V109L)
4th mutation (L44F/L73V /V109L/C117V)
5th mutation (L44F/L73V /V109L/C117V/L111I)
6th mutation (L44F/L73V /V109L/C117V/L111I/M67I)
ΔΔ double deletion mutant Δ104-106 and Δ120-122
NP – Not published

$\Delta\Delta G = \Delta G_{WT}-\Delta G_{mutant}$
(Brych et al., 2001)[1], (Kim et al., 2003)[2], (Brych et al., 2004)[3], (Kim et al., 2005)[4].

The stability resulting from combinatorial mutations cannot simply determined by the summation of the individual point mutations since the effect of a single mutation is not localized to a single position within the protein, but rather propagates itself throughout other portions of the structure. Therefore a selective combinatorial approach was used along with experimental measurements to assess the effect of multiple mutations.

Residues L44, L73, V109 and C117 (4^{th} mutation) in the structurally conserved core region of FGF were mutated to the consensus residue found in 2 of the 3 trefoils. The 2^{nd}, 3^{rd}, 4^{th} and the 5^{th} mutations in FGF were less stable than the WT, due to introduction of strain and creation of a cavity. The combination of the 6 mutations (6^{th}) that has the highest symmatry for the core residues (Table 1.1), is less stable ($\Delta\Delta G$ =+10.2) than the wild type due to the resulting distorted tertiary structure (Brych et al., 2003). Based on sequence comparison of the three trefoils, insertion regions were identified in two loop regions, residues 104-106 and 120-122. The reduction in the loop length is expected to stabilize the structure due to a decrease in entropic energies. The optimal symmetric packing interactions within the core region of FGF were achieved by including double loop deletion mutations 104-106 and 120-122 designed to increase the tertiary structure symmetry (Dubey et al., 2005).

Turn point mutation A103G in the turn resulted in an increase in stability since glycine at i+3 position in type I β-hairpin turns either 3:5 or 4:6 type is a key contributor towards increasing the rate of folding of the native protein (Brych et al., 2004; Lee et al., 2006). Three different combinations of glycine mutations (E49G, S50G, 51G) and (E49G, S50G, V51G), in the β4/β5 turn, were more stable than the WT protein and increased the rate of folding (Kim et al., 2003). The results indicate that the residues within the turns are not a primary determinant of the turn structure, but do have a significant influence on protein stability and

folding. Substitution of the β8/β9 (E91, N92, H93) turn primary structure into a β4/β5 (S50, V51, G52) turn, which include the V51N and G52H point mutations S50E, V51N double mutants, had no effect on protein stability or folding kinetics (Kim et al., 2003).

The substitution of the β4/ β5 turn primary structure into β8/β9 turn, and the converse substitution of the β8/β9 turn primary structure into β4/ β5 turn were observed to destabilize the protein in both cases. This turn formation in proteins is influenced more by the local environment than by the turn sequence itself (Kim, et al., 2005), (Lee, et al., 2006).

However, in a recent study, protein folding could become faster if the six residue loop with an internal β-turn was transformed into a five or four residue turn with a sequence optimized to reverse the direction of the polypeptide chain (Deechongkit et al., 2006).

In interleukin 1-beta (IL-1β) mutation was done in some conserved residues (F101Y, F146Y, F101W, F146W, F101I, F146I, F101L, and F146L). All of these mutations severely destabilized of the IL-1β which is F in other beta-trefoil members. These observations indicate that aromatic residues at those positions have some folding and stability of beta-trefoil family (Adamek et al., 2005).

1.7 Folding and stability

1.7.1 Symmetry

β-Trefoil superfolds exhibit symmetric tertiary structures that have been postulated to have evolved via gene duplication and fusion events (Ponting and Russell, 2000) (Brych et al., 2003). The sequences of different superfamilies are very diverse, and there is also great divergence within the superfamilies. Although the primary amino acid sequences of hisactophilin, 1L-1β, and FGF have < 10% sequence identity, the residues responsible for stabilization of the barrel are conserved as hydrophobic amino acids as shown in (Figure 1.17) (Murzin et al., 1992).

The six stranded barrel is stabilized by the packing of large hydrophobic residues in the interior of the protein close to the barrel axis.

The goal of the experiments described in this thesis is to increase the tertiary structure symmetry within a β-trefoil to investigate whether this will lead to an increase in stability, by a series of point mutations that determine the effects of the residues on the stability of hisactophilin. Different types of mutations were made: one in a turn, one in a minicore and the others in hydrophobic core regions.

1.7.2 Turns

β-turns are common structural motifs in proteins, and they are short regions of non-α and non-β conformation. Turns are a central component of β-hairpin nucleation and stabilization. β-turns change the direction of the polypeptide chain and allow secondary structure elements to form tertiary interactions as shown in (Figure 1.18) **(**Wilmot and Thornton, 1988). Turns involving four residues are most common, hydrogen bonding occurs between the carbonyl of the first residue and the N-H of the fourth residue.

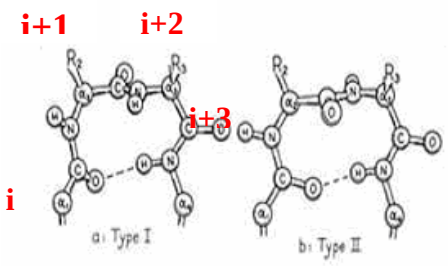


Figure 1. 18 Common β-turns conformation.
(i) is the first residue in the β-turn, (i+1) is the second residue in the β-turn, (i+2) is the third residue in the β-turn, and (i+3) is the fourth residue in the β-turn (Wilmot and Thornton, 1988).
.

Turns typically involves residues at the surface of peptides that can hydrogen bond to the solvent when not bonded to each other. β-turns are the most common type of non repetitive structure recognized in proteins, and comprise about ~25% of the residues. Despite

this abundance, less is understood regarding the contribution of turns to the structure, stability, and folding of proteins than for α-helices and β-sheets. One reason for this lack of knowledge is that β-turns appear to be far more varied and complex than either α-helices or β-sheets (Li W, 2005). There are preferences for particular amino acids at certain positions within the turn Proline and glycine are very common in turns (Hutchinson and Thornton, 1994; Wilmot and Thornton, 1988).

Glycine has different unique properties not inherent to other amino acids. Having the smallest side chain (H), glycine can assume conformations normally forbidden by close contacts of the β-carbons; thus it is more flexible than other residues, making it valuable for pieces of backbone that need to move or hinge (Parrini et al., 2005). All glycines and most histidines in WT hisactophilin are located in the loops and turns (Habazettl et al., 1992) (Figure 1.9).

From the multiple sequence alignment, hydrophilic residues are more likely to occur in turns than hydrophobic residues, because turns tend to be solvent exposed. There are 8 β-turns in WT hisactophilin (Habazettl et al., 1992). Turn β1-β2 (connecting β strands 1 and 2), β2-β3, β4-β5, β5-β6, β6-β7, β8-β9, β9-β10, and β10-β11. The types of turns found in WT hisactophilin as determined by NMR (NOE patterns) are the most common types possessed by all proteins, mainly types I`and II, (not all the turns in hisactophilin are actually defined by the NMR data). Glycine is commonly found in the first residue of the β-turns, and almost exclusively at the second position of the turn. In both type I`and type II`the hydrophobic amino acids are under-represented at all turn positions (hydrophobic amino acids contain branched $C_\beta$). In type I`glycine is the preferred residue at position i+2, while in type II`glycine shows a significant preference at position i+1 (Blandl et al., 2003).

1.7.3 Loops

Loops are an irregular secondary structure motif in proteins, consisting of 6 to 16 residues, with various lengths and irregular shape. When they connect two adjacent antiparallel β-strands they called hairpin loops, and short hairpin loops are called turns. Long loops are also sometimes called “random coils”, although they are not random because they exhibit some amino acid and some structural preferences. Furthermore, amino acids in loops often are not conserved. The main chain CO and NH groups do not hydrogen bond with each other, but are exposed to the solvent and can form hydrogen bonds with water molecules. Loops are the most mobile parts of protein structure, as shown by NMR solution and X-ray temperature factors. They play an important role in protein function, and they are often the most difficult structure to model. The loop regions connecting secondary structures demonstrate less regularity in their conformations, even though short loops linking specific secondary structures can be classified into distinct groups (Panchenko and Madej, 2005) . The loops in proteins are less well characterized than the secondary structural elements that they connect (Nagi and Regan, 1997). Long loops are classified based on the distance between the ends of the loop and of adjoining secondary structure into: long closed loops connect adjacent regions of secondary structure, and long open loops connect distant secondary structure.

Mutational study of selected proteins reveals that the presence of proline and glycine residues in the loops of β-helical structures generally enhances the structural stability of the system. The proline residues are very rigid and make the loop region more stable, while the glycine residues are flexible and thus sustain most of the loop movements and stabilize the structure (Haspel et al., 2006). Hisactophilin loops differ from other β-trefoil proteins like FGF and IL-1β, which are considerably shorter and more hydrophilic. Hisactophilin loops

have a sequence variability as shown by NMR, this variability is believed the cause of the high flexibility of the loops (Habazettl et al., 1992).

1.7.4 Main hydrophobic core

In a globular single domain protein, hydrophilic residues tend to be on the surface of the molecule. Surface residues are relatively flexible and can adopt alternative conformation. Hydrophobic side chains of residues are generally packed together inside the molecule, excluding water to form anhydrous cores.

Hisactophilin possesses a centrally located core, characteristic of all β-trefoil family proteins (Priestle et al., 1989) with a common hydrophobic core volume of 3000 +/- 120 Å. As described earlier in Section 1.5.1, there are 18 conserved hydrophobic residues that form the main hydrophobic core in hisactophilin core: 6 from each symmetry unit, three residues from the lower hairpin layer (L), three residues from the upper hairpin layer (U), and twelve residues from the bottom and middle layer of the β-barrel (B) (Figure1.11). The bottom (B) and middle (M) layers of the β-barrel contribute to the packing of the barrel core, and large hydrophobic residues occupy these positions. These hydrophobic residues have low solvent accessibilities and are in close contact with each other.

1.7.5 Mini-core

FGF contains an outer mini-core within each of the three β-trefoil folds that includes a pair of buried hydrophobic residues, these residues are not part of central core. In WT FGF protein contains Y or F at these positions (F22, F64, F108, packing against I42, C83, and I130), both point and double mutations were designed to increase primary structure symmetry within the minicore region(Dubey et al., 2005). In hisactophilin there is a absence of this

mini-core like FGF, but a point mutation was made in F13Ywhich correspond to F22Y in FGF.

1.8 Thesis objectives

This study focuses on depth biophysical analysis of the WT and mutant hisactophilin. This analysis was performed as an attempt to characterize the effects mutations have on the stability and folding of WT hisactophilin. More specifically, the role that specific turn residues play in protein folding was investigated. These studies were done in an effort to bring new insights into protein folding.

The aim of chapter 1 is to focus on recent efforts to understand protein folding, starting by discussing the nature of the protein folding problem and how the of folding of small single domain proteins is analyzed.

In this study, site directed mutagenesis as described in chapter 2 was performed to introduce a primary structure within hisactophilin that reflects the threefold tertiary symmetry present within the β-trefoil superfold. The goal of the study was to identify an appropriate solution to such a symmetric constraint within the hydrophobic core of the protein. To achieve this, specific point mutations in hisactophilin (core mutations & minicore and β-turn mutations) were created to establish primary structure symmetry within hisactophilin that reveals the threefold tertiary symmetry present within the β-trefoil superfold family. The mutations are shown in (Figures 1.8 and Figure 1.19 and Table 2.1).

Chapter 3 describes the stability results of the unfolding pathway of WT and mutant hisactophilin determined by fluorescence and CD monitored chemical denaturation. To further characterize the unfolding pathway, the rates of folding and unfolding were measured by kinetic renaturation and denaturation methods monitored by fluorescence.

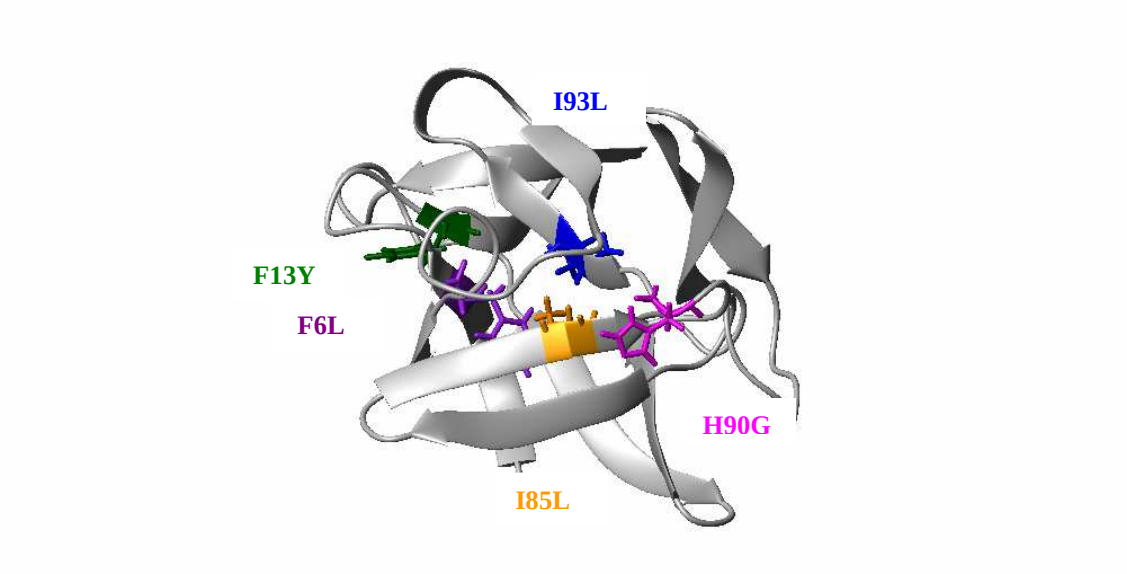


Figure 1. 19 Hisactophilin Mutations.
Ribbon representation of the hisactophilin structure with mutated residues shown in stick representation and labeled with wild type residue and amino acid number followed by mutated residue, in single letter code. The ribbon diagram was prepared using the program MolMol and PDB code 1HCD (Koradi et al., 1996).

Chapter 2: Methods

2.1 Hisactophilin plasmid (pHisI)

The hisactophilin gene (cDH516) was initially subcloned into the plasmid pIMS5 to generate pHisI (Simon et al., 1988). This plasmid codes for 4 additional amino acids (Gly-Glu-Phe-Gly) at the N-terminus of the protein (Habazettl, et al., 1992). The inserted amino acid codons were deleted from the original plasmid by site-directed mutagenesis yielding a plasmid designated pHW (Figure 2.1) (Wong et al., 2004). pHW1 also contains the β-lactamase gene for ampicillin resistance and an isopropyl-β-D-thio-galactosidase (IPTG) Ptacpromoter.

2.1.1 Plasmid preparation

Plasmid DNA was isolated using the HiYieldTM Plasmid Miniprep Kit (United Bioinformatica Inc, Canada). The DNA transformed to TG2 by Electroporation, from smear, isolated colonies are picked and streaked on 500 mg/ml amp plates. From isolated that plate, colonies were selected and used to inoculate amp LB media streaked on a final amp plate. 10 ml LB media grown overnight @ 37ºC and used to inoculate 5 ml amp LB media (1/100) using 50µl of the starter culture. Cells grown to midlog phase OD_{600} 0.7 add 34µg/ml chloramphenicol, to a final concentration 170µg/ml. Used HiYieldTM Plasmid MiniKit-High Copy Number Protocol according to manufacture's directions. The concentration of purified plasmid DNA was determined by measuring the OD_{260}, where one unit optical density at OD_{260} corresponds to approximately 50 µg /ml of DNA.

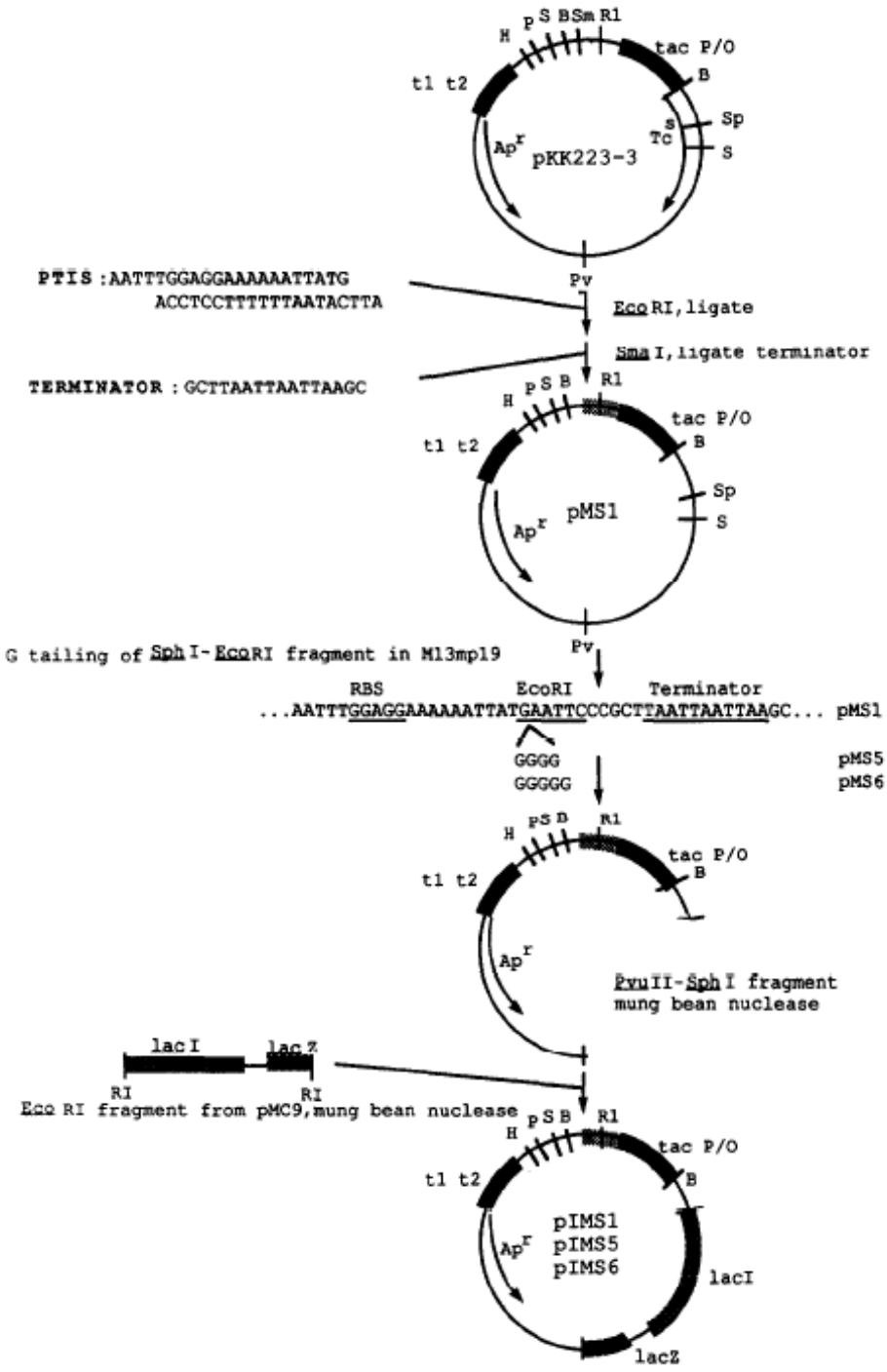


Figure 2. 1 Construction of the pIMS vectors.
tac P/O, tac promoter/operator; t1 t2, rrnB transcription terminator; PTIS, portabletranslation initiation sifters, ribosome binding site; T, translation terminator. Restriction sites are denoted as RI, EcoRI, Sm, SmaI; B, BamHI; S, SalI; P, PstI; H, Handier, Sp, SphI; and PV, PvuII (Simon et al., 1988).

2.2 Designing hisactophilin mutations

Hisactophilin mutations were designed based on sequence analysis between structurally related trefoil subdomain within WT hisactophilin, as well as between homologous positions for other members of the beta trefoil superfold. This analysis was done by comparing multiple sequence alignment in Pfam (Bateman et al., 2004), and Psiblast (Altschul et al., 1990) .

When designing the core mutations, attention was paid to substitute hydrophobic residues led with another hydrophobic residue or a hydrophilic residue with another hydrophilic residue to avoid change in polarity. (Figure 1.7) shows the multiple alignments of β-trefoil proteins sequences (Table 2.1) illustrated the point substitution mutations that have been made in hisactophilin. Based on conserved residues in multiple sequence alignment between β-trefoil proteins and statistical analysis of β-turn formation, H90G was designed. This mutation is located in the type I' β-turn of the protein structure.

Based on Psi-BLAST analysis, hisactophilin sequence shared clear sequence similarity with Fascin. This similarity can be attributed to the buried cores of their β-trefoil structures (Ponting and Russell, 2000). Related protein sequences were found using Pfam (Bateman et al., 2004), Psiblast (Altschul et al., 1990), servers.

```
His  G N R A F6KSH   HG    H F13L SAEG   EA     EAVKTH     HGHHDHH     THFHVEN
Fac  R T Y V A AMD    NG    K F T IGFP          HPEGEG     PNPEEI      FALVKTP
FGF  R P K L L YCS    NG    Y F L RILP    D     GTVDGT     RDRSDQH     IQLQLCAESV

His  G K V A L KTH    CG    K Y L SIGD   HK     QVYLSHH    LHGDH       SLFHLEHHG
Fac  S K I S L KTG    FG    R Y V GVDS   EY     QLVAMAE    AIGSR       EQFVLVFQE
FGF  G E V Y I KST   ETG    Q F L AMDT          DGLLYGS    NEE         CLFLERLEEN

His  G K V S I85KGH  HH90   H Y I95SAD   HHG    HVSTK      EHHDHD      TTFGEIII
Fac  G K T A F QAV   SS      P L F LSN    KEG    HIYVA      SRTATE      NEMVNIR
FGF  H Y N T Y ISK   HA      W F V GLK    KNG    RSKLG      HYGQ        LFLPLPVSSD
```

Figure 2.2 Sequence alignment of the hisactophilin mutations.
The F6 and I85 symmetric residues are shown in purpule and orange color, F13 symmetric residues are shown in green color, H90 symmetric residues are shown in magenta color, and I93 symmetric residues are shown in blue color. Strand residues are in black color. Hisactophilin (HIS), Fascin (Fac), and Fibroblast growth factor (FGF).

