NMR Characterization Guides the Design of Beta Hairpins and Sheets While Providing Insights into Folding Cooperativity and Dynamics

Frederick Michael Lewis Hudson

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

Doctor of Philosophy

University of Washington

2006

Program authorized to offer Degree:
Department of Chemistry

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

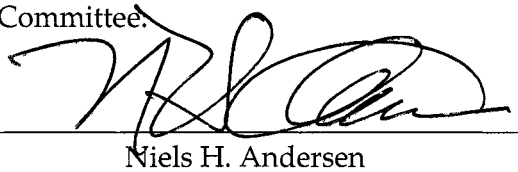
University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

Frederick Michael Lewis Hudson

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of the Supervisory Committee:

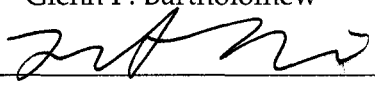


Niels H. Andersen

Reading Committee:

Niels H. Andersen

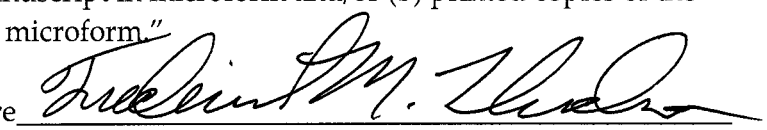
Glenn P. Bartholomew



Forrest Michael

Date: 7/31/06

In presenting this dissertation in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to Proquest Information and Learning, 300 North Zeeb Road, Ann Arbor, MI 48106-1346, 1-800-521-0600, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature 

Date 8/17/06

University of Washington

Abstract

NMR Characterization Guides Design of β Hairpins and Sheets While Providing Insights into Folding Cooperativity and Dynamics

by Frederick Michael Lewis Hudson

Chair of the Supervisory Committee:
Professor Niels H. Andersen
Department of Chemistry

Peptides populating the β hairpin and sheet folding motifs commonly found in proteins have been designed using peptide design criteria and these sequences have been characterized with NMR spectroscopy. The use of chemical shift deviations (CSDs are obtained by subtracting random coil reference chemical shifts from observed shifts) of backbone protons (Hα and H_N) is shown to be of significant value for the determination of the fold populations of these structures, as well as providing information allowing for detailed structural analysis. For β hairpin systetms, detailed structural analysis led to multiple important findings: 1) strand alignment in hairpins is driven towards the most energetically favorable association of the strands which results in similar strand alignment regardless of loop length and turn type; (2) when strands are joined with loops of various lengths, hairpin folds are favored with particular loop lengths; and (3) turn types have characteristic CSD patterns which can be used for their identification. When the CSD technique was applied for the characterization of a 3-stranded sheet (double hairpin), determination of the relative stability of two-residue turn forming sequences was possible. Measurement of the fold propensity of individual hairpins within three-stranded sheet sequences allowed for a determination of the cooperativity of the folding of the sheet sequences using a strict four-state model which provided larger estimates of cooperativity than had been previously reported using two-state analyses. At this point, the development of a new technique in this laboratory for

measuring the folding and unfolding rates of peptides using an NMR line-broadening analysis, allowed for the measurement of these rates for a hairpin system in an isolated context and when folding in a three-stranded sheet. These rate measurements further suggest that a four-state model is appropriate for the description of the behavior of these sequences.

Table of Contents

Page

List of Figures iv

List of Tables vi

List of Abbreviations vii

Chapter 1 Introduction 1

Proteins and Polypeptides 1

β Hairpins and Sheets 3

Peptide Folding and Cooperativity 6

Dynamics of Hairpin and Three-Stranded Sheet Formation 7

Trp-cage Peptides 9

Circular Dichroism 10

Nuclear Magnetic Resonance 11

Overview of Research 14

Chapter 2 Materials and Methods 18

Synthesis and Purification of Peptides 18

Circular Dichroism 19

Nuclear Magnetic Resonance 20

Preparation of Samples 20

Proton Spectra Collection 21

Chemical Shift Deviations 22

Folding Rates from 1D Spectra 23

Overview 23

Determination of Folding Rates 24

1D Spectral Data Collection 26

Chapter 3 Model Hairpins 27

Introduction 27

Hairpin Classification 28
Determination of Hairpin Populations 29
Characteristics of β Strands 31
Strand Modification and Optimization 35
MrH Series Strand Modifications 35
GB1p Strand Modifications 38
Turn Sequence Effects on Conserved Strands 42
CSD Characterization of Turns in β Hairpins 45
Conclusion 47
Chapter 4 Cage and Half-Cage Models 56
Introduction 56
Characterization of Trp-Cage Peptides 58
Characterization of Exenatide 59
^{1}HNMR of Exenatide in Aqueous Environments 60
^{1}HNMR of Exenatide in Water/DMSO Mixtures 61
CD Melting and Concentration Jump Studies 61
Structure of Exenatide and TrEx4 64
Conclusion 68
Chapter 5 Model Sheets 78
Introduction 78
Three-Stranded Sheet Mutants 82
Determination of Hairpin Populations in Three-Stranded Sheets 83
Correlation of CSD Measures with % Fold 89
Turn Types in β Sheets 90
Folding Pathway of the SG Peptides 93
Conclusion 95
Chapter 6 Minimization of a WW Domain 103
Introduction 103

Description of WW Domains 104
Minimization of WW Domain 106
N-cap Motif 106
Minimized WW Domains 108
Conclusion 109
Chapter 7 Folding Cooperativity and Dynamics of Three-Stranded Sheet Models 117
Introduction 117
Equilibrium Measures and Cooperativity 118
Dynamics of Hairpin and Sheet Formation 120
Cooperativity of SG Peptides 122
Four-State Analysis 124
Results 127
Dynamics of SG Peptide Mutants 130
Determination of Rates by ^{1}HNMR 131
Design of SG Systems 134
Application of Linewidth Analysis to SG W2 Peptides 136
Conclusion 140
Bibliography 148
Appendix A: Chemical Shift Assignments 162

List of Figures

Figure Number Page

Figure 1-1. Folding funnel cross-section for a hypothetical protein..... 15

Figure 1-2. Ribbon diagram for hpin1 adapted from Jager et al. 2001..... 16

Figure 1-3. Turn types indicating hairpin nomenclature..... 16

Figure 3-1: Diagrams showing strand orientation and nomenclature in β hairpins..... 48

Figure 3-2: CSDs of a well-folded β hairpin. 49

Figure 3-3: Aqueous CSDs for MrH2x peptides..... 50

Figure 3-4: CSDs for MrH2x peptides in 8%HFIP 51

Figure 3-5: CSDs of GB1p inspired, Trp-zip peptides..... 52

Figure 3-6: Aromatic region of ^{1}HNMR TOCSY spectra of HP7 peptides. 53

Figure 3-7: H_N CSDs of conserved strand turn mutants. 54

Figure 3-8: H_N CSDs from turn forming region of hairpin peptides..... 55

Figure 4-1: Structure of TC5b showing Trp-cage motif..... 69

Figure 4-2: Difference CSD plot for Exendin-4 in 30% TFE at pH 5.9 and 4.4. 70

Figure 4-3: NH-Hα region of TOCSY spectra for Exendin-4..... 71

Figure 4-4: Regions of TOCSY and NOESY spectra of exendin-4 in aqueous buffer. 72

Figure 4-5: CSDs of Exendin-4 in various media. 73

Figure 4-6: Circular Dichroism spectra for exendin-4 in .1% acetic acid at 280K..... 74

Figure 4-7: CD melts of exendin-4 in .1% acetic acid at high and low concentration..... 76

Figure 4-8: Proposed model of solution states of exendin-4..... 77

Figure 4-9: Concentration and medium dependence of melting behavior of exendin-4. 77

Figure 5-1: Schematic of SG three-stranded sheet..... 97

Figure 5-2: Hα and H_N CSDs for SG wild type and turn 1 mutants..... 98

Figure 5-3: Correlations between various fraction folded measures. 99

Figure 5-4: Loop CSD signatures for 4 and 5 residue loops. 100

Figure 5-5: Hα and H_N CSDs of SG and T2 mutants. 101

Figure 5-6: Four-state equilibrium model of SG three-stranded sheet. 102
Figure 6-1: Representation of WW domain. 111
Figure 6-2: Hydrophobic Clusters of Trp-Cage Miniproteins and WW Domain..... 112
Figure 6-3: Hα and H_N CSDs of Trp-Cage miniproteins..... 113
Figure 6-4: Hα and H_N CSDs for WW N-cap motifs..... 114
Figure 6-5: H_N CSDs of WWm1b. 115
Figure 6-6: Hα and H_N CSDs for WWm2. 116
Figure 7-1. Four-state model of three-stranded sheet formation. 142
Figure 7-2. CSDs of W2 mutants and wild type sequence..... 143
Figure 7-3. Aromatic regions of 1D ^{1}HNMR spectra of W2 mutants. 144
Figure 7-4. Correlation between Hε3 and % fold for W2 mutants. 146
Figure 7-5. Arrhenius plots showing folding and unfolding limbs for W2 mutants..... 146
Figure 7-6. 4-state and 2-state reaction diagrams..... 147

List of Tables

Table 3-1 Sequence of MrH4e and MrβHm 32
Table 3-2 MrH Series Mutants 35
Table 3-3 GB1p Strand Mutants 39
Table 3-4 Minimized Hairpin Systems 41
Table 3-5 MrH3 Turn Mutations 43
Table 3-6 HP7 Turn Mutants 44
Table 3-7 Sequences of Peptides Containing Different Loop Lengths 46
Table 4-1 Sequences of GLP-1 and Trp-cage forming peptides 57
Table 4-2 CD spectral data recorded for Exenatide and TrEX4 in various media 63
Table 5-1 Mutants of the SG Three-stranded Sheet 82
Table 5-2 CSD Data for SG Sheet and Mutants 86
Table 5-3 Fraction Folded of Hairpins 1 & 2 in SG Peptides 88
Table 5-4 Relative Turn Propensities of Two-Residue Sequences 92
Table 6-1 Sequences of hPin1 and Mutant WW Domains[†] 105
Table 6-2: WW N-cap motifs. 107
Table 6-3 Sequences of WW Domain Mutants 108
Table 7-1: SG and Mutants: 123
Table 7-2: Fraction folded values for hairpins in SG peptide constructs. 124
Table 7-3 Mutants of the SG Series for Rate Analysis 135

List of Abbreviations

CD, Circular Dichroism

CSD, chemical shift deviation

EtF, edge to face (perpendicular)

Hα, the α proton of an amino acid

H$_N$, the backbone amide proton of an amino acid

HFIP, 1,1,1,3,3,3-hexafluoroisopropanol

HPLC, high performance liquid chromatography

NMP, N-Methyl-2-Pyrrolidone

NMR, nuclear magnetic resonance

NOE, nuclear Overhauser effect

NOESY, nuclear Overhauser effect spectroscopy

R1, $[\theta]_{191}/[\theta]_{208}$

R2, $[\theta]_{222}/[\theta]_{208}$

TFA, trifluoracetic acid

TFE, 2,2,2-trifluoroethanol

TIS, triisopropyl silane

T-jump, temperature jump

TOCSY, total correlation spectroscopy

UV-Vis, ultraviolet-visual spectroscopy

Acknowledgements

Upon arriving at the University of Washington in 1999 I was presented with what was to me a frustrating array of projects suitable for doctoral study. After some reflection, I determined that Professor Niels Andersen had several projects that I found sufficiently interesting for study. In addition to these, I found his demeanor and absolutely remarkable grasp of the material he was presenting made me feel perfectly comfortable with the choice of joining his laboratory. As it happened, I did not find myself working on the project I initially thought I would find the most interesting. Luckily, the following years showed me that I had been directed in the proper direction, for the project I did find myself working on soon revealed itself to be both challenging and extremely rewarding. This was the first, and perhaps most important moment for which I must thank Professor Andersen. In the following years, there would be innumerable moments which he provided me with his insights into the mysteries of protein folding as well as many other pertinent areas of organic chemistry without which I would have found myself undoubtedly lost. Because of this, I provide these words as a most heartfelt thanks.

I would also find myself unable to have achieved this goal without the selfless work of my supervisory committee. I must especially thank Profs. Bartholomew and Michael who took on the obligation of reading this manuscript.

There were many people with whom I had the pleasure of working with on a daily basis and whose expertise with the complicated techniques required for this research was invaluable. These members of the Andersen Lab, Drs. Jonathan W. Neidigh and R. Matthew Fesinmeyer taught me many of the techniques that I came to use for the studies presented here. Other members of the Andersen Lab, were in the unique position of

directly sympathizing with the pain that can come from scientific research and helping to overcome these daily obstacles. In this regard, I must thank Katherine Olsen, and Jasper Lin in particular for this invaluable companionship. Without the help them as well as all other members of the Andersen lab, this volume would never have been created and I thank you all.

I would also like to thank my friends and family without whom my sanity would have almost certainly been compromised, resulting in my not completing this volume. Specifically I would like to thank Matthew Phillips, Paul Villa, Sebastian Derosia, Curtis Bigelow, Matthew Woodburn, Richard Terek, Brad Page, Mark Haines, Robert Gaumond, Michael Warren, Alex K. Brown, Agustin Gonzalez, Dr. Matt Garvert, John Ball, Robert Shirack, Donald Dill, Dave Sero, Justin Cranford, E. Alex Hudson, Peggy Schachter, Nelle Hudson, Beth McCullough Russell, Will Russell, Fred W. Hudson and Betty B. Hudson.

Dedication

For my grandfather, Dr. Isaac Frederick Hudson.

Chapter 1 Introduction

Proteins and Polypeptides

Proteins are macromolecular, polymeric structures that fold into defined, three-dimensional conformations and provide biological systems with the necessary machinery to perform the functions required for life. The conformation a protein adopts is directly related to its function; a fact which has been known for some time (Fischer 1894; Pauling 1940). There are 20, naturally occurring amino acids whose sequence is responsible for the conformation of the protein. These amino acids are linked together through the formation of amide bonds, resulting in the termini of the chain culminating in either an amino group (N-terminus) or a carboxylic acid group (C-terminus). Since amino acids differ only by the side chain attached to their α-carbon, the great diversity and functionality proteins possess is due solely to the length of the polypeptide chain and its sequence (Anfinsen et al. 1961; Anfinsen et al. 1962; Anfinsen 1973).

The structure proteins display is described from four perspectives, ranging from the finest structural features up to the description of the organization of multi-protein complexes. The "primary structural" description of a protein, which tells the least about its structure, is defined by its amino acid sequence. Secondary structures—for example α-helices and β-sheets—are three-dimensional sub-structures which are formed from the regular association of amino acids, located locally along the chain. The tertiary structure of a protein describes the way α helices, β sheets and/or other structures associate with each other to form the global fold of the protein. Quaternary structure, when applicable, is a description of the way proteins bind together, either domain/domain interaction geometry or organization within multi-protein complexes. The interactions within a protein and those a protein has with other molecules in its environment are

indicative of the functionality of that protein. Since proteins are responsible for enzymatic catalysis, metabolism, signal transduction, cell morphology and replication, their sequence and structure possess a direct relationship with human health. It is therefore fundamental to the resolution of these issues that the relationship between the structure, folding mechanism, and sequence of a protein be known.

The folding of a protein is currently thought to follow an energy landscape that can be represented by an irregularly surfaced funnel (Wolynes et al. 1995; Dill and Chan 1997; Dobson and Karplus 1999) (Figure 1-1). At the inlet of the funnel, the "random" multitude of unstructured, coil conformations is represented by the large width of the device at this point. As the protein forms native contacts, lower energy structures are formed which constrict the conformational entropy of the polypeptide chain, leading to fewer available structures populated by the chain which is represented by the proteins slide down into the funnel. This process culminates in the formation of the native state at the base of the tube at the bottom of the funnel. This location represents the enthalpic stability of the native state and the hydrophobic effect that favors compact structures with apolar sidechains internalized.

Because proteins are very large, complicated molecules, direct observation of their folding pathway in great detail is not currently possible. In the "protein engineering" approach to the elucidation of folding thermodynamics and transition state features, point mutations are used to map the energetics of the folding pathway and the extent of structuring oat specific sites. This so-called ϕ analysis (Fersht 1995) involves measuring the $\Delta\Delta G^{\ddagger}$ and the $\Delta\Delta G_{F\text{-}U}$ between a wild type system and one which contains a point mutation. When the values are the same (ϕ=1), the mutation site must be in a native-like conformation in the transition state. As ϕ approaches zero, the indication is that the site is not folded at the transition state. These studies, in the protein context, can be complicated due to the many interactions that lead to the stability of the folded state. It

is, therefore, convenient to study isolated secondary structural elements for insights into the relationship between primary structure, secondary structure and folding mechanism. These peptidic systems are are not proteins because they are without tertiary structure and function. These systems, referred to as peptides, are generally less than 50 amino acids in length. They may contain more than one secondary structural element but commonly consist of only one α helix or one β hairpin as these are the most common and simplest elements of secondary structure. One peptide motif that contains multiple secondary structural elements is the Trp-cage. This small system is said to have tertiary structure because it is made up of the collapse of several secondary structural features about a central tryptophan residue and because of this, is commonly referred to as a mini-protein. In the studies chronicled by this work, hairpins and sheets receive much attention but some work was done on Trp-cage species as well.

β Hairpins and Sheets

Consisting of two anti-parallel strands connected by a chain-reversing, turn-forming sequence, β hairpins are the simplest form of β secondary structure. Since α helices have been successfully designed for some time, much of the current emphasis in peptide design has been on the design of β-sheet structures. These structures, because of their simplicity, provide the framework for the exploration of interactions necessary for the formation of β sheets as well as providing insights into the folding mechanisms of such structures. The *de novo* design of polypeptides displaying elements of β secondary structure has been a challenging goal in peptide design since the subject was introduced. Experimentally, β structures are difficult to work with because they trend towards low solubility and aggregation (Osterman 1985). Unlike α helices, these structures eluded researchers for some time because they result from multiple interactions such as hydrophobic core formation and cation/pi interactions, in addition to hydrogen bond formation. The design of these structures has also been complicated by the fact that the

simple excision of sequences from proteins generally does not yield hairpins with a detectable, native-like, folded state in water. No structures that adopt a β hairpin outside of the protein context were known until hairpins from the B1 domain of protein G , GB1p (Kobayashi et al. 1993; Blanco et al. 1994b; Honda et al. 2000; Fesinmeyer et al. 2004), and Ubiquitin (1-17) (Cox et al. 1993) were examined. Instances of *designed* β hairpins that fold in water eluded researchers until the early 90's (Blanco et al. 1993; de Alba et al. 1996; Ramírez-Alvarado et al. 1996). Once it became clear that, in contrast to proteins, β-hairpin formation requires residues in the turn region that are particularly favorable for chain-direction reversal (Sibanda and Thornton 1991; Blanco et al. 1993; Searle et al. 1995; Karle et al. 1996; Ramírez-Alvarado et al. 1996; Haque and Gellman 1997; Maynard et al. 1998; Syud et al. 1999) , cross-strand hydrophobic interactions (Maynard et al. 1998; Andersen et al. 1999), and β-branched residues in the strands, rapid progress was made in this field. Since then, designed three-stranded (Sharman and Searle 1997; Kortemme et al. 1998; Schenck and Gellman 1998; de Alba et al. 1999b; Griffiths-Jones and Searle 2000; López de la Paz et al. 2001; Santiveri et al. 2003) and four-stranded (Carulla et al. 2002; Syud et al. 2003) sheets have been made.

Previously, β hairpins were designed by using a naturally occurring sequence containing the desired folding motif and making substitutions to improve fold stability. Recently, the focus was on elucidating interactions necessary for hairpin and sheet formation which was only possible after these naturally occurring systems were found. These studies have led to a much greater understanding some of these interactions though the quantitation of many remains incomplete and difficult. These studies have also increased the interest in more complicated systems such as naturally occurring three-stranded sheets and such systems have since been designed. The successful designs of these more complicated systems have provided opportunities for the study of the interactions necessary for the formation of these more complicated structures and have allowed for the probing of cooperativity in sheet formation. Others have focused

on the design of these complicated structures and extended these design criteria to natural systems with the goal of creating model systems emulating these folds. One naturally occurring system that has been of particular interest to the peptide folding and design community are a group of homologous sequences collectively known as WW domains.

WW Domains

WW domains are naturally occurring motifs that are approximately 40 amino acids long and derive their name from two highly conserved tryptophan residues (Chen 1995; Sudol 1995a; Sudol 1995b). In many cases, they are thought to autonomously fold into their native, three-stranded, antiparallel-sheet conformation (Macias et al. 1996; Koepf et al. 1999). The function of the domains is thought to be the binding of proline rich regions of partner proteins (Sudol 1996; Otte 2003) which makes them important components in bio-recognition phenomena. These peptides have come under intense study of late because their size lends them to mutational (Kraemer-Pecore 2003) and mechanistic (Jager et al. 2001) studies. WW domains have also been important for the determination of factors relevant to the kinetics of β-sheet folding (Crane et al. 1999; Koepf et al. 1999; Ferguson et al. 2001a; Ferguson et al. 2001b; Jager et al. 2001). Other studies have evaluated WW sequences by ϕ analysis, showing which residues are possess a native-like conformation in the transition state (Jager et al. 2001).

While many studies have concentrated on the folding dynamics of WW domains, others have focused on the equilibrium characteristics of these systems. NMR characterization of WW domains has been a topic of great interest of late and these studies have found the key interactions necessary for the folding of these domains (Macias et al. 2000; Jager et al. 2001). These studies have also shown that the domains are stabilized by the formation of two distinct hydrophobic clusters (Figure 1-2). The relative simplicity of the three-stranded β-sheet formed has led to work on designing WW-like peptides

(Macias et al. 2000; Kraemer-Pecore 2003) which has been meet with some success. In one study, the residue frequency at key sites was surveyed to define a statistically probable sequence (Macias et al. 2000). Other efforts have focused on the *a priori* design of a WW sequence (Kraemer-Pecore 2003) using a predictive algorithm. Both efforts have yielded peptides that fold into WW like three-stranded sheets. While this work is promising, very little has been done to minimize an existing WW domain. Such a project is sure to be challenging due to the very specific hydrophobic clustering these domains exhibit but is a worthwhile exercise to demonstrate the advanced state of the art of peptide design.

Peptide Folding and Cooperativity

The understanding of the factors that stabilize β structures has allowed for the design of systems for the study of folding cooperativity and dynamics. A cooperatively folding structure is defined in the classical sense as having a two-state folding mechanism. Thermodynamically, one would observe only two-states (F and U) when an intermediate concentration of denaturant is present or at intermediate temperatures along the melting curve. A cooperatively folding peptide or protein should have a well defined, highly populated, folded-conformation and lack partially folded intermediates (Creighton 1993). Kinetically, this means that there are no partially folded states having a sufficient lifetime for detection. In the case of peptides such as hairpins and three-stranded sheets, the experimental evidence of cooperativity can be more difficult to interpret than that for a protein because the melting point and chemical denaturation transitions are broadened and the folded state is in rapid, rather than slow, equilibrium with the unfolded ensemble. Because of this, experimentally assessing the cooperativity, or two-state behavior of a peptide is challenging and is the focus of much scientific attention (Blanco et al. 1994b; Searle et al. 1995; Sharman and Searle 1997; Schenck and Gellman 1998; de Alba et al. 1999b; Stanger et al. 2001). While these efforts have yielded insight into the modes of cooperativity of β sheets (lengthening of the strands vs.

increasing the number of strands), they have left many questions unanswered and provided opportunities for the development of new methodologies.

Dynamics of Hairpin and Three-Stranded Sheet Formation

Fluorescence or IR monitored T-jump experiments have been used to measure folding rates in the 10^3 to 10^7 s^{-1} timescale. The first microsecond timescale kinetic measurements of hairpin formation were carried out on a 16-residue peptide derived from GB1 (Muñoz et al. 1997). This study also utilized a fluorescence monitored T-jump technique to observe the unfolding of the peptide and determined a folding time constant of 6μs at 297K. These researchers later proposed a model to describe these observed apparent two-state hairpin formation kinetics (Muñoz et al. 1998). This model begins with nucleation of the turn followed by the sequential formation of inter-strand hydrogen bonds. The process of hydrogen bond formation, because the angles about two peptide bonds must be fixed per H-bond, was said to be uphill until sufficient favorable cross-strand enthalpic interactions accumulate (Muñoz and Serrano 1997). However, a computational study (Dinner et al. 1999) of this GB1 derived peptide indicated that hydrophobic collapse was the first event in the hairpin folding process. This study, utilizing multi-canonical Monte Carlo simulations, showed that the rate limiting step involved the rearrangement of the collapsed structures to form the native hydrophobic contacts and H-bonds.

Subsequent studies have implicated turn formation in the transition state of β-structure formation: the turn forming propensity of the loop region affects hairpin folding rates much more than strand optimization (Du et al. 2004) and the length of the region between the hydrophobic clustering residues affects the folding rate of the hairpin (Dyer et al. 2004). The report from Du et al. suggested that the formation of a hairpin proceeds through turn nucleation followed by inter-strand hydrogen bond formation; a so called

"zippering" model. The reports from Dyer et al. however, suggest a folding mechanism initiated by the initial collapse of residues into a non-native compact state followed by reorganization.

These kinetic studies of hairpin formation have all used temperature jump experiments monitored in the ultraviolet or infrared spectral regions but others have recognized the ability of dynamic NMR measurements to determine folding and unfolding rates in the microsecond to millisecond timescale (Huang and Oas 1995; Myers and Oas 2001; Wang et al. 2003). In this method, the exchange between folded and unfolded states that peptides exhibit leads to broadening of the proton resonances with a large chemical shift difference between the unfolded to the folded state. This method has traditionally relied on simulations of observed peak shapes to determine the degree of broadening due to exchange. The peak shape equations require fitting to several variables including the intrinsic linewidth and the broadening due to exchange. Since the effects of these variables on the simulated peak shape are the same, the technique is somewhat questionable. Because the procedure requires fitting over many variables, it is computationally intensive and subject to errors due to required assumptions. This led to the development, in this laboratory, of a method which greatly simplifies the procedure through the use of an internal standard (Olsen et al. 2005b). Currently, kinetic experiments probing the effects of mutation on the rates of designed β-structure formation are limited to the hairpin context. No studies to date have looked at the degree to which mutation affects the folding and unfolding rates of three-stranded sheet formation or the effect of mutation on the individual hairpins within the sheet context when the mutations are made within or remote from that hairpin.

Trp-cage Peptides

Other than β hairpins and sheets, Trp-cage species are also of interest here. The so called Trp-cage motif is formed when hydrophobic residues collapse around a centrally located tryptophan sidechain. This collapse involves the formation of multiple secondary structural elements including an α helix and a 3_{10} helix and produces a distinct ^{1}HNMR spectroscopic signature. The Trp-cage fold was first observed (Neidigh et al. 2001) at the time of the NMR characterization of exendin-4. Exendin-4 is a naturally occurring peptide isolated from the salivary secretions of a Gila monster (Eng et al. 1992). The NMR studies of this peptide (Neidigh et al. 2001) showed that the monomeric state was not observable in aqueous solution so the characterization was performed in a TFE containing medium. This study reported the notable medium dependence of the structure of the peptide, and the observation of a new fold in 30% TFE led to the design of Trp-cage mini-proteins which are stable in water (Neidigh et al. 2002).

Studies of exendin-4 in purely aqueous media have been hindered by aggregation. The study of a truncated analog of exendin-4, TrEX4, indicated that the Trp-cage portion of the peptide is destabilized by association with DPC micelles due to an energetically favorable interaction between the tryptophan residue and the phosphocholine head groups. This interaction prevents a key association involving a proline residue and the tryptophan.

Exenatide (synthetic exendin-4) is the first member of the incretin mimetic class of therapeutic agents and is in phase 2 clinical trials for the treatment type 2 diabetes. It has been shown to decreases post-meal blood glucose concentrations; an effect associated with weight loss. While the therapeutic formulation of this peptide is aqueous, no studies have been able to characterize it in this medium. Because this

peptide is in advanced drug testing stages, characterization of the agent under near physiological conditions has become an issue that should not be further delayed.

Circular Dichroism

Circular dichroism can be used for the spectroscopic characterization of proteins and peptide systems. A circular dichroism spectrometer generates circularly polarized light of both left and right handedness and a signal is produced when a sample selectively absorbs the light polarized in one direction or the other. This technique is commonly applied to the study of chiral molecules and lends itself to protein and peptide science because chiral peptide chains selectively absorb circularly polarized light; an effect which is dependent upon secondary structure as well. The far-UV region (260-188nm) of the spectrum is of generally used for this purpose because the amide n-π^* transitions which occur here, are sensitive to conformation. The precise wavelength and intensity of these transitions is related to the backbone angles, ϕ and ψ, adopted by the peptide chain. This leads to a correlation between the secondary structure of peptides and the observed spectrum; an observation that has been made for some time (Greenfield and Fasman 1969).

The spectrum observed for peptide systems is the sum of the spectra that would observed for each of the states populated by the peptide, allowing for the determination of the folded fraction and the structure preference of a polypeptide. For the case of purely α-helical peptides, the determination of fraction folded is straightforward due to a characteristic minimum at ~222nm. When both α and β conformations are present, deconvolution of the spectrum can provide the fractions of each that are present. CD spectroscopic analysis has also been applied to Trp-cage structures with the results indicating such systems produce a diagnostic spectroscopic signature (Neidigh et al. 2002). While circular dichroism has been used in some cases to estimate hairpin fold

populations (Ramírez-Alvarado et al. 1996; Maynard et al. 1998; Andersen et al. 1999; Griffiths-Jones and Searle 2000; López de la Paz et al. 2001; Blandl et al. 2003; Ciani et al. 2003; Fesinmeyer et al. 2004), the failure of many hairpin peptides to display a classic antiparallel β-sheet signature (Blanco et al. 1994b; Chen et al. 2001; Gibbs et al. 2002; Fesinmeyer et al. 2004) often precludes this type of analysis. This has led to the use of other spectroscopies in conjunction with CD for the characterization of these systems.

Nuclear Magnetic Resonance

The use of NMR for the determination of peptide conformations and populations has been reported for some time. The relationship between peptide structure and backbone chemical shifts leads to diagnostic chemical shifts for these resonances that reflect local secondary structure (Wishart et al. 1991). This study indicated the Hα resonance for a residue in an α helix are shifted, on average, 0.39ppm upfield from random coil and the average shift observed for Hα resonances in β sheets is 0.37ppm downfield from random coil. Similar observations have been made for $^{13}C\alpha$ chemical shifts (Williamson 1990). The earliest observations that the chemical shifts of backbone protons in β hairpins had characteristic chemical shifts led to the estimation of hairpin populations by the magnitude of the shift deviations from random coil (CSDs) (Blanco et al. 1994b; Searle et al. 1995). Since β-hairpin and sheet fold lifetimes ($1/k_u$) are short, less than 50μs (Muñoz and Serrano 1997; Xu et al. 2003; Dyer et al. 2004; Dyer et al. 2005), chemical shift deviations *are* population weighted averages and provide fold populations when the unfolded state approximates random coil norms and the chemical shifts for the fully folded state can be measured or accurately approximated. As peptides with well populated hairpin states were examined by NMR, an alternating pattern was observed in plots of CSDs of α protons (Searle et al. 1995; de Alba et al. 1997). This was shown to result from the orientation of strands in hairpins which causes alternate Hα protons to be directed either towards or away from the other strand. Amide proton chemical shifts

have rarely figured in the assignment of secondary structure because observations have shown that it is more difficult to calculate NH shifts based upon secondary structure (Osapay and Case 1991; Xu and Case 2001).

Efforts to estimate hairpin populations by using CSDs are complicated in the presence of aromatic residues that superimpose large ring current effects on the secondary structure shifts which are present. The Gellman lab (Espinosa and Gellman 2000; Syud et al. 2001; Syud et al. 2003) has used the Hα shifts of the hydrogen-bonded residues (S±1 and S±3, Fig. 1-3) because these sites avoid ring current effects, even though a survey of β structures in proteins determined that the non-hydrogen-bonding positions (S±2 and S±4) produced larger structuring shifts for these resonances (Sharman et al. 2001). To employ the CSD method for the determination of hairpin populations, most studies have relied on one or more assumptions about the fully folded state; a point that will be discussed in greater detail subsequently in this work.

The CSD technique is used in this laboratory for the determination of conformation state as well as fold populations of peptides. To do this, knowledge of the chemical shifts of peptide backbone protons when they are found in peptides that have no secondary structure (random coil chemical shifts) are needed. The random coil values are determined experimentally from peptides known to be unstructured, but the value is also dependent upon other interactions independent of folding. These values, which can be found in the literature, are now considered well characterized (Merutka et al. 1995; Wishart et al. 1995; Andersen et al. 1997) and are used in this lab to calculate CSDs with an in-house, automated online program, *CSDb* (Andersen et al. 2004). To estimate the chemical shifts of the fully folded state, we use the best folded mutant of a homologous series in the most structuring medium employed and have found this technique of significant value for determinations of peptide structure.

The characterization of turns in β-hairpin peptide structures by NMR has, in the past, relied on their classification by diagnostic NOE assignments and strand characterization by cross-strand NOE's (Wüthrich 1986; Searle et al. 1995; de Alba et al. 1996). The population of folded states has been determined through the integration of key NOE cross-peaks but, since NOE intensities are not simple population weighted averages, this technique is questionable and points to the need for reliable classification schemes. As will be shown herein, the chemical shifts of backbone protons of β-structures are generally downfield shifted and provide a wealth of information about the conformation of the peptide. The relationship between the strands of a hairpin or sheet results in an alternating large/small downfield shift for these protons, a feature called the "shift register". The shift register is a result of alternate Hα or H_N protons being inwardly or outwardly directed in an edge strand of a sheet structure. For each residue in a strand, either the Hα or the H_N is directed towards another strand. For the next residue, the opposite relationship is observed. The chemical shifts of inwardly directed and outwardly directed protons are quite different; when a plot of the difference between these shifts and random coil norms (CSD) is generated, the inwardly directed hydrogen's appear on the graph as more downfield shifted (large CSD) and the outwardly directed protons appear relatively upfield in comparison (small CSD) (see Figure 3-2 for reference). The CSD method is extended to turn type definition and the determination of hairpin populations in this work.

Overview of Research

The research presented in this document applies ^{1}HNMR NMR to the study of β hairpins and three-stranded sheets for the determination of peptide design criteria, thermodynamics, folding pathway and cooperativity. Also, NMR methods and circular dichroism were applied to the study of a naturally occurring therapeutic peptide, exendin-4 for characterization of its state in aqueous media.

In the course of study of β hairpins, key interactions that are commonly found in these peptides were probed. Investigations were also performed to determine whether particular strand pairings prefer specific types of turns or reversing loops. During these studies, characteristic CSD patterns were found that are diagnostic of turn formation and type. It was also observed that S±odd H_N CSDs provide more reliable estimates of folded state population than either S±even H_N CSDs or S±even Hα CSDs. These studies were extended to a three-stranded-sheet model which determined the turn-forming propensity of many 4 residue turn forming sequences. This research also allowed for the determination of the folding mechanism of a three-stranded-sheet model and for the calculation of the cooperativity of this peptide and a mutant containing different turn sequences. Dynamic NMR methods were also applied to an analog of the three-stranded sheet which provided insights into the effects on folding and unfolding rates of a hairpin when in an isolated vs. three-stranded sheet context.

The studies performed on exendin-4 characterized this peptide for the first time in strictly aqueous media. These studies indicate that this material forms stable oligomers at higher concentrations and that the monomeric state can be observed at low concentrations. This monomeric state is stabilized by the formation of a Trp-cage fold though it was not established whether formation of this fold is necessary for receptor activity.

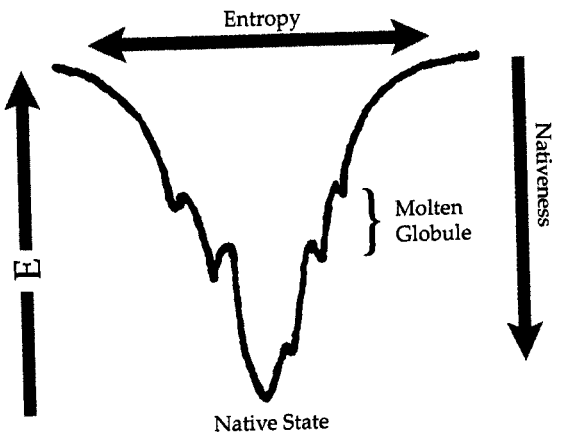


Figure 1-1. Folding funnel cross-section for a hypothetical protein.

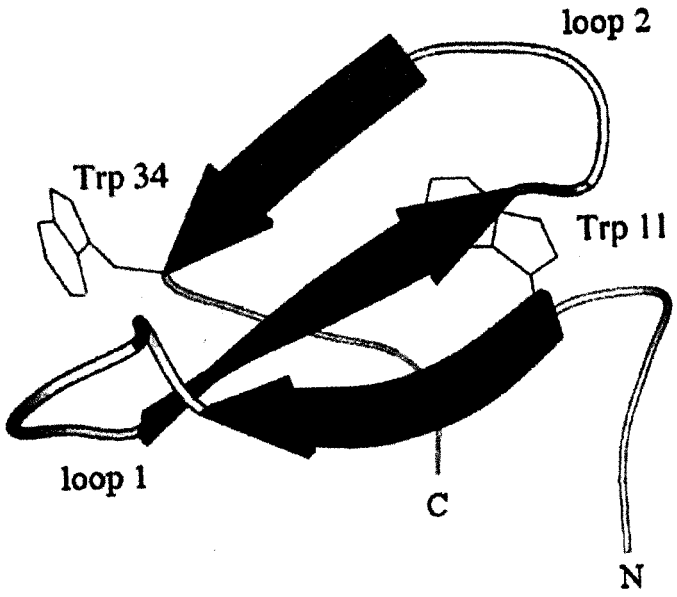


Figure 1-2. Ribbon diagram for hpin1 adapted from Jager et al. 2001.

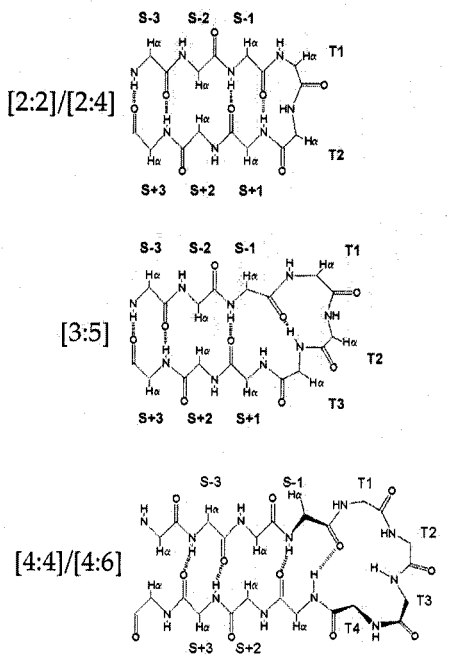


Figure 1-3. Turn types indicating hairpin nomenclature.

Chapter 2 Materials and Methods

Synthesis and Purification of Peptides

All peptides presented in this work were synthesized at the University of Washington, with the exception of Exenatide (synthetic exendin-4) lots (gifted from Amylin Pharmaceuticals Inc.); MrH5b, Gb1m4, and WWm1b (Synpep Corp.); and MrH6a (JPT Peptide Technologies GmbH).

Production of peptides in this lab uses standard Fmoc chemistry on an Applied Biosystems model 433A peptide synthesizer. Peptides with amidated C-termini were generated on Rink amide MBHA resin while peptides with free C-termini were constructed on Wang resin, preloaded with the C-terminal residue. Peptides with N-terminal acetyl groups were acetylated with 3% acetic anhydride and 5% Triethylamine in DMF for 1 hour. Cleavage of peptides from resin was achieved in TFA with 2.5% TIPS and 2.5% water as radical scavengers. Reaction times of 1.5 to 2 hours were generally sufficient for deprotection and cleavage from the resin to yield crude peptide.

Purification of crude peptide was accomplished using reverse phase HPLC using a C_{18} preparatory-scale column. Elution of peptide products was achieved with water/acetonitrile (with 0.1% and 0.085% TFA respectively) gradients which generally provided adequate separation of the desired peptide from the side products. Peptide elution was monitored by UV/Vis detection: the peptide backbone absorbance was monitored at 215 nm and/or UV active side-chain chromophore absorbance at 276 nm. When additional purification was necessary, a C_8 semi-prep scale column was used with the same solvent system mentioned above.

Isolation of purified peptides was accomplished by rotary evaporation of the organic phase followed by lyophilization to yield the purified peptide solid. This material was assumed to be the TFA salt of the desired product and *ca.* 85% pure; the remainder consisting of counter-ions and water. Verification of the product mass is achieved by chemical ionization of the peptide in 10% acetic acid/methanol and injection onto a Bruker Esquire electro-spray atomization, ion-trap mass spectrometer. Sequence verification is provided by NMR.

Circular Dichroism

Circular Dichroism spectra were collected on a Jasco J-720 spectropolarimeter calibrated to d-10-camphorsulfonic acid, $[\theta]_{192.5}$ = 15,600 (Yang et al. 1986) and equipped with a Peltier temperature control device. Spectra were acquired in .1mm quartz cells, .01mm quartz cell at high peptide concentrations, or a 1cm quartz cell for low concentrations. The cells were polarimetrically tested and certified as low birefringence. Blank spectra were acquired on the same experimental buffer formulation and cell used to acquire the sample spectra. These were then subtracted from the sample spectra to correct for any instrumental or impurity-induced signal. Samples for CD analyses were prepared by dissolving lyophilized peptide in sufficient solvent to yield a concentration of ≈ 250μM. The exact concentration was then determined by UV/Vis spectroscopy and the application of Beer's Law to the absorbance of aromatic residue side chains. For tyrosine containing peptides, an extinction coefficient (ε) of 1420 cm^2/mmol at 276 nm was used; for tryptophan containing peptides, the absorbance at 280 nm was measured and an ε value of 5580 cm^2/mmol was used while 1280 cm^2/mmol was added for each tyrosine present. Once the exact concentration was determined, the appropriate aliquot was added to a volumetric flask for standard dilution to 2, 20 or 30 μM as desired. Spectra were then acquired directly on these solutions. Blank spectra were routinely acquired at 25°C while spectra of peptide samples were acquired every 10°C from 5° to 95°C.

CD spectra were acquired as the average of 10 scans using a scan rate of 100 nm/min and a step resolution of .1 nm. Scans were generally acquired from 270 nm to 190 nm. Spectra were trimmed at the shortest wavelength which still provided a dynode voltage less than 600 V. The trimmed spectra were then subjected to a data processing routine beginning with noise reduction by reverse Fourier transform followed by subtraction of the blank spectrum. The blank corrected spectra were then zeroed at 260 nm and the units converted from millidegrees to molar ellipticity (deg*cm^2/dmol). The units were routinely converted to residue molar ellipticity by dividing the molar ellipticity by the number of amide bonds in the peptide.

Nuclear Magnetic Resonance

All NMR spectra were acquired on a Bruker DRX-500 or a Bruker DMX-750 spectrometer.

Preparation of Samples

Solutions for NMR experiments were made by dissolving lyophilized peptide in the desired aqueous buffer with cosolvents added when necessary. Sample volumes were routinely 450 μL and peptide concentrations were generally ≈1 mg/ml. The samples always contained 10 vol-% D_2O for instrument locking purposes and DSS (2,2-dimethyl-2-silapentane-5-sulfonate) as a chemical shift reference for proton spectra. The chemical shift of the –$SiMe_3$ groups was set to 0 ppm for all proton experiments. 20-50 mM potassium phosphate buffer was used for all samples with a desired pH within the 5-7 range. For samples around pH 3, 20 mM deutero-formate buffer was routinely used. The pH of the resulting sample solution was measured and adjusted to the desired value with small quantities of 1M HCl or 1M sodium hydroxide. When deuterated solvent samples were needed, the sample would first be prepared as described previously. This

sample would then be lyophilized to yield a material which contained both the peptide and the buffer salt. This material would then be reconstituted in 99.9+% D_2O.

Proton Spectra Collection

Proton data collection in non-deuterated aqueous buffer requires suppression of the solvent proton signal for it is orders of magnitude larger than the signal of the molecule of interest. The suppression of solvent protons is achieved with the WATERGATE (Piotto et al. 1992) pulse sequence. 1D and 2D experiments were acquired using this solvent suppression technique. 2 dimensional experiments were employed for the purpose of assigning the chemical shift of all protons resonances in each peptide. The 2D TOCSY experiment provides correlations based upon through-bond couplings in molecules. This information is usually sufficient for identification and assignment of all amino acid spin systems. 2D NOESY spectra provide through space couplings which are necessary for a sequential assignment. TOCSY experiments used a 60 ms MLEV-17 (Bax and Davis 1985) spinlock which was the appropriate length for propagation of magnetization throughout the spin-system without being of sufficient length for relaxation of the initial correlation to occur. NOESY experiments routinely used a 200 ms mixing time; an amount found sufficiently long to provide adequate NOE intensity for assignment while minimizing (but not completely eliminating) higher order signals.

Chemical shifts for all protons were recorded to a precision of .001 ppm despite the uncertainty of their value being on the order of ± .005 ppm. 1D spectra were commonly acquired with 16k points and 16 scans. When higher quality spectra were needed, 200-400 scans were acquired. TOCSY spectra were routinely recorded with 400 increments in the indirect dimension while 2k points and 8 scans were commonly acquired in the direct dimension. NOESY spectra were generally acquired with 512 increments in the indirect dimension while 2k points and 16 scans were generally required in the direct

dimension. All spectra utilized a sweep width of 11-12 ppm. Because the experimental resolution is not equal in the two dimensions, the chemical shifts, as reported, are determined from the directly observed dimension for greatest accuracy. The greatest sources of error are probably due to instrument temperature controller miscalibration and cosolvent variance.

Chemical Shift Deviations

The relationship between the proton chemical shift and the structure of a peptide can be illustrated by calculating the difference in chemical shift of the backbone hydrogen atoms from the chemical shift one would expect for the analogous atom in an unstructured, random coil peptide. This difference is referred to as the chemical shift deviation (CSD) and is defined as:

Equation 2-1: Definition of CSD

$$\delta_{observed} - \delta_{reference} = \delta_{deviation} = CSD$$

All CSDs reported here were calculated with the application, CSD*b* (Fesinmeyer et al. 2004; Fesinmeyer et al. 2005). The CSD*b* is an automated, online program for the storage of chemical shift data and provides on the fly calculation of CSDs by various methods as well as a graphical presentation of the data.

Folding Rates from 1D Spectra

Overview

NMR line shape has been used to measure conformer interconversion rates (Sandstrom 1982; Gunther 1995) in the past. The standard equations used to describe NMR line shapes are given below:

Equation 2-2: Lineshape Equations

$$g(\nu) = [(1+\tau\pi\Delta)P + QR]/(4\pi^2P^2 + R^2), where:$$
$$P = (0.25\Delta^2 - \nu^2 + 0.25\delta\nu^2)\tau + \Delta/4\pi$$
$$Q = [-\nu - 0.5(\chi_A - \chi_B)\delta\nu]\tau$$
$$R = 0.5(\chi_A - \chi_B)\delta\nu - \nu(1+2\pi\tau\Delta)$$

These equations were used to model peak shapes as described previously (Olsen et al. 2005b). In this method, the peptide must have a chemical shift probe which usually consists of two protons of the same type; one must change greatly upon folding and the other must have essentially no change in chemical shift upon folding. In the reference provided, the researchers used the two equivalent methyl groups (Hγ1 and Hγ2) of a valine residue as well as the Hδ and Hε of a tyrosine to calculate the folding and unfolding rates of GB1 and some analogues. This can be done because there are only three sources of linewidth in an NMR spectrum, the intrinsic linewidth (Δ°), the exchange broadening (Δ^{ex}) and that due to coupling (j). Since coupling constants of resonances in amino acids are known, and there is a resonance with virtually no exchange broadening which provides the intrinsic linewidth, the amount of exchange broadening can be known. In the case of two resonances with the same coupling constant, one of which (A) has a larger CSD in the folded state, exchange broadening is calculated via these equations:

Equation 2-3: Calculation of Exchange Broadening.

$$\Delta\Delta^{ex} = \Delta^{tot}_{A} - \Delta^{tot}_{ref}$$
$$\Delta^{ex}_{A} = \Delta\Delta^{ex}\,[R^2/(R^2\text{-}1)]$$
$$R^2 = (CSD_A)^2/(CSD_{ref})^2 = \Delta^{ex}_{A} / \Delta^{ex}_{ref}$$

Δ^{ex}_{A} can then be used to calculate folding and unfolding rates from the equations below:

Equation 2-4: Calculation of rates from Δ^{ex}.

$$k_F = (4\pi\chi_U\chi_F{}^2\Delta\upsilon^2) / \Delta^{ex}_{A}$$
$$k_U = (4\pi\chi_F\chi_U{}^2\Delta\upsilon^2) / \Delta^{ex}_{A}$$

Determination of Folding Rates

Folding rates were determined from peak broadening as measured in 500 MHz 1D ^{1}HNMR spectra. These measurements can only be performed on peptides exhibiting a greatly shifted resonance because the broadening observed in these cases can be measured precisely (Olsen et al. 2005a). In the specific cases discussed here, the calculation is possible due to a cross-strand Trp/Trp pair. Because of the orientation of the indole rings, the Hε3 of one of the indole rings feels a large ring current and exhibits a large upfield chemical shift deviation. The Hε3 proton on the other indole ring experiences a minimal ring current and has essentially no CSD. This resonance is used as a linewidth reference to determine how much of the broadening of the other Hε3 is intrinsic since for it, $\Delta^{total} \approx \Delta^{\circ}$. The sources of broadening measured here are due to experimental conditions like magnet shimming and temperature, etc. Data from these experiments was subjected to minimal processing, Fourier transformation and phase correction, then converted to ASCII format and imported into Excel for peak property

determination. Peak heights are automatically calculated by normalizing the spectral baseline to zero and, in the case of doublets, using the average height of the two peaks divided by two for the half-height. Half-height linewidths are determined by finding the closest data points to the calculated half heights and finding the difference between them, in hertz. These are then modeled assuming a coupling constant of 7.6Hz (observed for a tryptophan Hε3 proton) to determine the total amount of broadening, Δ^{total}, that is independent of the coupling constant. The Δ^{total} of the reference peak is determined by the same methodology and $\Delta^{ex}{}_A$ is the difference between the two Δ^{total}s. Δ^{ex} is then related to the folding and unfolding rates as indicated by equation 2-4.

Fraction folded and unfolded in these systems is determined by the CSD of the shifted Hε3 resonance via these equations:

Equation 2-5: Calculation of Fraction Folded.

$$X_F = CSD_{obs}/CSD_{100\%}$$

$$X_U = 1 - X_F$$

The determination of 100% folded is based mostly on literature values (Cochran et al. 2001a; Andersen et al. 2006) but this technique can be difficult to implement because different systems have different 100% folded values. In the case presented in this work, one peptide exhibited behavior that was consistent with literature values for this resonance while the other did not. In this case, the difference in the folded fraction of the two peptides could be known with certainty. This allowed for the relationship between the two different Hε3 CSDs to be known and a new 100% value for this resonance could be determined and used in further calculations.

1D Spectral Data Collection

1D ^{1}HNMR spectra were acquired for linewidth analysis at a range of temperatures. The spectra were acquired with the standard proton experiment pulse sequence, 90° pulse followed by FID acquisition. No WATERGATE or other solvent suppression is required as the samples are prepared in D_2O. The spectra, due to the necessity of having a very large signal to noise ratio, are acquired with 1500 to 2000 scans, depending on sample concentration, and with 16k points. No other experimental parameters need be considered.

Chapter 3 Model Hairpins

Introduction

Since the discovery that the B1 domain of protein G , GB1p (Kobayashi et al. 1993; Blanco et al. 1994a; Blanco et al. 1994b; Honda et al. 2000; Fesinmeyer et al. 2004) and Ubiquitin (1-17) (Cox et al. 1993) contains an autonomously folding β-hairpin, the field of β-hairpin and sheet design has maintained considerable interest. This first example of an isolated β-hairpin incited a great effort to understand the interactions important for hairpin formation and, eventually, to create designed peptide sequences that formed well converged β-hairpin and sheet structures. The C-terminal hairpin (GB1p) became the focus of many studies (Honda et al. 2000; Cochran et al. 2001a; Blandl et al. 2003; Fesinmeyer et al. 2004)and remains of considerable interest to this day. Another sequence, derived from the DNA binding interface of the Met-repressor protein complex, was published by the Searle group (Maynard et al. 1998). Since the DNA binding interface of the Met-repressor protein complex is the association of two strands from different proteins and not a hairpin unit, the sequence they designed featured the insertion of a loop between strands modeled after those of the protein complex. This sequence, MrβHm, like GB1p and Ubiquitin (1-17), folded into a hairpin in water in the absence of a protein context but is distinguished by the presence of an INGK loop, two N-terminal lysine residues and, contains only one aromatic residue in the sequence, making it more amenable to analysis using NMR chemical shifts. The report from the Searle lab on MrβHm indicated that the hairpin was 50% folded in water but had a tendency to aggregate. Subsequent studies (Andersen et al. 1999) confirmed these observations and found a large stabilization of the hairpin fold in fluoroalcohol containing media.

These and many other studies have led to advancements in the understanding of the interactions important to β-hairpin formation and current work in this field has established protocols for hairpin design and characterization by ^{1}HNMR. The design of β-hairpins and their subsequent spectroscopic characterization is challenging because characterization of these structures requires the a priori knowledge of the degree to which they are folded *and* intimate knowledge of the structure. This represented a paradox in the field for many years and led to the inaccurate estimation of fold populations of systems for some time. Today, we have model systems that exhibit large hairpin populations which allows for the standardization of spectroscopic protocols of structure determination and fold population quantitation. The work outlined here shows peptides made and indicates that these have been useful in refining our ability to spectroscopically characterize β-hairpin peptides as well as to advance the knowledge of interactions and other factors that are important to hairpin formation.

Hairpin Classification

Hairpins are generally classified by the number of residues contained in the turn. Turn residues are defined as those that do not contain the proper dihedral bond angles to be classified as β-strand residues. Alternatively, strand residues have been described as those whose amide proton *and* carbonyl oxygen participate in the appropriate cross-strand hydrogen bonds. The nomenclature used here was proposed by Sibanda and Thornton (Sibanda and Thornton 1991) and is compatible with both of the previous definitions. Loops are designated by the number of residues involved therein followed by the number of residues, counted from the center of the turn sequence, that are required before both cross-strand hydrogen bonds are observed. The shortest loops found in β-structures belong to a class known as [2:2] hairpins and, in this nomenclature system, [2:2] turns have 2 residues which do not have the proper dihedral angles to be classified as either α-helix or β-strand while both hydrogen bonds are formed by the two

turn-flanking (S+1 and S-1 positions, see Figure 3-1) residues. [2:4] loops differ because the two S+1 and S-1 residues form only one of the two possible hydrogen bonds. This classification system is applied to larger loops as well; insertion of another residue into the turn results in the [3:5] turn or Gly bulge, while the largest turns that will be discussed are formed by adding one more residue. These turns are of the [4:4]/4:6] class and a survey of residue preferences in protein hairpins (Fesinmeyer et al. 2004) provides design concepts. The [2:2]/[2:4] class of turns can be further characterized as type I, II, I′ and II′ by the dihedral angles found in the turn residues. These turn types (and the larger reversing loops for that matter), are defined by the backbone dihedral angles (ψ and ϕ) of the residues that form them, leading to a residue preference for each turn site. In proteins, the most common turns are the type I and II while in β-hairpins (including those in proteins), their mirror images, types I′ and II′, are most common.

Determination of Hairpin Populations

The characterization of β-hairpin peptide structures has, in the past, relied on the classification of turn type by diagnostic NOE assignments and strand characterization by cross-strand NOE's (Searle et al. 1995; de Alba et al. 1996). The population of folded states was determined through the integration of key NOE cross-peaks but, since NOE intensities are not simple population weighted averages, this technique is questionable. While circular dichroism has also been used to estimate hairpin fold populations (Ramírez-Alvarado et al. 1996; Maynard et al. 1998; Andersen et al. 1999; Griffiths-Jones and Searle 2000; López de la Paz et al. 2001; Blandl et al. 2003; Ciani et al. 2003; Fesinmeyer et al. 2004), the failure of many hairpin peptides to display a classic antiparallel β-sheet signature (Blanco et al. 1994b; Chen et al. 2001; Gibbs et al. 2002; Fesinmeyer et al. 2004) often precludes this type of analysis. Since β-hairpin fold lifetimes ($1/k_u$) are short, less than 50μs (Muñoz and Serrano 1997; Xu et al. 2003; Dyer et al. 2004; Dyer et al. 2005), chemical shift deviations (CSDs) *are* population weighted

averages and provide fold populations when the unfolded state approximates random coil norms and the chemical shifts for the fully folded state can be measured or accurately approximated. Work here demonstrates that NMR backbone proton resonance CSDs can, in some contrast to NOE determinations, provide peptide conformation and population estimates quickly, reliably and accurately.

The earliest observations that chemical shifts of backbone protons in β-hairpins had characteristic chemical shifts led to the estimation of hairpin populations by the magnitude of the shift deviations from random coil (CSD) (Blanco et al. 1994b; Searle et al. 1995). Efforts to estimate hairpin populations by using CSDs are complicated in the presence of aromatic residues that superimpose large ring current effects on the structuring shifts which are present. Even though a survey of β structures in proteins determined that the non-hydrogen-bonding positions (S±2 and S±4) produced larger structuring shifts (Sharman et al. 2001), the Gellman lab (Espinosa and Gellman 2000; Syud et al. 2001; Syud et al. 2003) has used the Hα shifts of the hydrogen-bonded residues (S±1 and S±3, see Fig. 3-1) because these sites avoid ring current complications. To employ the CSD method for the determination of hairpin populations, most studies have relied on one or more assumptions about the fully folded state: that 30-40% trifluoroethanol (TFE) gives the fully folded species (Ramírez-Alvarado et al. 1996; Ramírez-Alvarado et al. 1997; Syud et al. 2001); that the average Hα CSDs in β strands should be +0.40 ppm (Santiveri et al. 2001); or that the shifts observed for the ^{D}Pro-Gly turn species, particularly cyclic ones (Espinosa et al. 2002; Tatko and Waters 2004), corresponds to 100% folded.

During the course of our efforts to determine the thermodynamics and kinetics of haipin folding (Andersen et al. 2002a), researchers from this lab observed diagnostic patterns in both Hα and backbone amide NH (H_N) CSDs that provided independent and unambiguous assignment of hairpin register. We further found the magnitude of the H_N

CSDs to be useful for providing population estimates even though H_N shift trends have rarely figured in the assignment of secondary structure because it has generally been observed that it is more difficult to predict or calculate NH shifts based on polypeptide structure (Osapay and Case 1991; Xu and Case 2001). Currently, members of this lab have succeeded in synthesizing hairpin peptides that are well folded and contain no aromatic residues, which provides a convenient starting point from which to show the CSD characteristics that are indicative of hairpin formation. These will then be applied to the systems I made and studied in an effort to understand the factors important to hairpin formation and the spectroscopic signatures that are indicative of their formation.

Characteristics of β Strands

β hairpins never achieve measurable hairpin populations in water without the inclusion of β-branched residues in the strands and the presence of cross-strand hydrophobic clustering. Experimentally, the presence of strand residues in proteins can be determined from the extraction of ϕ and ψ backbone angles and cross-strand distances from X-ray crystal structures (Sibanda 1984). NOE determination of distances can also provide the orientation of strands relative to each other. While cross-strand Hα/Hα and Hα/H_N NOEs can provide the register in both parallel and anti-parallel strand alignments, the determination of hairpin populations using the volumes of the cross-peaks is a questionable practice since NOEs are not simple population weighted averages of the folded and unfolded states.

In this lab, we use CSDs to determine the components of hairpin structure. To illustrate the spectroscopic characteristics of strands in well folded β hairpins, it is convenient to show data from a sequence synthesized by another member of the Andersen lab. This peptide, MrH4e, is one of a series of peptides modeled after MrβHm, first reported by Searle and coworkers, (Maynard et al. 1998) and was synthesized and characterized by

Kelly Huggins, from our group. The peptide contains no aromatic residues throughout, and has an amino-isobutyric acid (abbreviated Aib but given the sequence code U) residue in the T1 position. The aromatic residue, Y3, was mutated from the original sequence to eliminate the possibility of ring current effects obscuring the β signature and the Aib residue was used because it has been shown to have a high β-turn propensity when used in the T1 position (Aravinda 2002). The sequence of this peptide is given, as well as that for MrβHm, in Table 3-1.

Table 3-1 Sequence of MrH4e and MrβHm

MrβHm†	KKYTVS	INGK	KITVSI
MrH4e‡	KKLTVS	IUGK	KITVSA

† Literature sequence provided for purposes of comparison
‡ Made and characterized by coworker Kelly Huggins

Figure 3-2 shows the Hα (a) and H_N (b) CSD histograms for MrH4e in aqueous and fluoroalcohol containing media (20% HFIP). In the Hα panel, the shift register of the peptide strands is quite apparent when looking at the data acquired in fluoroalcohol containing media; the CSDs of Lys^2, Thr^4, Ser^6, Lys^{11}, Thr^{13} and Ser^{15} (S±2, 4 and 6) are much larger than the others, indicating that these are the α protons that are cross-strand directed. In the absence of added fluoroalcohol, somewhat smaller CSDs are observed at most sites but the register indicator is retained; the Hα CSDs from the cross-strand directed sites (S±2, 4 and 6) are 30-40% smaller but still larger than those for the S±1, 3 and 5 sites. Since peptide hairpin populations have been shown to be higher in fluoroalcohol, the indication of this data is that only the cross-strand directed sites reflect the folded-state population increase. Under any circumstances, the use of these smaller CSDs would lead to greater uncertainty in any population estimates. There is also a clear relationship between the magnitudes of the CSDs in one strand relative to the other: the N- and C-terminal strands of a hairpin, because of the interactions involved in

the folded state, must be aligned concurrently for hairpin formation to occur yet the magnitudes of the structuring Hα CSDs at equivalent positions in each strand are not equal. The Hα CSDs of the S+even (strand 2) sites are all smaller than those in the S-even (strand 1) sites, in both aqueous and fluoroalcohol containing media. While the diminishing magnitude of the CSDs going from the turn to the termini clearly indicate fraying near the strand termini, the diminution of the C-terminal strand Hα CSDs cannot be explained. This observation will be referred to as the "C-terminal effect" on Hα CSDs.

Panel (b) of Figure 3-2 shows the amide proton CSDs (H_N) for MrH4e. The histogram, upon first inspection, is very similar to that which reflects the Hα CSDs. Here we find that the S±odd sites are those that have the large CSDs. This indicates which NH protons are involved in cross-strand hydrogen bonds and shows the expected phase shift between the strand register of the Hα and NH protons. Addition of HFIP causes a large increase in the magnitude of these CSDs and has little impact on the non-hydrogen-bonding (NHB) (S±even) CSDs, again indicating that the NHB H_N values are of questionable value for population estimation. The hydrogen-bonding CSDs of the C-terminal strand have similar values to those in the N-terminal strand, indicating the "C-terminal effect" applies only to Hα CSDs. Comparison of panels (a) and (b) shows that the average magnitude of the H_N CSDs is somewhat larger than the Hα values which should provide greater precision in population determination. Also, because of the magnitude of the H-bonded CSDs in HFIP, we will assume that this peptide is nearly fully folded and such an assumption provides expectation values for strand NH CSDs in these systems. While this peptide is not ideal for the characterization of the turn forming region of the peptide, the location of the turn is still apparent due to the lack of large CSDs throughout that region (I7-K10). The characterization of the various turn types is demonstrated in the following section.