Table 2. 1 The point substitution mutations that had been made in hisactophilin.

Residue	mutation	Locations	Trefoil number	Explanations
F6	L	Major core	Trefoil 1 strand 1 Bottom β-barrel layer	To increase symmetry in β-barrel
F13	Y	Mini-core	Trefoil 1 strand 2 β-hairpin	To increase symmetry in β-hairpin
I85	L	Major core	Trefoil 3 strand 9 Bottom β- barrel layer	To increase symmetry in β-barrel
H90	G	Type I' β-turn	Tight turn between strand 9 and strand 10	To identify the role of turn in protein stability, and increase symmetry
I93	L	Major core	Trefoil 3 strand 10 Upper β-hairpin layer	To increase symmetry in β-hairpin

2.3 Site-Directed Mutagenesis

The hisactophilin mutations were designed to introduce a three fold symmetric constraint in the primary sequence, consistent with the symmetry of the β-trefoil tertiary structure. The mutagenesis primers were designed to be complementary to the hisactophilin cDNA (Scheel, 1989) using the modified pHisI plasmid (Simon et al., 1988) pHW as template DNA (Wong et al., 2004). The mutations were generated using the QuikChange®Site-Directed Mutagenesis method; silent mutations were incorporated into the primer so the correct (stratagene) PCR product could be identified by the pattern of restriction enzyme digestion restriction site to facilitate identification of mutants. The hisactophilin mutations F6L, F13Y, I85L, H90G, and I93L were engineered into pHW (Wong et al., 2004) using criteria described

in the QuikChange^{TM} site-directed mutagenesis kit (Instruction Manual, cat. No. 200518, Stratagene). The primers were synthesized at (Sigma Genosys Oakville, ON), on a Scale: 3.0 D select, and purified by desalting. The primers were designed to have a minimum guanine and cytosine content (gc) content of ~40%, where the desired substitution was flanked by a minimum of 10-15 bases of correct sequence on either side. The melting temperature of the primers was designed to be close to 78°C (WatCut, University of Waterloo website). (Table 2.2) lists the primer sequences, nucleotide substitution of interest and melting temperatures of the primers for the various mutants engineered.

Table 2. 2 The forward and reverse Primer sequences and melting temperatures for the engineered hisactophilin mutants.

Mutation	Codon change	Primer Sequence*	Melting Temp
F6L	TtC→CtG	5' g ggt aac aga gcT CtG aaa tca cat cac ggt cac ttt tta ace gc 3` 3' gc gct taa aaa gtg acc gtg atg tga ttt CaG Agc tct gtt acc c 5'	77 °C
F13Y	tTt→tAt	5' c aaa tca cat cac ggt cac tAC tta agc gct gaa ggc g 3' 3' c gcc ttc agc gct taa TGa gtg acc gtg atg tga ttt g 5'	77. 7 °C
I85L	AtC→CtG	5' cat cat ggc ggt aaa gtc tca CtG aaG ggt cat cac cac cac tac 3' 3' gta gtg gtg atg acc Ctt CaG tga gac ttt acc gcc atg atg 5'	77.9 °C
H90G	CAc→GGc	5'gtc tca atc aaa ggt cat caT GGc cac tac att tcc gct g 3' 3' c agc gga aat gta gtg gGG Gtg atg acc ttt gat tga gac 5'	78.1°C
I93L	AtT→CtG	5' ggt cat cac cac cac tac CtG tcc gcG gat cat cat ggt c 3' 3' g acc atg atg atc Cgc gga CaG gta gtg gtg gtg atg acc 5'	77.6 °C

The primer melting temperatures were calculated from the equation t= 81.5+0.41(%gc)-675/n-%mismatch, where %gc is the percent guanine plus cytosine composition, % mismatch is the number of substitution divided by n. n is the total number of nucleotides in the primer *Mismatches are given in capital letters.

In the polymerase chain reaction (PCR) the concentration of reagents was 10x buffer 1 Expand Long Template with $MgCl_2$ (from Roche), dNTPs (from UBI, United Bioinformatica Inc, Canada), 25 mM, 1.75 mM Template DNA, and MilliQ H_2O to give a final volume of

24.25 µl in each PCR tube. Expand Long Template PCR polymerase (Cat. No.11 681 834 001, Roche Diagnostics GmbH, Germany) was used for DNA amplification during mutagenesis.

Initial Elongation cycles were preformed with two reactions generate single strand amplification product. The remainder of the PCR was performed with the two reactions combined to generate, nicked, double-strand product.

Table 2. 3 PCR Programming protocols

Steps	**Temperature (ºC)**	**Time**
Initial Denaturation	92	5 min
Pause	[Add 0.5 µl Expand Long Template PCR System DNA Polymerase mix]	
Denaturation Annealing Elongation	92 55 68	30 sec 1min 20 min } Cycle 5 times
Pause	Complementary tubes pooled together	
Denaturation Annealing Elongation	92 55 68	30 sec 1 min 20 min } Cycle 15 times
Final Elongation	68	20 min

Thermocycling was performed using a Techgene thermocycler (Techne. Inc. Barloworld Scientific) using the parameters described in Table 2.3. The PCR plasmids were incubated with DpnI to remove methylated template DNA prior to transformation. *E. coli* cells (strain

XL1-blue) were transformed with the PCR products by electroporation followed by growth on Amp containing agar plates overnight at 37ºC. Transformants were selected to grow culture for plasmid isolation. Transformants containing plasmids with desired mutations were identified by DNA sequencing using the sequence primer for hisactophilin expression vector. The concentration of purified DNA was calculated by measuring the OD_{260}, DNA samples were sent for sequencing when appropriate (Mobix Central Facility, DNA Synthesis Laboratory, The Institute for Molecular Biology and Biotechnology, McMaster University, Hamilton, ON).

2.4 Electroporation

Electroporation is a highly efficient method (nearly 80 %) of transforming plasmid DNA into bacterial cells (Miller and Nickoloff, 1995). Electroporation involves a structural rearrangement of the cell membrane caused by the application of a short electric field pulse. This electric field causes pore formation and the driving force to transport plasmids into the cell (Weaver, 1995).

Electrocompetent cells were thawed on ice. To a clean epedorff tube, 40 uL of electrocompetent cells were added. To these cells, 1 µL of plasmid was added. The mixture was kept on ice, and transferred to a chilled electroporation cuvette (VWR Scientific). The cuvette was inserted into an electroporation rig (E. coli Pulser, Bio-Rad), and 1.80 kV of current run through the cuvette for 4 milliseconds. The transformed cells were then immediately added to 1 mL of SOC media (2% Bacto-tryptone w/v, 0.5% Bacto yeast exract w/v, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM Glucose), and let stand at 37 °C for 1 hour. 100 uL of the transformation media was then plated on selective plates and incubated overnight to obtain putative transformants.

2.5 Protein engineering

Luria-Broth (LB) media was prepared by dissolving 10 g tryptone, 5 g yeast extract, and 5 g NaCL (Bishop Canada Inc., Burlington, ON) in 1 L distilled deionized water (ddH2O) and autoclaved.

2.5.1 Antibiotic agar plate preparation

Antibiotic agar plates were prepared by dissolving 7.5 g agar (Bishop Canada Inc) in 1 L LB and autoclaving to dissolve the agar. The mixture was cooled down to ~ 50 ºC using a water bath before adding 100 µg /ml Ampicillin. The solution was then poured onto plastic Petri dishes, and the gel was allowed to settle for 2-3 hours before being stored at 4 ºC.

2.5.2 Preparation of competent cells and transformation

The stock cell lines (*E Coli* strains XL-1blue for cloning processes and BL21 (DE3) for protein expression) were inoculated in10 ml of LB media and incubated at 37ºC with 200 rpm of shaking overnight. 5 ml of this overnight culture was then used to inculcate 500 ml of LB media which was allowed to grow to mid-log phase with an OD_{600} was reached 0.4-0.6 at 37ºC with shacking at 200 rpm. Once the cells had reached mid log phase they were placed on ice for approximately 20 minutes. Cells were harvested by centrifugation at 4000g for 15 minutes at 4ºC. The super was removed and the cells were washed by resuspending in 500 ml of ice cold sterile 10% glycerol (EMD pharmaceuticals, Durham, NC) .The cells were collected again by centrifugation and the wash was repeated with 250 ml and 20 ml of the glycerol. The competent cells were resuspended in a final volume of 2 ml of 10 % glycerol; flash frozen in dry ice as 80µl aliquots and stored at -80 ºC.

2.5.3 Hisactophilin and mutant hisactophilin expression and purification

Hisactophilin and mutant hisactophilin plasmid was transformed into *E.coli* BL21 cells by electroporation. Cells with the plasmid were resistant to ampicillin, the cells were grown in Luria Broth to an optical density between 0.5-0.8, and induced with isopropyl-thio-β-D-galactoside (IPTG) at a final concentration of 1 mM and grown for 4-8 hours at 37°C before harvesting by centrifugation. As an exception F6L cells were grown at 25°C for 16 hours.

The purification of WT and mutant hisactophilin begin with harvesting the cells by centrifugation. The harvested cells from WT and mutant (from 6 L of media) were suspended on ice in 25-50 ml of buffer A (20 mM Tris pH 8.0, 1 mM EDTA). Cells were lysed using an EmulsiFlex-C5 homogenizer (Avestin Inc, Ottawa, ON) at a pressure of at least 17000 psi according to the manufacture's directions. The lysate was incubated with 10 mg/ml DNase and 5 mM $MgCl_2$ on ice for one hour. After incubation the cell debris was collected by centrifuging the lysate at 20000 g for 20 minutes at 4ºC.

The supernatant filtered using a 0.4 μm filter (from Life Sciences) and loaded onto a DEAE Sepharose anion-exchange column (Sigma). Equilibrated with Buffer A. Anion exchange chromatography was performed at 4ºC. The 50 mL cell lysate was loaded at 3 ml/min onto a self-packed 2.5 x 50cm DEAE-Sepharose Fast Flow (Sigma chemical company, St. Louis, MO) anion exchange column pre-equilibrated with 2 L of buffer A (20 mM Tris pH 8.0, 1 mM EDTA) using the BioLogic LP chromatography system (Bio-Rad Laboratories Hercules, A). The column was washed with 90 ml of buffer A at 2 ml/min. The protein was eluted using a linear gradient of 0-400 mM NaCL over 350 min at flow rate of 2 ml/min, WT and variants from mutagenesis eluted between 140-240 mM NaCl. Fractions containing hisactophilin identified by SDS-PAGE to contain hisactophilin were collected and

concentrated to approximately 7.5 ml using a 3 kDa cutoff YM-3 Amicon (Millipore, Billerica, MA) filter.

The second purification step employed size exclusion chromatography. A 5x 60 cm G75-Superdex Size Exclusion column (Pharmacia, Piscataway, NJ) was equilibrated with 1 column volume (approximately 320 ml) of exclusion buffer (150 mM Ammonium Sulphate, 1 mM EDTA, 1mM DTT) using a BioCad Sprint perfusion chromatography System (Perceptive BioSystem, Ramsey,MN). The fractions from anion exchange chromatography were diluted to 15 ml with 2 x size exclusion buffer and injected onto the column over 3 injections using a 5 ml loop. The protein was eluted with size exclusion buffer at 3 ml/min. Fractions containing hisactophilin were identified by SDS-PAGE; hisactophilin containing fractions were combined and dialyzed against 25 mM ammonium carbonate, and concentrated with a 3 kDa cutoff YM-3 Amicon filter. The protein concentration was determined by absorbance measurements at 280 nm using an extinction coefficient of 0.33 for a 1 mg/ml protein solution (Liu et al., 2001). The WT and variant proteins were divided into aliquots with desired concentration, flash frozen in liquid nitrogen, lyophilized and stored at -80°C.

2.5.4 SDS-PAGE

Polyacrylamide gels were made from 40% 37.5:1 acrylamide:bis-Acrylamide (Bioshop), with the stacking gel being 5% polyacrylamide (5% 37.5:1 Acrylamide:bis-Acrylamide v/v, 0.125 M Tris-HCl pH 6.8 v/v, 0.1% SDS w/v, 0.1% ammonium persulfate (APS) w/v, 0.1% Tetramethylethylenediamine (TEMED) w/w), and the resolving gel being 15% polyacrylamide (15% 37.5:1 acrylamide:bis-acrylamide v/v, 0.375 M Tris-HCl pH 8.8 v/v, 0.1% SDS w/v, 0.1% APS w/v, 0.05% TEMED v/v).

To each sample well, 10 uL of sample was added. Gels were run with running buffer (1.5% Tris-Base w/v, 7.2% Glycine w/v, 0.5% SDS w/v) at a constant voltage of 140 Volts until the marker dye had begun to leave the gel. Gels were stained using coomassie staining solution (50% MeOH v/v, 5% glacial acetic acid v/v, and 0.1% coomassie brilliant blue R-250) for 3 hours with mild shaking at room temperature. Gels were then destained with destain solution (50% MeOH v/v, 5% glacial acetic acid v/v) for 6 hours with mild shaking at room temperature.

2.5.6 Mass spectrometry

The protein samples of WT and H90G mutant hisactophilin were prepared for electrospray mass spectroscopy. The solution used for MS contained 50:50 MeCN:H_2O and 0.2% (v/v) formic acid. Samples were analyzed using a Micromass Q-TOF Ultima^{TM} Global instrument (supplied by the University of Waterloo mass Spectrometry facility). The calculated mass was determined by mass calculation using the amino acid sequent and then the observed mass was compared to WT mass.

2.6 Spectroscopic probes used to monitor urea denaturation curves

2.6.1 Circular Dichroism spectroscopy (CD)

Protein secondary structural changes occurring during urea denaturation may be monitored by CD spectroscopy. Monitoring denaturation by CD utilizes circular polarized light to determine the secondary structure of a peptide. Circular polarized light, unlike planar polarized light, rotates with the frequency of the radiation (Woody, 1996). This type of polarized light can be rotating in either a clockwise or counterclockwise direction and is labeled by convention as right-circular polarized and left-circular polarized light, respectively (Woody, 1996). During an experiment, a sample is exposed to alternating left-and right-

circular polarized light and the absorbance of each is measured (Manning, 1989). The difference in the absorbance is defined as the circular dichroism (ΔA). When the CD is collected over a range of wavelengths in the far-UV, (~ 180-260 nm) it will produce a CD spectrum which is unique to the secondary structure of that protein (Berndt, 1996). The far-UV region can report on both tertiary and secondary structure. The CD spectrum in the amide region gives information about the secondary structure of the protein as α-helices, β-sheets, turns and random coils have characteristic CD spectra in this region (Johnson, 1990).

2.6.2 Fluorescence measurements

The second optical probe to be used for monitoring urea denaturation of (WT and mutant hisactophilin) is fluorescence spectroscopy. Fluorescence emissions are observed when an excited electron returns to the ground state by releasing energy as light. The fluorescence emissions will be limited to certain species within the sample known as fluorophores. In the case of proteins, fluorescence originates mainly from the aromatic amino acids phenylalanine, tyrosine or tryptophan (Schmid, 1989). The fluorescence that is emitted from a fluorophore is highly dependent on the physio-chemical properties of the surrounding environment (Szabo, 2000). This condition exists because the long lifetime of the excited state allows for a variety of interactions to occur prior to emission. These possible interactions can greatly modify the fluorescence emissions of a protein (Schmid, 1989). since the local environment of the fluorophore is highly dependent upon the tertiary structure of a protein, a dramatic change in tertiary structure, such as denaturation, will cause a significant change in the fluorescence spectrum (Schmid, 1989). These unique properties of fluorescence was utilize to monitor the tertiary structure of WT and mutant hisactophilin during urea denaturation. Tryptophan is the most sensitive fluorophore yielding a very good signal whereas tyrosine yields much less and phenylalanine is rarely observed in native proteins

(Schmid, 1989). As well, fluorescence intensity tends to be decreased as residues become buried within the interior of a protein. WT hisactophilin has relatively weak fluorescence, because it does not contain tryptophan and fluorescence originates mainly from three tyrosine residues: Y62 is exposed to solvent, while Y52 and Y92 are partially buried near the surface of the protein (Figure 2.2) (Habazettl et al., 1992).

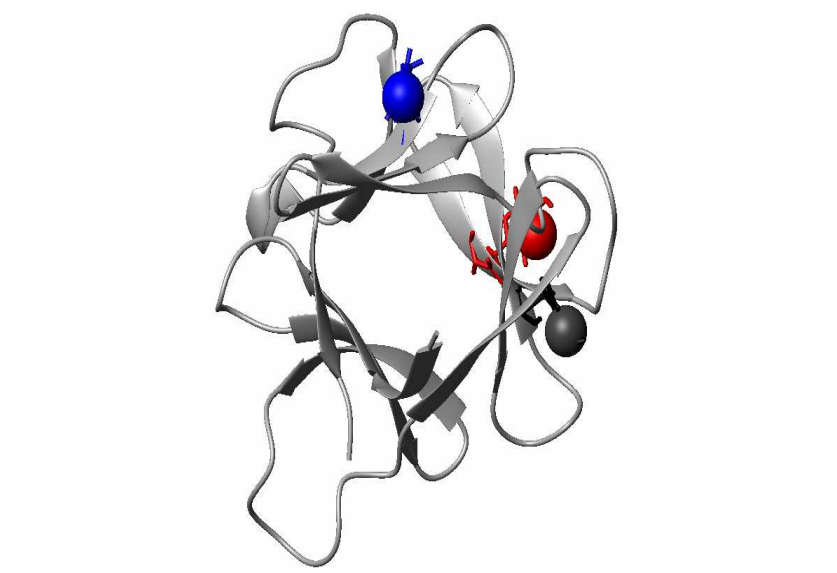


Figure 2. 3 Location of tyrosine fluorophores in WT hisactophilin.
The tyrosine 52 colored in black sticks, Y62 colored in red sticks and Y92 colored in blue sticks. Ribbon diagrams were generated using MolMol (Koradi et al., 1996) using PDB accession code 1HCD.

To determine the effects of the mutations on hisactophilin stability, urea equilibrium denaturation curves was measured at different pHs (6.7, 7.7, and 8.7). At pH 8.7 where the protein has maximal stability and pH 6.7 has lowest stability. It has been determined previously that denaturation by urea is reversible from pH 5.7-9.7 and data fit well to a 2-state transition between native and denatured states (Liu et al., 2001).

2.7 Equilibrium studies

2.7.1 Denaturation curve sample preparation

Urea denaturation curves were obtained by fluorescence measurements using a 1 cm pathlength cuvette in a Fluoro-22 spectrfluorometer (Jobin Y Von-Spex, Instruments S.A., Inc., NJ) that was interfaced to the fluoro-22 /Neslab water bath system (RTE-211, NESLAB Instruments Inc., NH). Fluorescence measurements were made with an excitation wavelength of 277 nm and emission wavelength of 306 nm with slit width set to 1 and 5 nm, respectively. For denaturation curve measurements, the samples were incubated in 25ºC water bath during the measurement time. Samples were incubated in the cuvette for at least one minute prior to taking readings. Each urea denaturation curve was obtained by following the fluorescence signal of 30-31 samples at pH 8.7 and pH 7.7 containing 2-3 mg/ml lyophilized WT or hisactophilin variant protein in 500 mM glycine/sodium glycinate, pH 8.7, 500 mM KH_2PO_4/K_2HPO_4, pH 7.7, and pH 6.7, 10 m M DTT and 10 mM EDTA and a concentration of urea varying from 0 to 10 M. First filtered ultrapure 10 M urea (Sigma) and filtered MilliQ water were aliquoted in 1.5 ml Eppendorf tubes according to desired urea concentration. Stock protein solutions was added last using Hamilton syringe .

CD measurements were preformed with a 1 mm path length cuvette using a J715 spectropolarimeter (Jasco) to assess secondary structure before and after denaturation. Temperature was controlled using a peltier cell (ELFIN model ELDC5D4, Japan Servo Co.Ltd). CD measurements were acquired at 227 nm with slit widths of 5 nm. Equilibrium and incubation prior to measurement of the CD signal for all samples was followed using the same procedure as in fluorescence measurements as mentioned above.

The change in the fluorescence intensity and the change in ellipticity were plotted as a function of denaturant concentration.

2.7.2 Denaturation curve data analysis

The free energy of transfer of the side chains and polypeptide backbone from water to solution of denaturant is not linearly proportional to the concentration of denaturant (Liu et al., 2002). Denaturation curve data were fit using the binomial extrapolation method (BEM) to a two transition between the native and the denatured state, with a non sloping pretranstion baseline and a sloping post-transition baseline using Equation 2.1. Fitting of the experimental curves was preformed using OriginPro7.5 (OriginLab). The m-value allowed to vary, then fixed, which an average value is obtained by calculating m based on the change in solvent accessible surface area for the unfolding. The m-value is related to how cooperative the transition is, how much structure remains in the denatured state, perhaps how much denaturant binds to the unfolded state.

The denaturation curves were fit to a two state transition using Equation (2.1).

$$Y = \frac{Y_N - [Y_N - (Y_U + S_U[\text{urea}])]e^{-(\Delta G_u^\circ + m_1[\text{urea}] - m_2[\text{urea}]^2)/RT}}{1 + e^{-(\Delta G_u^\circ + m_1[\text{urea}] - m_2[\text{urea}]^2)/RT}} \quad (2.1)$$

where Y is the measured optical signal at a given urea concentration, Y_N is the signal for the native state, Y_U is the signal for the unfolded state in the absence of urea, S_U is the denaturant dependence of the signal for the unfolded state, m_1 and m_2 are constants describing the denaturant dependence of Gibbs free energy of unfolding, R is the gas constant, T is the temperature in Kelvin, $\Delta G°_u$ is the Gibbs free energy of unfolding in the absence of

denaturant, m_2 was fixed to 0.072 based on combined analysis of urea and thermal denaturation data (Liu et al., 2002).