Characteristics of β Turns

Surveys of hairpins in proteins have found many examples of these secondary structural elements. The classification of the turn regions initially required the x-ray coordinates of structures (Sibanda 1984) such that backbone ϕ and ψ angles and atom distances could be extracted. These allowed for the identification of antiparallel strands and the location of turns. Since the determination of turn type requires the knowledge of the hydrogen bonding pattern, distances were also extracted from NOE data to determine which amide/carbonyl pairs were involved in a hydrogen bond. That study found that many of the hairpins found in proteins had tight, two-residue turns, but that three-, four- and five-residue loop sequences could also be found. They also found that, in some contrast to other β turns in proteins, which are commonly the type I and II 2-residue turns, β-hairpins usually contain the mirror image type I′ and II′ turns.

The search for NMR based criteria for turn determination led to the NOE based characteristics of different turn types (Wüthrich 1986). In his text, the author outlined the NOEs that are observed in type I, II, I′ and II′ turns. The difficulty of this type of analysis is that many of the turn types exhibit the same diagnostic NOEs [ie d_{NN} (i,i+2), $d_{\alpha N}$(i, i+2), and d_{NN}] with the differences being in relative intensities. As many experimental features lead to changes in the intensity of NOEs (including fold population), the application of this technique proved challenging.

Currently, many optimized turn sequences have been introduced which allow for the insertion of high propensity turn sequences into peptides. Of the two-residue turn sequences, pG (Karle et al. 1996; Haque and Gellman 1997; Syud et al. 1999), NG (Sibanda and Thornton 1991; Ramírez-Alvarado et al. 1996; Maynard et al. 1998) and UG (Aib-Gly) (Aravinda 2002) are notable for promoting hairpin formation. Of [3:5] turns, many sequences have been extensively investigated (Blandl et al. 2003), leading to the

conclusion that PDG has a high propensity for turn formation. There is a relative paucity of data on loops longer than three residues, see however Fesinmeyer et al. 2004. Because of this, my initial attention was focused on [4:4]/[4:6] turns though [3:5] and [2:2]/[2:4] turns do receive some attention as well. The presence of many turn types, along with the extended strands, highlights the challenges of spectroscopic characterization of these peptides. It will be demonstrated in this chapter that β strands and turns have unique spectroscopic signatures that allow for their full characterization, as well as a determination of the extent to which they are formed (fold population).

Strand Modification and Optimization

MrH Series Strand Modifications

The report from the Searle lab (Maynard et al. 1998) that a stable hairpin based upon the Met-repressor dimer interface had been designed, started the efforts to improve strand propensity in this series. Because very little was known about the interactions in strands that led to high hairpin fold populations, many mutations were made concurrently by several members of the lab. The mutations made by me, and their relationship to MrβHm are shown in Table 3-2. The reports from the Searle lab indicated that

Table 3-2 MrH Series Mutants

MrβHm†	Ac-KKYTVS	INGK	KITVSI
MrH2a	RKYTVS	INGK	KITVSE-NH_2
MrH2b	RKWTVS	INGK	KITVSE-NH_2
MrH2b′	RKWTVS	INGK	KITVSE
MrH3a‡	KKYTVS	INGK	KITVSA

† Literature sequence from Maynard et. al.

‡ Sequence synthesized and characterized by coworker Matthew Fesinmeyer

the wild-type peptide, MrβHm, was 50% folded in aqueous solution but also had a tendency to aggregate (Maynard et al. 1998). Studies in this lab confirmed the peptides tendency to aggregate. The poor solubility behavior of this peptide made it an unacceptable host for turn characterization studies. Therefore, members of this lab set out to improve the peptides solubility characteristics, as well as to increase the fold population so that both turn and strand optimization studies could be performed.

It was theorized that the aggregation problem could be caused by Ile^{16} which was expected to be relatively unstructured, due to its location at the C-terminus. The MrH3a peptide, made by Matt Fesinmeyer, attempted to correct this problem by mutating this Ile residue to an Ala. This peptide, as it turned out, had improved solubility characteristics and an improved folded state population (Fesinmeyer et al. 2005), as hoped. The peptides I was preparing concurrently attempted to solve the aggregation problem as well, but with an emphasis on coulombic interactions at the termini of the strands.

The MrH2x peptides were designed to probe the importance of the coulombic interactions at the strand termini as well as the effect of a conservative change to the hydrophobic cluster. The first mutation that was made investigated the impact of a sidechain-sidechain coulombic interaction, in the absence of charged strand termini. This was done by keeping the N-terminus free, amidating the C-terminus, and mutating the Ala^{16} to glutamate (MrH2a). The CSDs of the MrH sequences given in Table 3-2, are are shown in Figure 3-3, as determined in aqueous media. The magnitude of both the structuring $H\alpha$ (S±even) and H_N (S±odd) CSDs of MrH2a are less than those of MrβHm. The sidechain coulombic interaction that was attempted between Lys^1 and Glu^{16} or the C-terminal amide group was either not structure favoring or less favorable than the interactions that were replaced. The MrH2b peptide utilized the MrH2a sequence but had K1R and Y3W mutations. The K1R mutation increased the size of the cationic

moiety at the end of the sidechain of residue 1 while the Y3W mutation changed the size and hydrophobicity of one of the key sidechains assumed to be involved in the hydrophobic cluster. The CSD plot of this peptide shows that it has gained a small increment in all of its structuring CSDs, relative to MrH2a. Figure 3-3 also shows that MrH2b is still not as structured as MrβHm but the gains in stability realized by these mutations are probably mostly attributable to the increased size and hydrophobicity of the Trp^3 sidechain.

Because it was believed that the gains upon going from MrH2a to MrH2b were largely driven by hydrophobicity, it was proposed that increasing the coulombic interaction at the termini may realize further gains in stability as well as help decrease the strand fraying at the termini that is evidenced by near random coil CSDs. The peptide MrH2b′ has the same sequence as MrH2b but has a free C-terminus. In Figure 3-3 we see that, in water, except at the extreme C-terminus, there was no gain in CSD magnitude relative to MrH2b. The terminal effect that was desired appears to be present, as evidenced by the large increase in both the Hα and H_N CSDs of E16. If we look at the CSDs of these peptides in water with HFIP added to 8% (Figure 3-4), we find that the relative hairpin populations have changed. The fold stability of MrH2b′ now is near that of MrH2a. The apparent gains in fold population observed for the MrH2b′ peptide are in accordance with the accepted views of the mechanism of fluoroalcohol stabilization of peptide structures. Because HFIP disfavors the solvation of amide bonds, intramolecular H-bonds are favored. Since this peptide has free termini HFIP may also be favorable for the formation of coulombic interactions. Because the MrH2x series of peptides was less well-folded in water, MrH3a was selected as the host for turn mutation and this was undertaken by myself and other members of the lab. This will be discussed in a later section of this chapter because at this point, a report out of the protein engineering department at Genentech changed my focus to mutations of the Gb1p sequence.

GB1p Strand Modifications

The report from the Genentech labs that the GB1p sequence was dramatically stabilized by the introduction of 2 sets of cross-strand Trp/Trp interactions at non-hydrogen bonding (S±2 and S±4) loci (Cochran et al. 2001a), led to the study of GB1p-like peptides that displayed this interaction (Table 3-3). The peptides that were reported to display this novel interaction were said to be dramatically destabilized when any of the Trp sites were reverted to the original GB1p residues (Cochran et al. 2001a). Spectroscopic characterization, including melting curves, of these Trp-zip peptides was possible by CD due to a large exciton couplet that resulted from the defined conformation of the cross-strand tryptophan pairs (Cochran et al. 2001a). The studies of these systems in this laboratory were done by me, Matthew Fesinmeyer and Brandon Kier. While my goal was to determine whether hairpin stabilization was in fact due to Trp/Trp interaction (and which one?) and whether this could be substituted for a different kind of interaction (namely cation/pi), Matt set out to optimize the turn region. The utilization of both Matt and Brandon's peptides, along with mine, allows for the determination of the individual contribution of each of the Trp/Trp interactions to the stability of the resulting hairpin.

The first peptide of the HP series that was made, HP5W4, conserved the four tryptophan residues but introduced a new [4:4] turn, two lysine residues at the N-terminus, and had a T15Q mutation near the C-terminus. This peptide was used to probe the stability of a different turn sequence along with the effects of a favorable coulombic interaction at the termini. The peptide proved to be spectacularly stable (T_m=85°C) and, in addition to the presence of an exciton couplet in the CD spectrum, also exhibited a greatly upfield shifted W5 Hε3 proton resonance. The greatly shifted Hε3 resonance is due to an edge-to-face relationship between the cross-strand tryptophans which results in the thrusting of this proton into the face of the indole ring of the cross-strand tryptophan; in this case W12. The shielding that results from close proximity to the face of an aromatic ring is

what leads to such a large upfield chemical shift (for reference see Figure 3-6 panel a). Of added interest, the extreme upfield shift of Hε3 of W5 of Trpzip 4 was not seen in HP5W4. Peptide HP5W4 then became the point from which future studies began.

Table 3-3 GB1p Strand Mutants

GB1p[1]	GEWTY	DDATKT	FTVTE	F_f=49%*
Trpzip4[1,2]	GEWTW	DDATKT	WTWTE	T_m=70
HP5A	KKYTW	NPATGK	ATVQE	F_f=15%*
HP5R	KKYTW	NPATGK	RTVQE	F_f=22%*
HP5W[3]	KKYTW	NPATGK	WTVQE	T_m=66
HP5W3[4]	KKATW	NPATGK	WTWQE	T_m=58
HP5W4[3]	KKWTW	NPATGK	WTWQE	T_m=85

[1]Literature sequences shown for comparison
[2]From Cochran et al. 2001.
[3]Synthesized and characterized by coworker Matthew Fesinmeyer
[4]Synthesized and characterized by coworker Brandon Kier
*At 4°±1°C

Initially, the relative contribution of the two cross-strand Trp/Trp pairs were not known and the possibility that a Trp/Arg, π-cation interaction could also work was considered. The mutations made to HP5W4 were aimed at the elucidation of three things: (1) the nature of the Trp-Trp interactions in the peptide; (2) the determination of whether the two cross-strand pairs had different effects on the stability of the peptide and; (3) to see if they could be replaced by a different kind of interaction.

To determine whether the S±2 W/W interaction can be replaced with a cation/pi interaction, a peptide was made with a W12R mutation. This sequence, HP5R, has relatively small CSDs (Fig 3-5) which indicates that the cation/pi interaction is far less stabilizing than the W/W interaction it replaced. This leads to the conclusion that the Trp/Trp interaction closest to the loop must be uniquely stabilizing in these systems.

The quantitation of the interactions can be determined relatively when we look at many of the mutants. The determination of the contribution of the pair near the termini is readily determined from the peptides HP5W4 and HP5W3. HP5W3 has an identical sequence to HP5W4 except that Trp^3 has been mutated to alanine; a mutation that causes a ≈27°C drop in melting point. When this change in melting point is compared to changes in a similar system (Andersen et al. 2006), a 3 kJ/mol decrease in stability can be calculated for these peptides. If we look at peptide HP5W, it has a cross-strand W/W interaction at the S±2 sites but has a Y/V interaction near the termini and has a melting point of 66°C. This peptide can be directly compared to HP5A which has the same sequence except the S+2 W has been mutated to alanine. The mutation is highly destabilizing and results in a peptide which is perhaps 15% folded at the low temperature experimental limit. This equates to a destabilization of 8.5±1 kJ/mol. This is significant for the goal of hairpin minimization because it indicates that a very stabilizing interaction can be engineered near the turn while the strand terminal W/W interaction is apparently much less stabilizing to the hairpin fold, and could perhaps, be eliminated. The quantitative data also indicates that the hairpin fold is made much more stable by the introduction of an interaction which prevents the unfolding of the turn, rather than prevention of the fraying of the strand ends. This knowledge was then used to create systems with minimal strand length that still formed stable β-hairpins.

Strand Minimization in Trp-zip Peptides

The study of Trp-zip peptides provided information indicating the possibility of designing very stable minimal hairpin systems. Design of minimalist systems is of interest because their successful design requires a fundamental understanding of the interactions necessary for secondary structure formation. It was also thought that the creation of very short, simple strands could act as a very simple template for turn mutations that could lead to a simple method for determining the spectroscopic signature of different turn types. To design a minimalist hairpin system, a turn

sequence that has a significant propensity for chain reversal as well strands which have the minimum number of cross-strand interactions needed to stabilize the hairpin fold are needed. Since the Gb1 sequence, and ultimately the HP series peptides (Table 3-3), provided knowledge of a very stabilizing cross-strand interaction, a series of minimalist hairpins was envisioned. The truncation strategy adopted, as a result, was to remove as much of the strands as possible while keeping the S±2 tryptophan pair. The HP7 series (Table 3-4) of peptides originated from the removal of the hydrophobic residues at the S±4 positions of the HP5W strands and removing an N-terminal lysine and the glutamine from the C-terminal strand residues and substituting a shorter turn sequence. The species made, HP7, contained a [4:4] turn because this loop was shown to be particularly favorable for turn formation. This peptide proved to be well folded in water and had the characteristic upfield Hε3 for the Trp in the N-terminal strand. The edge-to-face interaction that is characteristic of the favorable trp-trp cross-strand pair in the longer systems is present in the truncated systems as well (See Figure 3-6a for reference).

Table 3-4 Minimized Hairpin Systems

				T_m (°C)†
HP5W‡	KKYTW	NPATGK	WTVQE	66
HP7*	KTW	NPATGK	WTE	66
HP7INGK	KTW	INGK	WTE	54
HP7NPDGT	KTW	NPDGT	WTE	
HP7EPDGK	KTW	EPDGK	WTE	

†Determined by monitoring the $[\theta]_{228}$ CD band
‡Made and characterized by coworker Matthew Fesinmeyer
*Made and characterized by coworker Katherine Olsen

This led to the introduction of other loop sequences into the HP7 strand context. This resulted in the HP7INGK, NPDGT and EPDGK peptides. These peptides were determined to be relatively less folded than HP7 but still had measurable fold populations. The investigation by ^{1}HNMR led to the determination that the [3:5] loop

sequences were not supportive of the edge-to-face interaction between the tryptophan residues, while the [2:2] loop was (figure 3-6). Figure 3-6 shows the aromatic region of TOCSY spectra for HP7INGK (a), HP7NPDGT (b) and HP7EPDGK (c). Looking at panel a for reference, this spectrum shows the Hε3 proton of W3 has a resonance at ~5.9 ppm while the random coil value is 7.65 (Wüthrich 1986). In the [3:5] peptides (b and c), the W3 Hε3 resonances are near 6.9 ppm indicating a signal much closer to random coil values and the relative lack of an edge-to-face interaction. These peptides, as well as other conserved strand studies by other lab members, allowed the focus of my research to shift to determination of the relative hairpin stabilization available from different turn/loop sequences and whether strand alignment can be retained with different loop sizes.

Turn Sequence Effects on Conserved Strands

Optimization of the turn region of β-hairpins has been pursued in many labs. A [2:2] turn was reported by the Gellman lab that was particularly good for chain reversal (Haque and Gellman 1997). This IpGK loop contained the non-natural amino acid ^{D}proline (p) which was shown to have the highest turn propensity of any other known two-residue turn loci. An INGK loop was put into the MrβHm peptide by Searle et al (Maynard et al. 1998) because a statistical analysis of the PDB determined that NG is a very common loop sequence in hairpin forming sequences (Sibanda and Thornton 1991) and that beta branched residues and charged residues are often found at the S-1 and S+1 positions, respectively, of an NG turn (Ramírez-Alvarado et al. 1996). Turning to [3:5] turns, extensive studies of sequences forming this type of turn (Blanco et al. 1993; Searle et al. 1995) were performed, leading to the PDG sequence which has the highest propensity for hairpin formation of the sequences examined. The literature available about designing β hairpins with longer loops was, however, less well developed. This prompted us to look at longer loops to insert into the hairpin sequences we had already

developed. Since hairpin systems based on GB1p and MrβHm had been explored, resulting in strands that form stable hairpins, a context for determining the ability of various turn sequences to provide for the association of strands in two different systems was present. These two hairpin models are of particular interest for turn sequence effects because they contain hydrophobic clustering on opposite faces of the hairpin: GB1p has hydrophobic residues at the S±odd sites while MrβHm has an S±even cluster. This leads to the hydrophobic association of residues on the opposite faces of the hairpins.

Because of the substantial fold populations afforded by the strands of MrH3 peptides, as well as their relative stability in solution, these strands were chosen for a turn mutational study. The purpose of the mutations is to spectroscopically determine whether the length of turn has effects on the register of the strands, as well as to determine the relative turn propensity of the loop sequences within this strand context.

Table 3-5 MrH3 Turn Mutations

MrH3a†	KKYTVS	INGK	KITVSA
MrH3b†	KKYTVS	IpGK	KITVSA
MrH6a	KKYTVS	NPDGT	KITVSA
MrH5a	KKYTVS	DPATGR	KITVSA
MrH5b	KKYTVS	HPATGR	KITVSA

†Made and characterized by coworker Matthew Fesinmeyer.

Table 3-6 HP7 Turn Mutants

HP7†	KTW	NPATGK	WTE
HP7 INGK	KTW	INGK	WTE
HP7 NPDGT	KTW	NPDGT	WTE
HP7 EPDGK	KTW	EPDGK	WTE

†Made and characterized by fellow coworker Katherine Olsen.

The mutations made are given in Table 3-5. These show that two, three and four-residue loops were inserted between these strands. The other strands that were chosen for loop effects studies came from the GB1p based HP7 series. The sequences made are given in Table 3-6. Here, the original sequence had a [4:4]/[4:6] loop sequence so shorter loops were inserted. The H_N CSDs of the strands of these peptides in aqueous media are given in Figure 3-7; panel (a) shows MrH3 peptides while panel (b) gives HP7 CSDs. The length of the strands of the MrH3 peptides is much longer than those of the HP7 series and will be discussed first. While tight, two-residue loops lead to the highest hairpin populations for these strands, the longer loops afford measurable hairpin populations. For these strands, a DPATGR loop (MrH5a) sequence leads to a hairpin population of about 60% of the INGK (MrH3a) loop peptide. Moving to the [3:5] loop (MrH6a), the hairpin population is reduced to perhaps half of that populated by the INGK peptide. The HPATGR (MrH5b) is further destabilizing and shows that His in the S-1 position is less favorable than Asp. In all of these peptides, though they have reduced populations relative to the INGK peptide, the residues with the largest H_N CSDs are still S±odd, leading to the conclusion that the loop sequence does not affect strand register. Similar conclusions were reached from an examination of Hα CSD plots (data not shown).

Turning to the HP7 series, a different trend is observed. Here we see that the loop sequence that leads to the greatest hairpin propensity is the NPATGK (HP7) turn. The shorter loops also provide for measurable hairpin populations but are somewhat reduced relative to HP7. The two residue turn NG leads to a ≈25% loss of hairpin population relative to HP7 while the better of the [3:5] loops (NPDGT) leads to a ≈65% decrease. The relative magnitude of CSDs within the same sequence are essentially unchanged leading to the conclusion that loop length in this hairpin construct also does not effect shift register. The results from both hairpin constructs lead to the conclusion that strands do have loop length preference though the loop sequence affects only hairpin population and does not shift the register of well-designed hairpin strands. It

can also be proposed that strands with hydrophobic residues in the S±odd positions prefer short, two-residue loop sequences while strands with hydrophobic residues at S±even positions prefer four residue loops. Many of the peptides designed for this study led to the discovery that reversing loops of different lengths have a diagnostic CSD signature. This data is presented below.

CSD Characterization of Turns in β Hairpins

Having multiple turn sequences that form the same turn type, as well as the presence of the same types of turns in different strands lends confidence that the spectroscopic patterns seen in CSD plots are indeed conserved across most members of a class of turns regardless of their strand environment. A table of sequences that demonstrate the norms of each turn type has been compiled and can be seen below (Table 3-9). As previously indicated, the location of the turn residues are generally identified by the disruption of the shift register. The H_N CSDs from the turn forming regions, as well as the turn flanking residues (S±1) of many of the peptides presented in Table 3-9 (residues shown in the center column), are graphed versus their position in the turn or strand in Figure 3-8. The X-axis of this figure is labeled with the nomenclature that has been used throughout and each turn type will be shown to have diagnostic CSDs, allowing for convenient identification. It is worthy of note that due to the length of some loops, all turn sites are not occupied (such as [2:2] turns have no T3 and T4 residues) which results in the alignment of data seen in Figure 3-8. All turn types studied here have a common, negative, S+1 CSD which heralds the end of the turn and the start of the second strand. Also, many of the peptides, regardless of turn type, have a large positive S-1 CSD because this is the first position (counting away from the turn) that forms a cross-strand hydrogen bond. Those S-1 sites that do not display a large positive CSD are preceded by an aromatic residue, as can be seen by cross reference with Table 3-9. Looking at the T1 site, the [2:2]/[2:4] turn plotted has the large downfield CSD that is typical of these turn

Table 3-7 Sequences of Select Peptides Containing Different Loop Lengths

[2:2]/[2:4] Peptides			
MrH3a‡	KKYTVS	INGK	KITVSA
MrH4a‡	KKLTVS	INGK	KITVSA
HP7 INGK	KTW	INGK	WTE
[3:5] Peptides			
MrH6a	KKYTVS	NPDGT	KITVSA
UTH3†	SEIYS	NPDGT	WTVTE
H7 EPDGK	KTW	EPDGK	WTE
[4:4]/[4:6] Peptides			
MrH5a	KKYTVS	DPATGR	KITVSA
HP7 P5A*	KTW	NAATGK	WTE
GB1 m3‡	KKWTY	NPATGK	FTVQE

† Literature sequence (de Alba et al. 1999a)
‡ Made and characterized by colleague R. Matthew Fesinmeyer
* Made and characterized by colleague Katherine A. Olsen

types. All examples of [3:5] turns examined have a proline at this location, making analysis impossible but one [4:4]/[4:6] turn was successfully made that has an alanine at this position. The CSD at this site was small and positive but the relative lack of other examples containing amide resonances for this position provides uncertainty about whether this would be typically observed. The T2 position of the β-turns distinguishes the [2:2]/[2:4] type from the others: the two residue turns have a small positive CSD here while the three and four residue loops have significant negative deviations. The T3 position provides the differentiation between three and four residue loops because the four residue loops have large negative CSDs here while the Gly-bulge turns have negligible CSDs. In this manner, the turn type that is present in the peptide can be determined.

Conclusion

These studies of hairpin strand and loop sequences have increased the general knowledge about the interactions that are important to strand formation and association. Additionally, they have significantly increased the state of knowledge regarding the effects of different turn types on the ability of hairpin peptides to fold. Also, they have shown that certain turn types can be disadvantageous to the folding of particular systems. Finally, these studies have demonstrated that turn type in hairpin peptides can be determined from their H_N NMR chemical shift deviations rather than NOEs. Throughout, H_N CSDs provide the most reliable measure of fold population independent of whether cross-strand H-bonded strand sites or turn sites are employed.

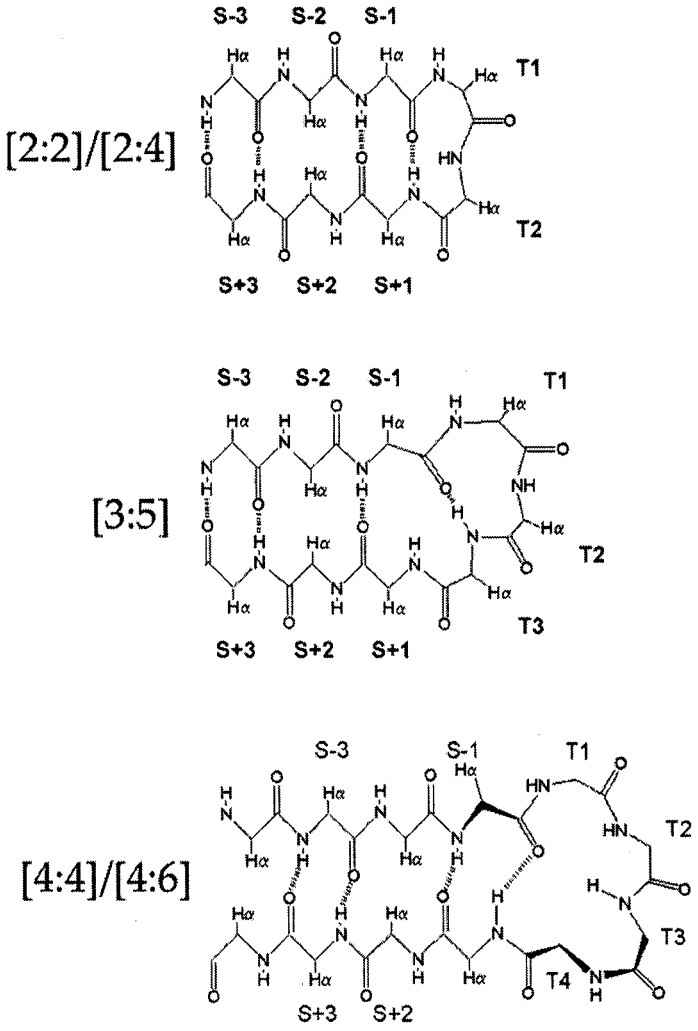


Figure 3-1: Diagrams showing strand orientation and nomenclature in β hairpins.

The nomenclature of each strand and turn site is given at that respective location. Hydrogen bonds shown in red indicate which are necessary for the formation of the shorter turn indicated.

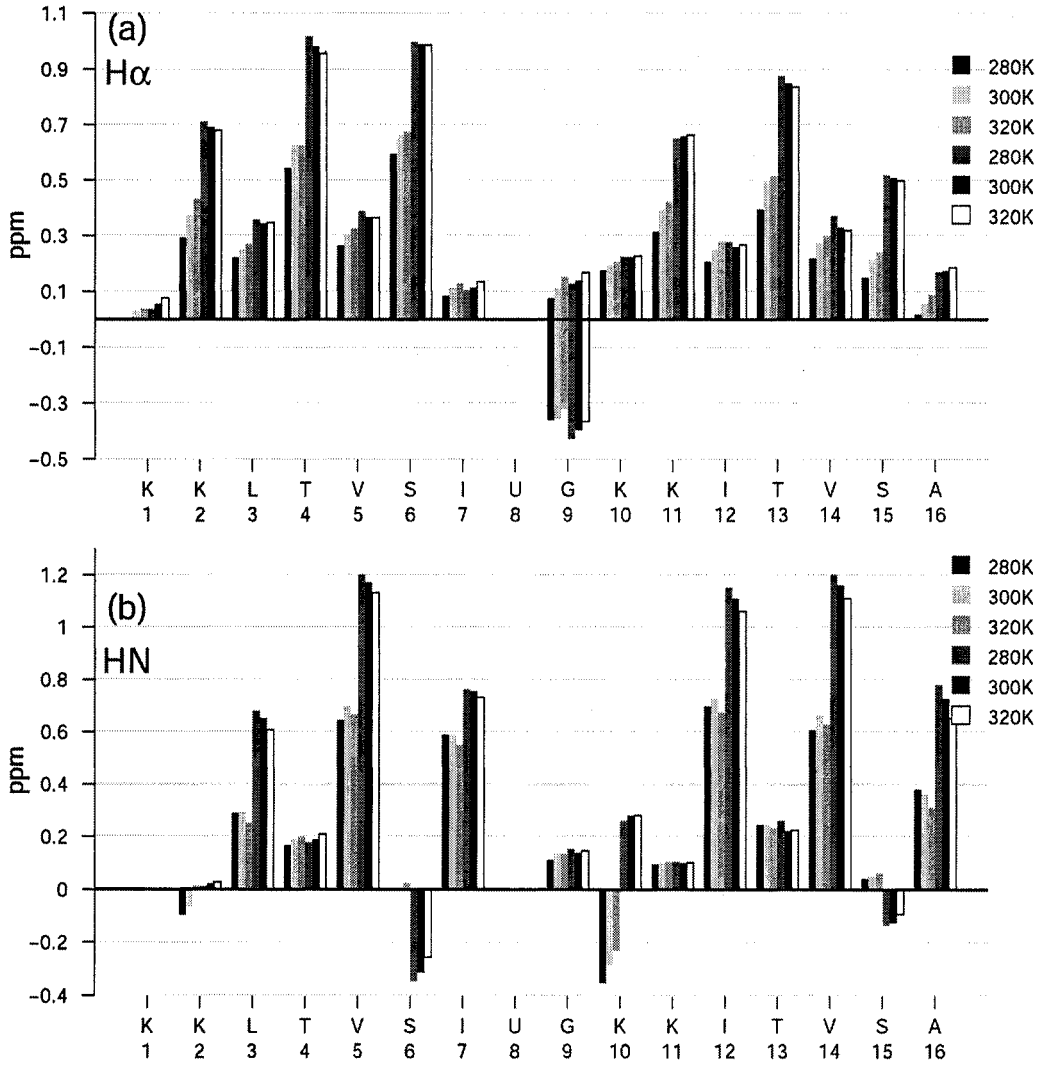


Figure 3-2: CSDs of a well-folded β hairpin.

The Hα (a) and H_N (b) CSDs for MrH4e in aqueous and fluoroalcohol (larger CSDs) containing media. This data demonstrates the shift register of a well-folded β hairpin and, in fluoroalcohol, provides 100% folded expectation CSD values for cross-strand directed protons.