2.8 Kinetic studies

2.8.1 Urea-induced unfolding kinetics

Unfolding rates were sufficiently slow that they could be measured by manual mixing experiments, and monitored by fluorescence using the fluorlog-3 (HORIBA; JOBIN YVON-SPEX). Kinetics of unfolding of WT and mutant hisactophilin proteins was preformed by diluting the native protein stock solution 0.2 mg/ml 10-fold in 50 mM glycine/sodium glycinate (10 mM DTT, 10 mM EDTA) pH 8.7. Different concentrations of urea were buffered with unfolding buffer. The unfolding buffer was composed of 50 mM glycine/sodium glycinate (10 mM DTT, 10 mM EDTA) pH 8.7, filtered 10 M urea. 900 uL of unfolding buffer was placed in a quartz cuvette (1 cm pathlength), to which 100 uL of native protein was added and mixed by pipetting. The unfolding signal which is an increase in fluorescence was monitored by fluorescence using Fluorolog-3 (HORIBA; JOBIN YVON-SPEX) with the excitation wavelength 277 nm and emission wavelength 306 nm with slit widths set to 1 and 5 nm, respectively. To minimize temperature artifacts, both protein and urea solutions were pre-equilibrated at 25ºC prior to initiation of unfolding. The dead time of each unfolding experiment was determined using a stop-watch and was on average approximately 10 seconds in duration. The final concentration of urea was routinely checked by refractive index with the equation (Pace, 1986).

$$[\text{urea}] = 117.66\Delta n + 29.753\Delta n^2 + 185.56\Delta n^3$$

where Δn is the difference in refractive index readings between the urea solution and a blank (buffer).

2.8.2 Refolding kinetics measurement

Refolding kinetics at moderate urea concentrations were sufficiently slow that they could be measured using manual mixing techniques and monitored by fluorescence. Refolding kinetics preformed by diluting a stock solution of 0.2 mg/ml unfolded protein 10-fold in 20 mM Acetic acid, pH 4.41, 10 mM DTT. The unfolded protein was then diluted so that the urea concentration was varied for each refolding condition , 900 uL of refolding buffer was placed in a quartz cuvette (1 cm pathlength) then mixed with 100 uL of denatured protein by pipetting. The refolding signal was recorded by fluorescence using Fluorolog-3 (HORIBA; JOBIN YVON-SPEX) with the excitation wavelength 277 nm and emission wavelength 306 nm with slit widths set to 1 and 5 nm, respectively. To minimize temperature artifacts, both protein and urea solutions were pre-equilibrated at 25ºC prior to initiation of refolding. The dead time of each unfolding experiment was determined using stop-watch and was on average 10 seconds. The final concentration of urea was routinely checked by refractive index as indicated in section 2.8.1 (Pace, 1986).

2.8.3 Stopped-flow fluorescence kinetics

Refolding kinetics at very low urea concentrations is too fast to be measured using manual mixing techniques. The kinetic of WT and hisactophilin mutant refolding were measured by stopped-flow, using an Olis RSM 1000 instrument (On-line instrument systems, Bogart, Georgia). The temperature was maintained at 25ºC using a circulating water bath connected to the stopped-flow instrument. Fluorescence measurements were made with an excitation wavelength of 277 nm and emission wavelength of 306 nm. Kinetics were measured at least two times under identical conditions, averaged, and analyzed using Biokine software 32 (Biologic).

2.8.4 Fitting of the unfolding and refolding traces

Raw kinetic data of the WT and mutant hisactophilin were fit to a single or double exponential with a linear drift using Biokine 32 (Biologic, Version 2.10, Molecular Kinetics) and OriginPro7.5 (OriginLab) software. For the 2-state model, the time course should follow a simple exponential, with a small linear drift being observed in some cases due to expermental factors such as diffusion of reagents or lamp instability.

The rate constants of folding and unfolding of WT and mutant hisactophilin were fit to 2-state model as described by (Liu et al., 2002) using a quadratic equation.

$$\ln k_{unf}\,[urea] = \ln k^{o}_{unf} (m_{unf,1}/RT)[Urea] - (m_{unf,2}/RT)[Urea]^2 \qquad (2.2)$$

Refolding rates were fit using equation:

$$\ln k_{ref}\,[urea] = \ln k^{o}_{ref} (m_{ref,1}/RT)[urea] - (m_{ref,2}/RT)[Urea]^2 \qquad (2.3)$$

Complete kinetic data were fit using:

$$\ln k_{obs}([urea]) = \ln(k^{o}_{unf} \exp\{(m_{unf,1}/RT)[Urea] - (m_{unf,2}/RT)[Urea]^2\} + (k^{o}_{ref} \exp\{(-m_{ref,1}/RT)[urea]\}) \qquad (2.4)$$

where k^{o}_{ref} and k^{o}_{unf} are the folding and unfolding rate constants at 0 M urea, $m_{unf,1}$ and $m_{unf,2}$ describes a quadratic denaturant dependence for the natural logarithm of the refolding rate. The value of $m_{unf,2}$ was fixed to 0.028, which is the average value for curvature of unfolding rates obtained from repeated experiment. R is the gas constant, T is absolute temperature and m_{ref} and m_{unf} are constants of proportionality which describes the sensitivity of the refolding and unfolding rates, respectively, on denaturant concentration.

Chapter 3: Results

3.1 Mutagenesis results

All of the constructs (F6L, F13Y, I85L, H90G, and I93L) were confirmed first by restriction enzyme analysis and then by DNA sequence analysis (Mobix Central Facility, DNA Synthesis Laboratory, The Institute for Molecular Biology and Biotechnology, McMaster University, Hamilton, ON). Except for H90G, the mass of the purified protein was also verified by mass spectrometry. The results of the sequence analysis of the variants and the WT as well as H90G mass spectroscopy data are presented in Appendices 1-6. All sequences were compared with the WT sequence confirming successful mutagenesis.

F13Y, I85L, I93L mutants protein were obtained at a similar level to that of the WT protein (~28 mg/L) and exhibited similar chromatography profiles during purification. However, H90G was obtained at a higher level (~ 48 mg/L). In contrast, the F6L point mutation exhibited aggregation during purification resulting in much lower yields of purified protein (~ 12 mg/L). This protein was also significantly less stable than WT (see section 3.2.1.3). Thus, the relative expression levels are correlated with the relative stabilities of the proteins.

As indicated in chapter 1 and 2, hisactophilin does not contain tryptophan. Fluorescence therefore, arises mainly from three tyrosine residues (Figure 2.3). The fluorescence for WT hisactophilin increases by ~30 % upon unfolding (Figure 3.1). Mutating phenylalanine 13 to tyrosine (4 tyrosine) causes a substantial increase in the fluorescence quantum yield. As a consequence, the difference between the fluorescence of the native and denatured state becomes very small, so that the unfolding transition can no longer be accurately monitored by fluorescence.

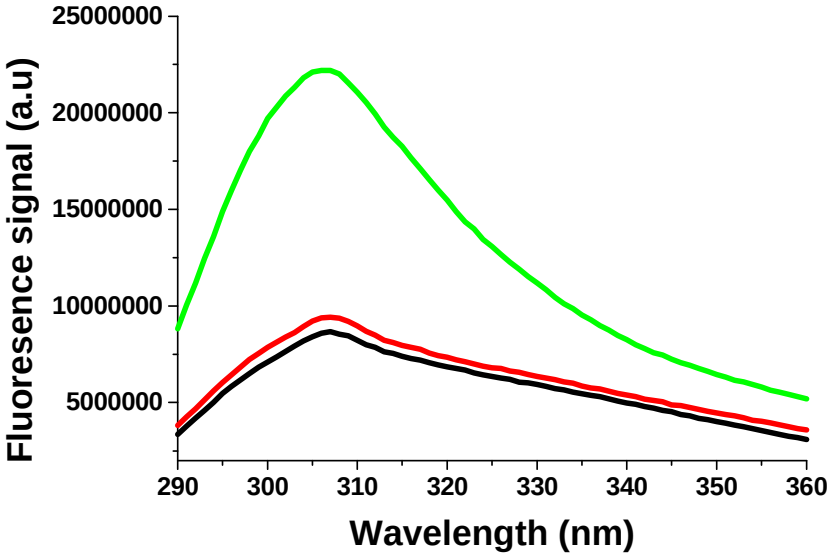


Figure 3.1 Hisactophilin fluorescence spectra as a function of denaturant concentration.
The fluorescence monitored scans were measured at 2mg ml^{-1} in different urea concentrations: 0M (black line), 3 M (red line), and 9 M (green line). All the scans were performed at 25°C in glycine buffer.

3.2 Equilibrium denaturation curves

Chemical denaturation of WT and H90G, I85L, I93L, F6L hisactophilin mutants by urea at pH 6.7, 7.7 and 8.7, monitored by fluorescence and Circular dichroism (CD) spectroscopy is highly reversible (Figures 3.1, 3.2, and 3.3). Chemical denaturation of WT hisactophilin has previously been found to be highly reversible at pH 6.7, 7.7 and 8.7 (Meissner, 2007); (Marty.T.J.Smith and E. Meiering, unpublished data). The results obtained by fluorescence and CD were fit using OriginPro7.5 and are consistent, within experimental error, with a two-state model of unfolding between the native and denatured states (Figure 3.2). The denaturation curves were fit to a two state transition using Equation 2.1.

The thermodynamic analysis requires that chemically induced unfolding be reversible and at equilibrium. Urea denaturation of the WT hisactophilin is highly reversible (Liu et al., 2001, Meissner, 2007, Marty .T.J. Smith and E. Meiering, unpublished data) and the time to

reach equilibrium is short (less than 4 hours). The reversibile experiments were performed individually for each pH ranging from 6.7 to 8.7. According to the CD and fluorescence spectra, the unfolding and refolding of hisactophilin is more than 90% reversible at each of these pHs (Liu et al., 2002). Stability here was measured at several pHs: 6.7, 7.7 and 8.7, and compared with previous results obtained at these pHs. Many previous experiments on PWT hisactophilin were performed at pH 6.7 (Liu et al., 2002); however, the protein is relatively unstable at this pH and the stability is very highly dependent upon pH. At pH 8.7, the protein has the highest stability, and the stability varies much less with pH. However, the relatively high stability made characterization of the relatively more stable mutants hard to measure accurately at these conditions.

Urea denaturation curves fit to obtain ΔG^{o}_{u}, C_{mid} and m-value, the higher the Gibbs free energy change is in the absence of denaturants, or the higher the unfolding midpoint is in denaturants, the more stable the protein is. These thermodynamic parameters were obtained for WT and mutant hisactophilin at different pHs (6.7, 7.7 and 8.7), the values of these parameters monitored by CD and fluorescence spectroscopy agrees with each other within experimental error for each pH.

The measurements of equilibrium denaturation was done at pH 6.7 (Table 3.1) when the protein is not very stable and then measured at pH 7.7 (Table 3.2) when the protein is more stable. Urea denaturation curves were measured at pH 8.7 (Table 3.3), when the stability of hisactophilin is greatest.

3.2.1 Equilibrium denaturation curves of WT and mutant hisactophilin in urea

3.2.1.1 Stability at pH 6.7

The conformational stability of WT and H90G mutant at pH 6.7 was determined using urea denaturation curves. The urea denaturation curves exhibit a two-state transition between the folded and unfolded protein. The PWT and WT behave the same within the experimental error. The data were fit with Equation 2.1 and the values from the fit are given in Table 3.1.

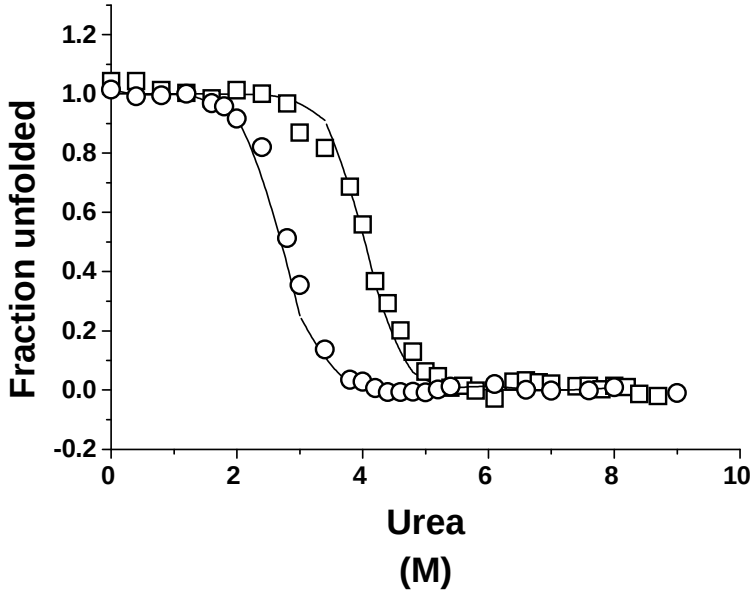


Figure 3. 2 Fluorescence-monitored urea denaturation curves in phosphate buffer at pH 6.7. Solution contained 50 mM KH_2PO_4/K_2HPO_4 pH 6.7, 1mM EDTA, 1mM DTT, and 0.2 mg ml^{-1} protein at 25°C. WT (○) C_{mid} of 2.78 M and H90G (□) with C_{mid} of 4.04 M, displayed in terms of the fraction of unfolded protein. The continuous line represents the best fit of the equilibrium denaturation data to Equation 2.1. WT results obtained from Marty. T. J. Smith.

Based on the fit of the experimental data measured by fluorescence and CD, the m-value obtained from the fitting (which is related to the extent of change in exposure of the hydrophobic side chains upon unfolding of native protein to the unfolded state) was

essentially the same for the WT and all mutants, within experimental error, so an average m-value was calculated and used to determine ΔG_u and $\Delta\Delta G_{u.}$ H90G is more stable than the WT at pH 6.7 with C_{mid} of 4.04 M, m-value of 2.14 kcal $mol^{-1}M^{-1}$and $\Delta\Delta G$ of -2.84 $kcalmol^{-1}$ as shown in (Table 3.1).

Table 3.1 Equilibrium parameters for WT and H90G mutant hisactophilin at pH 6.7.

Protein	pH	Probe	C_{mid} (M)	m ($kcalmol^{-1}$)	$\Delta G°_u$ ($kcal\ mol^{-1}$)	$\Delta\Delta G$ ($kcal\ mol^{-1}$)
H90G	6.7	FL	4.04±0.03	2.14±0.11	8.65±0.45	-2.84
WT[a]	6.7	FL	2.78±0.26	2.37±0.08	6.59±0.34	
PWT[c]	6.7	FL	3.15±0.09	1.96±0.26	6.17 ±0.82	
		CD	3.26 ±0.04	2.15± 0.14	7.0 ± 0.45	
PWT[b]	6.7	FL	2.91±0.01	2.51±0.07	7.30 ± 0.20	
WT FGF	6.7	FL	1.11±0.01	18.9±0.6	21.1±0.6	
V109L[F]	6.7	FL	0.86±0.02	20.2 ± 1.1	17.4 ± 0.5	4.9

WT Wild Type Hisactophilin
WT[a] Marty T.J. Smith and E.M Meiering unpublished data
PWT[b] (Wong, 2004) PWT data with 4 extra amino acids
PWT[c] (Liu et al., 2002), data obtained at 20ºC
V109 L[F] FGF data which V109L correspond to I93L in hisactophilin (Brych et al., 2003).
Equilibrium values (C_{mid} , m-value and ΔG_U) were obtained from fitting urea denaturation curve data in (Figure 3.2) to Equation 2.1, m_2 was fixed to 0.072
$\Delta\Delta G = (C_{mid\ WT} - C_{mid\ mutant})(m_{WT} + m_{mutant})/ 2$. A negative value indicates a more stable mutant.

3.2.1.2 Stability at pH 7.7

The denaturation curves at pH 7.7 for the WT and mutant hisactophilin were monitored by both CD and fluorescence. The CD and fluorescence denaturation curves at pH 7.7 exhibited two-state transition (Figure 3.3). The curves were fit with Equation 2.1 to obtain

the thermodynamic parameters which are summarized in (Table 3.2), an average m-value was calculated and used to determine ΔG_u.

The thermodynamic stability of WT increases from 6.6 Kcalmol^{-1} at pH 6.7 to 8.0 kcal mol^{-1} at pH 7.7. A comparison of the mutants with WT at pH 7.7 yields the same trend as that seen at pH 6.7 for H90G. H90G is the most stabilizing mutation relative to the WT with C_{mid} of 5.28 M, m-value of 2.03 kcalmol^{-1}M^{-1} and $\Delta\Delta G$ of -2.25 kcalmol^{-1}. I85L is more stable as well compared with the WT with C_{mid} of 4.46 M, m-value of 1.88 kcalmol^{-1}M^{-1} and $\Delta\Delta G$ of -0.66 kcalmol^{-1}. I93L is less stable relative to the WT with C_{mid} of 3.87 M, m-value of 0.94 kcalmol^{-1}M^{-1} and and $\Delta\Delta G$ of 0.49 kcalmol^{-1}. F6L is a very unstable mutant with C_{mid} of 2.44, m-value of 1.85 kcal mol^{-1}M^{-1} and and $\Delta\Delta G$ of 3.26 kcalmol^{-1} as indicated in (Table 3.2).

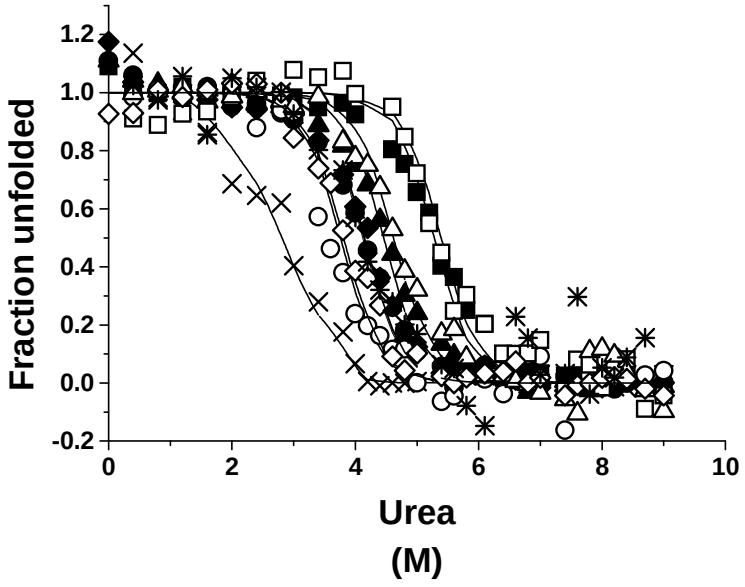


Figure 3. 3 Fluorescence and CD-monitored urea denaturation curves in phosphate buffer at pH 7.7. Solution contained 50 mM KH_2PO_4/K_2HPO_4 pH 7.7, 1 mM EDTA, 1 mM DTT, and 0.2 mg ml^{-1} protein at 25°C. Fuoresence in closed symbols and CD in open symbols.WT (●) C_{mid} of 4.12 M, WT **(**○**),** H90G (■) C_{mid} of 5.28 M, H90G **(**□**),** I85L (▲) C_{mid} of 4.46 M, I85L **(**△**),** I93L (◆) C_{mid} of 3.87 M, I93L **(**◇**),**F6L (×) C_{mid} of 2.44 M, and F13Y (∗), displayed in terms of the fraction of unfolded protein. The continuous line represents the best fit of the equilibrium denaturation data to Equation 2.1.

Table 3. 2 Equilibrium parameters for WT and mutant hisactophilin at pH 7.7.

Protein	pH	Probe	C_{mid} (M)	m (kcal mol $^{-1}$M^{-1})	ΔG_u (kcal mol^{-1})	$\Delta\Delta G$ (kcal mol^{-1})
WT	7.7	FL	4.12± 0.03	1.94±0.07	7.99±0.29	
	7.7	CD	3.76± 0.03	1.94±0.07	7.29± 0.26	
WT[a]	7.7	FL	4.34	2.14	7.92	
PWT[b]	7.7	FL	4.39± 0.02	2.13± 0.05	7.96± 0.19	
	7.7	CD	4.31± 0.02	2.23± 0.06	8.30± 0.23	
H90G	7.7	FL	5.28± 0.05	1.94±0.07	10.24± 0.37	-2.25
	7.7	CD	5.39± 0.14	1.94±0.07	10.46±0. 38	
I85L	7.7	FL	4.46± 0.04	1.94±0.07	8.65±0.31	-0.66
	7.7	CD	4.62± 0.06	1.94±0.07	8.96±0.33	
I93L	7.7	FL	3.87± 0.05	1.94±0.07	7.50±0.27	0.49
	7.7	CD	3.81± 0.06	1.94±0.07	7.39±0.27	
F6L	7.7	FL	2.44± 0.06	1.94±0.07	4.73±0.18	3.26

WT Wild Type Hisactophilin
WT[a] Marty T.J. Smith and E.M Meiering unpublished data
PWT[b] (Wong, 2004) PWT data with 4 extra amino acids
PWT[c] (Liu et al., 2002), data obtained at 20ºC
Equilibrium values (C_{mid} , m-value and ΔG_U) were obtained from fitting urea denaturation curve data in (Figure 3.3) to equation 2.1, m2 was fixed to 0.072
$\Delta\Delta G = (C_{mid\ WT} - C_{mid\ mutant})(m_{WT} + m_{mutant}) / 2$. A negative value indicates a more stable mutant.

3.2.1.3 Stability at pH 8.7

(Figure 3.4) shows the fraction of unfolded protein measured by fluorescence and CD as a function of urea concentration of the WT and mutant hisactophilin at pH 8.7.

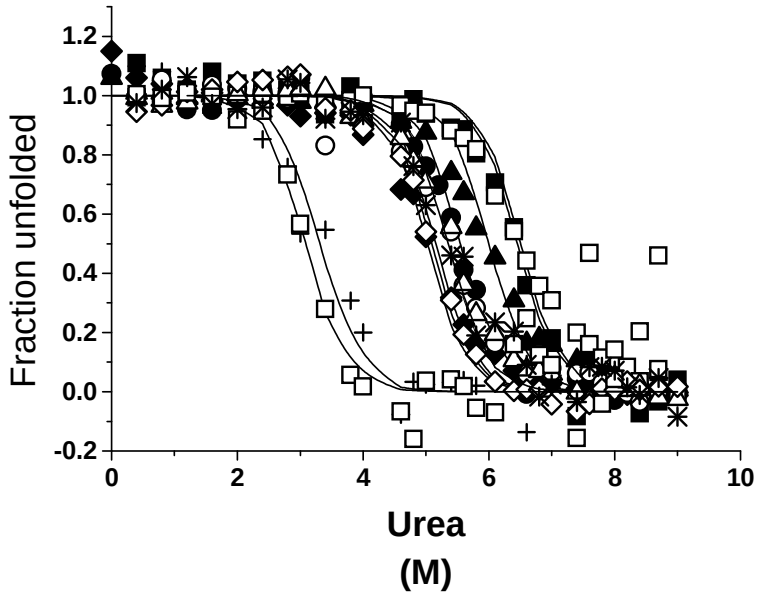


Figure 3. 4 Fluorescence and CD-monitored urea denaturation curves in glycine buffer at pH 8.7. Solution contained 50 mM, glycine/sodium glycinate pH 8.7, 1 mM EDTA, 1 mM DTT, and 0.2 mg ml^{-1} protein at 25°C. Fuoresence in closed symbols and CD in open symbols.WT (●) C_{mid} of 5.49 M, WT **(**○**)** C_{mid} of 5.36 M **,** H90G (■) C_{mid} of 6.39 M, H90G **(**□**)** C_{mid} of 6.51 M**,** I85L (▲) C_{mid} of 5.94 M, I85L **(**△**)** C_{mid} of 5.16 M**,** I93L (◆) C_{mid} of 5.02 M, I93L **(**◇**)**C_{mid} of 5.09 M**,** F6L (×) C_{mid} of 3.32 M, (-) C_{mid} of 3.09 M and F13Y (∗) CD data with C_{mid} of 5.3 M, displayed in terms of the fraction of unfolded protein. The continuous line represents the best fit of the equilibrium denaturation data to Equation 2.1.