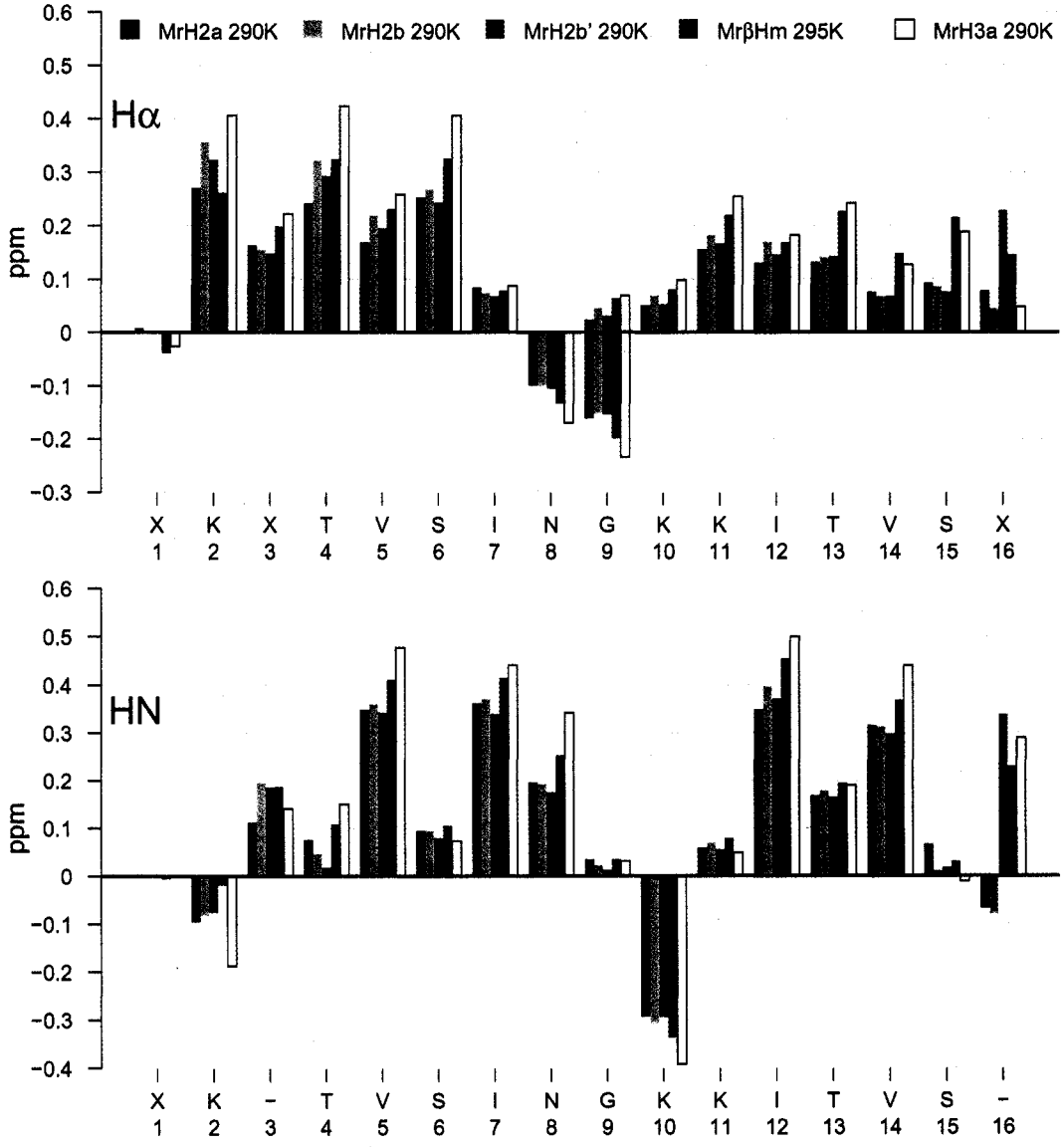


Figure 3-3: Aqueous CSDs for MrH2x peptides.
The Hα (a) and H_N (b) CSDs for MrH2x peptides, plotted versus MrβHm and MrH3a for reference.

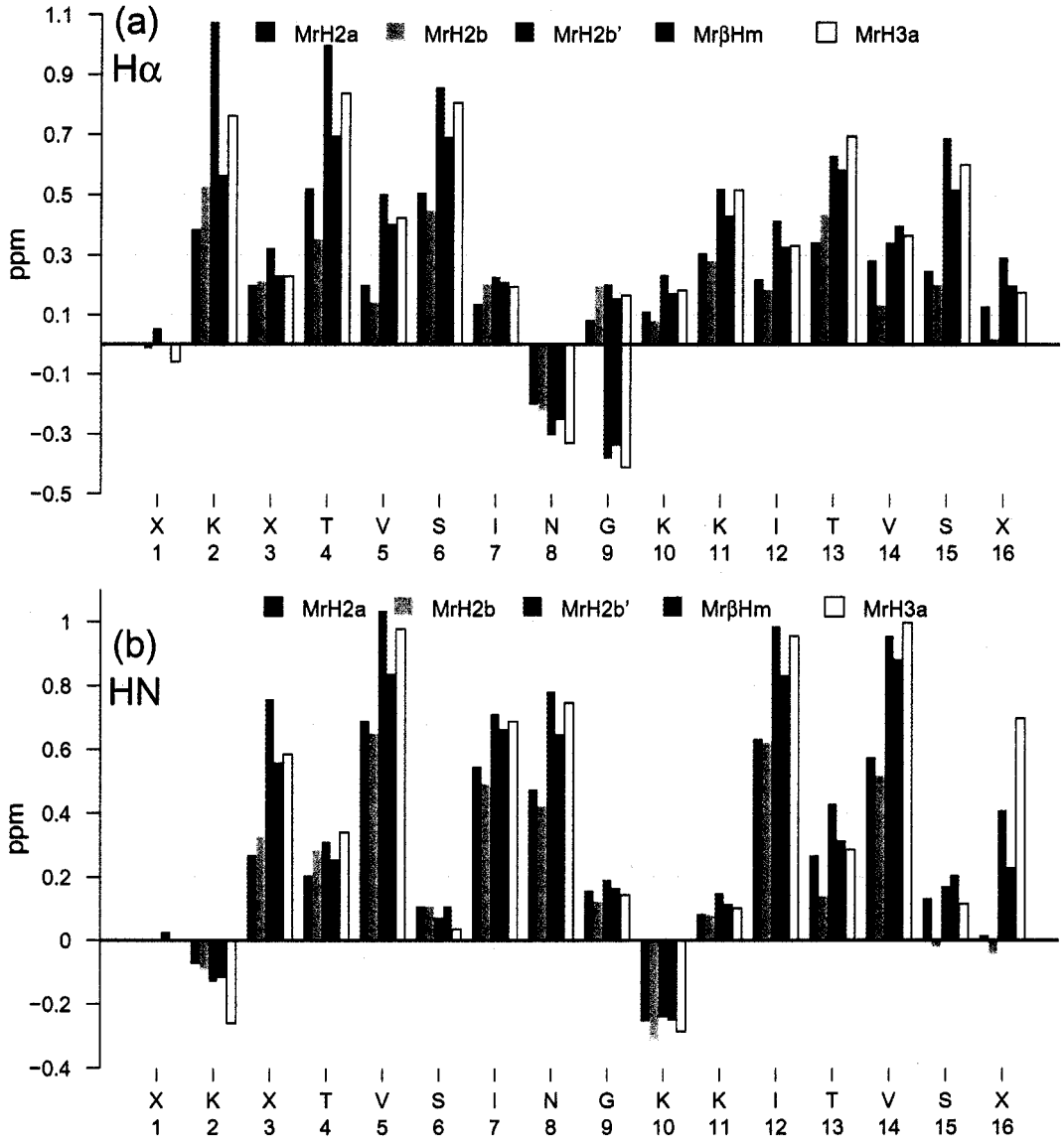


Figure 3-4: CSDs for MrH2x peptides in 8%HFIP
The Hα (a) and H_N (b) CSDs for MrH2x peptides in 8%HFIP at 290K.

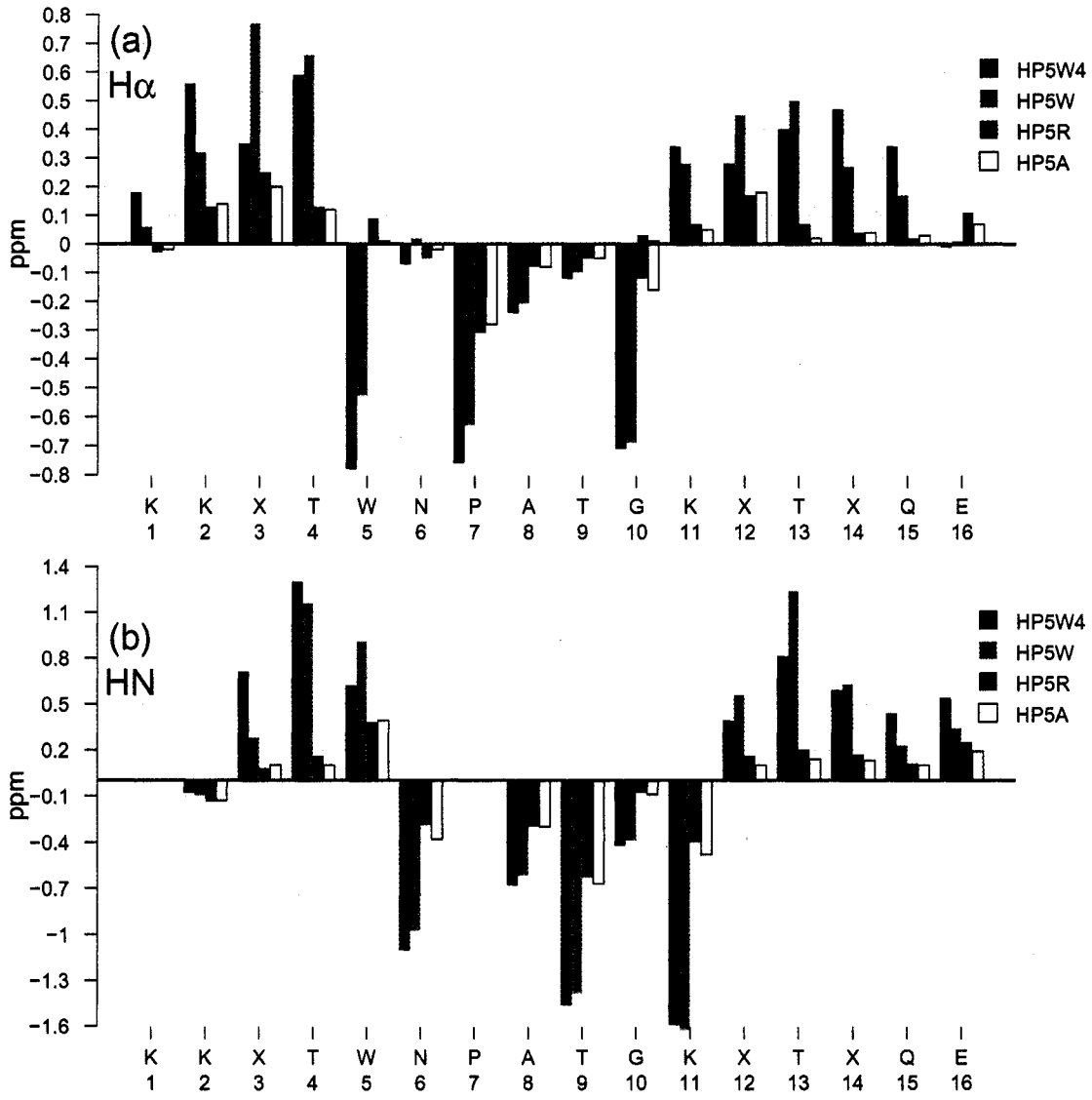


Figure 3-5: CSDs of GB1p inspired, Trp-zip peptides.
The Hα (a) and H_N (b) CSDs show the effects of strand mutations on the stability of Trp-zip peptides. All data is shown in aqueous solution at 290K.

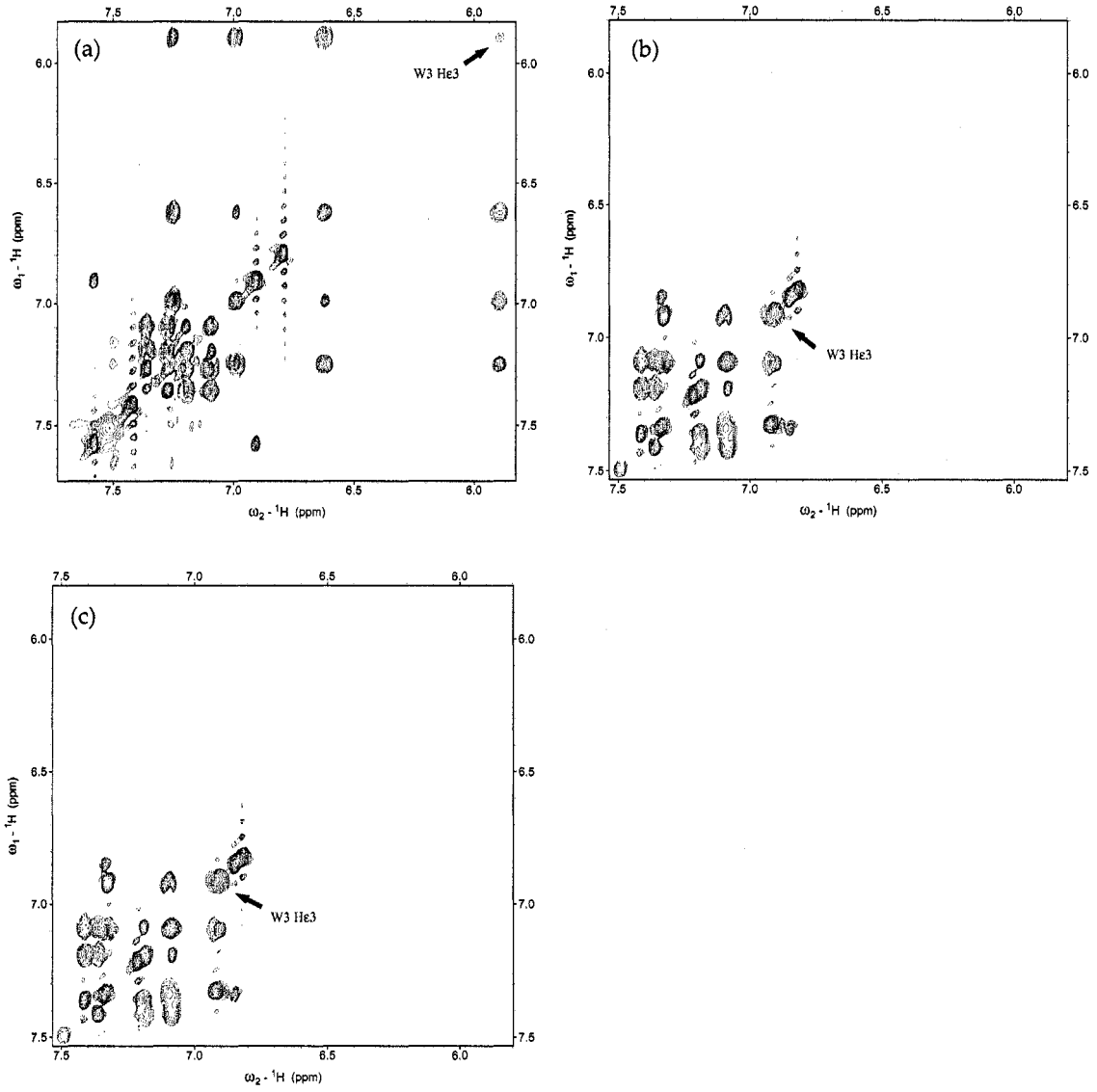


Figure 3-6: Aromatic region of ^{1}HNMR TOCSY spectra of HP7 peptides.
Aromatic region of TOCSY spectra of HP7 INGK (a), HP7 NPDGT (b), and HP7 EPDGK (c) with W3 Hε3 proton resonance indicated. Panel a demonstrates the NMR spectral signature of the edge-to-face Trp/Trp interaction.

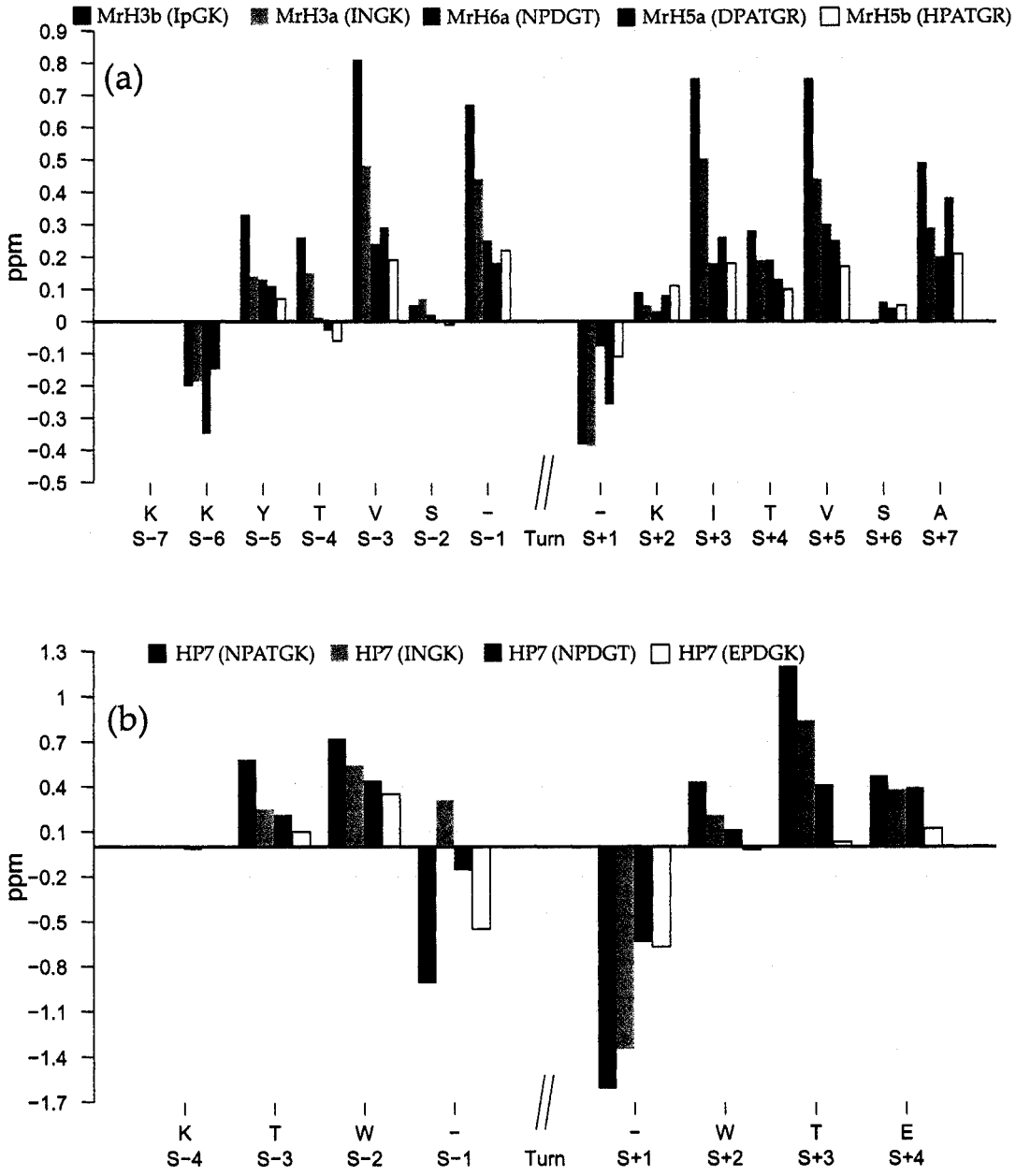


Figure 3-7: H_N CSDs of conserved strand turn mutants.
(a) CSDs of peptides containing the strands of MrH3 peptides and different turns.
(b) CSDs of HP7 peptides with different turn types.
All data shown was acquired in aqueous buffer at 290K.

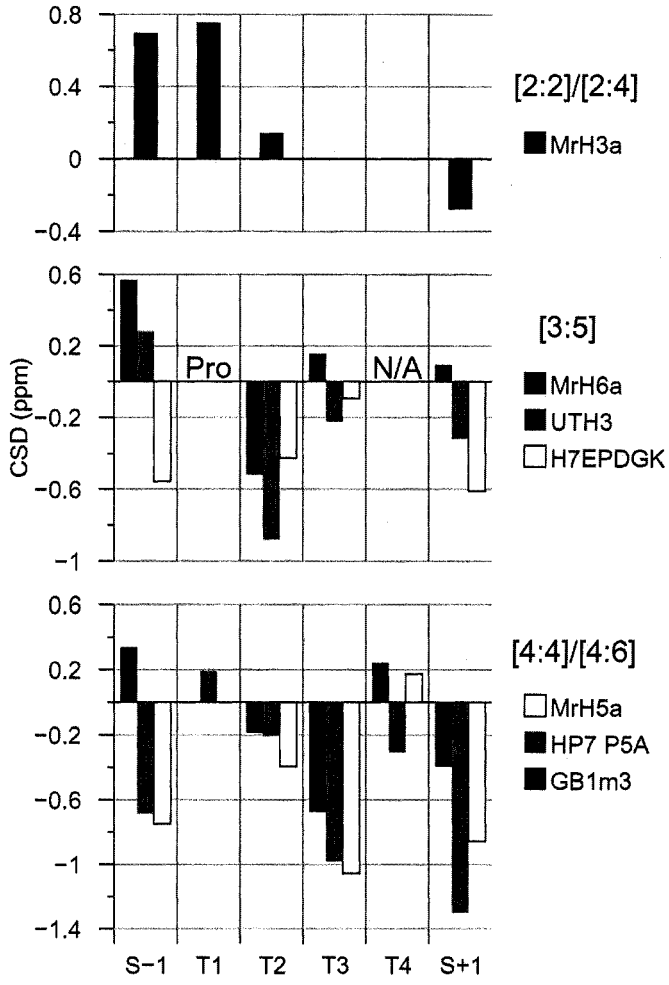


Figure 3-8: H_N CSDs from turn forming region of hairpin peptides.
All data shown at 290K. Panels indicate turn length and include S±1 positions for clarity.

Chapter 4 Cage and Half-Cage Models

Introduction

Exenatide, synthetic exendin-4, is the focus of this study of Trp-cage forming peptides. It is the first member of the incretin mimetic class of potential therapeutic agents and has been the subject of extensive clinical trials in people with type 2 diabetes. Results to date indicate that exenatide decreases post-meal blood glucose concentrations and that this effect is associated with weight loss. Exendin-4 (EX4) was first isolated and characterized from the salivary secretions of Gila monsters (Eng et al. 1992). This 39 amino acid peptide is 53% homologous with Glucagon-like Peptide-1 (GLP-1) and has been shown to bind and be a fully potent agonist at the GLP-1 receptor (Göke et al. 1993), making it of considerable interest as a therapeutic agent. There have been a number of spectroscopic comparisons of GLP-1 and EX4 (Neidigh et al. 2001; Andersen et al. 2002b); however only those in media containing organic co-solvents or added lipid micelles have yielded structures for a monomeric state (Neidigh et al. 2001). These and similar studies of a truncated exendin analog (TrEX4) (Neidigh and Andersen 2002) indicate that the structuring propensity of this sequence is notably medium-dependent. The addition of a proline-rich C-terminal extension in EX4 versus GLP-1 results in the formation of a novel tertiary structural motif, first characterized in the NMR structural ensemble for EX4 in 30 vol-% trifluoroethanol (TFE), which has been designated as the Trp-cage fold (Neidigh et al. 2001). The Trp-cage fold is characterized by a collapse of hydrophobic residues around a tryptophan residue (Figure 4-1) and is formed when a Trp sidechain is surrounded by: a poly proline segement, a 3_{10}-helical segment, and an α-helical segement. Becauase the trp-cage acts as a helix C-cap and the helix makes up the N-terminal portion of the structure, the length of the helix can be varied. Key

interactions with the indole ring include glycine Hα and proline Hα and δ protons (G30, P37α and P38δ in Table4-1) which allow for spectroscopic characterization by ^{1}H-NMR. These close associations lead to large CSDs due to ring currents and are used to diagnose Trp-cage formation. Prior to these studies, it was not known whether EX4 forms a Trp-cage in strictly aqueous media but it has been shown that this feature is retained in TrEX4 in the SDS micelle-associated, but not in the DPC micelle-associated, state (Neidigh and Andersen 2002).

The characterization of EX4 in strictly aqueous medium has been hindered by aggregation in the past, but further truncation and mutational optimization has afforded water soluble 20-mers (for example, TC5b) that display the same Trp-cage fold in both water and 30 vol-% TFE (Neidigh et al. 2002). However, equally soluble 20-mers which bear a closer sequence similarity to EX4 (such as TC3a) are only 40% folded in water. The study of TrEX4 concluded that the Trp-cage portion of the peptide is destabilized by association with DPC micelles due to an energetically favorable interaction between the tryptophan residue and the phosphocholine head groups. This interaction prevents the association of Pro^{37} with Trp^{25}, a key interaction in the Trp-cage. The sequences (and residue numbering) of these peptides are shown below.

Table 4-1 Sequences of GLP-1 and Trp-cage forming peptides

	1	11	21	31
GLP-1	**HAEGTFTSDV**	**SSYLEGQAAK**	**EFIAWLVKGR**-NH_2	
EX4	**HGEGTFTSDL**	**SKQMEEEAVR**	**LFIEWLKNGG**	**PSSGAPPPS**-NH_2
TrEX4		**Ac-SEDEAVR**	**LFIEWLKNGG**	**PSSGAPPPS**-NH_2
Trp-cage (TC5b)			**NLFIQWLKDGG**	**PSSGRPPPS**
(TC3a)			**DLFIEWLKNGG**	**PSSGAPPPS**
(TC10b P18L)			**DAYAQWLKDGG**	**PSSGRPLPS**

Characterization of Trp-Cage Peptides

The prior studies of exendin-4 and Trp-cage constructs have provided a wealth of CD and NMR diagnostics of structuring. These serve, in the present study, as expectation values for characterizing structuring preferences that are not as complete as those observed for analogs or for EX4 in more solubilizing and structuring media. In 30% TFE, the ^{1}H-NMR spectrum for exendin-4 includes a large number of significant chemical shift deviations (CSDs) for both backbone and side chain resonances, as well as long-range NOE's which are diagnostic of the formation of the Trp-cage fold. The NMR derived structure ensemble generated from this data showed that Ex4 consists of a helix spanning from Ser^{11} to Asn^{28} as well as a Trp-cage fold at the C-terminus. The same Trp-cage was observed for the truncated analog, TC5b, in water and the complete structure of this construct is illustrated in Figure 4-1 (Neidigh et al. 2002). The specific CSD diagnostics of the Trp-cage (as observed in EX4 in 30% TFE at 280K) (Neidigh et al. 2001) are: the Hα signals of Leu^{26} (0.65 ppm further upfield than expected for a helical state), Gly^{30} (CSDs = –3.02 & –1.44 ppm) and Pro^{37} (CSD = –1.82 ppm), the β3 signal of Pro^{37} (CSD = –1.52 ppm), and both δ signals of Pro^{38} (CSD = –0.56 ppm). These sites are shifted by the ring current of the indole ring and the CSDs are nearly the same for TC5b. The specific conformations of the Phe^{22} and Trp^{25} sidechains also produce upfield shifts for specific aryl-^{1}H signals of the indole ring.

The structure of TC5b is well characterized and is identical in both water and 30% TFE. NH protection factors indicate a 98% fold population in water at pH 6.6. The circular dichroic (CD) spectrum of TC5b is also identical in water and 30% TFE. Analysis of this CD spectrum using the magnitude of the 222nm minimum would imply a greater net helicity than that suggested by the fraction of the sequence that is helical based on the NMR structure. This has been rationalized as the consequence of a large negative contribution from the indole ring of Trp^{25}. The indole ring chromophore normally appears at 226 nm resulting in a red shift of the helix backbone amide n $\rightarrow \pi^*$ band to

223-4 nm, rather than the typical 220 –222 nm value. In the studies of TrEX4 (Neidigh and Andersen 2002), this same red shift, and a greater intensity of the 223-4 nm band relative to the 208 nm minimum (R2 ratio (Bruch et al. 1991; Muñoz and Serrano 1995)) was observed in 30% TFE and in water containing SDS micelles, but not in the DPC micelle-associated state. The only prior CD studies of EX4 in water (at pH 3.5 and 5.9) (Andersen et al. 2002b) found that the molar ellipticity at circa 221 nm, diagnostic of both net helicity and Trp-cage formation, was highly concentration dependent. In 40% TFE, the $-[\theta]_{221}$ value was as large as 23,700° at 21°C (Andersen et al. 2002b). The values reported for EX4 at low pH in water changed from 11,370° to 17,600° on increasing the concentration from 2.3 to 228 μM. This change was tentatively attributed to the formation of helical oligomers at the higher concentration. It is important to determine whether these apparent aggregation effects are also observed with exenatide because slow aggregation/disaggregation phenomena could influence both the bioavailability and long-term stability of a therapeutic peptide formulation.

Characterization of Exenatide

Since the current studies of exenatide are carried out at pH 4 – 4.4 (to mimic therapeutic formulation conditions) and the prior NMR structure elucidation for EX4 and TC5b utilized buffers at pH 5.9 – 6.6, I first examined exenatide in 20mM acetate buffer (pH 4.4) with 30% TFE to determine if there were any discernable spectroscopic differences between the two conditions. The spectra (one example appears as the right panel of Figure 4-3) were readily assigned and the chemical shifts indicated no change in structure associated with the pH change. For example, for Gly^{30}-Hα2, Pro^{37}-Hα,Hβ3, and Pro^{38}-Hδδ′ the CSDs at pH 4.4 and 298K were –2.85, –1.72, –1.30, and –0.55, (versus –3.02, -1.82, -1.52 and –0.56 for pD = 6.6 at 280K). These correspond to 90±4 % Trp-cage formation and indicate partial melting upon warming. A complete Hα CSD comparison

(pH 5.9 versus 4.4) at 298 K appears in Figure 4-2. The only differences are at, or one helix turn away from, the ionizing Glu residues.

^{1}HNMR of Exenatide in Aqueous Environments

One-dimensional spectra, as well as 2D NOESY and TOCSY spectra were recorded for exenatide in 20 mM acetate buffer (pH 4.4) at 300 μM, 1.5 and 5 mM peptide concentrations and in 0.1% HOAc at 1 and 7 mM concentrations without the addition of co-solvents. All of these spectra were essentially identical with, in all cases, the signals for the central portion of the sequence (residues 12 – 26) not observable. A comparison of the H_N to Hα and Hβ TOCSY spectra observed in pH 4.4 buffer with and without the addition of TFE is shown in Figure 4-3. This illustrates the relative paucity of data observed in the absence of fluoroalcohol cosolvents; this is also observed in the upfield region of the spectrum. Figure 4-4 shows the upfield NOESY spectrum recorded for 1.5 mM exenatide in pH 4.4 acetate buffer. Complete resonance assignments could be made for residues 1 – 9 (with a partial assignment for Leu^{10}) and residues 33 – 39. The lysine sidechains (residues 12 and 27) and at least one additional Glu sidechain (presumably in the residue 15 – 17 span) are broadened but still discernable. All of the definable proline resonances appear at their random coil values. The Pro-Hδ signals all appear at ≈3.62 and 3.80 ppm while the Pro-Hα signals appear at 4.44 (**Pro**-Xaa) and 4.69 (**Pro**-Pro). The NOESY connectivities establish that the circa 4.69 ppm line includes both Pro^{36} and Pro^{37}. There is no definitive evidence establishing that Pro^{31} contributes to the observed proline signals. Its Hα signal could be at 4.44 ppm where Pro^{38} is observed but, integration of the 3.78 ppm Pro-Hδ3 signal suggests that only 3 (not 4) protons are present at this position in the 1D spectrum.

^{1}HNMR of Exenatide in Water/DMSO Mixtures

NMR spectra of exenatide were also acquired at a range of DMSO/H_2O solvent ratios (98, 90, 75, 60 and 25% DMSO by volume). Unlike purely aqueous media, exenatide in 98% DMSO affords sharp, yet severely overlapped spectra, yielding an essentially complete assignment. The spectra remained assignable as water was added only to the 40:60 H_2O:DMSO point of the solvent titration. Thereafter, the spectra resemble those collected in purely aqueous buffers with regard to both line broadening and the disappearance of signals for the central portion of the sequences from the 2D spectra. While there were no significant structuring shifts near the C-terminus (specifically for residues 30 – 39) in 98% DMSO, non-random shifts do appear as water is added. This data lends itself to CSD analysis of α protons for most positions even though in Trp-cage species, the structuring shifts at $Pro^{31,38}$ are larger at the Hδ resonances. Figure 4-5 shows the CSDs for exenatide in 98 and 60% DMSO with the data for pH 4.4 acetate buffer (with 30% TFE added) and 30% aqueous glycol solution (pH 3.5 at 320K) added for comparison. The latter values were taken from Neidigh et al 2002. The scale change for residues 4 – 27 versus 30 – 39 is necessary to accommodate large ring-current shifts in the C-terminus associated with Trp-cage formation.