Analysis of urea denaturation curves for fluorescence for the WT at pH 8.7 yielded a ΔG_u of 10.70 kcal mol^{-1}, and a C_{mid} value of 5.49 M.

The fitted values obtained by fluorescence and CD are in agreement within error, providing further support for the 2-state model. From the results listed in Table 3.3, H90G with C_{mid} of 6.39 M, m-value of 2.12 $kcalmol^{-1}M^{-1}$ and $\Delta\Delta G$ of -1.76 $kcalmol^{-1}$ is considerably more stable than the WT. I85L with C_{mid} of 5.94 M, m-value of 2.09 kcal $mol^{-1}M^{-1}$ and $\Delta\Delta G$ of -0.88 $kcalmol^{-1}$ is a little more stable than the WT. I93L is slightly less stable than the WT with C_{mid} of 5.02 M, m-value of 1.94 $kcalmol^{-1}M^{-1}$ and of $\Delta\Delta G$ 0.92 kcal mol^{-1}. The F6L mutation is less stable than the WT with C_{mid} of 3.32 M, m-value of 1.45 $kcalmol^{-1}M^{-1}$ $\Delta\Delta G$ of 4.23 $kcalmol^{-1}$. From the data indicated above there is clear evidence

that H90G and I85L are more stable mutants than the WT based on the midpoint of unfolding and ΔΔG (Figure 3.4, and Table 3.3). In contrast, I93L and F6L are less stable relative to the WT based on the midpoint of unfolding and ΔΔG.

In summary, the equilibrium data for WT and mutant hisactophilin can be fit by a simple 2-state process. The origins of the changes in stability for the mutants were further characterized by measuring the folding and unfolding kinetics described in the next section.

Table 3. 3 Equilibrium parameters for WT and mutant hisactophilin at pH 8.7.

protein	pH	probe	C_{mid} (M)	m (kcal $mol^{-1}M^{-1}$)	ΔG (kcal mol^{-1})	ΔΔG (kcal mol^{-1})
WT	8.7	Fl	5.49±0.07	1.95±0.11	10.70±0.62	
	8.7	CD	5.36±0.06	1.95±0.11	10.45±0.60	
WT[a]	8.7	Fl	5.37± 0.99	2.09±0.19	11.22±0.9	
PWT[b]	8.7	Fl	5.39± 0.03	2.53± 0.14	11.56± 0.82	
PWT[c]	8.7	Fl	5.33±0.33	2.28±0.07	10.11±0.34	
	8.7	CD	5.21±0.04	2.14±0.06	9.21±0.27	
H90G	8.7	Fl	6.39±0.07	1.95±0.11	12.46±0.71	-1.76
	8.7	CD	6.51±0.07	1.95±0.11	12.69±0.72	
I85L	8.7	Fl	5.94±0.07	1.95±0.11	11.58±0.66	-0.88
	8.7	CD	5.16±0.10	1.95±0.11	10.06±0.60	
I93L	8.7	Fl	5.02±0.05	1.95±0.11	9.79±0.56	0.92
	8.7	CD	5.09±0.04	1.95±0.11	9.93±0.56	
F6L	8.7	Fl	3.32±0.10	1.95±0.11	6.47±0.41	4.23
	8.7	CD	3.09±0.17	1.95±0.11	6.03±0.47	
F13Y	8.7	CD	5.3±0.14	1.95±0.11	10.34±0.64	0.12

WT Wild Type Hisactophilin
WT[a] Marty T.J. Smith and E.M Meiering unpublished data
PWT[b] (Wong, 2004) PWT data with 4 extra amino acids
PWT[c] (Liu et al., 2002) data obtained at 20ºC
Equilibrium values (C_{mid} , m-value and ΔG_U) were obtained from fitting urea denaturation curve data in (Figure 3.4) to Equation 2.1, m_2 was fixed to 0.072
$\Delta\Delta G = (C_{mid\ WT} - C_{mid\ mutant})(m_{WT} + m_{mutant})/ 2$. A negative value indicates a more stable mutation

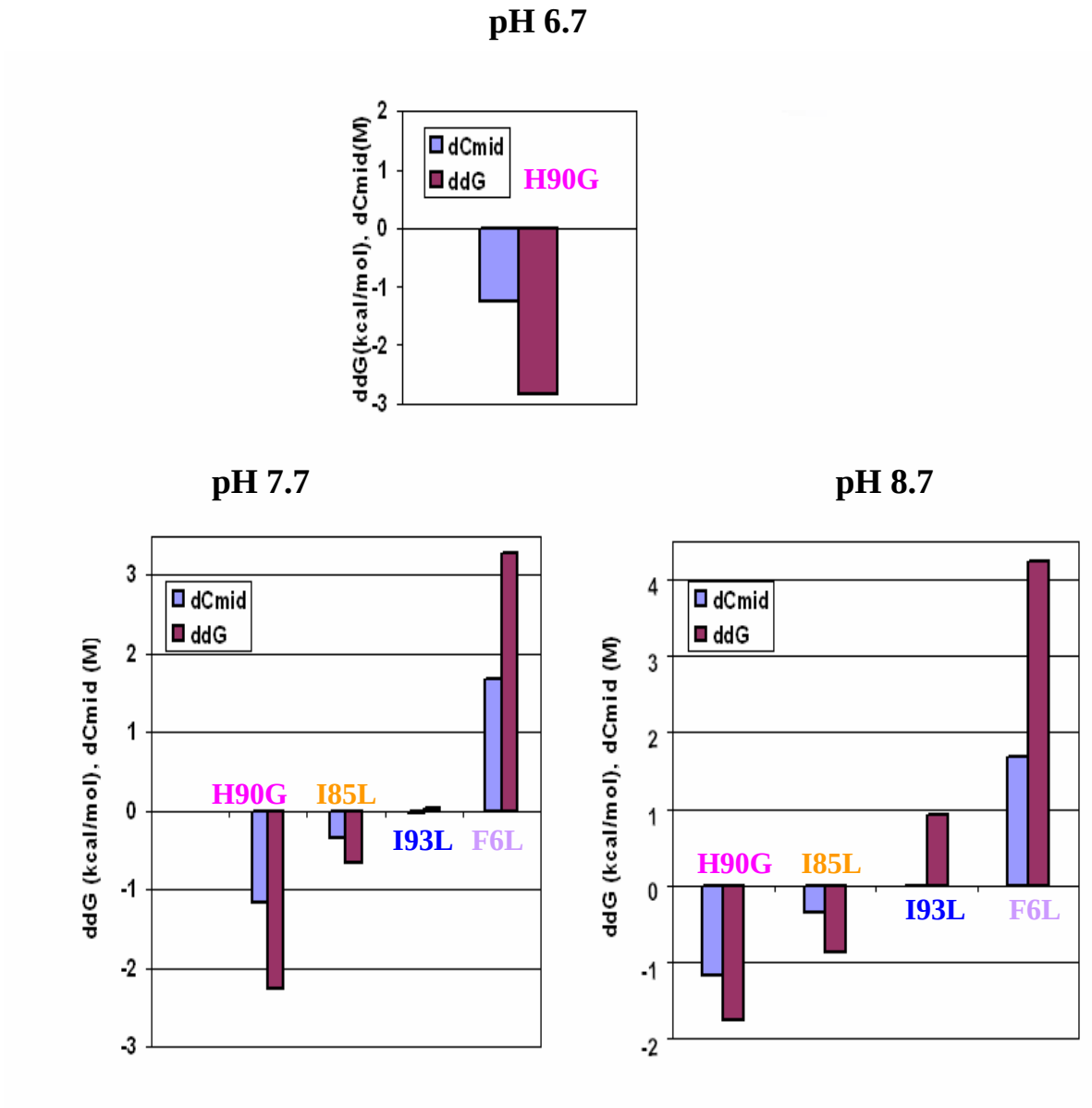


Figure 3. 5 $\Delta\Delta G$ and ΔC_{mid} as a function of pH for hisactophilin mutants.
ΔC_{mid} in blue color, and $\Delta\Delta G$ in purple color. $\Delta C_{mid} = C_{midWT} - C_{midmut}$, and $\Delta\Delta G = (C_{mid\ WT}-C_{mid\ mutant})(m_{WT}+m_{mutant})/2$, a negative value indicates a more stable mutation.

3.3 Kinetic results

3.3.1 Unfolding Kinetics

The origion of the difference stability for hisactophilin mutants H90G, I85L, and I93L compared with the WT were investigated using unfolding and refolding kinetic measurments monitored by manual mixing fluorescence from 6 to 9 M urea in (Figure 3.6).

For F6L the protein precipitated and the kinetic analysis was not performed for this mutation, and F13Y the increase in the fluorescence quantum yield lead to the difference between the fluorescence of the native and denatured state becomes very small, so that the unfolding kinetics can no longer be accurately monitored by fluorescence.

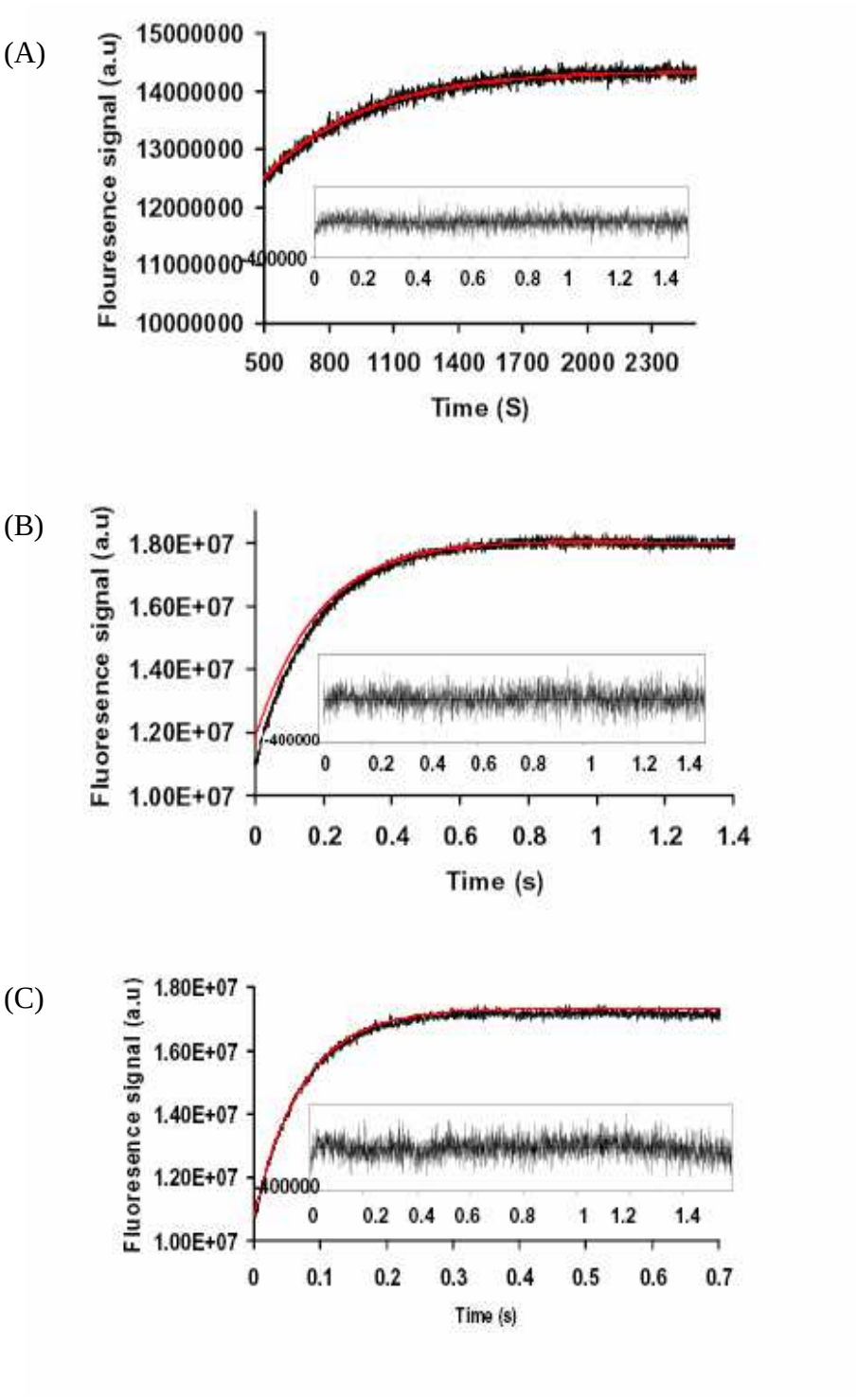


Figure 3. 6 Unfolding kinetics of WT hisactophilin.
A.Fluorescence-monitored unfolding kinetics of WT hisactophilin in 6.4 M urea fit to a single exponential equation with a linear drift. Shown inset the residual of the single exponential fit with k_u= 4.2 x$10^{-3}s^{-1}$. B.Fluorescence-monitored unfolding kinetics of WT hisactophilin in 7.2 M urea fit to single exponential equation with a linear drift. Shown inset the residual of the single exponential fit with k_u= 5.4 x$10^{-3}s^{-1}$. C. Fluorescence-monitored unfolding kinetics of WT hisactophilin in 9 M urea fit to single exponential equation with a linear drift. Shown inset the residual of the single exponential fit with k_u= 13.5 x$10^{-3}s^{-1}$. Residuals indicate the kinetics are well described by a single exponential.

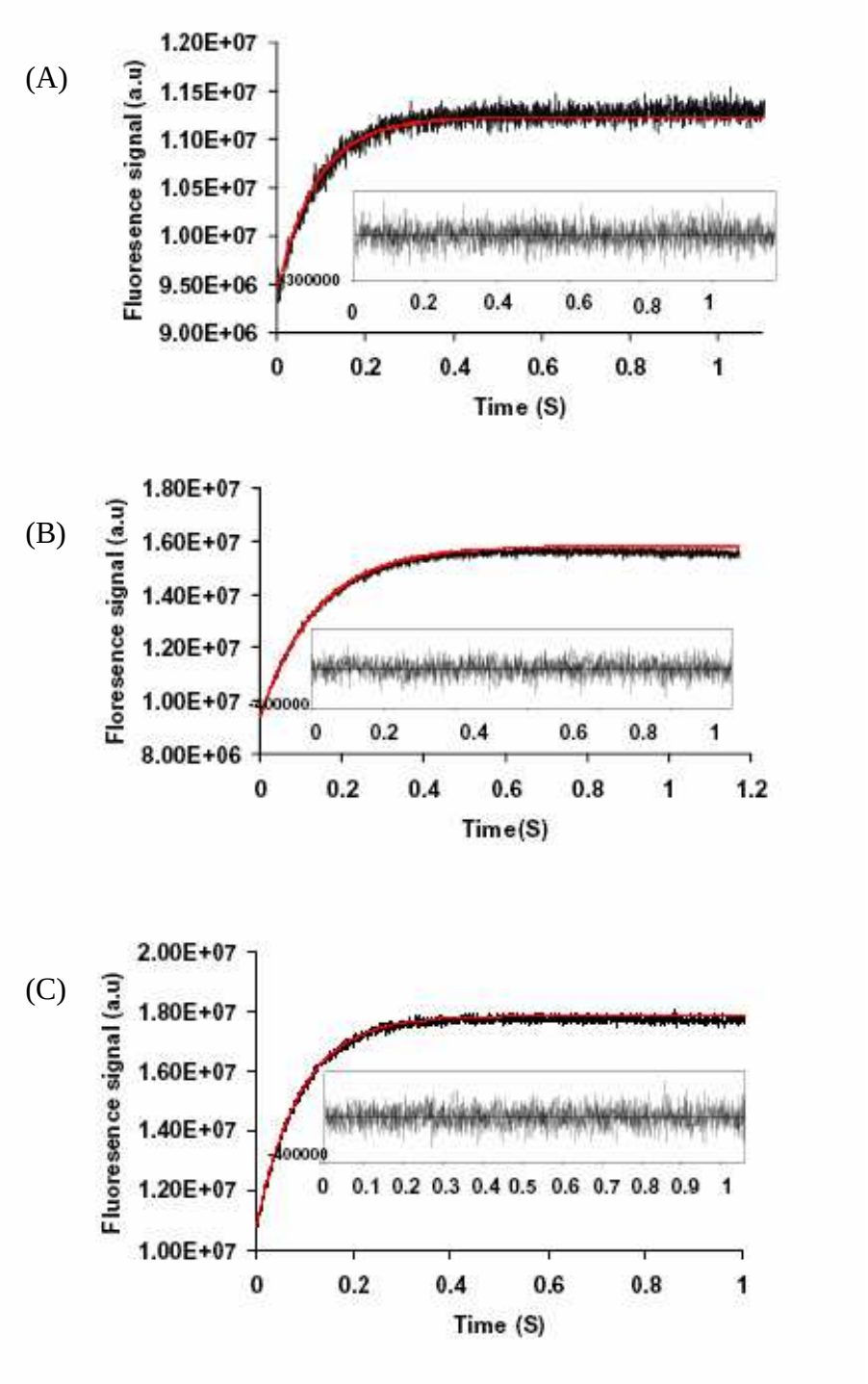


Figure 3. 7 Unfolding kinetics of H90G.
A. Fluorescence-monitored unfolding kinetics of H90G in 5.8 M urea fit to a single exponential equation with a linear drift. Shown inset the residual of the single exponential fit with $k_u=10.8x10^{-3}s^{-1}$. B.Fluorescence-monitored unfolding kinetics of H90G in 7.6 M urea fit to single exponential equation with a linear drift. Shown inset the residual of the single exponential fit with $k_u= 7.1s^{-1}$. C.Fluorescence-monitored unfolding kinetics of H90G in 8.6 M urea fit to single exponential equation with a linear drift. Shown inset the residuals of the single exponential fit with $k_u=10.9\ s^{-1}$. Residual indicate the kinetics are well described by a single exponential.

The kinetic data were fit by a single exponential equation, all unfolding reactions from pH range 6.7 to 8.7 were found to be single exponential (Figure 3.6). The amplitudes of the kinetic traces for the unfolding reactions correspond to the expected amplitude based on the fluorescence change in equilibrium curves for the WT and all the mutations being studied.

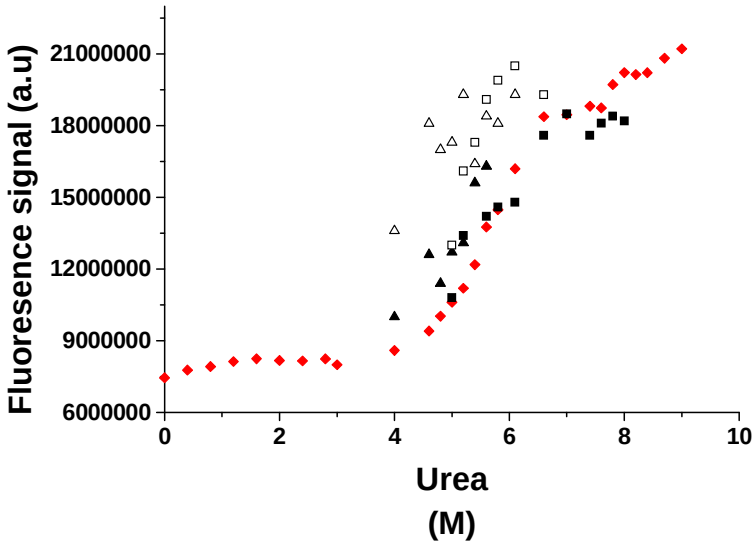


Figure 3. 8 Kinetic data of WT hisactophilin relative to equilibrium curve.
Fluorescence signal at pH 8.7**(♦).** Unfolding kinetic initial (□) and final (■) signals, and refolding initial (▲) and final (△) signals.

This is indicating that there is complete unfolding in both the WT and the mutants. The natural logarithm of the unfolding rates with urea concentration was found to vary linearly with denaturant concentration under strongly unfolding conditions, as illustrated in (Figure 3.10). Near the midpoint, the observed rate constants vary nonlinearly, H90G, aren't varying linearly because of the midpoint being so high. WT and I85L also start to curve. The slope of the plots is very similar for all the proteins, and not significantly different within experimental error. This indicates that the change in solvent exposure between the native and transition state is the same for all proteins. In order to facilitate comparison of the proteins, the

fitted m_{unf} values for each protein were averaged, and fixed for further fitting and calculations. Comparing the β_T of the WT and all mutant indicates that these mutant proteins all have native-like transition states for refolding.

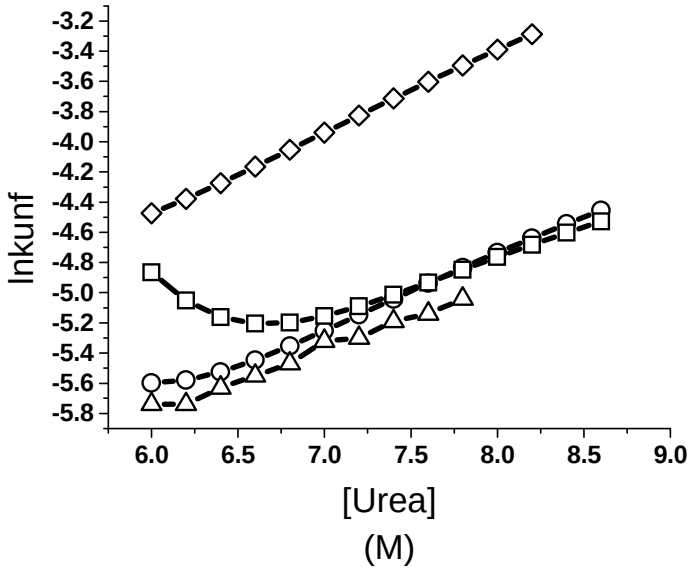


Figure 3. 9 Rate constants of unfolding of WT and mutant hisactophilin.
Dependence of the natural logarithm of the observed unfolding rate constant k_{unf} on denaturation concentration. WT (○), H90G (□), I85L (△), and I93L (◇). The continuous line represents a fit with single exponential for all the unfolding. Values for fits are summarized in Table 3.5.