CD Melting and Concentration Jump Studies

In contrast to previous samples of EX4 (Andersen et al. 2002b), lots of exenatide consistently displayed helical CD spectra that were not notably concentration dependent and the solutions prepared displayed the same spectrum for many days. The data are summarized in Table I, which also contains the literature values (Neidigh and Andersen 2002) for TrEX4. For each medium, the [θ] value at the minimum near 222 nm and the diagnostic intensity ratios (R2 = $[\theta]_{222}/[\theta]_{208}$ and R1 = $[\theta]_{191}/[\theta]_{208}$) (Bruch et al. 1991; Muñoz and Serrano 1995) are given for the low temperature observation and near room temperature. The intensity loss at 222 nm associated with dilution from circa 240 μM to

8 μM in acetate buffer was only 11 % in this study; in earlier studies with less pure EX4, a circa 35 % intensity loss was typically observed (Andersen et al. 2002b). The presence of impurities, or of other counterions, in the earlier samples may have encouraged the formation of non-helical aggregates at the lower concentration.

CD spectra recorded in 0.1% aqueous acetic acid displayed better signal to noise ratios and were fully reproducible. Typical CD spectra observed at 2 °C over the concentration range 2 – 600 μM are collected in Figure 4-6. These displayed even less concentration dependence with regard to the intensity of the circa 222 nm minimum but the low versus high concentration data could be differentiated by the R2 ratio. At low concentrations, the R2 value is larger and the apparent $n \rightarrow \pi^*$ band is consistently red-shifted by about 2 nm. The larger R1 values at the lower concentrations suggest that the extent of helicity in the sequence is greater under these conditions. Even larger differences are apparent in the CD melts. The full set of CD curves for melts at 2 and 240 μM are shown in Figure 4-7. Similarly precise and reproducible data was obtained at 40 μM in the 0.1% acetic acid medium. Comparable melts in 20 mM pH 4.4 acetate buffer provided spectra with more noise (due to increased absorbance) but still yielded usable melting curves.

Table 4-2 CD spectral data recorded for Exenatide and TrEX4 in various media.

Conditions	Peptide Conc. (μM)	Temp. (°C)	$-[\theta]_{221-3}$	R1†	R2‡
Aqueous Buffer pH 4.4	8	5	16,980	-2.11	0.94
Aqueous Buffer pH 4.4	8	25	15,200	-1.87	0.82
Aqueous Buffer pH 4.4	240	5	19,130	-1.94	0.83
Aqueous Buffer pH 4.4	240	25	17,600	-1.83	0.80
0.1% Acetic Acid	2	2	18,370	-2.06	1.01
0.1% Acetic Acid	2	20	15,500	-1.83	0.90
0.1% Acetic Acid	40	2	17,730	-1.85	0.89
0.1% Acetic Acid	40	20	16,050	-1.72	0.81
0.1% Acetic Acid	240	2	19,050	-1.81	0.88
0.1% Acetic Acid	240	20	17,650	-1.67	0.84
0.1% Acetic Acid	600	2	17,800	—	0.83
0.1% Acetic Acid	600	20	16,900	—	0.80
30% TFE/pH 4.4 Buffer	8 – 240	5	25,800	-2.27	1.065
30% TFE/pH 4.4 Buffer	8 – 240	25	20,530	-2.15	0.98
Prior Data for TrEX4*					
pH 7 water	10.7	25	18,370	-1.47	0.83
+ 30% TFE	10.7	25	22,320	-1.77	0.92
+ SDS micelles	10.7	25	17,950	-1.68	0.93
+ DPC micelles	10.7	25	18,800	-1.40	0.75

†$R1=[\theta]_{191}/[\theta]_{208}$(Bruch et al. 1991; Muñoz and Serrano 1995)
‡$R2=[\theta]_{222}/[\theta]_{208}$(Bruch et al. 1991; Muñoz and Serrano 1995)
*Literature data shown for comparison

Structure of Exenatide and TrEx4

This study indicates exenatide has a strong helical preference for the residue 9 – 27 span; an observation that extends to 98% DMSO, a solvent system typically viewed as a strongly denaturing, and can be readily seen in the Hα CSD profile in Figure 4-5. This stands as a rare instance of the use of CSDs in DMSO to detect a secondary structure preference. The conclusions regarding helicity in strictly aqueous media are based exclusively on CD data, due to the lack of observable NMR shifts. The residue selective broadening observed in NMR spectra collected in aqueous media implies conformer or state exchange phenomena that are occurring in the intermediate time regime. The significant concentration dependence of the CD melting data (Figure 4-7) for exenatide in aqueous media can only be explained by concentration dependent changes in conformation or association state. A conformational and aggregation state equilibrium model that can be used to explain this data appears in Figure 4-8. The monomeric species in this diagram are based on prior studies of Trp cage peptides (Barua and Andersen 2002; Neidigh and Andersen 2002; Neidigh et al. 2002).

In the case of N-terminally truncated exendin-4 analogs, the helix is not significantly populated unless it is capped by either a Trp^{25}/Pro^{31} hydrophobic staple or the complete formation of the Trp-cage. Complete Trp-cage formation serves as a very effective helix C-cap but there is also evidence for the contribution of the half-cage structure in partially melted Trp-cage species. Half-cage structures are those with the $Pro^{36,37,38}$ unit undocked. Trp-cage and half-cage structures have different diagnostic chemical shifts; only the complete Trp-cage structure produces upfield ring current shifts for Pro^{37}-Hα and Pro^{38}-Hδ2. In half-cage structures, the Pro^{31}-Hδ2 is only slightly shielded. Gly^{30}-Hα2 is shifted more than 3 ppm upfield shift in the full cage the two Hα protons have a large diastereotopic shift difference ($\Delta\delta_{\alpha 2-\alpha 3}$ = 1.6 – 2.5 ppm). In the half-cage structure, Pro^{31}-Hδ2 is shifted significantly upfield and both α protons of Gly^{30} show comparable but smaller shifts (circa 1.2 ppm upfield). Assuming the same extent of helix formation for

the two motifs, CD spectral differences for the capping interactions shown in Figure 4-8 are less distinct. The clearest difference is the red shift of the low energy band; R values will be used to illustrate this point. In the case of TrEX4, since Trp-cage formation is favored in SDS micelles and helix formation is favored in DPC micelles, a direct comparison of R values for the various structural moieties is possible. The R2 values are very different for Trp-cage forming species and non-cage forming species: 0.93 for the Trp-cage and only 0.75 for the comparably helical DPC state. The red shift of the low energy minimum is well illustrated by this example and appears at 221.9 nm in DPC micelle system, but at 224 nm in the presence of SDS micelles.

The time scales for the interconversion of the monomeric species in Figure 4-8 have also been investigated for the truncated analogs. For a wide variety of Trp-cage species, dynamic NMR studies indicate that none of the monomeric species display lifetimes in excess of 25 μs. As a result, these equilibria cannot contribute to the extensive line broadening observed in aqueous media. These line broadening phenomena and the concentration dependent phenomena observed by CD (see Figure 4-7) reflect formation of the oligomeric species, also shown in Figure 4-8.

Turning to exenatide, helical shifts for the α protons of residues 9 – 27 in 98% DMSO are observed, while the resonances from Gly^{30} to the C-terminus display no significant CSDs (see Figure 4-5), implying that helix formation can occur in exendin-4 without any C-capping interaction. This difference relative to Trp-cage 20-mers and TrEX4 likely reflects the additional length of the helical span, reflecting the fact that 8 – 13 residue helices require specific capping interactions for stabilization while the 19-residue span of the exendin-4 helix appears sufficiently stable for helix formation even in DMSO. Figure 4-5 provides evidence for additional structuring occurring upon titration with water. At 40% water, the upfield shifts at $Pro^{37,38}$ indicate some Trp-cage formation has occurred. The larger shifts at Pro^{31}-Hδ2 (rather than Pro^{38}-Hδ2) and the similar shifts for the

diastereotopic Hα's at Gly^{30}-Hδ2 indicate that the half-cage structure contributes to greater extent than the full-cage. Further water addition yields spectra in which essentially all resonances in the Gln^{13} to Leu^{26} span are broadened so extensively that 2-dimensional NMR analysis is impossible. Oligomerization resulting in the burial of the hydrophobic side of the helix over this span in interfacial contacts provides the simplest rationale for this observation. This is bolstered by the observation of resonances from glutamate sidechains which are apparently still solubilized. Trp-cage formation apparently does not provide as much hydrophobic surface burial as oligomerization.

The remaining questions are: 1) does the Trp-cage fold contribute to the exenatide ensemble observed in strictly aqueous conditions, 2) what is the nature of the oligomeric species formed, 3) what is the critical concentration required for oligomer formation and 4) is there a concentration dependence to the oligomerization state above the oligomerization threshold? The NMR data collected in aqueous media throughout the concentration range 300 μM to 7 mM clearly indicates that exchange broadening must involve oligomeric species. The 1D NMR spectra have identical linewidths indicating that the same species is/are dominant over this entire concentration range. In the 2D spectra, sharp signals are observed at random coil shifts for both the His^{1} – Ser^{8} and the Ser^{33} – Ser^{39} spans; these spans are therefore in rapid segmental motion which approximates random coil conditions meaning Trp-cage formation does not contribute under these conditions. Since the span form Asn^{28} to Pro^{31} cannot be assigned, the conformational features of this section cannot be determined by this technique. The CD melts shown in Figure 4-7 provide clear answers to both questions 1 and 2. For the helix $n\rightarrow\pi^*$ band a red shift is observed at 2 μM but not at 240 μM. The low concentration spectra also have a much larger R2 value which is quantified in Table 4-2. The red shift and the greater R2 value are explained by positive reinforcement of the $n\rightarrow\pi^*$ band due to the structured environment of the tryptophan sidechain chromophore in the Trp-cage. At 2 μM the melting behavior is monotonic and has a clear isodichroic, consistent with a

monomeric structure. This is strong evidence for a monomeric state with significant Trp-cage population.

In contrast, the 240 μM spectra do not display a clear isodichroic and melting occurs predominantly at higher temperatures. Sigmoidal melting is a diagnostic characteristic of the formation of coiled coils and other discreet helix bundle states in peptides. Figure 4-9 shows the melting behavior for exenatide under a variety of conditions and reinforces this observation. The melting behavior observed for exenatide in this figure provides an answer to question 3.

The monotonic melting behavior of the 2 and 8 μM solutions of exenatide clearly indicate that these conditions are favorable for monomer formation. 20 mM acetate buffer destabilizes the helical monomer and the quantity of helix bundle is relatively increased. The melting point of the oligomers appears to be over 80°C. Solutions were reexamined after minutes to weeks to determine the time dependence of helix bundle formation. The lack of change in CD signals of exenatide either at constant concentration or upon dilution suggests that all oligomer-monomer equilibration rates are relatively rapid and that even at quite high concentrations (up 600 μM) large aggregates that can associate to form particles or fibrils are not present. Also, there is no evidence of irreversible aggregation or the "higher aggregates" state postulated in Figure 4-8. The stability of NMR solutions extends these observations to the multi-mM concentration range which stands in contrast to the earlier data. This older data, acquired from samples of exendin-4 which were apparently less pure, suffered CD signal loss and eventual precipitation even at concentrations less than 1 mM . The extensive broadening of the NMR resonances in the Gln^{13} to Leu^{26} span implies that this is the region involved in the helix/helix interactions of the helix bundle state. Since such a structure would have multiple, different interactions, extensive broadening is expected. This does, however, lead to uncertainty about the structure of this state. The

helix bundle state could be fluxional or it could exist as multiple bundle states differing in either oligomeric state (dimer/trimer, etc.) or geometry (antiparallel *versus* parallel, or with different regions at the helix interface). Since chemical shift changes associated with the fluxionality of such species are quite small (< 0.3 ppm), the rates of conformer interconversion must be in the 20/sec to 20/min range.

Conclusion

This study has provided insights into the structure of exenatide in purely aqueous media. Though the data cannot provide a specific structure for the peptide at therapeutic formulation conditions, it clearly establishes the existence of a stable oligomeric state at higher concentrations. Also, it indicates that, at concentrations <10μM, the monomeric state is observable and that it is stabilized by Trp-cage formation. While it has not yet been determined what peptide conformation is necessary for the receptor interaction that leads to therapeutic activity, this study confirms that oligomers/monomer interconversion rates are rapid, indicating that the features required for the activity of the peptide will be available for presentation at the receptor site.

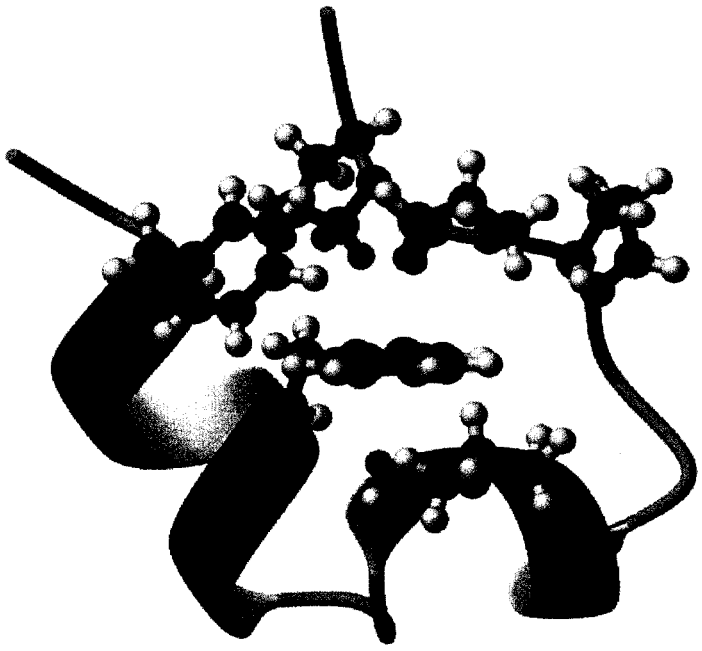


Figure 4-1: Structure of TC5b showing Trp-cage motif.
Protons with large upfield CSD's are shown in green.

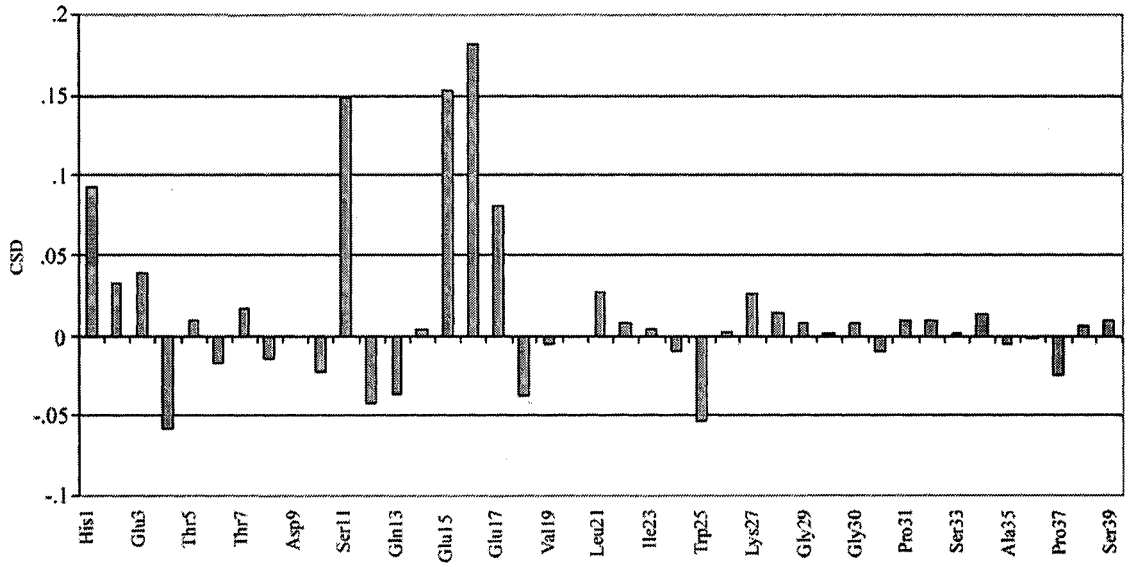


Figure 4-2: Difference CSD plot for Exendin-4 in 30% TFE at pH 5.9 and 4.4. Difference CSDs show the only changes in NMR spectral data are at or around the ionizing Glu residues.

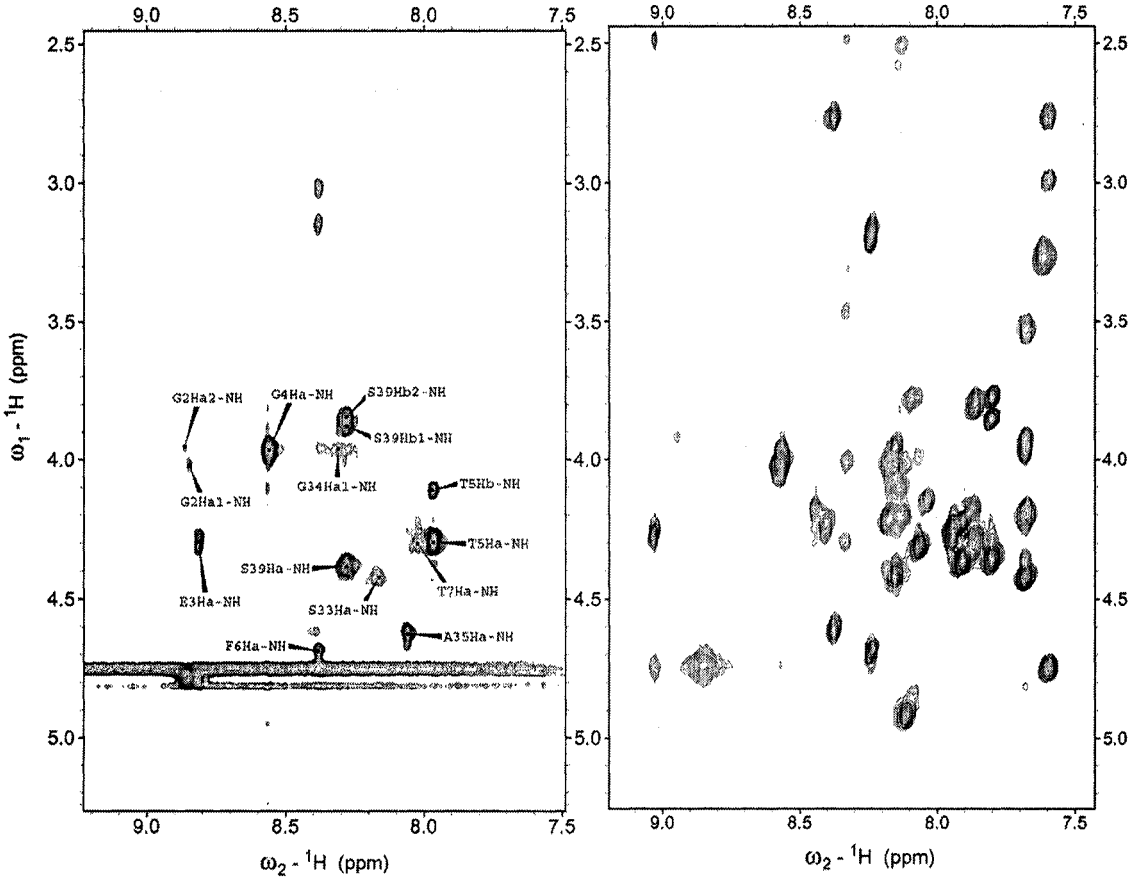


Figure 4-3: NH-Hα region of TOCSY spectra for Exendin-4.
Aqueous data presented in panel on left while data in right panel acquired on 30% TFE containing sample.

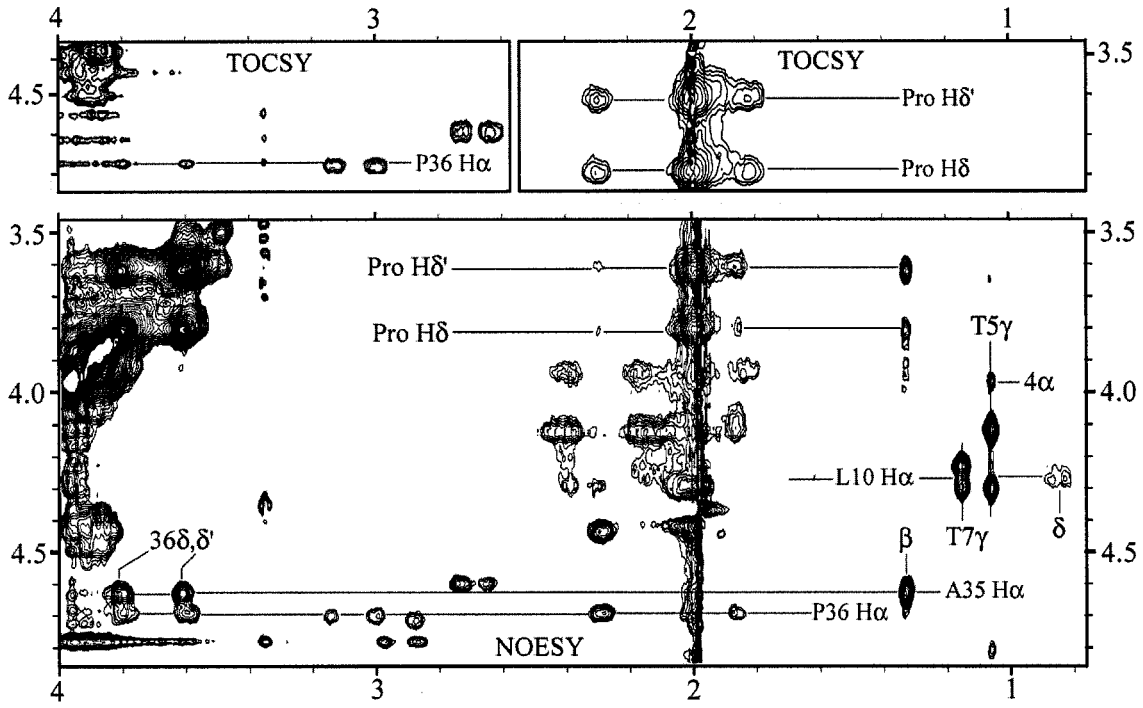


Figure 4-4: Regions of TOCSY and NOESY spectra of exendin-4 in aqueous buffer.

TOCSY given in upper panels and NOESY in lower panel.

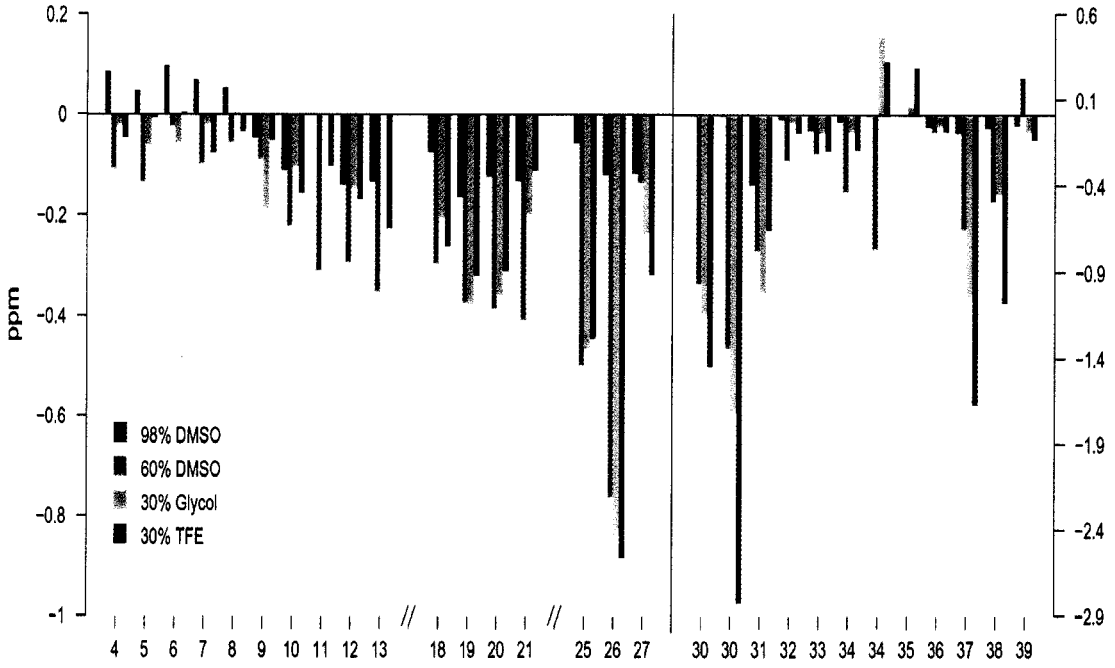


Figure 4-5: CSDs of Exendin-4 in various media.

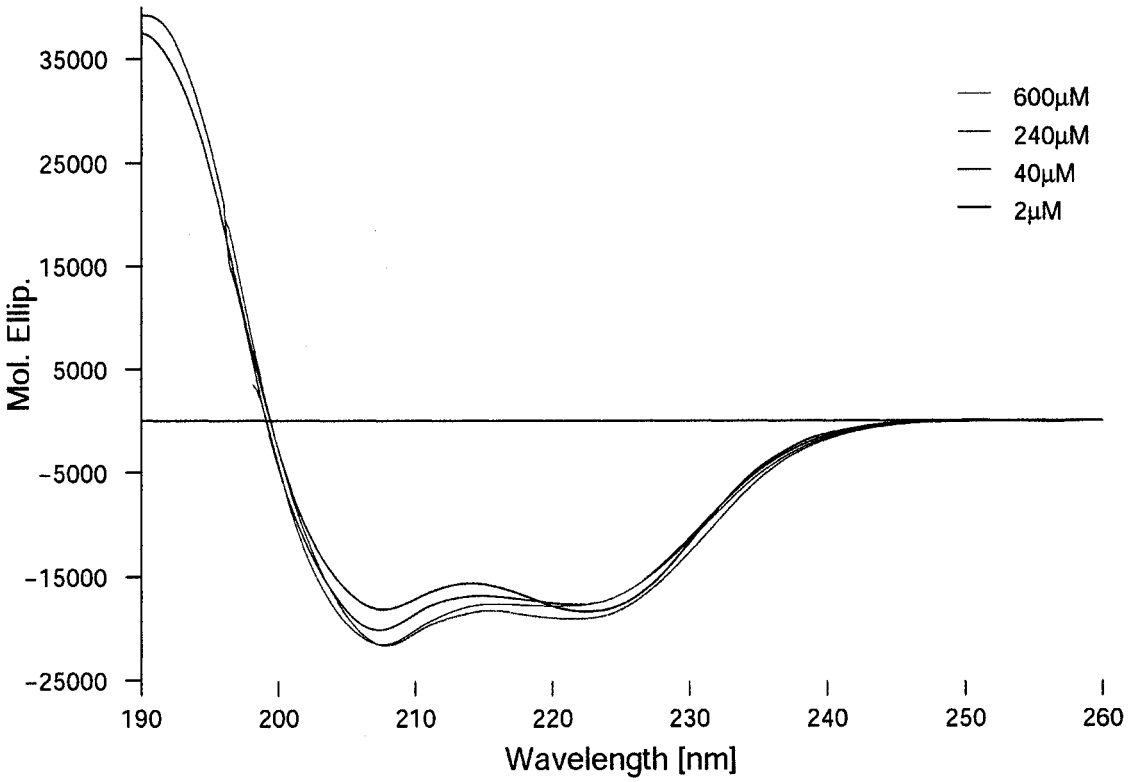


Figure 4-6: Circular Dichroism spectra for exendin-4 in .1% acetic acid at 280K.

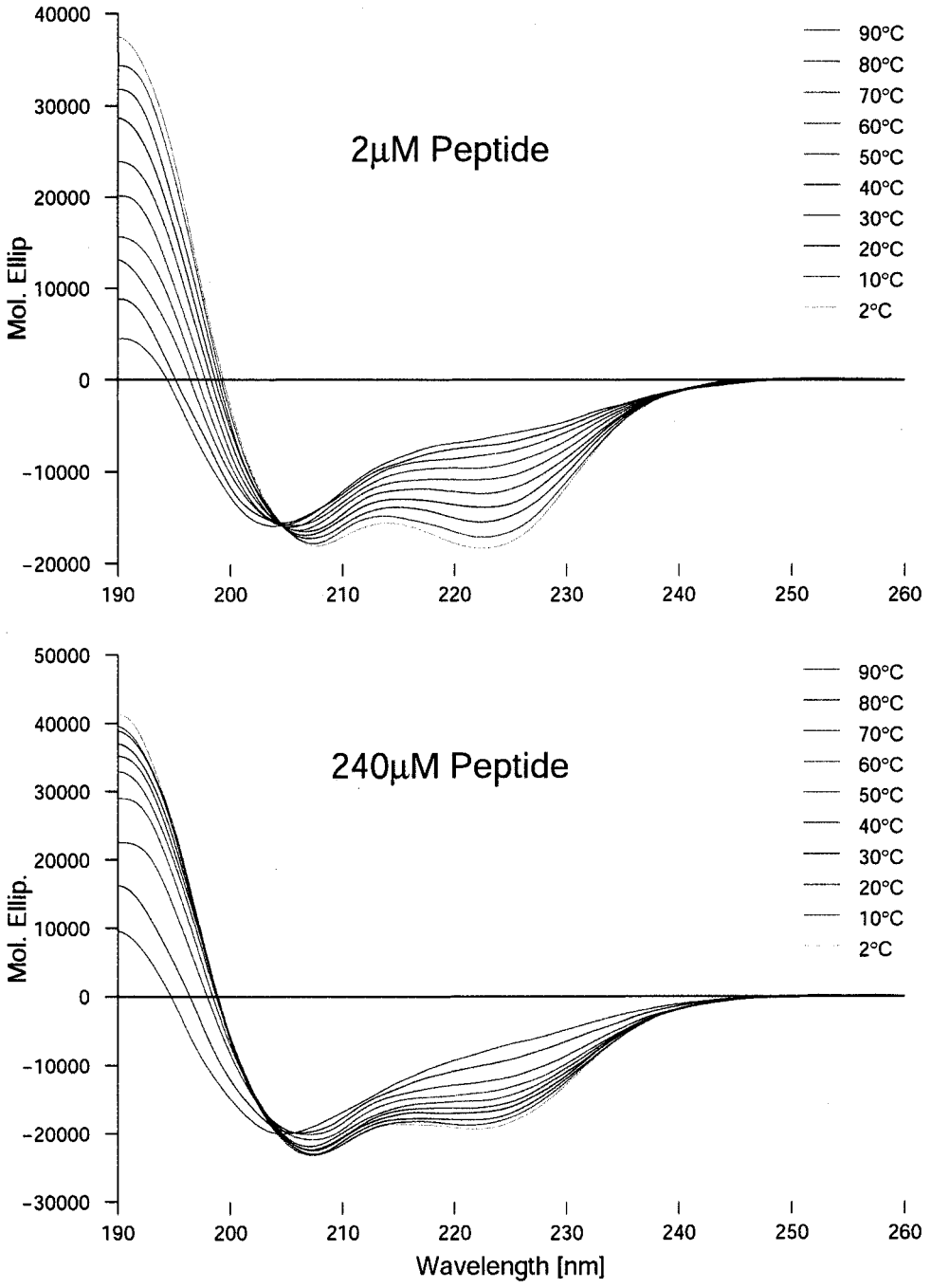


Figure 4-7: CD melts of exendin-4 in .1% acetic acid at high and low concentration.

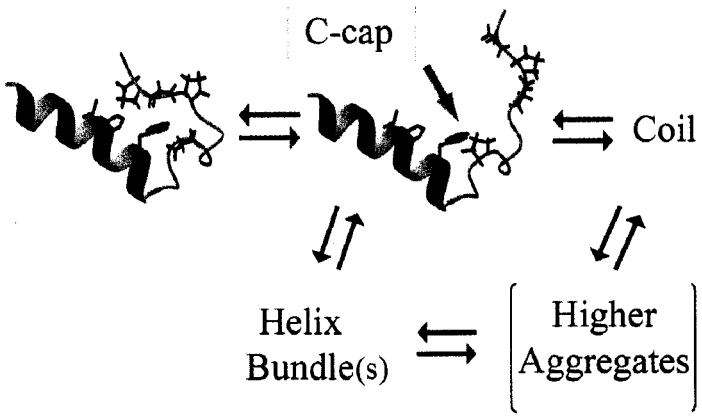


Figure 4-8: Proposed model of solution states of exendin-4.

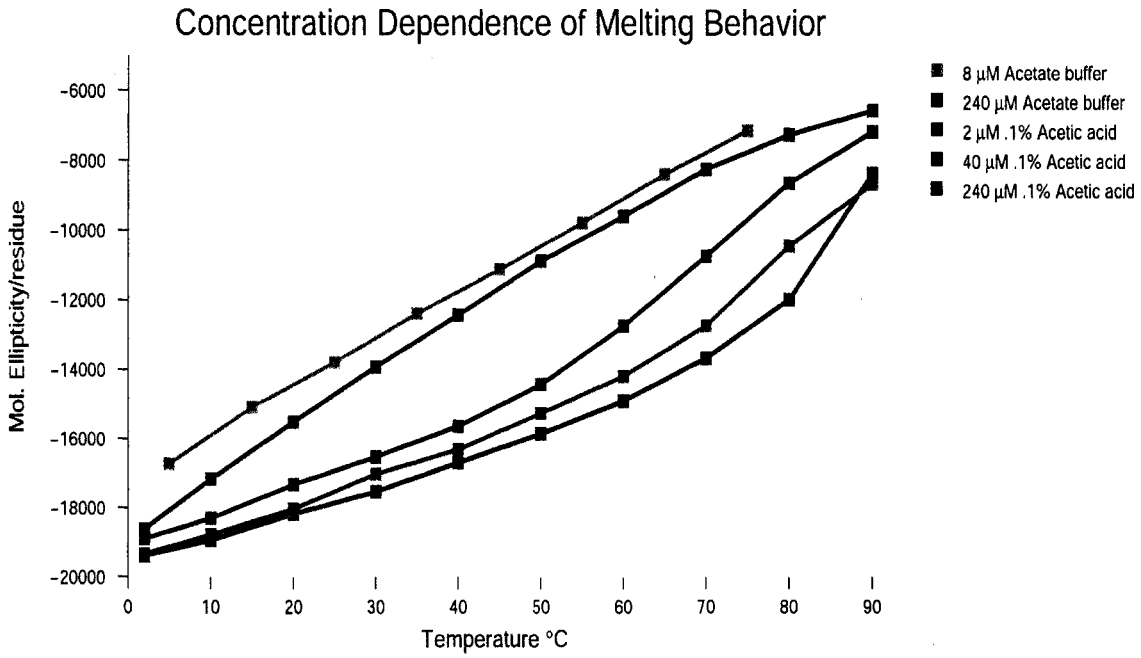


Figure 4-9: Concentration and medium dependence of melting behavior of exendin-4.

All data is aqueous.

Chapter 5 Model Sheets

Introduction

Designed peptide sequences that fold into stable three-stranded β-sheets have been sought for some time. Containing a novel central strand, this structural motif provides challenging problems in peptide design and structure elucidation. In principle, β-sheet peptides can be used to probe the cooperativity of the folding process but also serve to advance our knowledge of the relationship between the sequence and structure of peptides and proteins. The first designed, three-stranded sheet reported, *Betanova* (Kortemme et al. 1998), came some years after the first autonomously folding β-hairpins were reported. This peptide was designed by combining structural elements from 2 hairpins that were well folded in fluoroalcohol containing media, but lacked significant fold population in buffered, aqueous media. While *Betanova* was, at that time, a breakthrough in model sheet design, it was shown to be poorly folded (10-26%) in water by subsequent studies (Boyden and Asher 2001; López de la Paz et al. 2001; Kuznetsov et al. 2003).

These studies highlighted the fact that obtaining fold population estimates for double hairpin structures are particularly problematic. ^{1}H, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ CSDs, as well as far UV CD have been employed for structural characterization and %fold quantification. As with hairpins, the β-fold signature is characterized by downfield CSDs of backbone protons, upfield CSDs for $^{13}C_{\alpha}$ resonances, and downfield CSDs for C_{β} resonances, while circular dichroism spectra display a minimum at 216-218 nm and a maximum at 196-205 nm for a well-folded hairpin. In the determination of hairpin and sheet content by CD, the Serrano group favors the use of the ratio between 217nm and 204nm ($R_{217/204}$) (López de la Paz et al. 2001) while our UW group and the Searle group typically observe a 218nm minimum and a 196-202nm max. Correlations with NMR measures of folding

(Andersen et al. 1999; Dyer et al. 2005) prompted the UW group to use the ellipticity difference ($[\theta]_{216} - [\theta]_{-200}$) between the 214-218nm minimum and the maximum around 200nm. The Searle group, in measuring the stability of analogues of MrβHm and quantitating melts, relies on the intensity of the minimum at 216 nm (Maynard et al. 1998). Other methods of spectroscopic characterization have also been used to estimate β-sheet content. In one study of *Betanova* (Boyden and Asher 2001), UV Raman resonance spectroscopy was used to probe whether the unfolding mechanism of *Betanova* is cooperative. They found that the difference UV Raman resonance spectrum showed very little change at 206.5 nm from low temperature to high temperature (Boyden and Asher 2001). Th is wavelength is dominated by amide bands and is sensitive to secondary structure, indicating little or no dependence of the backbone Φ and Ψ angles on temperature. This suggests that the structure of the peptide changes little over the 77° range reported. Molecular dynamics simulations have also been performed on *Betanova* in an attempt to determine the folding thermodynamics (Bursulaya and Brooks III 1999). These indicated that the folding of the peptide occurred in two steps; the initial collapse involves the formation and 10-60% of the native contacts, followed by a second collapse in which the remaining contacts are formed, resulting in the native structure of the three-stranded sheet. These simulations also showed that there were a number of structures possible in the native energy basin, suggestive of a poorly converged structure. The conclusion from these studies is that the entire sheet is present in solution perhaps 10% of the time, while partially folded species make up approximately 25% of the equilibrium ensemble.

Because *Betanova* was shown to be such a poorly folded sheet model, efforts to improve its structural convergence and folded state population were undertaken (López de la Paz et al. 2001). These efforts relied on the application of *PERLA*, a protein design algorithm (Finsinger 2001), to the *Betanova* system. Mutations to the native sequence were generated by this algorithm at one to three sites per sequence. The peptide populations

were determined by circular dichroism and ^{1}HNMR, using the ratios of extrema and the CSD method, respectively. The authors calibrated their CSD scale with the spectroscopic characterization of unfolded and well-folded reference peptides. They relied on Hα chemical shifts for their analysis of peptide β-sheet content. As was shown in Chapter 3, the Hα CSDs in the C-terminal strand appear reduced in magnitude relative to those in the N-terminal strand ("C-terminal effect"). This may have led to the underestimation of hairpin or sheet content at these sites, which is often apparent when the HN CSDs are compared. The Hα based CSD interpretation used in that study led to the conclusion that the second hairpin had either less β content or increased C-terminal fraying (López de la Paz et al. 2001). This conclusion almost certainly underestimated the sheet content of *Betanova* and the mutants because of a falsely reduced C-terminal hairpin population. The topic of the terminal strand effect on Hα CSDs will be discussed in relation to three-stranded sheet peptides in further detail in this chapter.

Design of β structures has for years focused on elucidation of the interactions important to β-hairpin and sheet formation. Once it became clear what interactions were necessary to form these structures, rapid progress was made. Today, we know that to form a well folded β-hairpin or sheet, multiple interactions are required. The strands must contain β-branched residues because of their natural preference for an extended conformation. Strands must also contain hydrophobic or aromatic residues for their propensity to form cross-strand hydrophobic clusters, and such strands must be linked by two or more residues which prefer the proper φ/ψ geometries necessary for chain direction reversal (Sibanda and Thornton 1991; Ramírez-Alvarado et al. 1996). Also, oppositely charged residues are generally found at the termini of hairpins to enhance the coulombic attraction that exists naturally between the N and C termini and too ensure solubility. Since these design criteria for β-structures have been found, more complicated structures such as three and four-stranded sheets have been made (Schenck and Gellman 1998; Sharman and Searle 1998; de Alba et al. 1999b; Santiveri et al. 2003; Syud

et al. 2003). There is still, however, much to be learned from the simplest of hairpins and sheets; the cooperativity and kinetics of hairpin and sheet formation are still rather poorly understood and shall be discussed in Chapter 7 while here, the efforts made to characterize a model sheet with various turn mutations will be described.

The design of three-stranded sheet sequences that are well folded generally relies on the non-natural amino acid, D-proline (abbreviated with p throughout), to increases turn propensity and sheet population. Among these, the sequence designed by the Gellman group (Schenck and Gellman 1998) is perhaps the most stable and well studied. The sequence of this peptide, showing strand alignments, is given in Figure 5-1.

Turning specifically to the SG sheet, studies of the wild-type peptide and some derivatives indicated that the sequence adopts a highly-populated, three-stranded sheet in aqueous media which the researchers believed would allow them to probe the cooperativity of the folding of a β-sheet (Schenck and Gellman 1998). They also hoped studies of this peptide could distinguish between the modes of β-sheet cooperativity: folding can propagate along the strand direction or perpendicular to the strand (Schenck and Gellman 1998). Utilizing the fact that the L-Pro-Gly sequence prevents turn formation, they sequentially mutated each turn such that the resulting peptides would, presumably, not be able to form a turn about the mutated locus, leaving only the possibility of forming a hairpin about the unchanged turn. Hairpin fold populations were determined by using the outwardly directed Hα proton resonances. Since these have smaller CSDs than their inwardly directed counterparts, the error involved in these calculations is inherently greater. The studies presented here will characterize the SG three-stranded sheet and turn mutants which have implications in the field of peptide design.

Three-Stranded Sheet Mutants

For the present studies, mutations to the SG three-stranded sheet sequence were made. The mutations were focused about each turn locus (table 5-1) to systematically reduce the hairpin population about each turn. Ostensibly, these mutations were made to determine the folding mechanism and cooperativity of three-stranded sheet folding but they also provided other insights in sheet design. The mutants, listed below in Table 5-1, are organized into four categories by turn sequence: wild type with T1 mutations, wild type with T2 mutations, an NG-NG turn species with T1 mutations, and the NG-NG turn species with T2 mutations.

Table 5-1 Mutants of the SG Three-stranded Sheet

	1 10 20
SG	Ac-VFITSpGKTYTEVpGOKILQ-NH_2
SG NG-pG	Ac-VFITSNGKTYTEVpGOKILQ-NH_2
SG DG-pG	Ac-VFITSDGKTYTEVpGOKILQ-NH_2
SG GG-pG	Ac-VFITSGGKTYTEVpGOKILQ-NH_2
SG PG-pG	Ac-VFITSPGKTYTEVpGOKILQ-NH_2
SG GP-pG	Ac-VFITSGPKTYTEVpGOKILQ-NH_2
SG	Ac-VFITSpGKTYTEVpGOKILQ-NH_2
SG pG-NG	Ac-VFITSpGKTYTEVNGOKILQ-NH_2
SG pG-GG	Ac-VFITSpGKTYTEVGGOKILQ-NH_2
SG pG-PG	Ac-VFITSpGKTYTEVPGOKILQ-NH_2
SG NG-NG	Ac-VFITSNGKTYTEVNGOKILQ-NH_2
SG pG-NG	Ac-VFITSpGKTYTEVNGOKILQ-NH_2
SG PG-NG	Ac-VFITSPGKTYTEVNGOKILQ-NH_2
SG NG-NG	Ac-VFITSNGKTYTEVNGOKILQ-NH_2
SG NG-pG	Ac-VFITSNGKTYTEVpGOKILQ-NH_2
SG NG-PG	Ac-VFITSNGKTYTEVPGOKILQ-NH_2

Determination of Hairpin Populations in Three-Stranded Sheets

CSD analysis of three-stranded sheets allows for the rapid determination of the folded state conformation but the quantitative assignment of population to the individual parts of the structure requires a more in depth analysis. α proton CSD magnitudes for inwardly directed positions of the third (C-terminal) strand appear reduced compared to those in the first and second strands, even under the most fold favoring conditions. As previously noted, this characteristic has been observed (if not recognized) in other sheet models (López de la Paz et al. 2001) and contrasts with the cross-strand directed H_N CSDs which have similar magnitudes in the first and third strands. This feature is evident in Figure 5-2: the wild type species can be seen as having much larger Hα CSDs (panel a) in strand 1 relative to strand 3 (in both water and HFIP) but in the HN plot, the average magnitudes are much more similar. *A priori*, a glance at the Hα CSDs of mutants that are only somewhat destabilized aboutT1 (DG, NG and GG), suggests that both hairpins form to approximately the same extent. When looking at the H_N CSD plot, we see that the strand 1 CSDs are significantly diminished while, upon going from pG to DG, NG or GG, strand 3 has incurred only small changes in the magnitude of its CSDs, resulting in their having a greater overall magnitude than those in strand 1. This is expected since hairpin 2 has not been directly destabilized. Therefore, the reduction of hairpin 1 population relative to hairpin 2, due to reduced turn propensity at T1, is reflected more clearly in the HN CSDs. The central strand is not generally used in the assignment of hairpin populations because the protons in this strand, which are directed towards the less folded hairpin, may receive some contribution to their CSD from the structuring of this strand due to the formation of the other hairpin. However in Fig. 5-2, both the Hα and H_N CSD plots show accentuated losses at sites that are directed toward strand 1 (•).

Calibration of the Fraction Folded Scale

The determination of the folded state populations of secondary structure motifs is complicated by the lack of 0% and 100% folded reference values for calibration of a CSD scale. The quantitation of secondary structure populations is also hampered by the variance that exists among the structure probes available. The variance can be caused by portions of the structure being relatively unstructured (such as strand terminus fraying) or the presence of aromatic residues that otherwise change the magnitude of the observed CSD. In the case of the SG system, the Gellman Lab has used unfolded reference peptides (with ^{L}Pro-Gly "turns") to estimate 0% CSDs and cyclic reference peptides to estimate the 100% folded values (Syud et al. 2003). In this laboratory, we utilize well folded peptides in fluoroalcohols to generate the CSDs of a system that is highly populated. We also use unfolded reference peptides to correct for non-zero CSDs in the unstructured state; a point which will be discussed further in the next section. When the "100% reference CSD values" determined for the SG sequence in 8% HFIP, some disagreement between Gellman population estimates and ours, for the peptide in aqueous media, is observed. The Gellman lab finds that when their method is employed, the fraction folded of the SG peptide in aqueous media as determined by Orn^{16} is 0.83 at 277K and 0.76 at 297K and for Ile^{18} it is 0.75 at 277K and 0.66 at 297K (Syud et al. 2003). We find that only when our reference CSDs (from SG in 8% HFIP) are equated to 92% folded, is good agreement with the Gellman Lab estimates found.

The raw CSD data for the SG peptide and mutants made is provided in Table 5-2. The turn locus Gly residues $\Delta\alpha$ values are provided because they are used by other labs to quantitate hairpin formation (Tatko and Waters 2002). The CSD factor (CSDf) is the number which, when multiplied by the CSD, yields the fraction folded as determined at that site. The fraction folded of each hairpin for all mutants, determined from the CSDs in Table 5-2, are reported in Table 5-3. The folded state populations were calculated in a site specific manner, with weighted averages reported. Certain CSDs are doubly

weighted in the values reported here, because larger CSDs provide greater measurement precision; the resonances that are doubly weighted are named in the table legend. Even though the largest CSDs are weighted more heavily, the populations that appear in Table 5-3 are based solely on the cross-strand directed Hα and HN sites which display larger CSDs. Because the wild-type peptide in 8% HFIP showed the largest CSDs, both hairpins in this peptide were assumed to be 92% folded for reasons cited above, while the SG PG-NG and pG-PG/NG-PG peptides provided 0% reference values for the N- and C-terminal hairpins, respectively.

Table 5-2 CSD Data for SG Sheet and Mutants

Hairpin 1 reporting CSDs							
cmpd	**T4α**	**I3N**	S5N	T9α	Y10N	N6N	G7Δα
CSDf‡	0.875	1.0	1.257	1.227	0.875		2.606
8 % HFIP, 280K							
pG-pG	1.051	0.92	0.731	0.75	1.052	**n.a.**	0.351
8 % HFIP, 290K						***0.74†***	***0.65†***
NG-NG	0.369	0.307	0.357	0.390	0.511	0.216	0.272
NG-**pG**	0.579	0.499	0.504	0.556	0.710	0.409	0.353
GG-**pG**	0.559	0.477	0.372	0.517	0.660	0.538	0.267
pG-GG	0.935	0.783	0.644	0.662	0.947	**n.a.**	0.300
Water, 290K							
pG-pG	0.808	0.690	0.556	.655	0.771	**n.a.**	0.320
PG-**pG**	-0.033	0.048	0.121	0.097	0.15	**n.a.**	**0.0**
pG-PG	0.665	0.534	0.478	0.494	0.649	**n.a.**	0.218
GG-**pG**	0.393	0.363	0.265	0.387	0.499	(0.391)	0.185
pG-GG	0.800	0.681	0.565	0.603	0.793	**n.a.**	0.240
NG-**pG**	0.387	0.352	0.359	0.401	0.514	0.261	0.212
NG-PG	0.119	0.111	0.157	0.127	0.206	(0.045)	0.126
pG-NG*	0.779	0.666	0.540	0.581	0.770	n. r.?	0.311
PG-NG*	**-0.029**	**0.00**	**-0.066**	**0.051**	**0.126**	**n. r.?**	(0.254)
NG-NG	0.267	0.247	0.262	0.287	0.394	0.148	0.197
30% TFE, 290K							
NG-**pG**	0.586	0.406	0.472	(0.587)	0.693	0.393	0.373
GG-**pG**	0.370	(0.117)	0.162	0.405	0.373	--	0.224

(Table continued on following page)

(Table 5.2 continued from previous page)

Hairpin 2 reporting CSDs					
cmpd	K17α	**I18N**	E12α	N14N	G15Δα
CSDf‡	1.945	0.952	1.077		3.286
8 % HFIP, 280K					
pG-pG	0.473	0.966	0.854	**n.a.**	0.280
8 % HFIP, 290K				***1.09†***	***0.44†***
NG-NG	0.247	0.592	0.544	0.634	0.245
NG-**pG**	0.406	0.800	0.790	**n.a.**	0.243
GG-**pG**	0.407	0.822	0.797	**n.a.**	0.234
pG-GG	0.324	0.682	0.644	0.691	0.278
Water, 290K					
pG-pG	0.414	0.884	0.671	**n.a.**	0.265
PG-**pG**	0.276	0.629	0.576	**n.a.**	0.223
pG-PG	**0.037**	0.194	**0.03**	**n.a.**	0.037
GG-**pG**	0.350	0.764	0.665	**n.a.**	0.185
pG-GG	0.271	0.646	0.527	0.635	0.222
NG-**pG**	0.347	0.757	0.665	**n.a.**	0.195
NG-PG	0.045	**0.182**	**0.015**	**n.a.**	**0.0**
pG-NG*	0.271	0.678	(0.366)	0.700	0.27
PG-NG*	0.137	0.394	0.228	0.394	0.138
NG-NG	0.192	0.541	0.408	0.578	0.201
30% TFE, 290K					
NG-**pG**	0.412	0.744	0.809	**n.a.**	0.222
GG-**pG**	0.356	0.677	0.765	**n.a.**	0.256

*Literature values from Syud et al. 2003.

†NGNG 100% values

‡CSDf is the multiplication factor which converts the CSD value to fraction folded

Table 5-3 Fraction Folded of Hairpins 1 & 2 in SG Peptides

	Frac. folded 1st HP[a]	Frac. folded 2nd HP[b]
pG -- **pG**	0.718 ± 0.063	0.812 ± 0.055
NG -- **pG**	0.365 ± 0.038	0.682 ± 0.027
GG -- **pG**	0.346 ± 0.058	0.679 ± 0.042
PG -- **pG**	< 0.04 >	0.597 ± 0.083
pG -- PG	0.557 ± 0.023	< 0.06 >
pG -- GG	0.682 ± 0.030	0.552 ± 0.030
pG -- NG	0.668 ± 0.019	0.563 ± 0.091
NG -- NG	0.247 ± 0.029	0.461 ± 0.060
PG -- NG	0.00 ± 0.02	0.300 ± 0.032
NG -- PG	0.104 ± 0.024	0.040 ± 0.034
In 8% HFIP		
GG -- **pG**	0.485 ± 0.063	0.788 ± 0.042
pG -- GG	0.800 ± 0.017	0.640 ± 0.049
NG -- **pG**	0.550 ± 0.055	0.784 ± 0.049
NG -- NG	0.354 ± 0.050	0.536 ± 0.045
In 30% TFE		
NG -- **pG**	0.507 ± 0.066	0.736 ± 0.094
GG -- **pG**	0.315 ± 0.095	0.710 ± 0.117

[a] Since larger CSDs provide greater precision, T4α and I3N were doubly weighted in this calculation

[b]For the reason provided in [a], I18N was doubly weighted in the calculation of HP2 populations

Correlation of CSD Measures with % Fold

The SG mutants reported here have been tested to determine the ability of individual site CSDs to predict the folded population of the peptide and which reference values lead to accurate determinations of structure propensity. The T1 mutant sequences reported in Table 5-1 provide peptides with both partially and fully destabilized versions of hairpin 1 while the T2 mutants provided peptides with analogous populations of hairpin 2. Figure 5-3 illustrates how our structure probes correlate with reference values for all mutants of the SG sequence reported here. Both X- and Y-axes in all panels are percent folded with the 100% value provided by the SG wild type sequence in 8% HFIP. 100% was used instead of 92% for the reference to simplify the calculations. The 0% reference values are plotted in two different ways: using 0 CSD (panel b) and using the more specific unfolded reference peptides to estimate the CSDs for the unfolded state (panels a and c). The two methods for determining the 0% reference CSDs were compared because, looking at Table 5-2, the unfolded reference peptides (SG PG-pG, SG pG-PG and SG NG-PG) have some significant, non-zero CSDs. In panel (a), the various sites that report on structuring are identified, showing which sites do and do not correlate with fraction folded. The 0% reference for this data is provided by SG PG-pG for the N-terminal hairpin and SG pG-PG and SG NG-PG (smallest values between the two were chosen) for the C-terminal hairpin. Panels (b) and (c) employ only the data from panel (a) that correlates well with the percent folded scale, namely all data besides some of the S-1 sites. Panel b uses 0 CSD for the 0% reference on the Y-axis. Using these values for the 0% reference overestimates hairpin populations as evidenced by a Y-intercept at 19%. Since the linear correlation is nearly as good as that when compared to 0% CSD, and the Y-intercept is near zero, this method is shown to be estimating populations more accurately. Since the use of specially designed unstructured peptides provides a better correlation with the data and does not overestimate folded state population, the use of such references is a valid practice for this system.

Turn Types in β Sheets

The mutants outlined in Table 5-1 are expected to contain [2:2] or [2:4] turns at both loci, by peptide design criteria. Since it has been previously shown that turn type in β-hairpin peptides can be determined by the H_N CSD signature of the residues involved in that turn (Fesinmeyer et al. 2005), the same analysis is extended to three-stranded sheets here. Figure 5-2 presents the Hα (panel a) and H_N (panel b) CSDs for the turn 1 mutants that were made, with the wild type sequence shown for reference. The Hα CSDs are perhaps more compelling for this analysis, and indicate the changes in turn sequence are not changing the turn type and are merely leading to a decreased hairpin population about turn one. This can be seen because we expect the inwardly directed protons to have the larger downfield CSD. In the wild type peptide, the α protons of Phe^2 and Thr^4 have larger CSDs than Ile^3 and Ser^4, thus indicating the shift register of this peptide. When the turn is destabilized, this feature, while obscured in the case of the DG, NG and GG peptides, is still present. Phe^2 and Thr^4 still have larger CSDs than Ile^3, though the difference is much smaller now and the meaning more subject to interpretation. The indication is still that there is no change in which of these protons is inwardly directed. Additionally, the CSD of Ser^5 has been reduced to near random coil in all the mutants of the wild type, a probable indication of the loosening of the turn, directed at strand one, since no other marked changes in the CSD plot are observed.

While Hα CSDs at T9 and T11 likely reflect the decreased hairpin population, it has been reported that the SG DG-pG peptide actually forms a [3:5] turn at T1 (Chen et al. 2001; Chen 2004). This conclusion was based upon significant NOE evidence (1α/11α, 2N/11α, 2N/10N, and 3α/10N) which could only be explained by the loop sequence not forming the expected SDGK turn but a TSDGK loop instead (Chen et al. 2001). The study utilized a 300ms mixing time in their NOESY pulse sequence, an amount fully 50% longer than those routinely used for the analysis of peptides in this study. The NOESY

spectrum I acquired on this peptide contained none of the diagnostic NOESY connectivities cited by the prior study as indicative of [3:5] turn formation. I found the expected NOEs for a two-residue turn (such as the cross-strand 4α/9α, 2α/11α, 3N/10N and 4α/10N NOEs). Also, when looking at the CSDs, a shift from a two-residue to three-residue turn would involve a change in the shift register as the first strand is shortened to accommodate the longer turn and the CSD signature at the turn itself would change (Figure 5-2). While the shift register argument is not compelling because this feature is muted in this peptide, the amide CSD signature of the turn does present a compelling argument.

In the CSD plot of the turn (Figure 5-4), we would see the CSD of D6 change from positive to substantially negative, as is found in the T2 position of [3:5] turns. Because all of the examples of [3:5] turns have a proline at the T1 position, we do not know what to expect of an Ser^5 H_N CSD occupying this location of the turn. In the NG-turn locus [2:2]/[2:4] peptide, it has a large downfield chemical shift deviation which is consistent with a hydrogen bonded position. If the turn type had changed, we would expect Thr^4 to occupy this position, resulting in a greater downfield CSD. This is not seen (Fig. 5-4); we do see that Ile^3 and Ser^5 both have larger CSDs than Thr^4, although the differences are small. Thr^4 has an H_N CSD completely consistent with a non-hydrogen bonded strand position. It can then be concluded that all CSD evidence indicates that a p6D mutation does not change the turn type from [2:2]/[2:4] as previously reported (Chen et al. 2001); it adopts approximately the same conformation as all the other turn sequences examined in this study. It would appear that a NOESY mixing time of 300ms led to the observation of higher order NOE signals, resulting in the conclusion previously reported.

The data which led to the determination that all turn sequences made in this peptide scaffold form the same number of residue loops has also provided a means to assign the

relative turn propensity provided by these sequences. Since we can quantify the populations of individual hairpins accurately (at least relative to each other) (Table 5-3), the assignment of relative turn propensity can then be made based upon the intensity of strand CSDs. With the application of this methodology, the relative turn propensities of the sequences used in this study are found below.

Table 5-4 Relative Turn Propensities of Two-Residue Sequences

pG>>DG>NG≈GG>>PG≈GP

The fold populations presented in Table 5-3 also allow for a quantitative treatment of the turn propensity issue. The extent to which XG turns (X = G, N) are less favorable for hairpin formation than pG can be assessed by numerous comparisons. These will be discussed in the context of XG→pG comparisons only in SG XG-pG and SG pG-XG peptides, using the values given in Table 5-3. In water, the pG→GG and pG→NG mutations destabilize the 1st hairpin by 3.7 ± 0.1 kJ; smaller effects, 2.95 ± .05, are seen at the 2nd turn locus. The effect of pG→NG substitution has been measured in other hairpin systems (Cochran et al. 2001b; Espinosa et al. 2001; Fesinmeyer et al. 2005) as well. These studies found a destabilization, ranging from 2.7 kJ/mol to 3.1 kJ/mol in these systems. These measurements are completely consistent with what is observed in the 2nd hairpin on the SG series. The greater destabilization observed about the T1 locus is probably due to the lower β propensity of the residues in the 1st strand. The effects of turn mutations on hairpin stability in fluoroalcohol cosolvents have only been examined for the pG→GG mutation in 8% HFIP; the calculated ΔΔGs are 5.9 (1st turn) and 4.5 kJ (2nd turn) but the potential for error is greater since the folded population estimate (F_f = 0.92) for SG pG-pG may be too large. Estimates of ΔG are intrinsically less accurate near

the fully folded limit and fully unfolded limit because small changes near the extrema lead to large changes in ΔG.

Folding Pathway of the SG Peptides

The greater tendency for hairpin formation between strands 2 and 3, seen in MD simulations (Roe et al. 2005) is evident in many comparisons: i) the 2nd versus 1st hairpin populations in SG NG-NG which differ by 0.20 in both water ($\Delta\Delta G$ = 2.3 kJ) and 8 % HFIP ($\Delta\Delta G$ = 1.8 kJ), ii) the unequal populations about NG and GG turns in SG XG-pG versus SG pG-XG which favor the 2nd hairpin by 1.9 kJ and 1.5 kJ in water and 8% HFIP, respectively for X = G, and 2.0 kJ in water for X = N. All of these $\Delta\Delta G$ estimates are based on "2-state" comparisons focused on folding measures for a specific hairpin. Two-state assumptions are clearly unwarranted: when both turn sites are occupied by turn-permissive sequences (pG, NG, GG), the population estimates for the two hairpins can differ by as much as 0.33, particularly when the second turn is the more stable pG sequence.

Cooperativity, in the sense of proteins, is generally thought of as the increased stabilization of the folded state due to tertiary interactions. This leads to a sharp melting transition in spectroscopic analyses because of a cascade effect that begins once the first of these stabilizing interactions is lost on melting. Since peptide models have fewer tertiary interactions, we think of cooperativity as requiring a well converged folded state melting to an unfolded state at the same rate in all parts of the structure. This would be reflected in a uniform loss of structuring CSDs at all sites. Such an observation implies a two-state folding mechanism because the peptide converges towards one folded state and always has the same fraction folded across the entire sequence at all temperatures. In the case of a peptide which has individual parts of the structure selectively destabilized, the cooperativity of the peptide can be described as the amount by which

the native-like portion of the sequence is destabilized by mutation of a structurally isolated part. Turning to the SG series of peptides, both T1 and T2 destabilizing mutations cause some loss of structure in the unmodified hairpin. Efforts will be made to quantify this cooperativity effect in Chapter 7, but here the focus is on the observation that even complete destabilization of one turn or the other results in a peptide which is folded about the unmodified turn.

Since CSD analysis of peptides provides site specific information about conformation and population, an assignment of the relative hairpin propensities for both hairpins in the three-stranded sheet can be made. If we look at the effect of T1 mutants on the population of hairpin 2, we can ascribe no more than a 30% loss in population on going from pG to PG at T1 in aqueous media (see Figure 5-5 and Table 5-2). Since the PG sequence precludes hairpin formation about that site, there is essentially no formation of hairpin 1 in this peptide. Therefore, we have effectively determined that in the absence of hairpin 1, hairpin 2 is still significantly populated. In the other case, we find that the complete destabilization of T2 yields a peptide with a significant population of hairpin 1 remaining, though it is destabilized by as much as 35%. This data shows that both hairpins do form to a significant extent in the absence of the other when a pG turn is present. As a result, a strict, two-state folding mechanism can be ruled out entirely as this would preclude the isolated formation of either one of the individual hairpins. Since they both can form to a measurable extent autonomously, a four-state folding scheme can be applied and is described in Figure 5-6. It is clear that K_{UN} is less favorable than K_{UC} but a 4-state mechanism still allows for folding of the random coil peptide through either of two intermediates: the N or C terminal hairpin folded with the remaining strand dissociated.

Conclusion

The work presented here, characterizing a designed, three-stranded β-sheet, has led to some important conclusions regarding sheet formation and conformation. While this data shows the relative turn propensity of many two-residue turn-forming sequences, it also indicates that determination of the population of the folded state can be achieved through the use of the terminal strand CSDs, exclusively. It also shows that this peptide must exist in equilibrium between four states, as two conformations unique from the folded and unfolded states can be seen in analogs.