There is overall correlation between values of ΔG_u, C_{mid} and m-values obtained from equilibrium and kinetic data (Table 3.4, 3.5).

All equilibrium and kinetic results obtained for WT and mutant hisactophilin at pH 8.7 are summarized in (Table 3.5).

Table 3. 4 Equilibrium and kinetic parameters for WT and H90G mutant hisactophilin at pH 6.7.

Protein	pH	Probe	C_{mid} (M)	M (kcalmol^{-1}M^{-1})	ΔG_u (kcalmol^{-1})	k_{ref} (s^{-1})	m_{ref} (kcalmol^{-1}M^{-1})	k_{unf} (s^{-1})	m_{unf} (kcalmol^{-1}M^{-1})	C_{mid} (M)	m (kcalmol^{-1}M^{-1})	ΔG_u (kcal mol^{-1})	βT
H90G	6.7	FL	4.04±0.03	2.14 ±0.11	8.65 ±0.45	**9±2**	**0.98± 0.04**	**$6.0x10^{-5}$± 7**	**0.67 ± 0.01**	**4.68**	**1.65±0.04**	**7.12±2.01**	**0.59 (MM)**
						104±9	**1.14±0.02**	**$3.0x10^{-5}$± 7**	**0.72 ± 0.02**	**5.21**	**1.86±0.03**	**8.93±2.50**	**0.61 (SF)**
WT[a]	6.7	FL	2.78±0.26	2.374±0.08	6.59±0.34	**2±8**	**1.39**	**$2.3x10^{-4}$ 14**	**0.5632**		**1.9612**	**5.58**	**0.28**
PWT[c]	6.7	FL	3.15 ±0.09	1.96 ±0.26	6.17 ±0.82								
		CD	3.26	2.15± 0.14	7.0 ± 0.45								
PWT[b]	6.7	FL	2.91 ±0.01	2.51 ±0.07	7.30 ± 0.20	**4±1**	**2.33±0.07**	**1.54±0.60**	**1.11 ± 0.05**	**5.35±**	**2.01±0.05**	**9.95±**	**0.68**
V109L[F]	6.7	FL	0.86±0.02	20.2 ± 1.1	17.4 ± 0.5	**3.04**	**6.04**	**3.11**	**0.56**				**1.10**

WT Wild Type Hisactophilin
WT[a] Marty T.J. Smith and E.M Meiering unpublished data
PWT[b] (Wong, 2004) PWT data with 4 extra amino acids
PWT[c] (Liu et al., 2002) data obtained at 20ºC
V109 L[F] FGF data which V109L correspond to I93L in hisactophilin
Equilibrium values (C_{mid} , m-value and ΔG_U) were obtained from fitting urea denaturation curve data in (Figure 3.2) to Equation 2.1, m2 was fixed to 0.072
$\Delta\Delta G = (C_{mid\ WT}-C_{mid\ mutant})(m_{WT}+m_{mutant})/2$. A negative value indicates a more stable mutation
Text in bold typeface refers to the kinetic data
Observed unfolding and refolding rate constants were fit to equation 1.11, C_{mid} , m_1, ΔG_u and βT were calculated from the fitted kinetic parameters using equations,2.4, 1.13, 1.14 and 1.15 respectively. Error estimates were obtained from the fitting program (Origin).MM, manual mixing, SF stopped flow data.

Table 3. 5 Summary of equilibrium and kinetic parameters for WT and mutant hisactophilin at pH 8.7.

protein	pH	probe	C_{mid} (M)	M (kcal	ΔGu(kcal mol^{-1})	k_{ref} (s^{-1})	m_{ref}(kcal $mol^{-1)}$	k_{unf} (s^{-1})	m_{unf}(kcal $mol^{-1}M^{-1}$)	C_{mid} (M)	M(kcalm $ol^{-1}M^{-1}$)	ΔG_u(kcal mol^{-1})	ΔΔG(kc $almol^{-1}$)	βT
WT	8.7	Fl	5.49±0.07	1.95± 0.11	10.70± 0.62	**4419±1443**	**1.57±0.46**	**$1.0x10^{-5}$±7**	**0.71±0.05**	**5.53**	**2.29±46**	**11.79± 9.13**		**0.68 (SF)**
	8.7	CD	5.36±0.06	1.95± 0.11	10.45± 0.60	**73.24±23**	**1.10±0.04**	**$70x10^{-5}$ ±1**	**0.74±0.2**	**5.63**	**1.85±0.20**	**9.51±9.71**		**0.6 (MM)**
WT[a]	8.7	Fl	5.37	2.09± 0.19	11.22± 0. 9	**156±6**	**1.19**	**$1.0x10^{-5}$ ± 16**	**0.69**		**1.89**	**9.82**		**0.63**
PWT[b]	8.7	Fl	5.39± 0.03	2.53± 0.14	11.56± 0.82	**4.04±1**	**2.33±0.07**	**1±0.60**	**1.1±0.1**	**5.35± 0.21**	**2.01±0.05**	**9.95±0.27**		**0.68**
PWT[c]	8.7	Fl	5.33±0.33	2.28± 0.07	10.11± 0.34									
	8.7	CD	5.21±0.04	2.14± 0.06	9.21± 0.27									
H90G	8.7	Fl	6.39±0.07	1.95± 0.11	12.46± 0.71	**2583±98**	**1.30±0.04**	**$1.0x10^{-5}$ ±4**	**0.70±0.03**	**6.26**	**2.01±0.05**	**11.48± 5.45e+6**	**-1.76**	**0.65 (MM)**
	8.7	CD	6.51±0.07	1.95± 0.11	12.69± 0.72								**-2.24**	
I85L	8.7	Fl	5.94±0.07	1.95± 0.11	11.58± 0.66	**165±32**	**1.26±0.03**	**$2.0x10^{-5}$ ±5**	**0.64±0.02**	**5.40**	**1.90±0.04**	**9.44±3.06**	**-0.88**	**0.66 (MM)**
	8.7	CD	5.16±0.10	1.95± 0.11	10.06± 0.60	**9526±2621**	**1.72±0.04**	**$2.0x10^{-5}$ ±1**	**0.64±0.04**	**5.36**	**2.36**	**11.8**	**0.39**	**0.72 (SF)**
I93L	8.7	Fl	5.02±0.05	1.95± 0.11	9.79± 0.56	**203467± 102131**	**2.02±0.07**	**$3.0x10^{-5}$ ±7**	**0.72±0.02**	**5.16**	**2.74**	**13.4**	**0.92**	**0.74 (SF)**
	8.7	CD	5.09±0.04	1.95± 0.11	9.93± 0.56								**0.53**	
F6L	8.7	Fl	3.32±0.10	1.95± 0.11	6.47± 0.41								**4.23**	
	8.7	CD	3.09±0.17	1.95± 0.11	6.03± 0.47								**4.43**	

WT Wild Type Hisactophilin
WT[a] Marty T.J. Smith and E.M Meiering unpublished data
PWT[b] (Wong, 2004) PWT data with 4 extra amino acids
PWT[c] (Liu et al., 2002) data obtained at 20ºC
Equilibrium values (C_{mid} , m-value and ΔG_U) were obtained from fitting urea denaturation curve in (Figure 3.4) data to Equation 2.1, m2was fixed to 0.072
$\Delta\Delta G = (C_{mid\ WT}-C_{mid\ mutant})(m_{WT}+m_{mutant})/2$. A negative value indicates a more stable mutation
Text in bold typeface refers to the kinetic data
Observed unfolding and refolding rate constants were fit to equation 1.11, C_{mid} , m_1, ΔG_u and βT were calculated from the fitted kinetic parameters using equations,2.4, 1.13, 1.14 and 1.15 respectively. Error estimates were obtained from the fitting program (Origin)
MM, manual mixing, SF stopped flow data.

3.3.2 Refolding kinetics

Kinetic refolding traces of the WT and H90G, I85L, and I93L mutants were monitored by manual mixing fluorescence from 4 to 6 M urea, and by stopped flow fluorescence from 0.2 to 4.2 M urea both at pH 8.7. All the kinetic refolding traces were fit to a single exponential equation and an excellent fit to this model was only obtained at concentrations of denaturant greater than approximately 3.8 M, suggesting that the transition observed corresponds to a 2-state model. Previous studies of PWT and WT have indicated that the protein fold and unfold to 2-state model and data fit will to a single exponential equation (Wong, 2004).

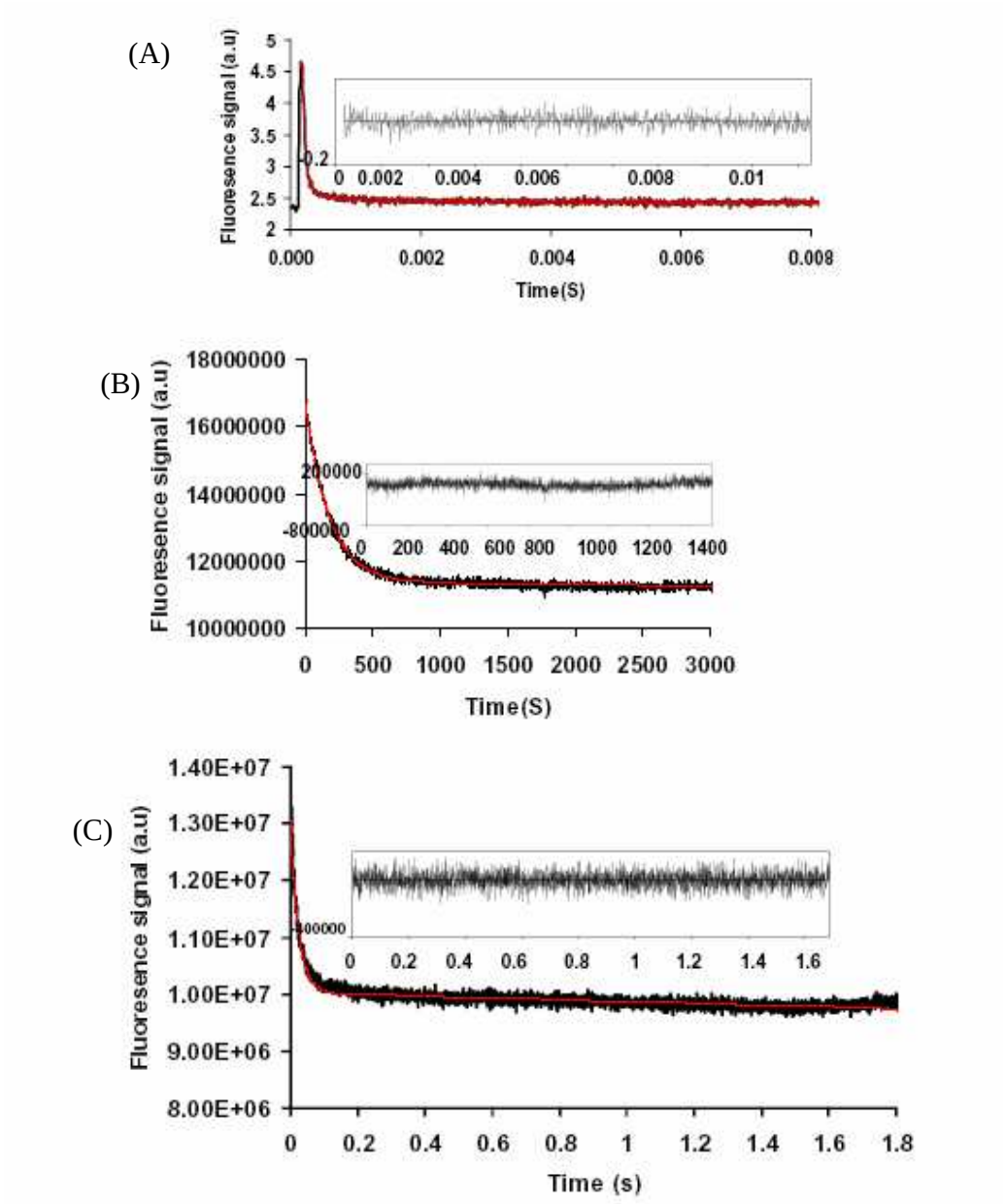


Figure 3. 10 Refolding kinetics of WT hisactophilin.
A. Stopped flow fluorescence-monitored refolding kinetics of WT hisactophilin at 1.2 M urea fit to a double exponential equation with a drift. Shown inset are the residuals of double exponential fit with k_f=19.7s^{-1}, which indicates the kinetics are well described by a double exponential. B.fluorescence-monitored refolding kinetics of WT hisactophilin at 4.8 M urea fit to a single exponential equation with a drift. Shown inset are the residuals of single exponential fit with k_f=11.4x10$^{-3}s^{-1}$. C. fluorescence- monitored refolding kinetics of WT hisactophilin at 4 M urea fit to a single exponential equation with a drift. Shown inset are the residuals of single exponential fit with k_f= 4.1x10$^{-3}s^{-1}$, which indicates the kinetics are well described by a single exponential.

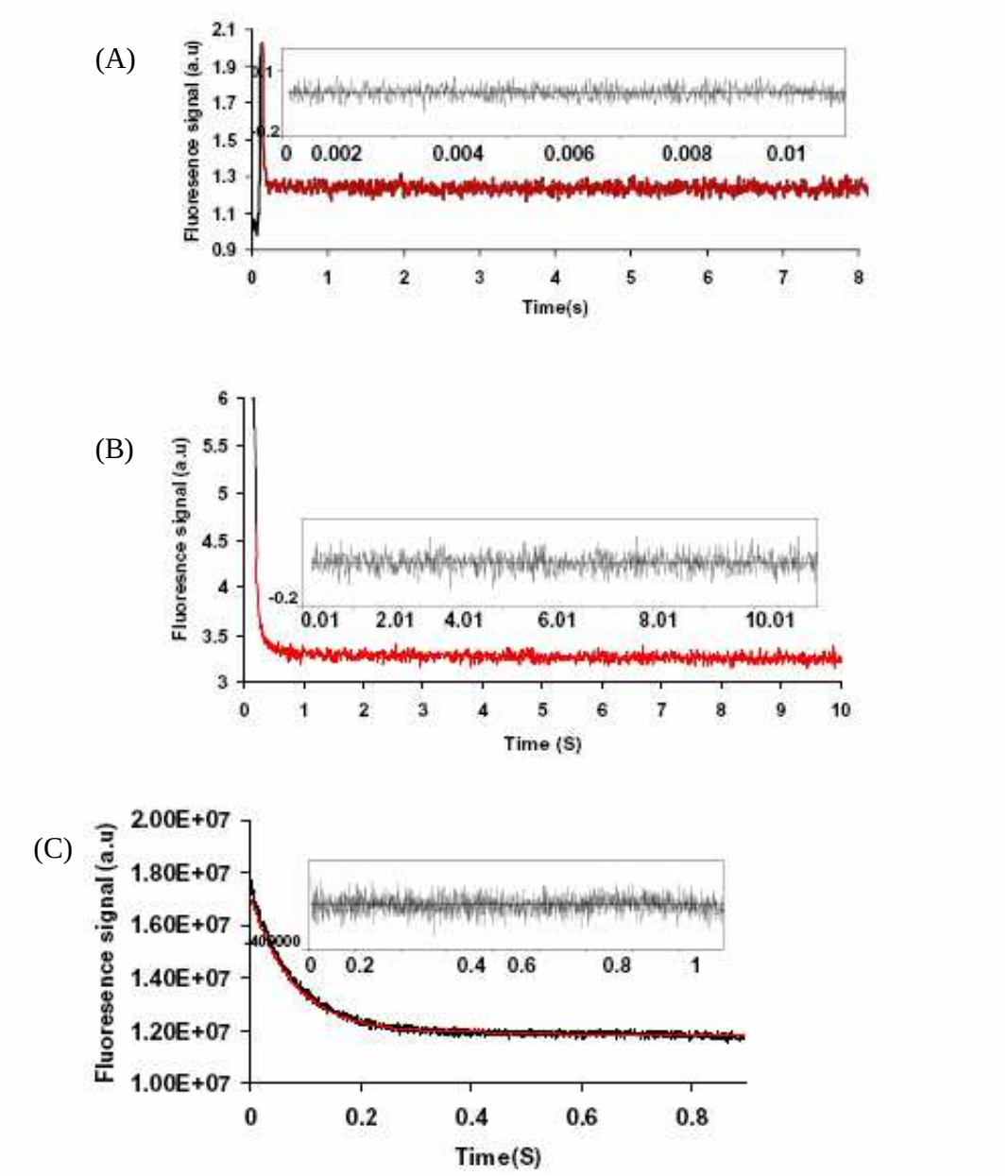


Figure 3. 11 Refolding kinetics of H90G. A. Stopped flow fluorescence-monitored refolding kinetics of H90G at 0.2 M urea fit to a double exponential equation with a drift. Shown inset are the residuals of double exponential fit with k_f= 82.5s^{-1}, which indicates the kinetics are well described by a double exponential. B. Stopped flow fluorescence-monitored refolding kinetics of H90G at 2.0 M urea fit to a double exponential equation with a drift. Shown inset are the residuals of double exponential fit with k_f=28.2s^{-1}. C. fluorescence-monitored refolding kinetics of H90G at 5.6 M urea fit to a single exponential equation with a drift. Shown inset are the residuals of single exponential fit with k_f= 13.4x10^{-3} s^{-1}, which indicates the kinetics are well described by a single exponential.

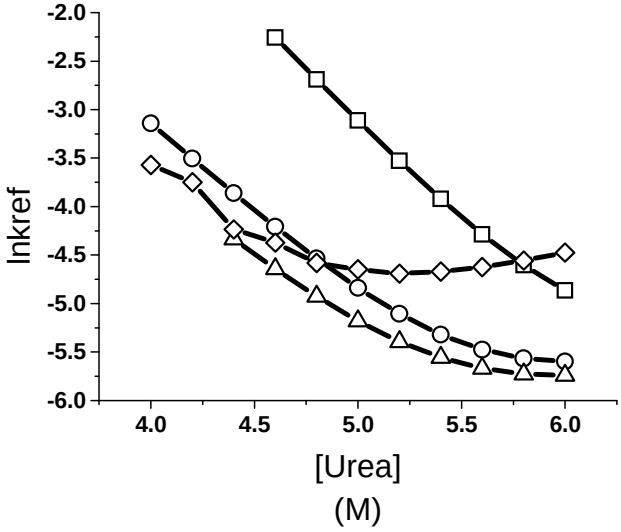


Figure 3. 12 Rate constants of refolding of the WT and mutant hisactophilin.
Dependence of the natural logarithm of the observed refolding rate constant k_{ref} on denaturation concentration. WT (○), H90G (□), I85L (△), and I93L (◇). The continuous line represents a fit with single exponential for the data. Values for fits are summarized in Table 3.5.

At denaturant concentrations below this critical level, there were systematic deviations for a single exponential fit, a double exponential model fit the data well. The fast phase of this double exponential represents the major amplitude and the slow phase has small amplitude.

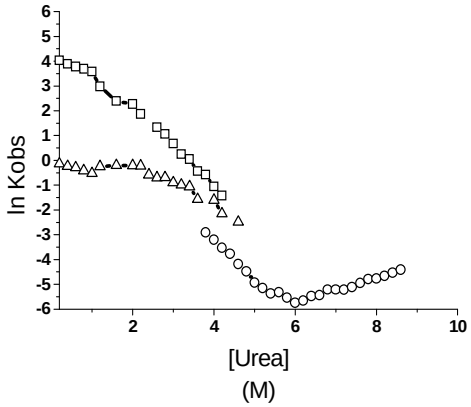


Figure 3. 13 Folding and unfolding manual mixing and stopped flow kinetics of WT hisactophilin at pH 8.7. Solution contained 50 mM Glycine/sodium glycinate pH 8.7, 1 mM EDTA, 1 mM DTT, and 0.2 mg ml^{-1} protein at 25°C. (○) manual mixing, (□) fast phase stopped flow, (△) slow phase stopped flow.

A plot of the natural logarithm of both the refolding and unfolding rate constants with urea concentration is shown in (Figure 3.13). Both rate constants vary with urea. The major amplitude phase has a considerably larger denaturant dependence. Minor amplitude phase that become apparent at lower denaturant varies less strongly with denaturant. The slow phase phase has been observed previously; it could indicate formation of an intermediate or perhaps arise from misfolding or aggregation. This was investigated further by measuring refolding as a function of protein concentration, as described below.

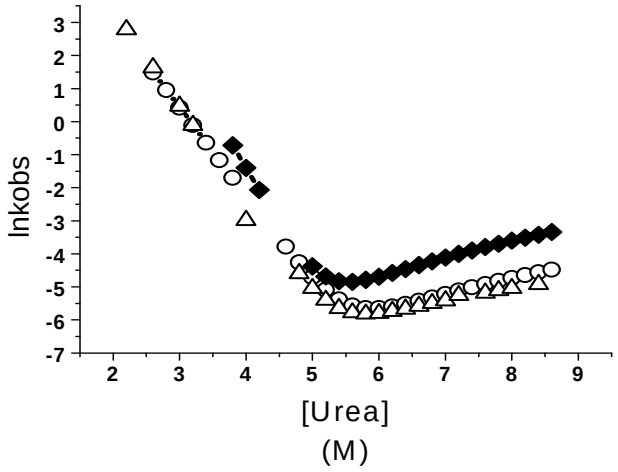


Figure 3. 14 Folding and unfolding manual mixing and stopped flow kinetics of WT and mutant hisactophilin at pH 8.7 .
Solution contained 50 mM Glycine/sodium glycinate pH 8.7, 1 mM EDTA, 1 mM DTT, and 0.2 mg ml^{-1} protein at 25°C. WT (○), I85L (△), and I93L (◆). The continuous line represents a fit with single exponential for the unfolding and double exponential for refolding data lower than 3.8M urea. Values for fits are summarized inTable 3.5.

In the refolding kinetics observed at pH 8.7, refolding rates deviate from the linear urea dependence expected for 2-state folding kinetics for the WT and all mutant at urea concentrations lower than 2 M, and there is discrepancy between the manual mixing and stopped flow data. This may indicate formation of a folding intermediate or refolding of aggregated protein. To distinguish between the processes of aggregation and intermediate formation in the refolding experiment, the protein concentration dependence of refolding was measured for H90G. Formation of folded intermediate should be independent of protein concentration, while aggregation would be dependent on protein concentration (Wong et al., 2004).