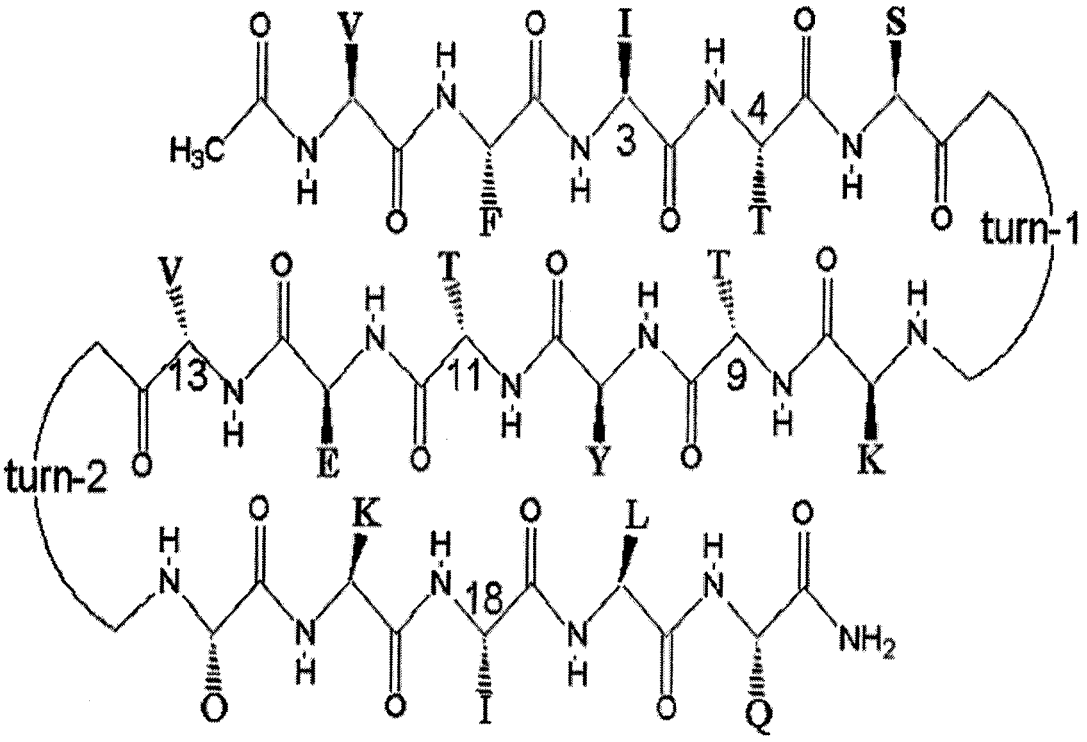


Figure 5-1: Schematic of SG three-stranded sheet.
Diagram shows strand alignment and sequence of wild type and turn mutants.

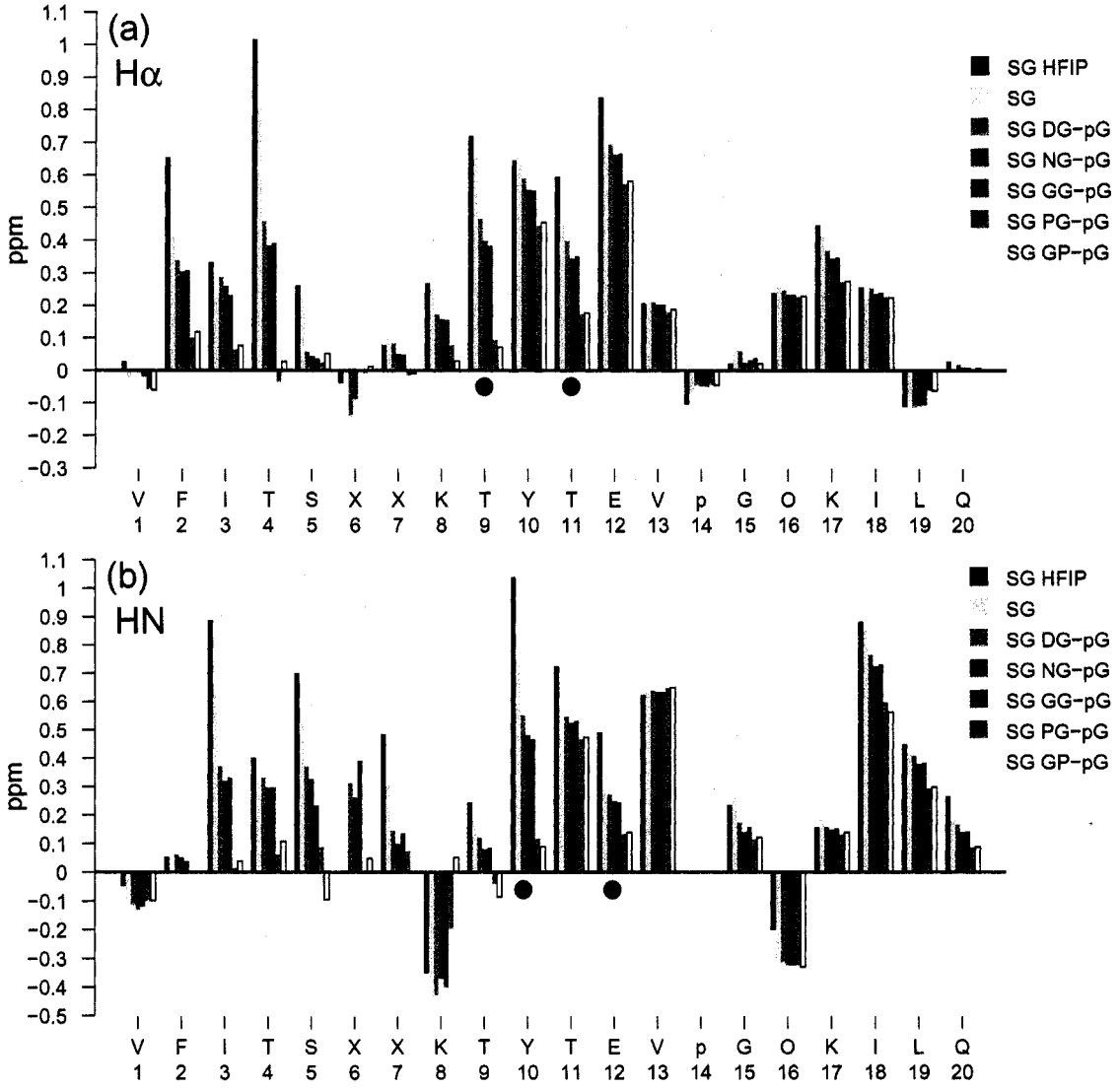


Figure 5-2: Hα and H_N CSDs for SG wild type and turn 1 mutants.
Data shows the effect of destabilization of turn one on the entire structure of the peptide. CSDs of resonances in the central strand that are directed toward strand 1 are indicated (•). All data shown at 290K

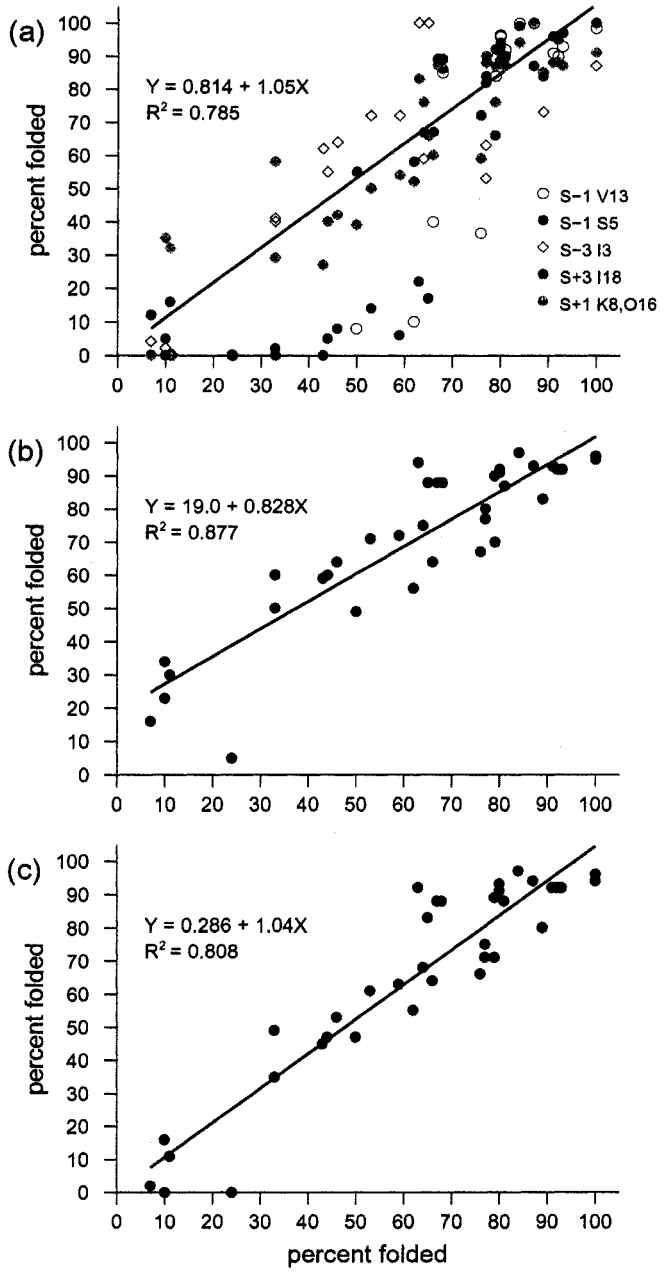


Figure 5-3: Correlations between various fraction folded measures.
Both axes use SG in 8%HFIP for the "100%" reference. Both panel (a) and panel (c) employ CSDSs from PG-pg and pg-PG for 0% reference but panel (c), employs only structuring shifts- I3 and K8 for the N-terminal hairpin, O16 and I18 for the C-terminal hairpin- for the correlation. Panel (b) shows the same data as panel (c) referenced to our random coil shifts for the 0% folded values.

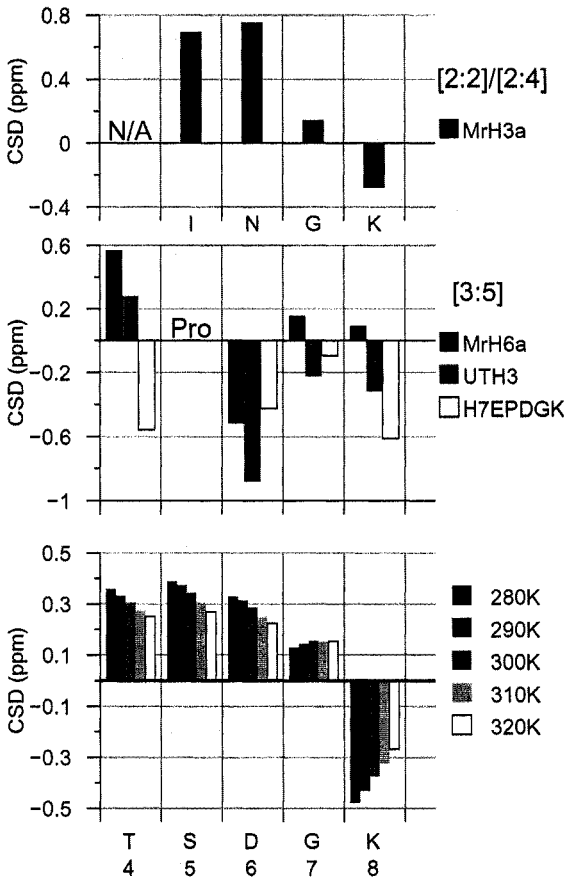


Figure 5-4: Loop CSD signatures for 4 and 5 residue loops.
Bottom panel shows loop 1 CSDs for SG DG-pG.

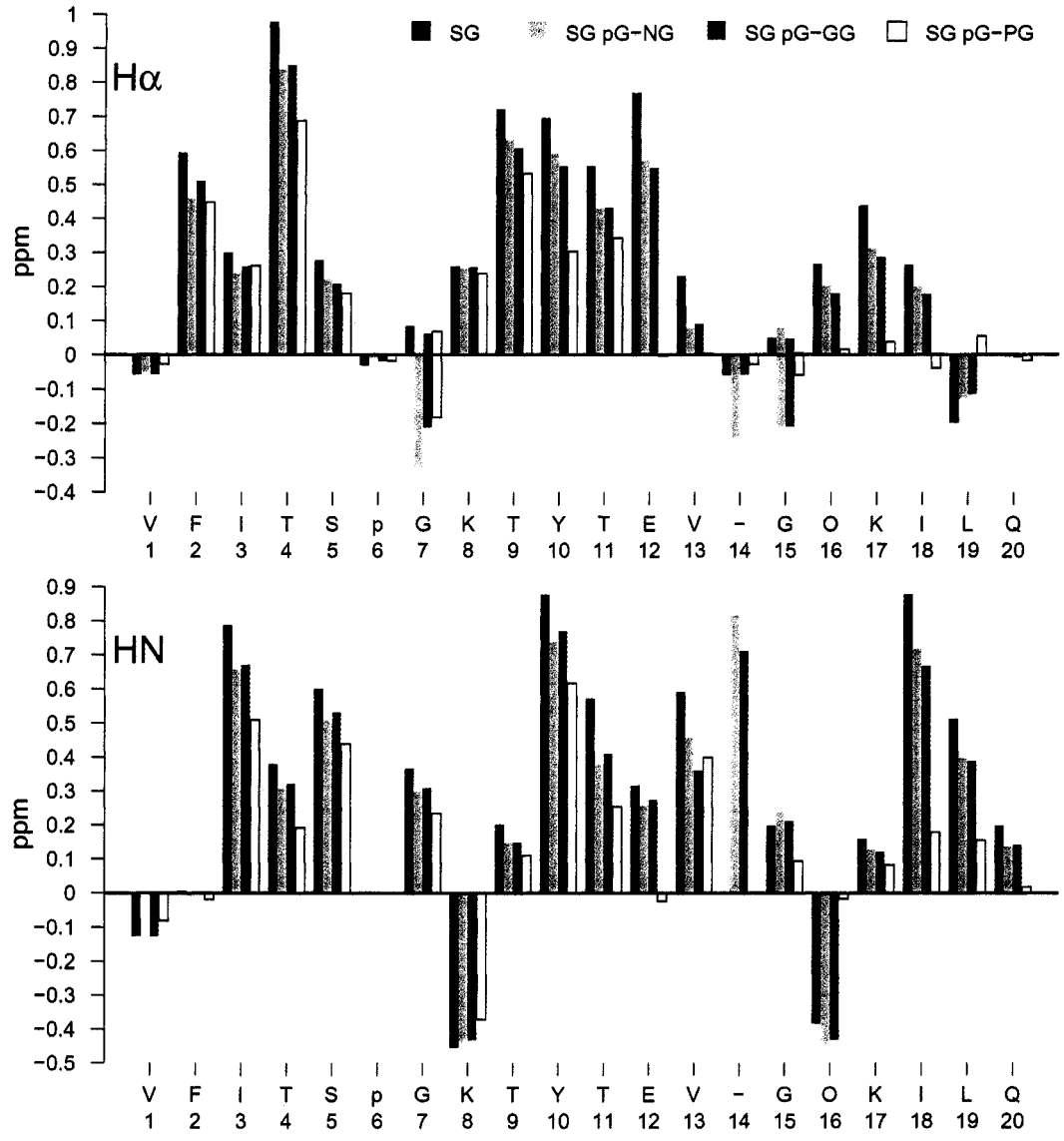


Figure 5-5: Hα and H_N CSDs of SG and T2 mutants.
Data presented at 290K.

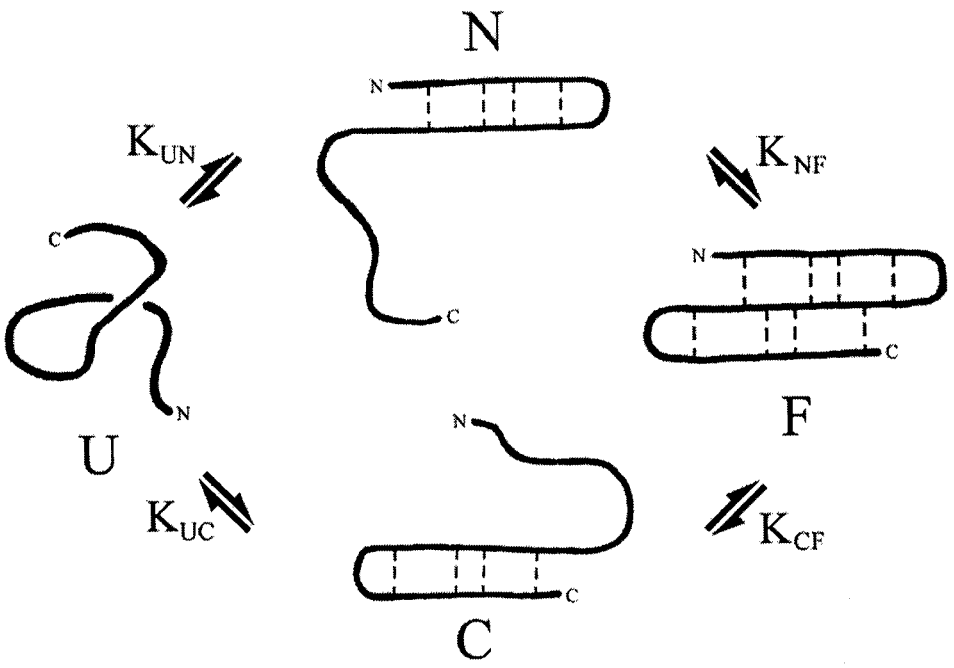


Figure 5-6: Four-state equilibrium model of SG three-stranded sheet.

Chapter 6 Minimization of a WW Domain

Introduction

WW domains are naturally occurring motifs that are approximately 40 amino acids long and derive their name from two highly conserved tryptophan residues (Chen 1995; Sudol 1995a; Sudol 1995b). In many cases, they are thought to autonomously fold into their native, three-stranded, antiparallel sheet conformation (Macias et al. 1996; Koepf et al. 1999). The purpose of the domains is thought to be the binding of proline rich regions of partner proteins (Sudol 1996; Otte 2003) which makes them important for recognition phenomena. These peptides have come under intense study of late because their size lends them to mutational (Kraemer-Pecore 2003) and mechanistic (Jager et al. 2001) studies. WW domains have also been important for the determination of factors relevant to the kinetics of β-sheet folding (Crane et al. 1999; Koepf et al. 1999; Ferguson et al. 2001a; Ferguson et al. 2001b; Jager et al. 2001). In the case of hPin1, an extensive mutational study, evaluated by ϕ analysis, has determined the portions of the sequence which are important for definition of the transition state (Jager et al. 2001). This study determined that the hairpin about loop one was formed in the transition state.

While many studies have concentrated on the folding dynamics of WW domains, others have focused on the equilibrium characteristics of these systems. NMR characterization of WW domains has been a topic of great interest of late and studies have determined the key interactions necessary for the folding of these domains (Macias et al. 2000; Jager et al. 2001). These studies have shown that the domains are stabilized by the formation of two distinct hydrophobic clusters. The relative simplicity of the three-stranded β-sheet formed has led to work on designing WW like peptides (Macias et al. 2000; Kraemer-Pecore 2003) which has been meet with some success. In one study, the

residue frequency at key sites was surveyed to define a statistically probable sequence (Macias et al. 2000). Other efforts have focused on the *a priori* design of a WW sequence (Kraemer-Pecore 2003) using a predictive algorithm. Both efforts have yielded peptides that fold into WW like three-stranded sheets. Very little has been done to minimize an existing WW domain and efforts here will demonstrate the difficulties of such a research plan.

Description of WW Domains

The broad class of WW domains takes the form of three-stranded, anti-parallel β-sheets. This broad class of sequences can be divided into three categories based upon which residues are conserved: 1) sequences containing both a C-terminal proline and an N-terminal tryptophan; (2) sequences lacking the C-terminal proline; and (3) sequences lacking the N-terminal tryptophan (Macias et al. 2000). Sequences containing both the C-terminal proline and the N-terminal tryptophan are discussed here. The sequences of two WW domains as well as a designed sequence are given in Table 6-1; all numbering referred to here applies to the hPin1 sequence but is used throughout for simplicity. The β sheet of WW domains is stabilized by the formation of two hydrophobic clusters (Jager et al. 2001); one of these clusters is formed by L7, W11, Y24 and P37 (bold) while the other involves R14, Y23, and F25 (underlined). The hydrophobic cluster involving L7 and P37 forms about the sidechain of W11 and serves to bind the N and C-termini together, preventing fraying of the strands of the sheet. In the sheet itself, Y23 and F25 are involved in a cross-strand interaction with R14 which stabilizes strands one and two.

Table 6-1 Sequences of hPin1 and Mutant WW Domains†

	10	15	20	25	30	35	40
hPin1	KLPPG	WEKRM	SRSSG	RVYYF	NHITN	ASQWE	RPSGN SSS
fbp28	GA TAVSE	WTEYK	TADG	KTYYY	NNRTL	ESTWE	KPQEL K
SPANS-WW2‡	SLPSG	WTQLT	KAPD	TTYYY	NKTTD	VVTNT	RPTD
		ββββββ	Loop	ββββ	Loop	ββββ	

†Numbering applies only to hPin1
‡Literature sequence from Kraemer-Pecore et al.

Due to the location of the second hydrophobic cluster, considerable twist is introduced into the sheet at the time of its collapse. This twist, combined with a relatively short third strand, creates a cleft in which the solvent exposed W34 lies (Figure 6-1). This characteristic is what leads to the binding activity of the domain.

Analogy to the Trp-Cage

Upon analysis of the figure from Jager et al. (Figure 6-1), the presence of an analogous fold geometry was seen between the hydrophobic cluster about W11 of the hPin1 WW domain and the tryptophan in a Trp-cage system such as exendin-4 (Chapter 4) and can be seen in Figure 6-2. The hPin1 hydrophobic cluster involves Trp^{11}, Leu^{7} and Pro^{37} with a flanking Tyr (Y24) and that of exendin-4 involves Trp^{25}, Pro^{37} and Pro^{38}, again with a flanking Tyr. While exendin-4 is a long sequence that is not monomeric in aqueous solution, the analogous Trp-cage miniproteins that were developed contain the same N-terminal structure in water as exendin-4 has in 30% TFE, and are well folded, monomeric, stable systems in this medium. Though it is not immediately obvious that these motifs may be more similar than a simple tryptophan/proline hydrophobic association, Figure 6-2 shows that the two clusters form near mirror image geometry. I accordingly took a Trp-cage miniprotein and mutated one of the prolines associated with the tryptophan to leucine. Since the WW domain hydrophobic cluster involves the association of a proline and a leucine about the indole ring of a tryptophan, this mutation was designed to determine w hether leucine is as favorable as proline for

hydrophobic association with an indole ring. The CSDs of this peptide are shown in Figure 6-3. This mutation was destabilizing to the folded state, though a well-folded peptide is still observed when looking at the H_N or $H\alpha$ CSDs. This indicates that perhaps the leucine in the C-terminal portion of the WW sequence might be favorably mutated to proline.

Minimization of WW Domain

The plan for minimization of the WW domain initially involves the loss of binding activity, due to the elimination of the unfavorably solvent exposed tryptophan. Once the goal of a stable, minimized peptide is realized, binding activity could be reengineered into the system. The first studies that were done recognized the importance of the hydrophobic cluster that involves residues near the N- and C-termini and W11. It was believed that the association of L7 and P37 with W11 is of particular importance for the stability of the sheet, so studies were undertaken to optimize this interaction. I turned my attention first to L7 and looked at its role in an N-capping motif.

N-cap Motif

Before full scale mutational studies of a WW domain were undertaken, smaller motifs that emulated the association of W11 with L7 near the N-terminus were made. These peptides were designed around the scaffold of a designed hairpin that contains an NPATGR turn and the key to determining the success of the interaction is a large CSD due to the close association of the indole ring with one of the side-chains of the N-capping motif. These N-cap motifs are given in Table 6-2.

Table 6-2: WW N-cap motifs.

WW Ncap 1	SPPPGG	WNPATGRYT
WW Ncap 2	KPLPPG	TWNPATGRYT
WW Ncap 3	KPLPPGG	TWNPATGRYT

The first sequence, WW Ncap 1 was designed to test for an interaction between proline, and the tryptophan of the hairpin, in some analogy to Trp-cage motifs. The other designs returned to the leucine containing motif and also increased the length of the N-terminal strand of the hairpin by addition of a threonine residue at the S-2 position. WW Ncap 2 also reduced the length of the linker to the N-capping moiety (glycines) to one residue but WW Ncap 3 retained a 2 glycine linker. Figure 6-4 presents the CSDs of these peptides and shows that none of these ideas produced an effective N-capping unit. The hairpin is well folded in all cases, as evidenced by significant downfield H_N CSDs for the Trp at S-2 and a well defined turn region, but in all cases there is a lack of significant CSDs in the N-capping portion. While the indication from Ncap 1 is that the proline substitution is not an effective way to create this motif, the lack of success in any of the variants shows that there must be a flaw in the design of all systems, independent of the hydrophobic residues. As the data from Ncap 1 and 2 became available, it was hypothesized that the reason for the failure to form the desired hydrophobic cluster was that the linker or the N-capping motif was of an insufficient length to allow for association with the tryptophan sidechain. This is the reason WW Ncap 3 was designed with a longer linker. Since this attempt also failed to produce a viable sequence, I switched to working with full length WW domain sequence.

Minimized WW Domains

After the N-capping interaction was explored, the design of a minimized WW system focused on a dramatic redesign of the hPin1 system. Based upon the observation that the loop forming regions of this domain are quite long and poorly converged and the strand residues are not optimized for extended conformation, much of the sequence was

changed to residues that I knew were favorable for three-stranded sheet formation. The first sequence, WWm1b (Table 6-3), has an NG turn at the T1 locus and a PATG turn at T2 while the C-terminal tag was removed. Strand 1 was also completely redesigned to provide for a higher sheet population. This included moving the C-terminal Trp closer to loop to to facilitate a cross strand interaction between Y24 and W31. The H_N CSDs of this peptide (Figure 6-5) in both aqueous solution and fluoroalcohol containing media show that the 2nd hairpin is well folded while the first strand does not associate even in the structure favoring media. It was hypothesized that the unfolding of the first strand was caused by both unfavorable turn geometry at T1 and the lack of a hydrophobic interaction between the termini, as is seen in the hPin1 sequence. In fluoroalcohol containing medium, loop 1 shows the characteristic CSDs of a [3:5] turn though strand alignment is still lacking. To rectify these problems, WWm2 was made.

Table 6-3 Sequences of WW Domain Mutants

	10	15	20	25	30	35	40
hPin1	KLPPG	WEKRM	SRSSG	RVYYF	NHITN	ASQWE	RPSGN SSS
WWm1b	KPLPGG	WTYRT	VPNGK	TYYYN	PATGT	WTSK	
WWm2	KLPGG	WTYRT	VADGK	TYYYN	PATGT	WTNER	PS
		ββββββ	Loop	ββββ	Loop	ββββ	

A statistical analysis of many WW domain sequences first showed that the location and identity of the residues of the hydrophobic clusters was highly conserved. This was done to determine which residues are highly conserved, besides those involved in the hydrophobic clusters. This analysis showed that the first turn is usually a loop of five residues which indicated that a two-residue loop may have caused unfavorable turn geometry. WWm2 builds on the WWm1b sequence but changes T1 from (P)NG to ADG and lengthens the N-terminal section while including a proline C-terminal unit. I hoped a three-residue turn would be more favorable for this strands alignment and that if the

termini could form a favorable hydrophobic interaction, then sheet formation would necessarily occur. The H_N CSDs of this peptide in aqueous and fluoroalcohol containing media are reported in Figure 6-6. NMR spectra of the peptide were not fully assignable. Many resonances were broadened beyond the experimental detection limits, implying partial tertiary structure formation on a >10μs timescale. In the aqueous data (filled bars), evidence for the alignment of the third strand is present and some resonances from the first strand were able to be assigned which indicate that there is perhaps some structuring of this strand. No resonances from the central strand could be assigned. The data acquired in aqueous media with 20% HFIP (open bars) show the structuring of the third strand, but here it is the first strand that could not be assigned. The central strand shows large CSDs as would be expected for a well formed hairpin but the random coil shifts about the turn 1 locus indicate that the first strand is not formed under these conditions. If the first strand were structured, the CSD of Y8 would be expected to be on the order of that measured for Y18, furthering the notion that the first strand is unstructured in the 20% HFIP medium. Because this data was not encouraging about the success of this project, I returned to designed three-stranded sheets to conclude my thesis project.

Conclusion

The minimization of a WW domain is a challenging but feasible project. The successful sequence will likely retain the incorporation of N- and C-terminal residues in the same hydrophobic cluster. This property is challenging to design into an optimized three-stranded β-sheet because these motifs generally result in the termini being far from each other in space. One possible solution is optimization of a fold after the introduction of a disulfide bond to pin the termini together in a near native conformation.

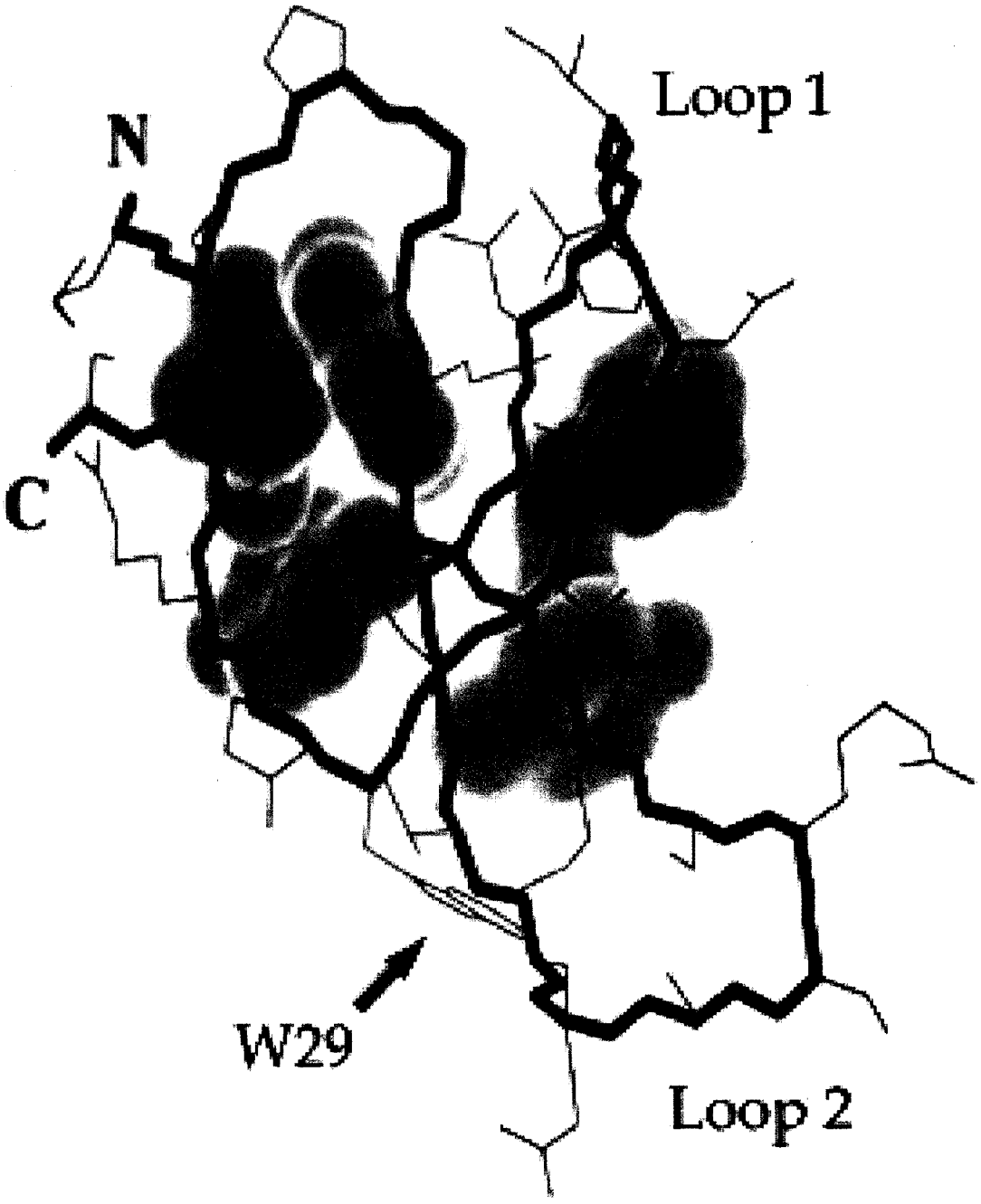


Figure 6-1: Representation of WW domain.
Hydrophobic clusters shown in blue and green. Adapted from Jager et. al.