Refolding kinetics were measured for H90G with protein concentrations ranging from 0.03 to 0.2 mg/ml with urea concentrations at 0.2 M and 3 M. The expected result if the slow phase is due to aggregation there should be an increase in the rate and the amplitude of this phase

relative to the fast phase with increasing protein concentration. To also examine if aggregates were accumulating in the unfolded protein stock solution with time, the refolding experiments were performed by incubating the protein freshly acid denatured and protein incubated in acid for 24 hours before taking refolding measurements. In this case, the amplitude of the fast phase should decrease and the amplitude of the slow phase should increase. The increase in the amplitude of the slow phase also tended to be bigger for higher protein concentration. The results, however, show that as protein concentration increased, the fast phase remained the predominant phase. There was some evidence in 0.2 M urea for some aggregation of stock protein solution with time, and these aggregates were removed by filtering. There was no clear trend in the relative amplitudes of the fast and slow phases for 3 M urea because the rates were not properly defined.

The slow phase seemed to be the smaller proportion of the total amplitude (~5 %) for 0.2 M urea, which is close to negligible. Further experiments on H90G mutation at pH 6.7 were performed where the protein is more unstable (Liu et al., 2001) to investigate if the curvature in the refolding limb is due to intermediate formation or due to protein aggregation. The result indicates that the kink and discrepancy between the manual mixing and stopped flow results are less obvious at pH 6.7 refolding limb than of pH 8.7 as shown in (Figure 3.15).

At pH 6.7 both phases are observed but only the fast phase is plotted since the slow phase is too slow to be accurately determined by fluorescence. The dependence of the natural logarithm of the observed refolding and unfolding rate constants of WT and mutant hisactophilin were plotted and fit together.

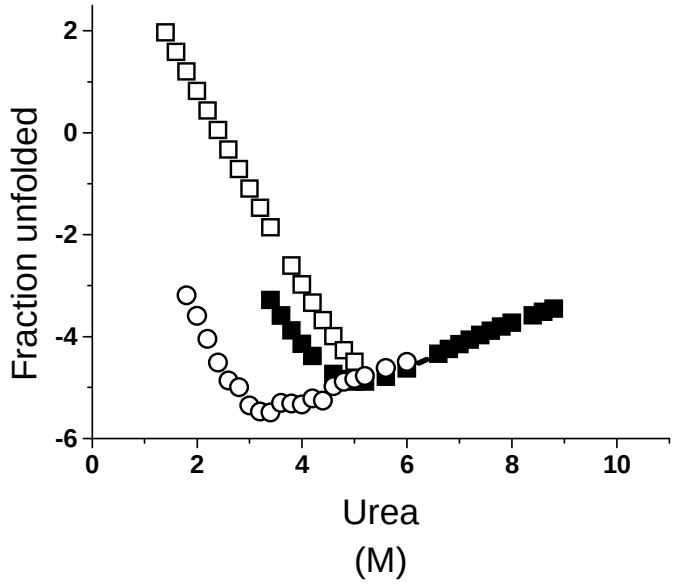


Figure 3.15 Folding and unfolding manual mixing and stopped flow kinetics of WT and H90G at pH 6.7. Solution contained 50 mM KH_2PO_4/K_2HPO_4 pH 6.7, 1 mM EDTA, 1 mM DTT, and 0.2 mg ml^{-1} protein at 25°C. WT(○) MM, H90G Stopped flow (□), H90G MM (■). The continuous line represents a fit with single exponential for the unfolding and double exponential for refolding data lower than 3.8 M urea. Values for fits are summarized in Table 3.4.

The discrepancy was observed for the rates measured by stopped-flow fluorescence below 2 M and linear above that concentration at pH 8.7 and 6.7. The reason of the curvature has been attributed to different causes. Firstly, the amplitude of the refolding transitions monitored by stopped-flow at low urea concentration (below 2 M) was very small compared to our expectations. The reaction was too fast to measure accurately as most of the reaction was completed in the dead time. This resulted in less accurate refolding rate constants, which could explain why these rates did not follow the same trends as the measured rates and thus were not taken into account for further analysis. Secondly, some mixing artifact of the stopped flow apparatus might be occurring due to the slow phase because the amplitude of this phase was very small so that the rate constants may not accurately determined or the dead time limitation of the

machine. Thirdly, this curvature may be due to transient protein aggregation during refolding (Bachmann and Kiefhaber, 2001) or may also be due to an inherent quadratic dependence of the free energy of transfer of amino acid constituents from water to denaturant (Wong et al., 2004). The fourth reason may be due to transition state movements (Hammond effect). The reason for the curvature is not completely clear at the present time and there is no clear interpretation for the available data.

The curvature in the refolding limb (at low urea concentrations) of the chevron plot has been observed for numerous apparent 2-state folders (Sanchez and Kiefhaber, 2003). Many proteins that show 2-state folding at moderate to high denaturant begin to populate an intermediate at very low denaturant concentrations (Sanchez and Kiefhaber, 2003). This is also likely to occur for hisactophilin (Wong, 2004). Accordingly, rates in the region of rollover were excluded, and the remaining rates fit to a 2-state model.

In this case the folding limb roll-over was ignored as it occurs at low urea concentrations, and the remaining data fit to a two- state model. The F6L mutant precipitated during concentration, and no kinetic data for this mutant was obtained.

3.3.3 Φ-value analysis

The basis of Φ-value analysis is to compare the free energy change for a mutation to the individual contributions of the folding and unfolding free energy change. From (Table 3.6) it is clear that the site of H90G and I93L in the transition state are closely resembles the structure in the native state. While the site of the I85L in the transition state is closely resembles the structure in the unfolded state.

Table 3. 6 **Φ-value analysis of hisactophilin mutants**

protein	pH	$\Delta\Delta G_{eq}$	$\Delta\Delta G_u^{\ddagger}$	Φ_F
H90G	8.7	-1.76	-1.40	0.80
I85L	8.7	-0.88	-0.32	0.18
I93L	8.7	0.92	-2.27	1.29

$\Delta\Delta G_{eq}$= -RT an kf/kf
$\Delta\Delta G_u^{\ddagger}$=-RTln ($k_{f'}/k_f$)
$\Phi_F = \Delta\Delta G_u^{\ddagger} / \Delta\Delta G_{eq}$

3.3.3 The chevron plot

The dependences of the natural logarithm of the observed unfolding and refolding rate constants of WT and mutant H90G, I85L, I93L with urea concentration are plotted in (Figure 3.16). The kinetic constants determined from the fitting of the chevron plot are summarized in Table 3.5. Equilibrium and kinetic values calculated from kinetic parameters (using Equation 3.1, values in Table 3.5) were compared to those obtained from equilibrium curve measurements determined in section 3.2. The ΔG_U and m-values calculated from the kinetic experiments of WT type and mutant hisactophilin were very similar to the values obtained from the equilibrium measurements.

The unfolding limb has relatively smaller urea dependence while the refolding limb has a stronger dependence on urea. The refolding and unfolding rate constants for each mutant at pH 8.7 are listed in (Table 3.5) and were determined using the entire unfolding data, in combination with the single exponential region of refolding, and some of the fast phase of the double

exponential region of refolding as shown in (Figure 3.16). From the of equilibrium curves monitored by different probes (fluorescence and CD), and agreement of equilibrium and kinetic parameters fit to 2-state model. There was evidence from previous results and results here for 2-state.

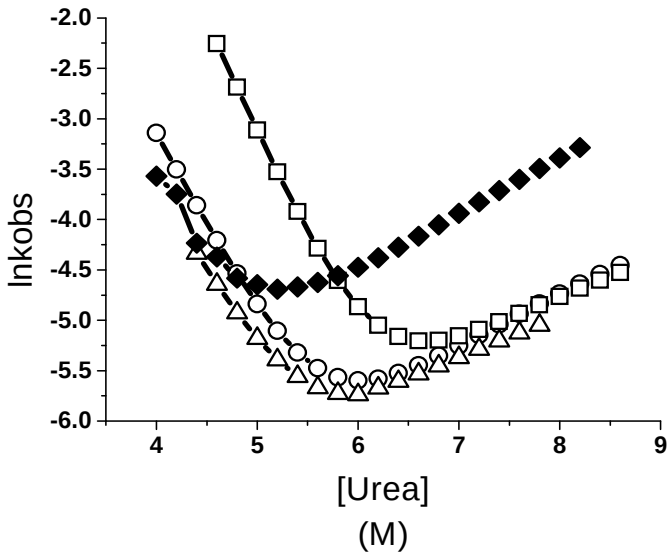


Figure 3. 16 Folding and unfolding manual mixing kinetics of WT and mutant hisactophilin.
Solution contained 50 mM Glycine/sodium glycinate pH 8.7, 1 mM EDTA, 1 mM DTT, and 0.2 mg ml^{-1} protein at 25°C. WT (○), H90G (□), I85L (△), and I93L (◆). The continuous line represents a fit with single exponential for the unfolding and refolding data. Values for fits are summarized in Table 3.5.

3.4 DSC preliminary results

Multiple attempts to obtain DSC data were performed for all the mutant proteins. Unfortunately, this did not yield analyzable, reversible thermal unfolding data because the mutant proteins tended to aggregate during the DSC experiment. The measurements were first performed at pH 8.7, the same pH used for the equilibrium and kinetic experiments using 0.5 mg ml^{-1} protein concentrations. However, the protein precipitated. Next experiment was performed at pH 9.7. These scans showed an additional peak at ~40°C for WT and all the mutant hisactophilin. This

extra peak has not been observed previously in DSC of hisactophilin. Further analysis is required to explain its presence in the thermogram.

Chapter 4: Discussion

From the denaturation curve results of WT and F6L, F13Y, I85L, H90G, and I93L hisactophilin mutants, it is evident that a decrease in pH causes a decrease in stability. This parallels to previous results obtained for the PWT and WT hisactophilin protein (Liu et al., 2001). There is only a minor change in stability from pH 8.7 to 7.7 and a much more substantial decrease when the pH is decreased to 6.7.

The WT and F6L, F13Y, I85L, H90G, and I93L mutant equilibrium and kinetic folding were investigated by measuring their structural transitions by fluorescence spectroscopy and CD. From the equilibrium and kinetic data, it can be seen that WT and all mutant hisactophilin can be fit to 2-state mechanism. In the kinetics result there was a rollover, which is evidence for formation of an intermediate under conditions of high protein stability. This has also been observed previously for PWT hisactophilin when the protein stability is increased in D_2O or due to addition of small molecule stabilizing agents (Wong, 2004). The stability and kinetics are considered for each mutant below.

4.1 Hisactophilin mutants

4.1.1 H90G

Position 90 corresponds to position 2 of a type I' β-turn see (Figure 1.19)**.** The kinetic data of the H90G mutation indicates that Gly residue at the i+2 position stabilizes the protein likely by eliminating the steric interactions between side chain C^{β} and main chain atoms in the native structure (Figure 1.18). These results are similar to those obtained for FGF (H93G type I β-

turn) (Kim et al., 2003) . The kinetic parameters of the H90G mutation indicates that the basis for the stability of the ΔΔG -1.76 kcalmol^{-1} increased in stability is due to an increase in the rate of refolding by ~35-fold while it has no effect on the unfolding rate compared with the WT (Table 3.5). This result supports the hypothesis that increased symmetry can give rise to increased stability symmetry hypothesis.

4.1.2 I85L

I85L is a core mutation located in strand 9 bottom layer of β-barrel. As indicated in (Figure 1.19) sequence alignment analysis indicated that I85 is a conserved residue see (Figure 1.11). The kinetic parameters of the I85L mutation shows that the basis of the -0.88 kcal mol^{-1} increase in stability is due to ~ 2-fold decrease and increase, respectively in the unfolding and refolding rates. This increase in stability suggests that the I85L mutation buries additional area within the core region or is able to pack better. This result also supports the symmetry hypothesis.

4.1.3 I93L

I93L is a core mutation located in strand 10 of the upper layer of β-barrel (Figure 1.19). This mutation exhibited the most significant effect on the unfolding rate resulting in a 3-fold increase in unfolding rate, (Table 3.5, Figure 3.16). The introduced leucine side chain may partially fill a centrally located cavity within the core, while an additional adjacent microcavity may be formed due to the change in the stereochemistry of the side chain. Isoleucine to leucine substitution could result in substantial formation of cavity within the core region. I and L have the same volume, but different stereochemistry, with I branched at the C^{β}, while L is branched at C^{γ}. This result is similar to that of FGF V109L mutation (Brych et al., 2004). In FGF no effect on the refolding rate was seen while there was an increase the unfolding rate by 4-fold. The Leucine

side chain was found to partially fill a centrally located cavity within the core. Furthermore, an additional adjacent microcavity was formed due to the deletion of the WT Val methyl group. Thus, overall the packing was not sufficiently improved to stabilize the protein. This suggests that for hisactophilin the destabilizing effect is smaller because the methyl group is not removed, and the L makes better packing interactions than the I.

4.1.4 F6L

F6L is a core mutation located in strand 1of the bottom layer of β-barrel as shown in (Figure 1.19). F6L is less stable than the WT as indicated by urea denaturation. The data for this mutant were generally of poor quality, complicated by protein precipitation and perhaps modification due to decreased stability. Therefore kinetic analysis was not performed for this mutation. F6 is a conserved residue, it appears that the larger F may be needed to maintain good packing interactions among residues in this layer of the hisactophilin structure. This may explain why substitution to the smaller L residue is destabilizing. There may also be some functional interactions that are modulating the results. Hisactophilin is naturally myristoylated and the myristoyl group of hisactophilin can pack into the central hydrophobic core of the protein (Meissner, 2007), or undergo a switch such that it extends from the protein and into the membrane lipid bilayer (Hanakam et al., 1996). Some of the residues in this layer may be involved in switching. Hisactophilin is more symmetric in sequence that fascins, suggesting that it is more ancestral as shown in (Figure 1.7).

4.1.5 F13Y

F13Y is a mini-core mutation located in strand 2 of the β-hairpin see (Figure 1.19). The thermodynamic and kinetic effects for this mutation were difficult to measure due to the increases

fluorescence intensity for the native protein resulting from an extra tyrosine in structure. The CD results at pH 7.7 and 8.7 indicated that F13Y is slightly less stable than the WT. Based on the previous studies for FGF results, F13Y was expected to be more stable than the WT compared F22Y in FGF (Dubey et al., 2005). For this mutation, further investigation is required using methods such as DSC, NMR to asses the stability.

4.2 Evidence for non 2-state behavior

The unfolding and refolding kinetics of the WT and mutant hisactophilin at pH 8.7 fit to 2-state transitionat medium to high urea concentration but this is not the case at very low denaturant concentration. The refolding kinetic traces fit with double exponential equation and a curvature observed at low denaturant concentration for the WT and mutant hisactophilin. This result was observed in our lab before at the same pH where the protein is very stable ((Wong, 2004), (Wong et al., 2004) more details in section 3.3.2. This may indicate formation of a folding intermediate or refolding aggregated protein. It will discussed more in the future work.

4.3 Relationship of folding to function

From FGF results the cavity-filling mutations are well known to stabilize protein structures (Eriksson et al., 1992), however, this may lead to a loss in the protein function as shown in FGF. A series of mutations in FGF N-and C-terminal (which hydrogen bonding to each other) β-strands were shown to stabilize the structure due to filling the local cavities present within this region. These mutations introduce a three-fold symmetric constraint upon the primary structure, (K12C, K12V, P134C, P134V, K12V/P134V, L46V/P134V, K12V/L46V, E87V/N95V/P134V). The results support the hypothesis that a symmetric primary structure within a symmetric superfold is a solution to, and not constraint upon, the protein folding

problem. In addition, the results also support the hypothesis that there exists a function-stability trade off. The multiple mutant (K12V, L46V, E87V, N95V, P134V) is very stable; however, it is functionally inactive (Dubey et al., 2007). From this result it would be interesting to investigate further the relationship between the stability and function of β-trefoil proteins and hisactophilin.

Folding of β-trefoil proteins was found to be slow, due to the presence of a folding intermediate. Hisactophilin is the smallest protein compared with other β-trefoil proteins e.g. FGF and IL-1β also folds quite slow. Hisactophilin is the smallest, has the lowest contact order and folds fastest. The longer, functional loops in the other proteins probably slow their folding. Further investigations were done and multiple routes were proposed to clarify the slow folding. Results of simulation of folding pathway show that hisactophilin folds by a direct route in which the formation of local contacts within turns gradually leads to the formation of secondary structure (Gosavi et al., 2008). In contrast IL-1β folds by two subsets of contacts that compete with each other, resulting in the unfolding of one subset of native contacts and their subsequent refolding later along the folding coordinate. This route is called backtracking. FGF folding initiates within the β-trefoil bringing its N and C termini together despite the high entropic cost (Roy et al., 2005).

IL-1 β folds on the time scale of minutes and the folding is accompanied by a broad and high free energy barrier. The kinetic experiments indicated that there is the presence of structural intermediate during folding. Hisactophilin folds much faster, but still relatively slowly compared with other single domain proteins (Liu et al., 2002) with lack of an intermediate in all but the most stabilizing conditions.

4.4 Possible functional features of the proteins that may affect the folding pathway

IL-1β is the only member of β-trefoil that folds by the backtracking route. This route occurs as a result of the topological frustration the functional β-bulge causes during folding (Capraro et al., 2008). The slow folding of IL-1β structure is related to the formation of β-bulge, which is directly involved in receptor binding. This result for IL-1β is consistent with the hypothesis that slow folding is often related to the functional obligation to conserving protein-protein or protein-ligand interaction sites (Gosavi et al., 2006). Several studies have indicated that the protein structure, function, and folding are inextricably related and that the structural differences between structurally similar proteins correlate not only with functional regions but also with topological folding traps (Hammond et al., 1998).

In conclusion, in order for the protein to function correctly, some residues need to be conserved and can not be optimized for folding. Furthermore, the protein must be stable enough to fold and in order to function well.

Chapter 5: Future work

The next step could be to investigate the slow phase in the kinetics. There was evidence from the previous studies in our lab that an intermediate starts to become populated under very stabilizing conditions at pH 8.7, also when we stabilized the protein in D_2O or upon addition of sodium sulphate (Wong, 2004) . From the results presented here the rollover more prominent in H90G decreases when dropping the pH from 8.7 to 6.7; this was also observed previously. It would also be useful to characterize the refolding kinetics with different protein concentrations to clarify whether the results indicate formation of folding intermediate or transient protein aggregation. Additionally, light scattering experiments may allow for further investigation of protein aggregation.

Additional DSC experiments of the WT and mutant hisactophilin should be conducted to better characterize the stability of these proteins. The preliminary results collected here indicated the presence of the second peak ~ 40°C of the scan. It would be interesting to investigate further if that peak is due to protein aggregation or due to formation of an intermediate or degradation product. EDTA may bind the protein and then associate at low temperature. DSC analysis in absence of EDTA may be useful. Which shown as a small peak. Furthermore it would be useful to perform DSC experiments at varying protein concentration to confirm whether the small peak arises from protein aggregation.

A detailed study on the designing SYM2, SYM3, SYM4, SYM5, and SYM6 mutations, which are double, triple, quadruple, quintuple, and hextuple mutations that make the protein symmetric to assess the folding and stability of the symmetric structure. This could provide a means to determine whether an increase in trefoil symmetry would result in increased protein

stability. Based on the previous studies for FGF, it is also necessary to carry out further experiments to study the effect the SYM6 on hisactophilin function, to compare with FGF. Another method to make additional hisactophilin turns. For such studies, more detailed information on the conformations of the turns would be desirable, since not all of the hisactophilin turns are well defined in the NMR structure of PWT. Thus, additional NMR studies should also be conducted to elucidate additional structural details of the turns.

References

Adamek, D. H., Guerrero, L., Blaber, M. and Caspar, D. L. (2005) Structural and energetic consequences of mutations in a solvated hydrophobic cavity. *J Mol Biol.* 346, 307-18. Epub 2004 Dec 24.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool. *J Mol Biol* 215, 403-10.

Andreeva, A., Howorth, D., Brenner, S. E., Hubbard, T. J., Chothia, C. and Murzin, A. G. (2004) SCOP database in 2004: refinements integrate structure and sequence family data. *Nucleic Acids Res* 32, D226-9.

Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* 181, 223-30.

Bachmann, A. and Kiefhaber, T. (2001) Apparent two-state tendamistat folding is a sequential process along a defined route. *J Mol Biol* 306, 375-86.

Bateman, A., Coin, L., Durbin, R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E. L., Studholme, D. J., Yeats, C. and Eddy, S. R. (2004) The Pfam protein families database. *Nucleic Acids Res* 32, D138-41.

Berndt, K. d. (1996) Ultraviolet Circular Dichroism Spectroscopy and application to Biological molecules.

Blandl, T., Cochran, A. G. and Skelton, N. J. (2003) Turn stability in beta-hairpin peptides: Investigation of peptides containing 3:5 type I G1 bulge turns. *Protein Sci* 12, 237-47.

Brych, S. R., Blaber, S. I., Logan, T. M. and Blaber, M. (2001) Structure and stability effects of mutations designed to increase the primary sequence symmetry within the core region of a beta-trefoil. *Protein Sci.* 10, 2587-99.

Brych, S. R., Dubey, V. K., Bienkiewicz, E., Lee, J., Logan, T. M. and Blaber, M. (2004) Symmetric primary and tertiary structure mutations within a symmetric superfold: a solution, not a constraint, to achieve a foldable polypeptide. *J Mol Biol.* 344, 769-80.

Brych, S. R., Kim, J., Logan, T. M. and Blaber, M. (2003) Accommodation of a highly symmetric core within a symmetric protein superfold. *Protein Sci.* 12, 2704-18.

Capraro, D. T., Roy, M., Onuchic, J. N. and Jennings, P. A. (2008) Backtracking on the folding landscape of the beta-trefoil protein interleukin-1beta? *Proc Natl Acad Sci U S A* 105, 14844-8.

Cordes, M. H., Walsh, N. P., McKnight, C. J. and Sauer, R. T. (1999) Evolution of a protein fold in vitro. *Science* 284, 325-8.

Deechongkit, S., Nguyen, H., Jager, M., Powers, E. T., Gruebele, M. and Kelly, J. W. (2006) beta-Sheet folding mechanisms from perturbation energetics. *Curr Opin Struct Biol.* 16, 94-101. Epub 2006 Jan 25.

Dubey, V. K., Lee, J. and Blaber, M. (2005) Redesigning symmetry-related "mini-core" regions of FGF-1 to increase primary structure symmetry: thermodynamic and functional consequences of structural symmetry. *Protein Sci.* 14, 2315-23. Epub 2005 Aug 4.

Dubey, V. K., Lee, J., Somasundaram, T., Blaber, S. and Blaber, M. (2007) Spackling the crack: stabilizing human fibroblast growth factor-1 by targeting the N and C terminus beta-strand interactions. *J Mol Biol* 371, 256-68.

Eriksson, A. E., Baase, W. A., Wozniak, J. A. and Matthews, B. W. (1992) A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene. *Nature* 355, 371-3.
Fersht, A. (1999) Enzyme Structure and Mechanism. Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding
Fersht.A. (1999) Enzyme kinetics. Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding
New York, W.H.Freeman.
Fresht, A. (1997) Structure and mechanism in protein science : a guide to enzyme catalysis and protein folding., (ed. G. L. Hadler), Freeman, New York.
Galani, D., Fersht, A. R. and Perrett, S. (2002) Folding of the yeast prion protein Ure2: kinetic evidence for folding and unfolding intermediates. *J Mol Biol.* 315, 213-27.
Gosavi, S., Chavez, L. L., Jennings, P. A. and Onuchic, J. N. (2006) Topological frustration and the folding of interleukin-1 beta. *J Mol Biol* 357, 986-96.
Gosavi, S., Whitford, P. C., Jennings, P. A. and Onuchic, J. N. (2008) Extracting function from a beta-trefoil folding motif. *Proc Natl Acad Sci U S A* 105, 10384-9.
Habazettl, J., Gondol, D., Wiltscheck, R., Otlewski, J., Schleicher, M. and Holak, T. A. (1992) Structure of hisactophilin is similar to interleukin-1 beta and fibroblast growth factor. *Nature.* 359, 855-8.
Hammond, M. S., Houliston, R. S. and Meiering, E. M. (1998) Two-dimensional 1H and 15N NMR titration studies of hisactophilin. *Biochem Cell Biol* 76, 294-301.
Hanakam, F., Gerisch, G., Lotz, S., Alt, T. and Seelig, A. (1996) Binding of hisactophilin I and II to lipid membranes is controlled by a pH-dependent myristoyl-histidine switch. *Biochemistry.* 35, 11036-44.
Haspel, N., Zanuy, D., Aleman, C., Wolfson, H. and Nussinov, R. (2006) De novo tubular nanostructure design based on self-assembly of beta-helical protein motifs. *Structure* 14, 1137-48.
Hutchinson, E. G. and Thornton, J. M. (1994) A revised set of potentials for beta-turn formation in proteins. *Protein Sci* 3, 2207-16.
Jackson, S. E. (1998) How do small single-domain proteins fold? *Fold Des* 3, R81-91.
Jaenicke, R. R., R. (1989) Folding Proteins, pp. 191-222, Oxford ,England.
Johnson, C. M. and Fersht, A. R. (1995) Protein stability as a function of denaturant concentration: the thermal stability of barnase in the presence of urea. *Biochemistry* 34, 6795-804.
Johnson, W. C., Jr. (1990) Protein secondary structure and circular dichroism: a practical guide. *Proteins* 7, 205-14.
Kim, J., Brych, S. R., Lee, J., Logan, T. M. and Blaber, M. (2003) Identification of a key structural element for protein folding within beta-hairpin turns. *J Mol Biol.* 328, 951-61.
Kim, J., Lee, J., Brych, S. R., Logan, T. M. and Blaber, M. (2005) Sequence swapping does not result in conformation swapping for the beta4/beta5 and beta8/beta9 beta-hairpin turns in human acidic fibroblast growth factor. *Protein Sci.* 14, 351-9. Epub 2005 Jan 4.
Koradi, R., Billeter, M. and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph.* 14, 51-5, 29-32.
Kuwajima. (1999) Old and New View of Protein Folding. In *International Congress Series,1994*, Tokyo,Japan.

Lee, J., Dubey, V. K., Somasundaram, T. and Blaber, M. (2006) Conversion of type I 4:6 to 3:5 beta-turn types in human acidic fibroblast growth factor: effects upon structure, stability, folding, and mitogenic function. *Proteins.* 62, 686-97.

Li W, a. L. Y. (2005) Prediction of beta-turn types using SVM and evolutionary information 453- 456.

Liu, C., Chu, D., Wideman, R. D., Houliston, R. S., Wong, H. J. and Meiering, E. M. (2001) Thermodynamics of denaturation of hisactophilin, a beta-trefoil protein. *Biochemistry* 40, 3817-27.

Liu, C., Gaspar, J. A., Wong, H. J. and Meiering, E. M. (2002) Conserved and nonconserved features of the folding pathway of hisactophilin, a beta-trefoil protein. *Protein Sci* 11, 669-79.

Manning, M. C. (1989) Underlying assumptions in the estimation of secondary structure content in proteins by circular dichroism spectroscopy--a critical review. *J Pharm Biomed Anal.* 7, 1103-19.

Matouschek, A. and Fersht, A. R. (1993) Application of physical organic chemistry to engineered mutants of proteins: Hammond postulate behavior in the transition state of protein folding. *Proc Natl Acad Sci U S A* 90, 7814-8.

Matouschek, A., Matthews, J. M., Johnson, C. M. and Fersht, A. R. (1994) Extrapolation to water of kinetic and equilibrium data for the unfolding of barnase in urea solutions. *Protein Eng* 7, 1089-95.

Meissner, J. (2007) Effects of myristoylation on the structure and stability hisactophilin Waterloo Waterloo.

Miller, E. M. and Nickoloff, J. A. (1995) Escherichia coli electrotransformation. *Methods Mol Biol.* 47, 105-13.

Murzin, A. G., Lesk, A. M. and Chothia, C. (1992) beta-Trefoil fold. Patterns of structure and sequence in the Kunitz inhibitors interleukins-1 beta and 1 alpha and fibroblast growth factors. *J Mol Biol* 223, 531-43.

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, 2138-48.

Nagi, A. D. and Regan, L. (1997) An inverse correlation between loop length and stability in a four-helix-bundle protein. *Fold Des* 2, 67-75.

Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* 131, 266-80.

Panchenko, A. R. and Madej, T. (2005) Structural similarity of loops in protein families: toward the understanding of protein evolution. *BMC Evol Biol* 5, 10.

Parrini, C., Taddei, N., Ramazzotti, M., Degl'Innocenti, D., Ramponi, G., Dobson, C. M. and Chiti, F. (2005) Glycine residues appear to be evolutionarily conserved for their ability to inhibit aggregation. *Structure (Camb).* 13, 1143-51.

Ponting, C. P. and Russell, R. B. (2000) Identification of distant homologues of fibroblast growth factors suggests a common ancestor for all beta-trefoil proteins. *J Mol Biol* 302, 1041-7.

Priestle, J. P., Schar, H. P. and Grutter, M. G. (1989) Crystallographic refinement of interleukin 1 beta at 2.0 A resolution. *Proc Natl Acad Sci U S A.* 86, 9667-71.

Roy, M., Chavez, L. L., Finke, J. M., Heidary, D. K., Onuchic, J. N. and Jennings, P. A. (2005) The native energy landscape for interleukin-1beta. Modulation of the population ensemble through native-state topology. *J Mol Biol* 348, 335-47.

Sanchez, I. E. and Kiefhaber, T. (2003) Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding. *J Mol Biol* 325, 367-76.

Scheel, J., Ziegelbauer, K., Kupke, T., Humbel, B. M., Noegel, A. A., Gerisch, G. and Schleicher, M. (1989) Hisactophilin, a histidine-rich actin-binding protein from Dictyostelium discoideum. *J Biol Chem.* 264, 2832-9.

Schmid, F. X. (1989) Pectral Methods of Characterizing protein Conformation and Conformational Changes . In *A practical Approach*, pp. 251-84, Oxford,England.

Simon, M. N., Mutzel, R., Mutzel, H. and Veron, M. (1988) Vectors for expression of truncated coding sequences in Escherichia coli. *Plasmid.* 19, 94-102.

Szabo, A. G. (2000) Fluorescence principles and measurment . pp. 33-66, Oxford Univeristy ,New York.

Thornton M, J. (2005) Protein Families and Their Evolution -A Structural Perspective. *Annual Revesion in Biochemistry* 74, 867-900.

Watson, J. D., Laskowski, R. A. and Thornton, J. M. (2005) Predicting protein function from sequence and structural data. *Curr Opin Struct Biol.* 15, 275-84.

Weaver, J. C. (1995) Electroporation theory. Concepts and mechanisms. *Methods Mol Biol.* 47, 1-26.

Wilmot, C. M. and Thornton, J. M. (1988) Analysis and prediction of the different types of beta-turn in proteins. *J Mol Biol.* 203, 221-32.

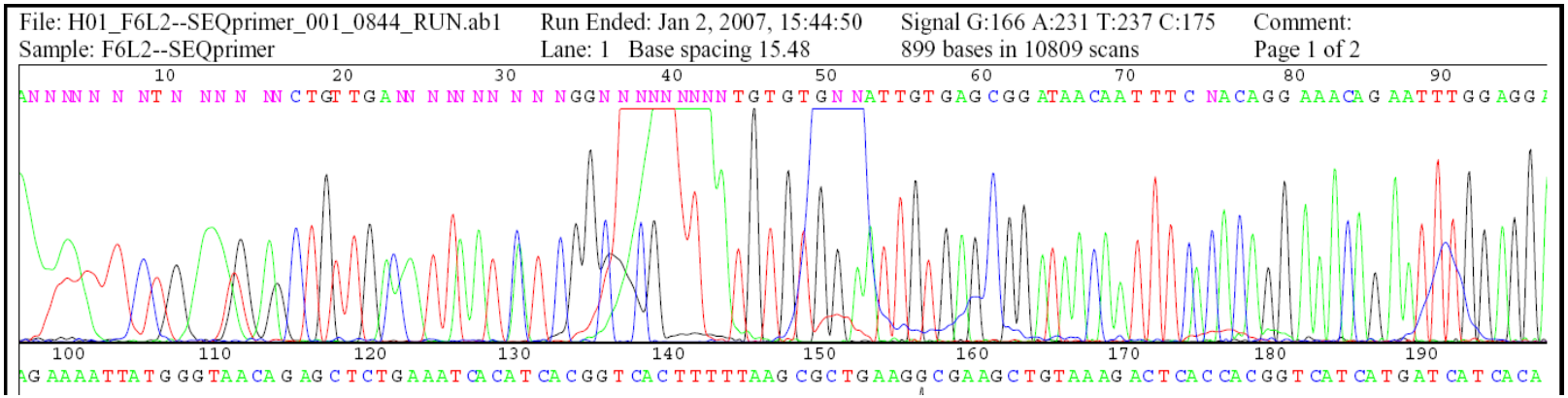
Wong, H. J. (2004) The Role of Temperture, Osmolytes and Myristoylation on the stability and folding of Hisactophilin. In *Chemistry*, Waterloo, Waterloo.

Wong, H. J., Stathopulos, P. B., Bonner, J. M., Sawyer, M. and Meiering, E. M. (2004) Non-linear effects of temperature and urea on the thermodynamics and kinetics of folding and unfolding of hisactophilin. *J Mol Biol* 344, 1089-107.

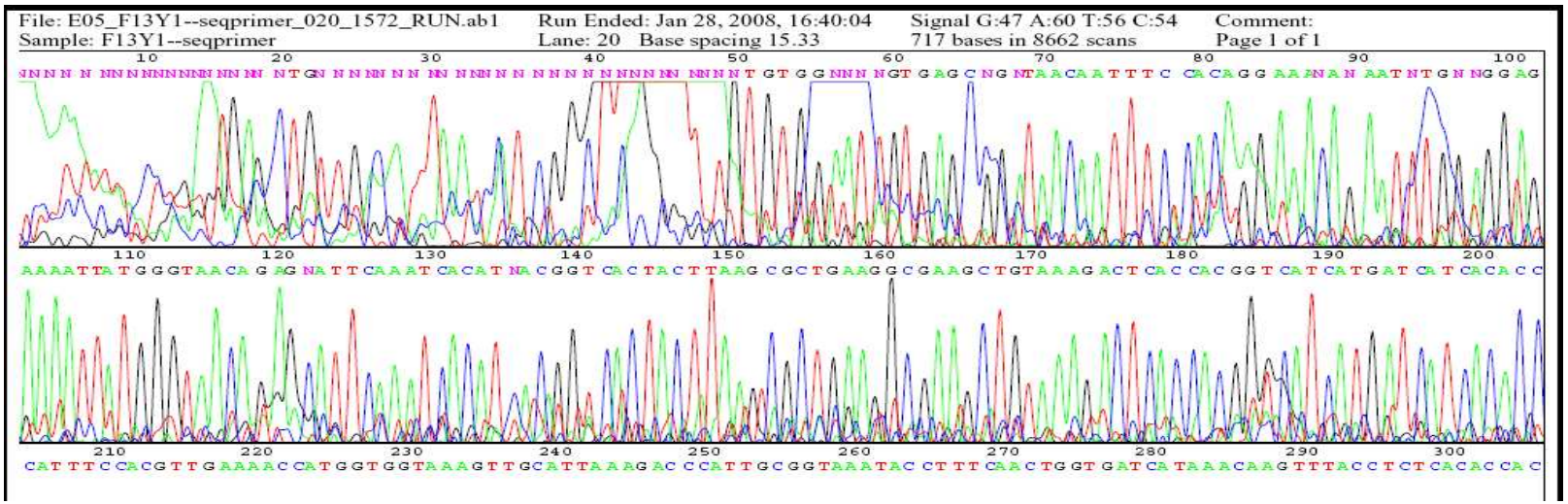
Woody, R. W. (1996) Theory of Circular Dichroism of Proteins . pp. 25-68, New York.

Zhu, L., Zhang, X. J., Wang, L. Y., Zhou, J. M. and Perrett, S. (2003) Relationship between stability of folding intermediates and amyloid formation for the yeast prion Ure2p: a quantitative analysis of the effects of pH and buffer system. *J Mol Biol.* 328, 235-54.

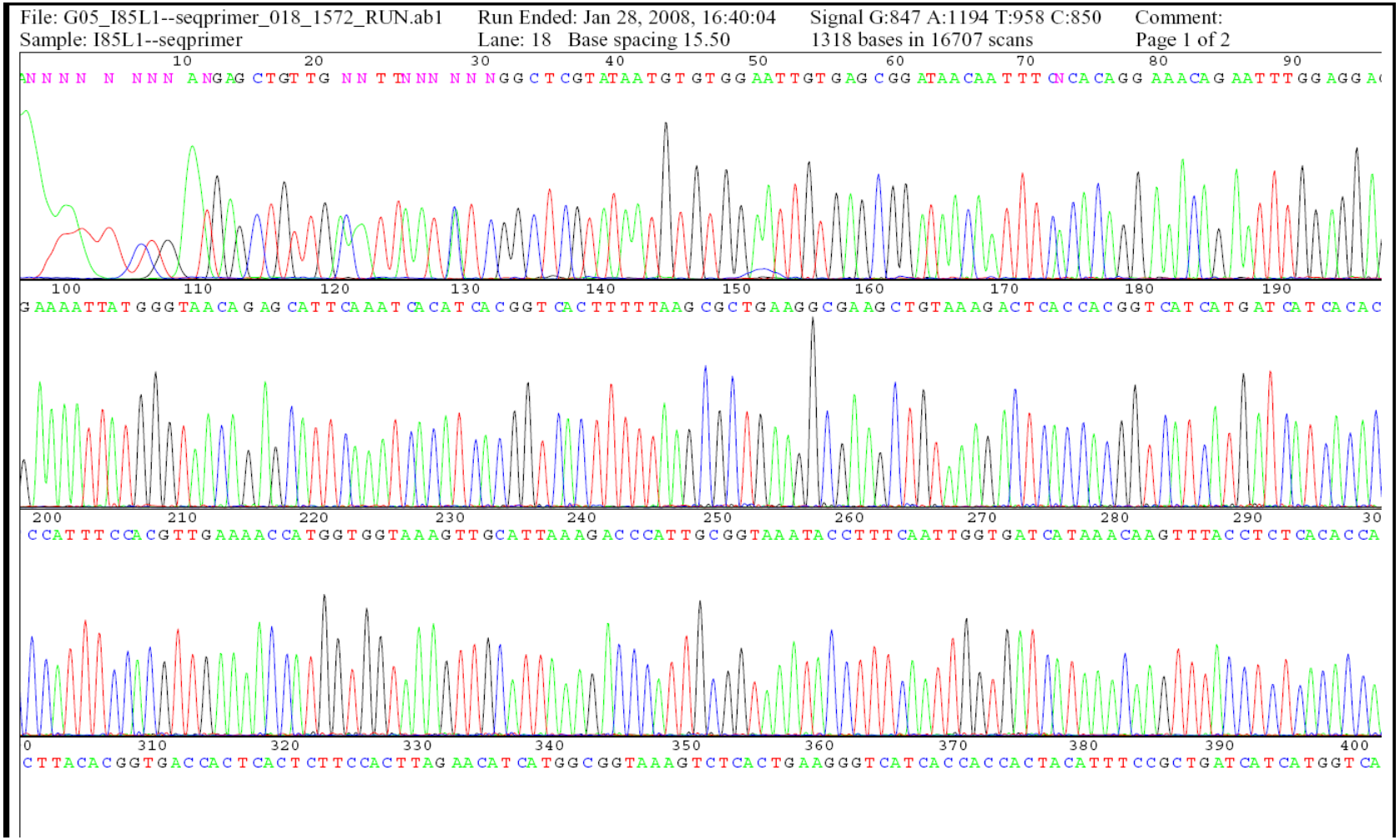
Appendix 1 : DNA sequencing for construct plasmid F6L



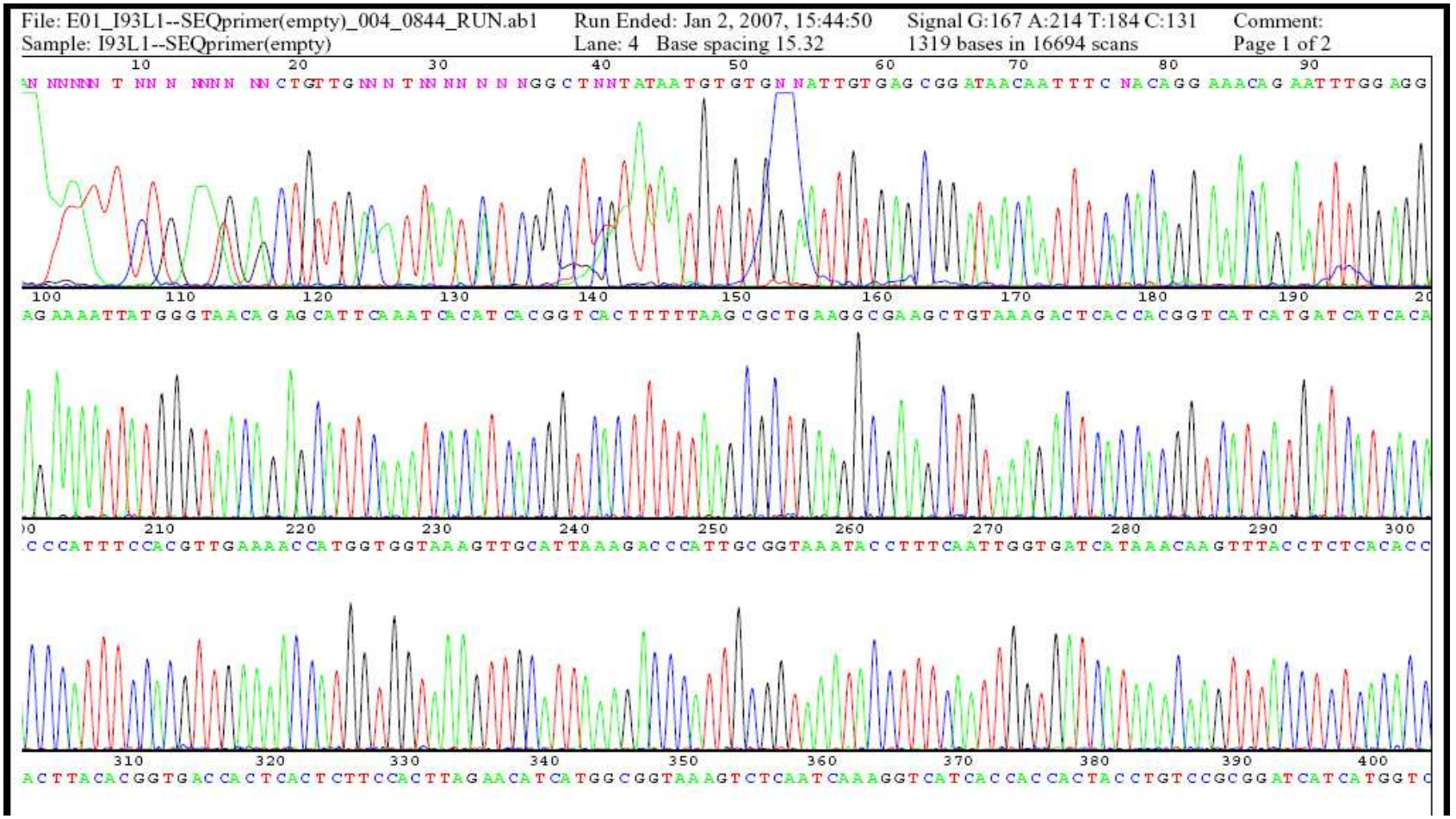
Appendix 2 : DNA sequencing for construct plasmid F13Y



Appendix 3 : DNA sequencing for construct plasmid I85L

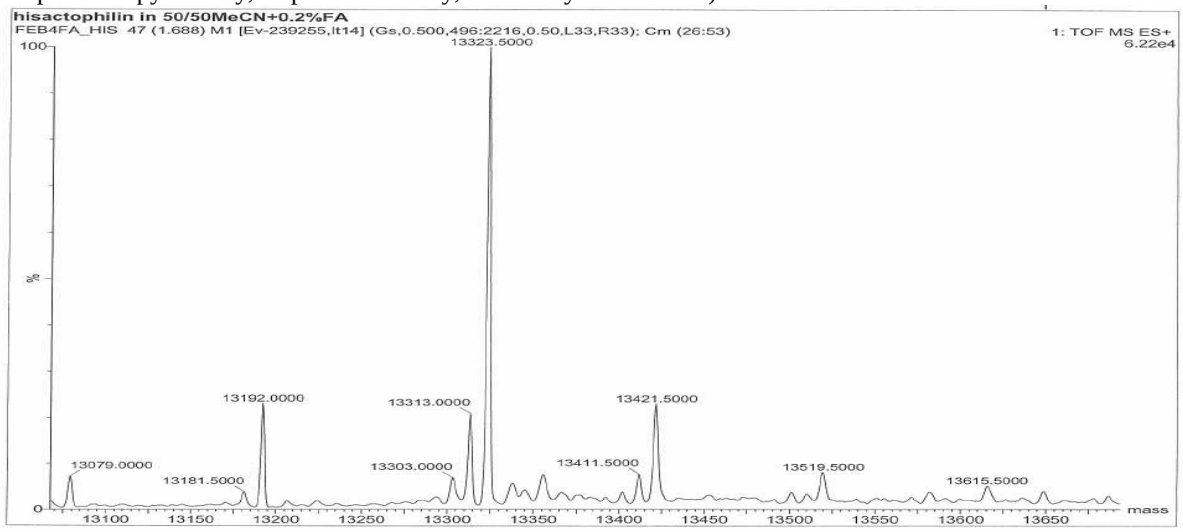


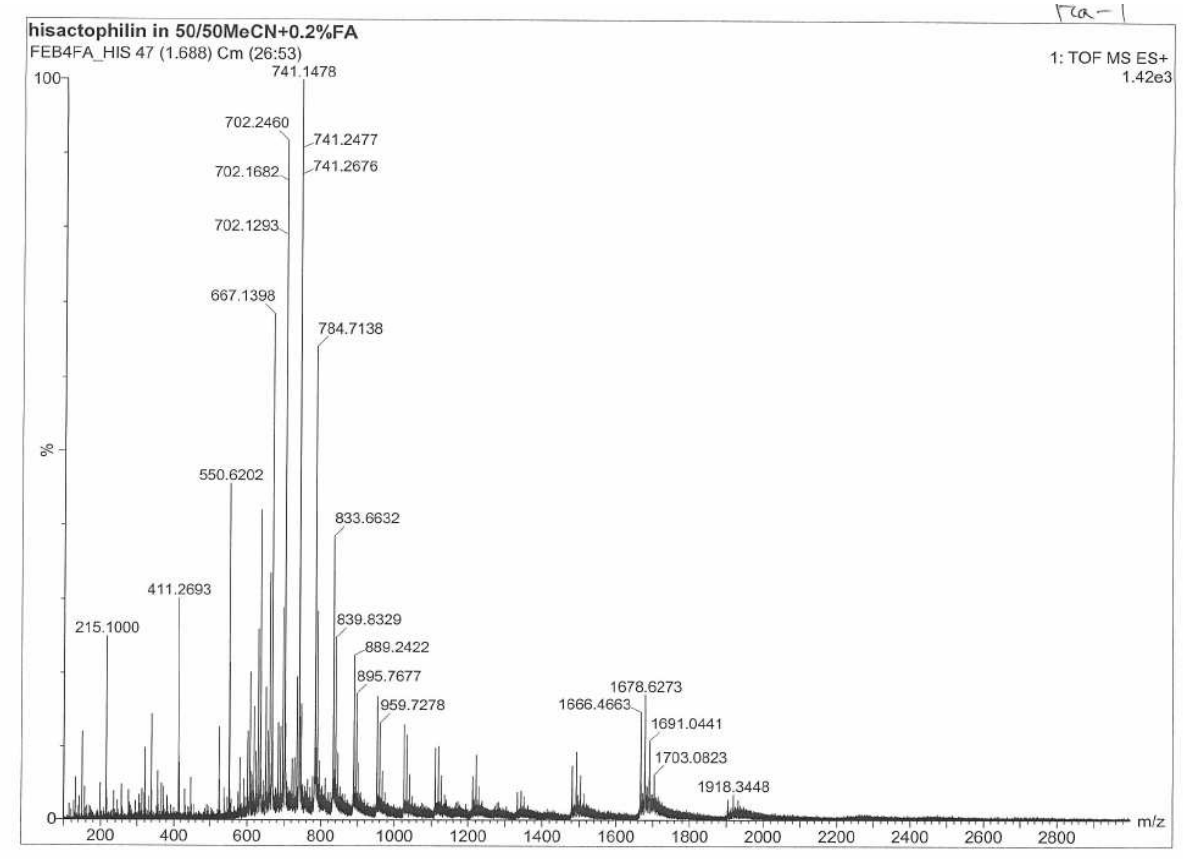
Appendix 4 : DNA sequencing for construct plasmid I93L



Appendix 5 (a) : Mass spectroscopy result of WT hisactophilin.

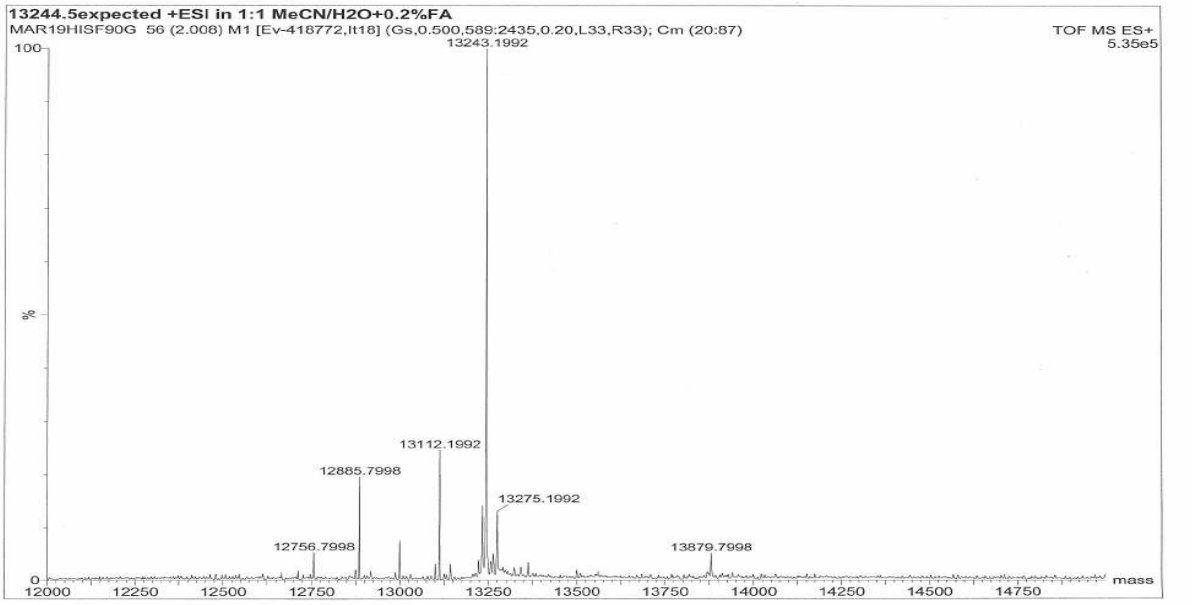
This determined the mass to be 13323.500 Da compared to the expected value of 13324.5 Da (Mass spectroscopy Facility, Dept of Chemistry, University of Waterloo).

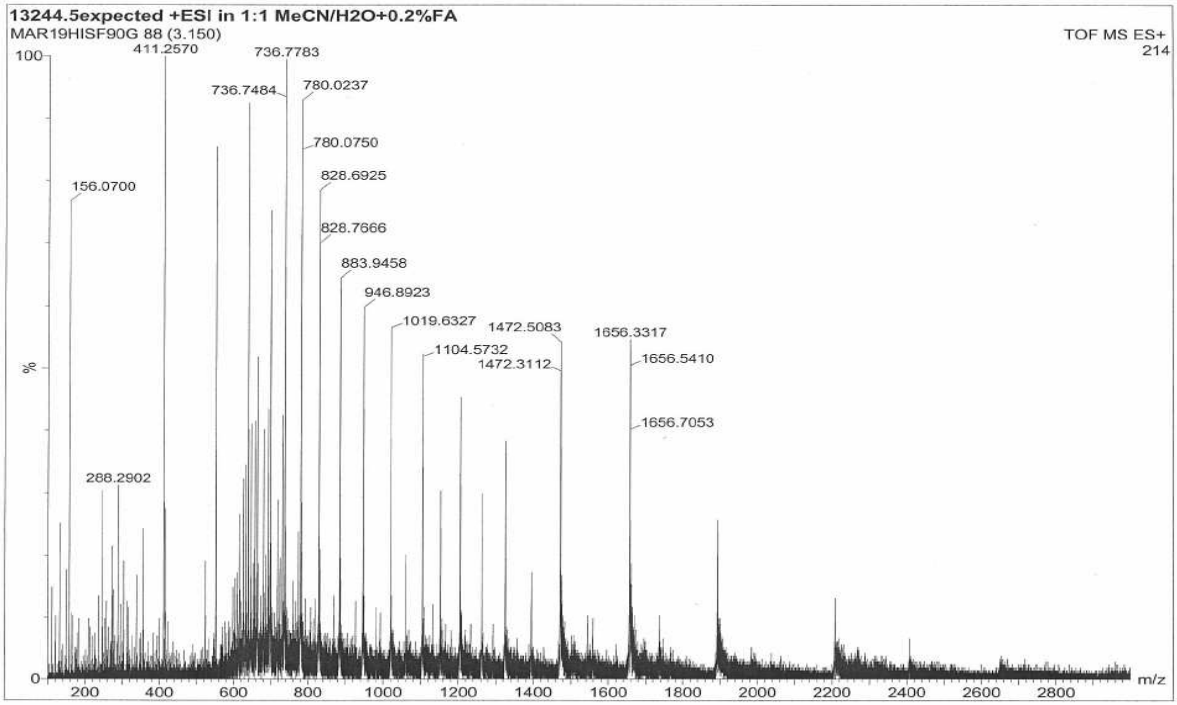




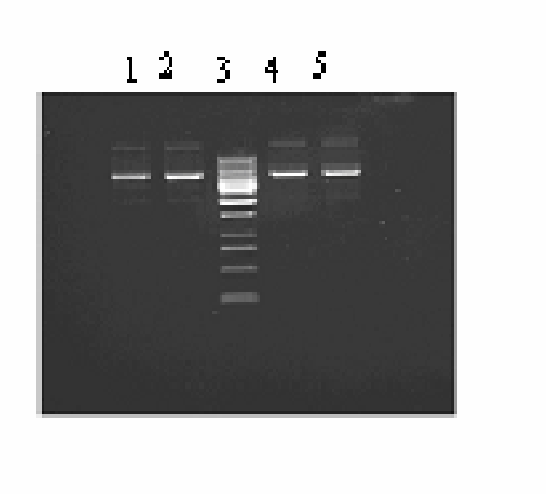
Appendix 5 (b) : Mass spectroscopy result of H90G.

This determined the mass to be 13243.1992 Da compared to the expected value of 13244.5 Da (Mass spectroscopy Facility, Dept of Chemistry, University of Waterloo).

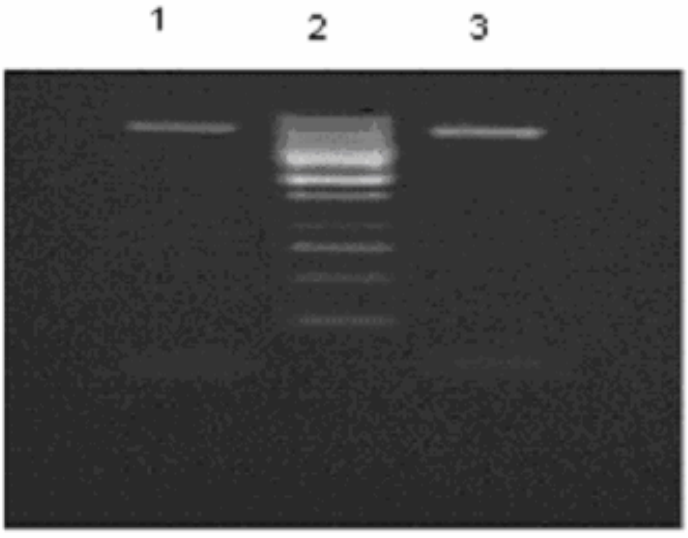




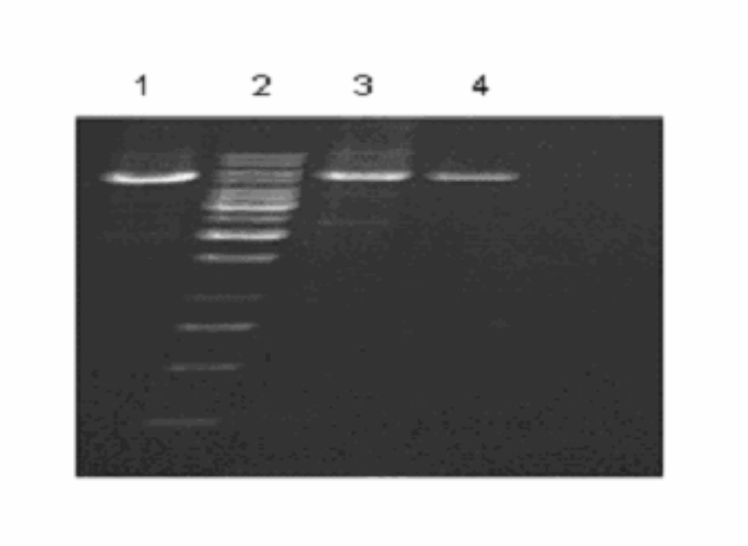
Appendix 6 : Restriction enzyme analysis of hisactophilin mutant



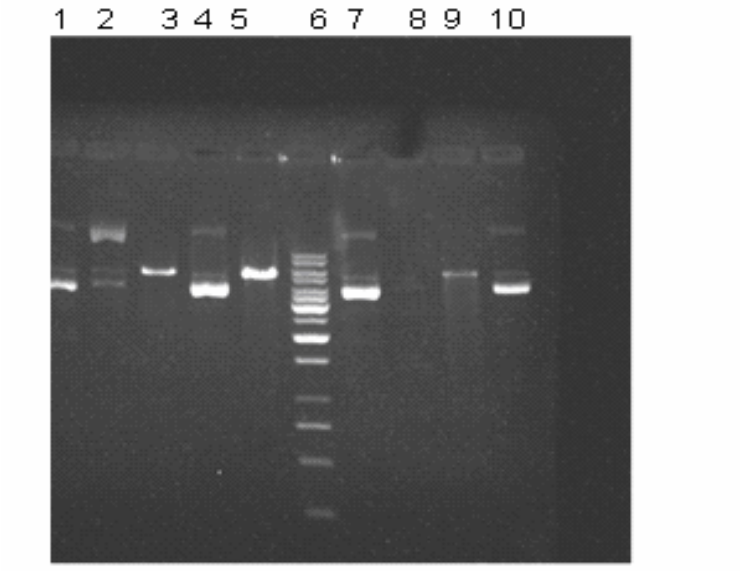
1. Hisactophilin plasmid in 0.7% agarose gel
Lane 1, 2, 4, and 5 are pHisI plasmid and lane 3 is the DNA ladder



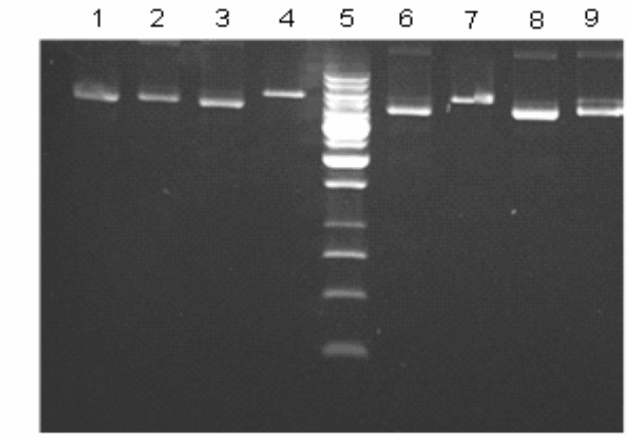
2. PCR products lane 1 F13Y, lane 2 DNA ladder, lane 3 H90G



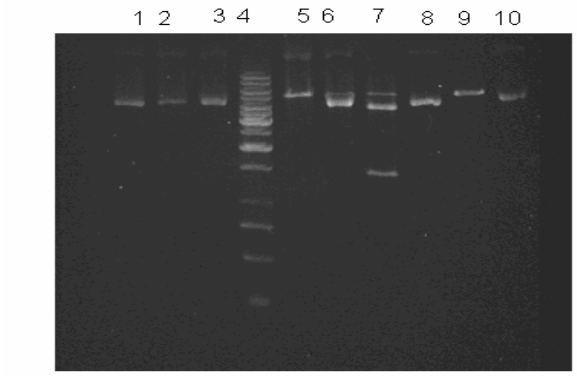
3. PCR products lane 1F6L, lane 2 DNA ladder, lane 3 I85L, lane 4 I93L



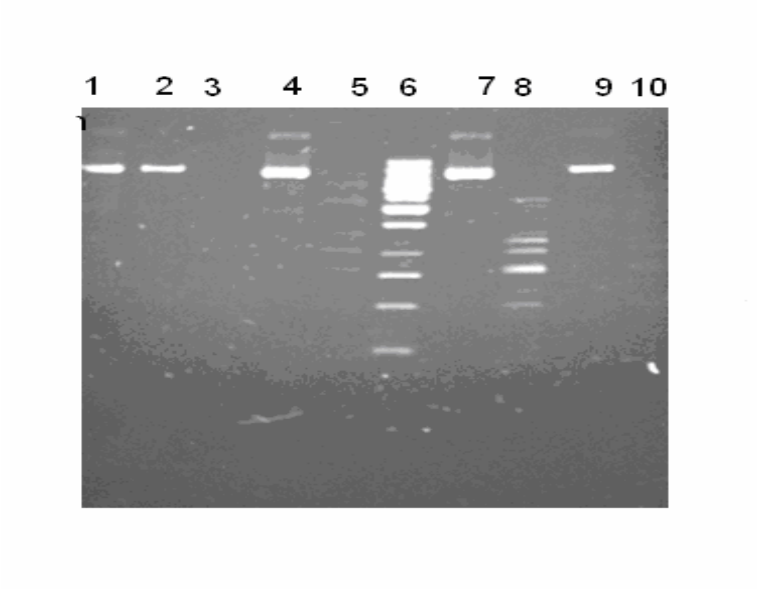
4. Restriction enzyme digests of plasmid F13Y on a 0.7% agarose gel.
Lane1 uncut pHisI plasmid run as spercoiled, lane 2 pHisI plasmid digested with AflII, lane 3 pHisI plasmid digested with EcoRv + AflII (Linear DNA, there is no restriction site in the plasmid), lane 4,7 and 10 are the uncut F13Y plasmid, lane 5,8 and 9 F13Y plasmid digested with EcoRv + AflII, there is no site for AflII just linear DNA from EcoRv digestion , lane 6 is the DNA ladder. Digestion solution contains 3µl plasmid, 2µl buffer, 4µH_2O, and 0.5µl EcoRv enzyme,0.5µl AflII enzyme incubated at 37°C for overnight.



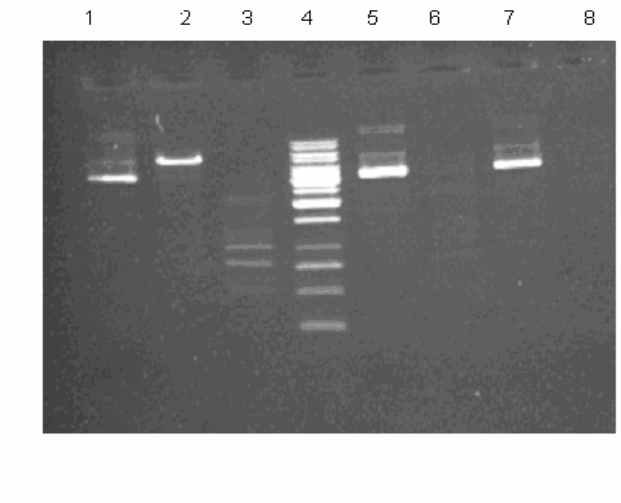
5. Restriction enzyme digests of plasmid F6L on a 0.7% agarose gel.
Lane 1 uncut pHisI plasmid, lane 2 pHisI plasmid digested with SacI there is no site in the original plasmid run as expected, lane 3,6 and 8 are the uncut F6L plasmid, lane 4,7, and 9 F6L plasmid digested with SacI, the enzyme cut once DNA this is the expected mutation, lane 5 is the DNA ladder. Digestion solution contains 3µl plasmid, 2µl buffer, 4.5 µH_2O, and 0.5µl SacI enzyme, incubated at 37°C for overnight.



6. Restriction enzyme digests of plasmid F13Y and F6L on a 0.7% agarose gel.
Lane 1 uncut pHisI plasmid, lane 2 pHisI plasmid digested with SacI there is no site in the original plasmid, lane 3 and 10 are pHisI plasmid digested with EcoRV, Lane 4 is the DNA ladder, lane 5 the uncut F13Y plasmid digested with AfIII, lane 6 uncut F13Y plasmid digested with EcoRV, lane 7 uncut F13Y plasmid digested with AfIII and EcoRV, lane 8 uncut plasmid F6L, and lane 9 F6L plasmid digested with SacI the enzyme cut once that confirm the mutation. Digestion solution contains 3µl plasmid, 2µl buffer, 4 µH_2O, and 0.5µl enzyme, incubated at 37°C for overnight.



7. Restriction enzyme digests of plasmid I85L on a 0.7% agarose gel.
Lane 1 uncut pHisI plasmid, lane 2 pHisI plasmid digested with EcoRV run around 6 Kbp as expected , lane 3 pHisI plasmid digested with Eco57MI, Lane 4,7, and 9 are uncut I85L plasmid, lane 5, 8, and 10 are the I85L plasmid digested with Eco57MI, lane 6 is the DNA ladder. Digestion solution contains 3µl plasmid, 2µl buffer, 4.5 µH_2O, and 0.5µl Eco57MI enzyme, incubated at 37°C for overnight.



8. Restriction enzyme digests of plasmid I85L on a 0.7% agarose gel.
Lane 1 uncut pHisI plasmid, lane 2 pHisI plasmid digested with EcoRV, lane 3 pHisI plasmid digested with Eco57MI. There was 5 faint bands as expected, but hard to know the exact size for each band, Lane 4 is the DNA ladder, lane 5 and 7 are uncut I85L plasmid, lane 6 and 8 are I85L plasmid digested with Eco57MI. Digestion solution contains 3µl plasmid, 2µl buffer, 4.5 µH_2O, and 0.5µl Eco57MI enzyme, incubated at 37°C for overnight.