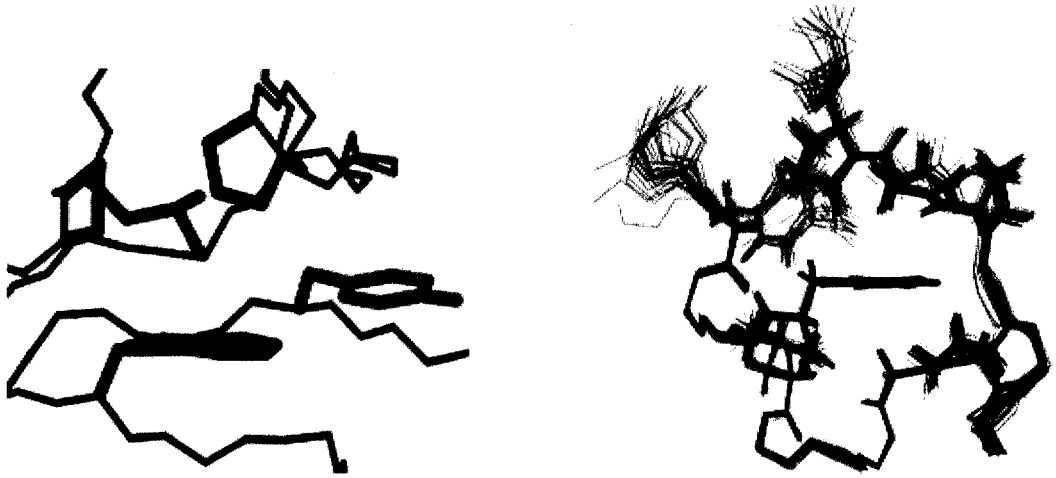


Figure 6-2: Hydrophobic Clusters of Trp-Cage Miniproteins and WW Domain. Figures rendered to show similar sidechain locations in both systems. Analagous positions shown in like colors to indicate mirror-image geometry.

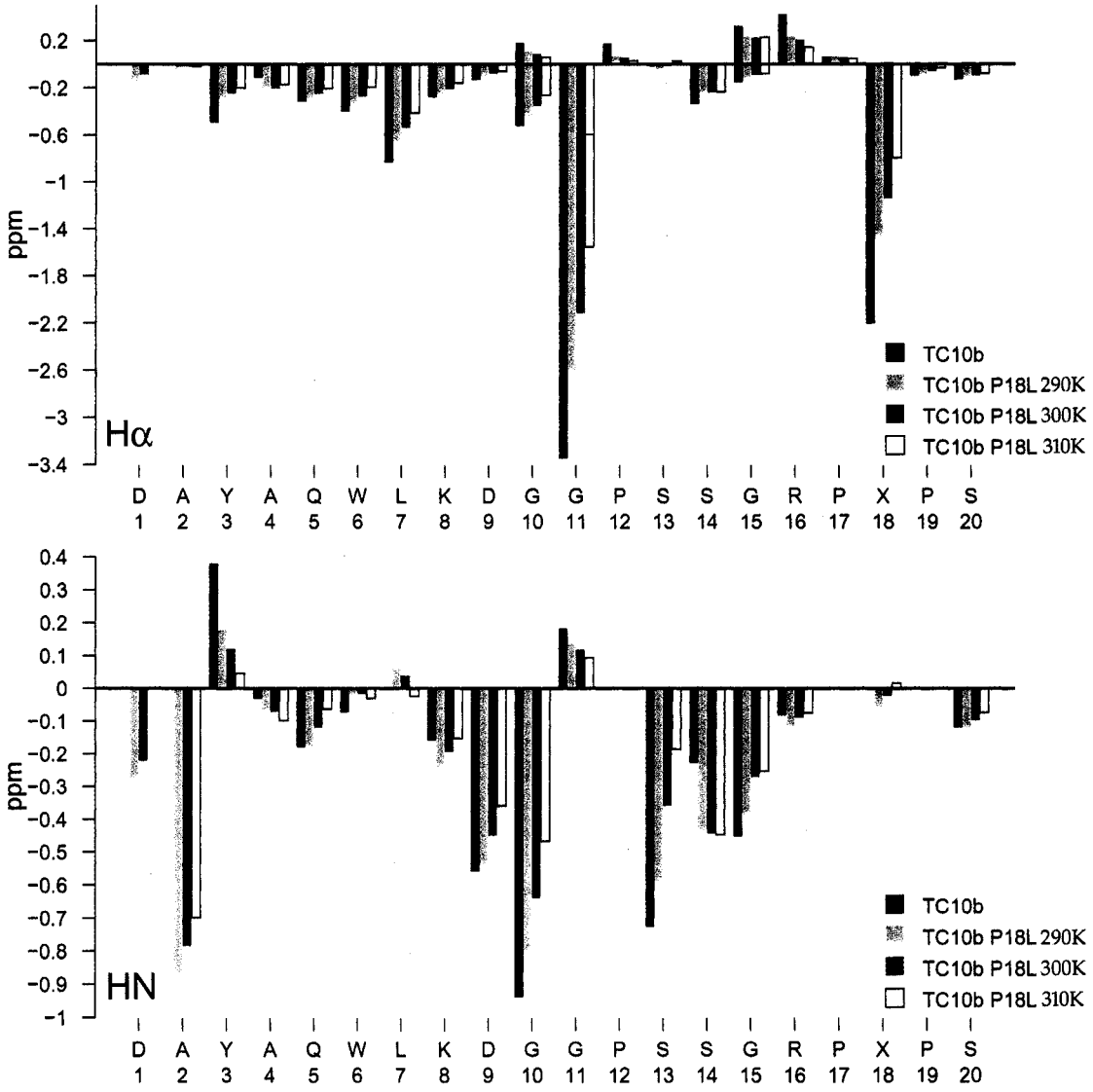


Figure 6-3: Hα and H_N CSDs of Trp-Cage miniproteins.
All data shown in aqueous solution. TC10b shows larger CSDs at virtually every site, compared to TC10bP18L.

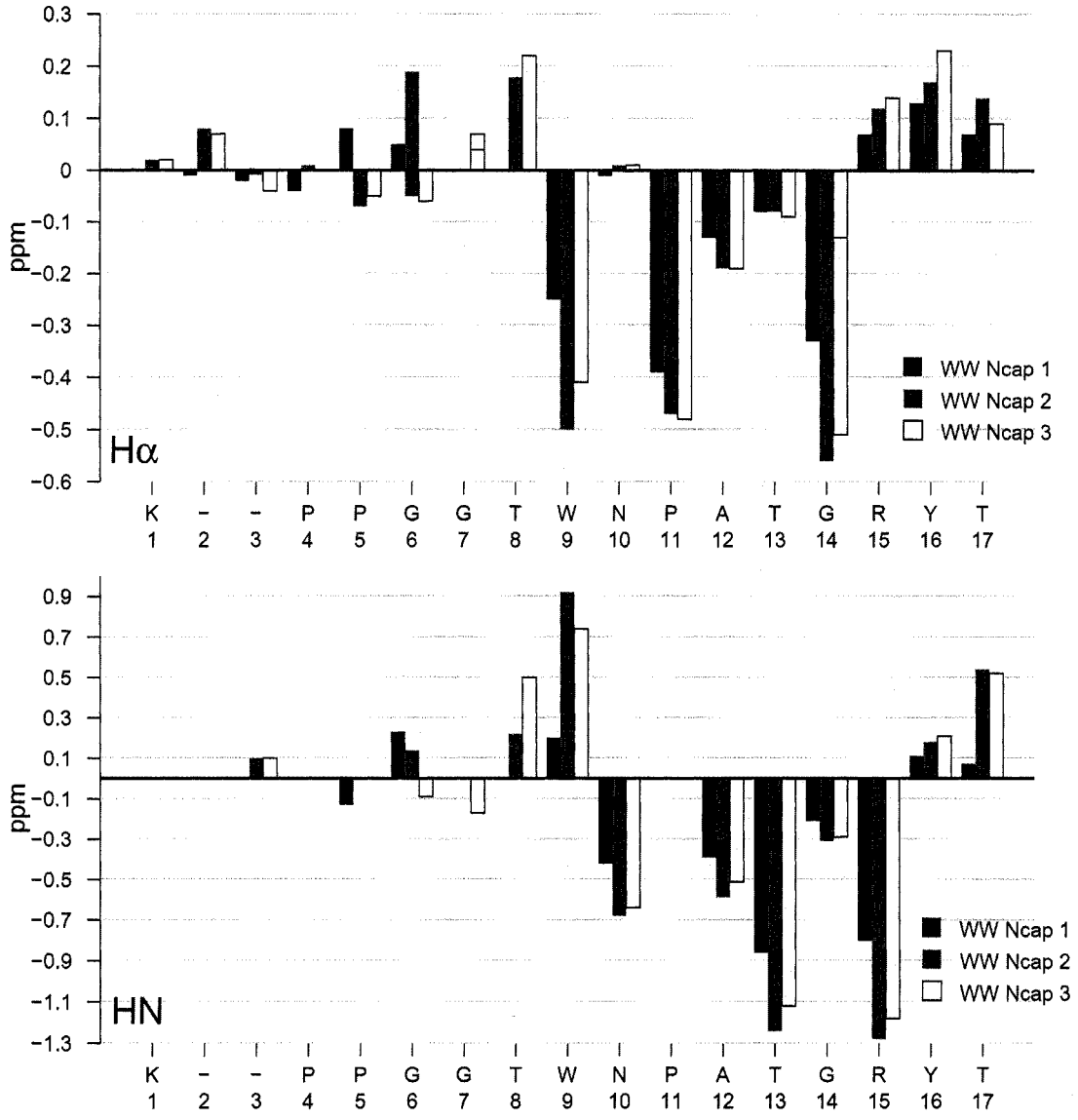


Figure 6-4: Hα and H_N CSDs for WW N-cap motifs.
All data shown at 290K in aqueous solution.

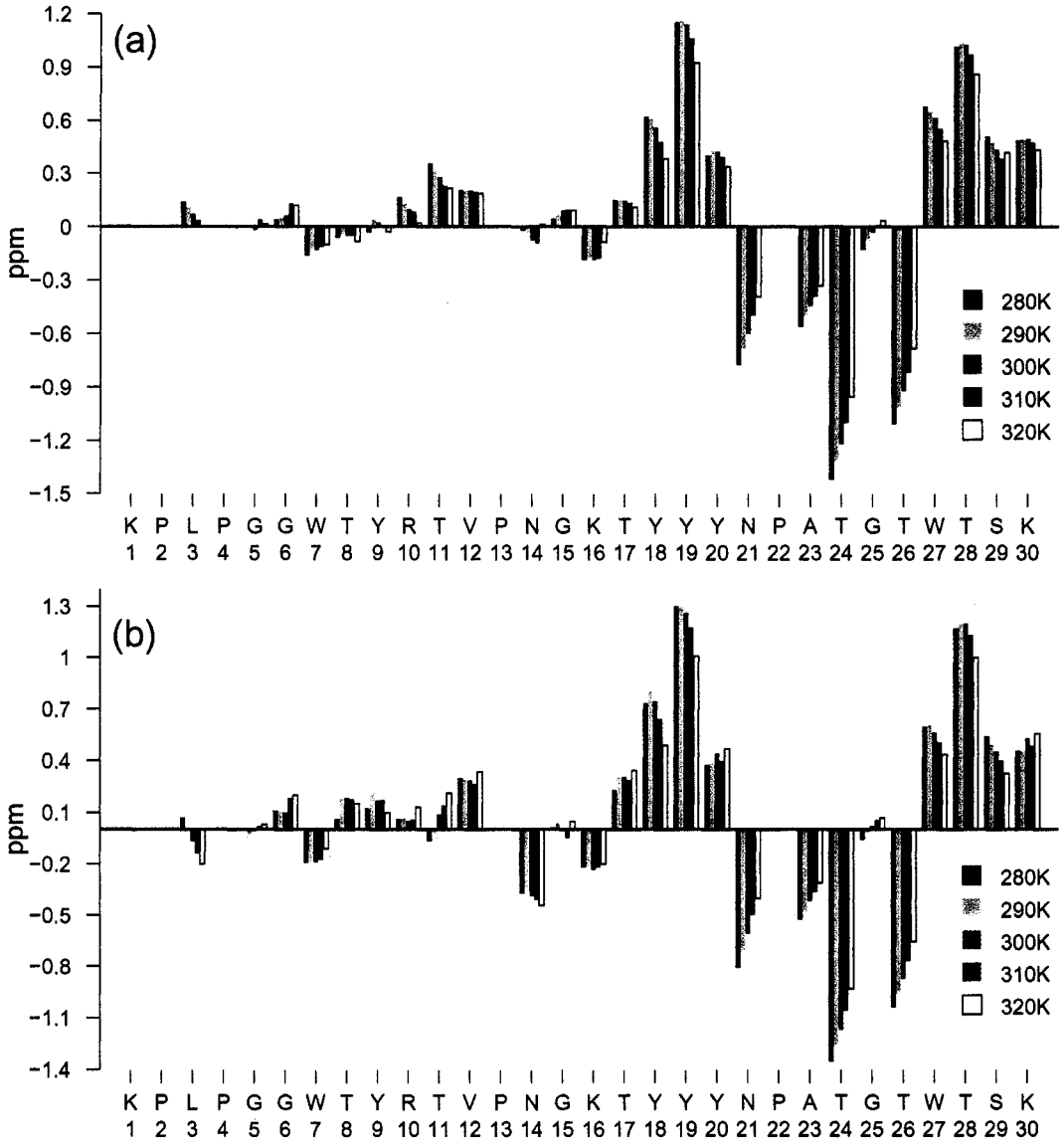


Figure 6-5: H_N CSDs of WWm1b.
CSDs shown in aqueous medium (a) and fluoroalcohol containing medium (b). Panel (b) shows diagnostic CSD for [3:5] turn formation and indicates turn sequence.

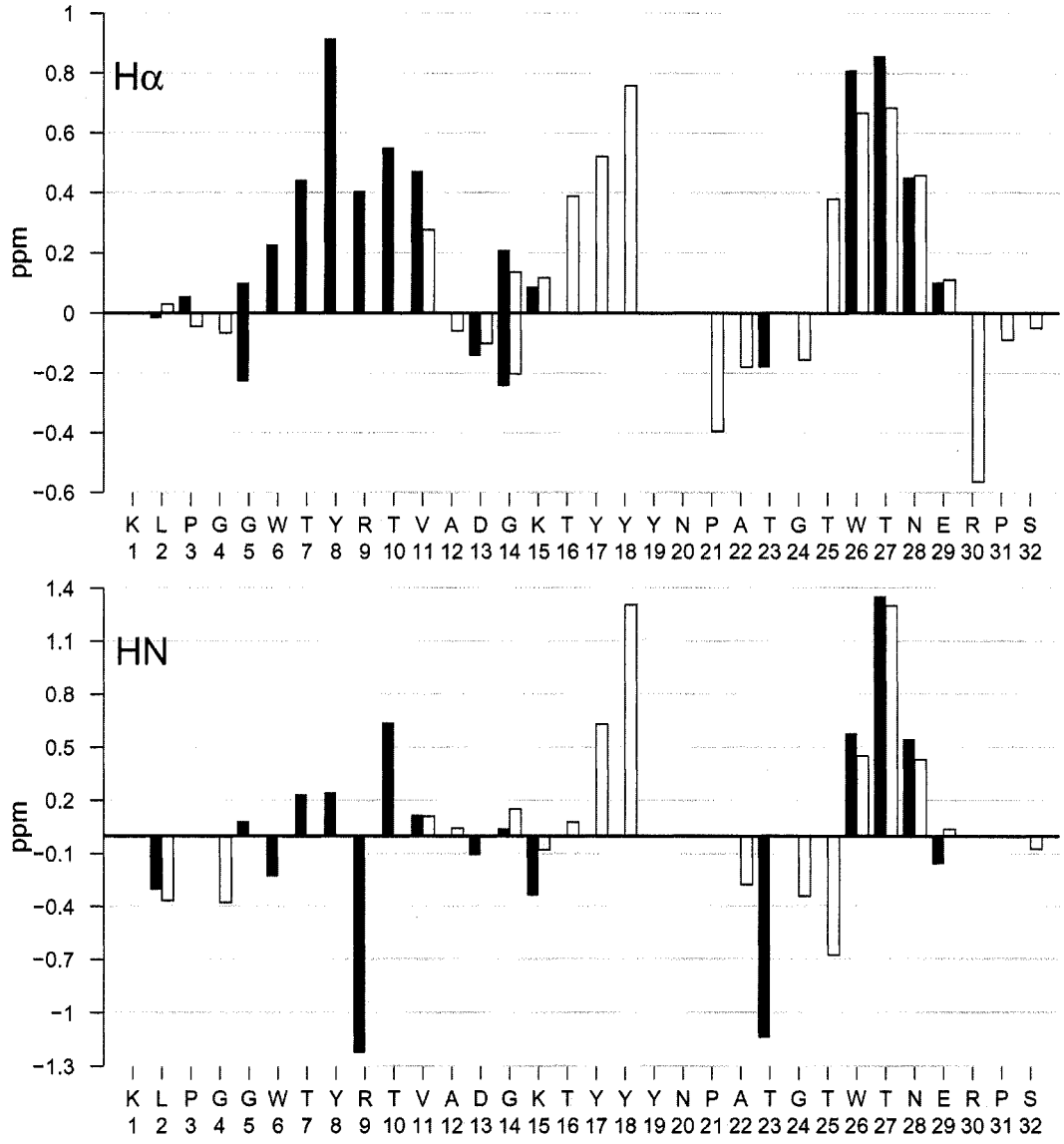


Figure 6-6: Hα and H_N CSDs for WWm2.
Data shown for both aqueous (filled bars) and fluoroalcohol containing (open bars) media at 290K.

Chapter 7 Folding Cooperativity and Dynamics of Three-Stranded Sheet Models

Introduction

Sequences that fold into highly populated β hairpins and sheets have been elusive to researchers in peptide design until recently. Only a few hairpins that fold outside the protein context were found, *eg.* the ubiquitin(1-17) sequence (Cox et al. 1993) and the second hairpin (residues 41-56) of the B1 domain of protein G, GB1p(Kobayashi et al. 1993; Blanco et al. 1994b; Honda et al. 2000; Fesinmeyer et al. 2004), and a priori peptides that populate β structures in water were not designed until the early '90's (Blanco et al. 1993; de Alba et al. 1996; Ramírez-Alvarado et al. 1996). Studies have since shown that significant hairpin populations in water are not possible without the incorporation of (turn) sequences that are particularly favorable for chain reversal, notably Asn-Gly (NG) (Sibanda and Thornton 1991; Ramírez-Alvarado et al. 1996; Maynard et al. 1998), ^{D}Pro-Gly (pG) (Karle et al. 1996; Haque and Gellman 1997; Syud et al. 1999) and Pro-Asp-Gly (PDG) (Blanco et al. 1993; Searle et al. 1995), *and* the inclusion of residues favoring an extended conformation, allowing for progress in this field. These discoveries also allowed for the design of sequences that formed measurable populations of three-stranded (Sharman and Searle 1997; Kortemme et al. 1998; Schenck and Gellman 1998; de Alba et al. 1999b; Griffiths-Jones and Searle 2000; López de la Paz et al. 2001; Santiveri et al. 2003) and four-stranded (Carulla et al. 2002; Syud et al. 2003) β-sheet models.

Equilibrium Measures and Cooperativity

Equilibrium measures of folding can be applied to problems in peptide folding cooperativity because the determination of cooperativity involves measuring the

populations of the various states accessible to the peptide. The site-specific nature of NMR chemical shift data allows for the determination of the effects of mutation on a portion of the sequence remote from the mutation site, providing information on how equilibria change in response to mutation. Since β-hairpin fold lifetimes ($1/k_u$) are short, less than 50μs (Muñoz and Serrano 1997; Xu et al. 2003; Dyer et al. 2004; Dyer et al. 2005), chemical shift deviations (CSDs) *are* population weighted averages and provide fold populations when the unfolded state approximates random coil norms and the chemical shifts for the fully folded state can be measured or accurately approximated. This property of the chemical shift of backbone protons has been used to estimate the populations of hairpin systems since the earliest observations (Blanco et al. 1994b; Searle et al. 1995) and has recently been applied to β sheets, with the results suggesting some degree of cooperativity. The effects of lengthening β-strands (Stanger et al. 2001) and increasing the number of strands have been examined, indicating the distinct effects of these two modes. Initial reports indicated the stability of a hairpin was improved upon strand lengthening from five to seven residues (Stanger et al. 2001) but no further gains were realized upon addition of two more residues to each strand. A significant effect due to an increase in the number of strands was also reported (Sharman and Searle 1997; Schenck and Gellman 1998); the population of a hairpin was greater in the context of a three-stranded system than isolated. Others however, observed little or no stabilization upon addition of a third strand (de Alba et al. 1999b).

Hairpin and sheet systems have been used to test hypotheses regarding the cooperativity of sheet formation for some time. The Searle group has measured the ΔΔG resulting from hairpins going from an isolated to a three-stranded sheet context with the intent of quantifying the cooperativity effect. In the original systems, both the original hairpin and the three-stranded sheet were poorly folded in water, allowing for comparison only in 50% methanol. They measured 3.2 kJ (Sharman and Searle 1998) of stabilization upon going to the three-stranded sheet context. Another of their designed

systems had a reasonably populated folded state in water yet displayed only 1.1 kJ of stabilization upon addition of a third strand (Griffiths-Jones and Searle 2000). Hairpin populations in these studies were determined by the difference in chemical shifts ($\Delta\delta$) of the diastereotopic α-protons of the Gly residue in the turn. This method for determining cooperativity, because it assumes the differential folding of two hairpins in the system, showed the need for a four-state equilibrium model (Griffiths-Jones and Searle 2000) but still used a two-state analysis to provide estimates of the ΔG_F contribution of strand addition. Other quantitative studies (Espinosa and Gellman 2000; Syud et al. 2001; Espinosa et al. 2002; Syud et al. 2003) have relied on the hydrogen-bonded residue (S±odd) Hα CSDs for measures of folding. Instead of measuring the change in population of a hairpin upon adding a third strand, this study estimated hairpin formation in three- and four-stranded model sheets before and after the disruption of the turn required for the alignment of the N-terminal strand. The change in fold propensity was determined to be 1.3-1.8 kJ/mol when the unaltered hairpin contained a pG turn (Syud et al. 2003). The effect was reported to be 2.1±0.5 kJ when the turn of the unaltered hairpin is NG.

While 2-state melting of model β-sheets has been suggested (Santiveri et al. 2004), differential melting of the two hairpins has also been observed (Griffiths-Jones and Searle 2000; Fesinmeyer et al. 2005). Searle and Gellman group publications (Griffiths-Jones and Searle 2000; Syud et al. 2003; Searle and Ciani 2004) agree that three-stranded sheet folding should, in principle, be viewed as a four-state process (Figure 7-1) but have not measured the increase in equilibrium constant associated with the preorganization of a strand in a stable hairpin in this context. Molecular dynamics simulations have been performed on the SG sequence and confirm the validity of a three or four-state equilibrium model but also suggest a larger cooperativity effect (Roe et al. 2005). The incremental stabilization of the hairpin, upon structuring of the remaining strand, was 3 kcal/mol in the case of either hairpin 1 or 2, suggesting sequence independence of the

cooperativity effect. This study also indicated that hairpin 1 does not form in the absence of hairpin 2, concluding that the folding pathway must proceed through hairpin 2.

The studies that relied on a two-state assumption determined changes in stability that are smaller than one would expect based upon entropic considerations which prompted this examination of the SG three-stranded sheet in the context of a four-state assumption. Dynamic NMR measurements of hairpin formation in the sheet were also conducted to determine the effects on the kinetics of hairpin formation in the presence and absence of the three-stranded sheet context.

Dynamics of Hairpin and Sheet Formation

Experiments in protein engineering have shown that the stabilization of a β-hairpin can both accelerate (Nauli et al. 2001; Kuhlman et al. 2002; Bofill et al. 2005) and retard (Kuhlman et al. 2002; Platt et al. 2003) the folding rate of the host protein. Since the implications of this are that these structures are involved in the transition states of folding pathways, it is critical to understand the effects of mutation on the folding rates of these structures. The dynamics of hairpin folding have garnered much interest and the body of work in this field (Muñoz et al. 1997; Andersen 2001; Xu et al. 2003; Du et al. 2004; Dyer et al. 2004; Snow et al. 2004; Dyer et al. 2005) has grown extensively of late. These studies have searched for the factors affecting the folding rate of hairpin systems.

The first microsecond timescale kinetic measurements of hairpin formation were carried out on a 16-residue peptide derived from GB1 (Muñoz et al. 1997). This study utilized fluorescence monitored T-jump technique to watch the unfolding of the peptide and determined a folding time constant of 6μs at 297K. These researchers later proposed a model to describe the apparent two-state kinetics of hairpin formation (Muñoz et al.

1998). In this model, hairpin formation begins with nucleation of the turn and is followed by the sequential formation of inter-strand hydrogen bonds. The process of hydrogen bond formation, because the angles about two peptide bonds must be fixed per H-bond, was said to be continuously uphill, unlike helix propagation which is limited only by the rate at which a single residue can be added to the growing structure (Muñoz and Serrano 1997). A computational study (Dinner et al. 1999) of this GB1 derived peptide however, suggested that hydrophobic collapse was the first event in the hairpin folding process. In this study, utilizing multi-canonical Monte Carlo simulations, the rate limiting step was associated with the rearrangement of the collapsed structures to form the native hydrophobic contacts.

Subsequently, other studies have implicated turn formation in the transition state of β-structure formation: the turn forming propensity of the loop region affects hairpin folding rates much more than strand optimization (Du et al. 2004) and the length of the region between the hydrophobic clustering residues affects the folding rate of the hairpin (Dyer et al. 2004). The report from Du et al. suggested that the formation of a hairpin proceeds through turn nucleation followed by inter-strand hydrogen bond formation; a so called "zippering" model. The reports from Dyer et al. however, were rationalized by a folding mechanism initiated by hydrophobic collapse placing big fatty and aromatic residues in a native-like arrangement without a well-structured turn or a full complement of cross-strand H-bonds present. Another study suggested that either initial hydrophobic collapse followed by turn formation, *or* turn formation followed by strand hydrophobic collapse can be favored depending upon sequence (Olsen et al. 2005a).

Currently, no experiments have been conducted to determine the degree to which the effects on the stability of one portion of a secondary structural unit affect the dynamics of the unaltered portion of the structure. As was seen in Chapter 5, the SG three-

stranded-sheet model retains significant hairpin formation upon destabilization of the other turn of the sequence. It is, therefore, an ideal candidate for the measurement of the change in folding and unfolding rates in an unaltered hairpin, when the propensity for the formation of the second hairpin is varied. In the present study, the degree to which the folding of the SG three-stranded sheet is cooperative is probed by equilibrium measures of folding, and a dynamic NMR technique is used to measure folding and unfolding rates of an analog of the 1st hairpin of the sequence, upon going from an isolated hairpin to the three-stranded sheet context.

Cooperativity of SG Peptides

The cooperativity of the original SG peptide and its NG-NG turn mutant were probed by the utilization of the CSD method for determining equilibrium fold populations applied to individual hairpins within three-stranded-sheet peptides. The peptides required for this study are given in Table 7-1. The mutational strategy pursued for both the SG and SG (NG-NG) peptides was one which provided peptides forming exclusively one hairpin or the other. In addition, mutants of the SG (NG-NG) peptide were made to sequentially place the highly favorable pG sequence or the non-turn forming PG sequence at each turn locus. The fraction folded (f_F) values for the 1st and 2nd hairpin of these peptides were then determined from the appropriate terminal strand CSDs and these values are reported in Table 7-2. These fraction folded values were determined relative to the SG (pG-pG) peptide in 8% HFIP, which was assumed to be 92% folded about both turns; providing peptide populations that are consistent with Gellman reports (Syud et al. 2003).

Table 7-1: SG and Mutants:

SG	Ac-VFITSpGKTYTEVpGOKILQ-NH_2
SG PG-pG	Ac-VFITS**PG**KTYTEVpGOKILQ-NH_2
SG pG-PG	Ac-VFITSpGKTYTEV**PG**OKILQ-NH_2
SG NG-NG	Ac-VFITSNGKTYTEVNGOKILQ-NH_2
SG PG-NG†	Ac-VFITS**PG**KTYTEVNGOKILQ-NH_2
SG NG-PG	Ac-VFITSNGKTYTEV**PG**OKILQ-NH_2
SG pG-NG†	Ac-VFITS**pG**KTYTEVNGOKILQ-NH_2
SG NG-pG	Ac-VFITSNGKTYTEV**pG**OKILQ-NH_2

†Literature sequences from Syud et al. 2003.

In the case of the pG-pG system (SG), Table 7-2 shows that destabilization of turn 1 yields a peptide (PG-pG) that forms a hairpin about turn 2 that is 60% folded while destabilization of T2 results in a hairpin (pG-PG) that is 56% folded about T1. Data from the NG-NG system shows that destabilization of hairpin 1 leads to f_F of 0.30 at hairpin 2 while prohibition of turn 2 leads to a f_F of 0.10 about turn 1. This data, which shows that either hairpin 1 or 2 can form when the docking of a third strand is precluded, contrasts with MD simulations for SG(pG-pG) (Roe et al. 2005) that indicated a 3-state model is appropriate. The authors drew this conclusion based upon their observation that hairpin 1 could not form in the absence of the three-stranded-sheet context. Additional support for my use of a four-state model is provided by the pG-GG peptide which has a higher population of hairpin 1 than hairpin 2. Since the current data shows a measurable population of hairpin 1 in both T2=PG mutants, the need for a four-state model to describe these systems is confirmed and will be applied to the analysis of data from both pG and NG turn containing peptides.

Table 7-2: Fraction folded values for hairpins in SG peptide constructs.

	$\chi_{(N+F)}$	$\chi_{(C+F)}$
	weighted ave.[a]	weighted ave.[a]
pG -- pG	0.718 ± 0.063	0.812 ± 0.055
PG -- pG	< 0.04 >	0.597 ± 0.083
pG -- PG	0.557 ± 0.023	< 0.06 >
NG -- NG	0.247 ± 0.029	0.461 ± 0.060
PG -- NG	0.00 ± 0.02	0.300 ± 0.032
NG -- PG	0.104 ± 0.024	0.040 ± 0.034
pG -- NG	0.668 ± 0.019	0.563 ± 0.091
NG -- pG	0.365 ± 0.038	0.682 ± 0.027
pG -- GG	0.682 ± 0.030	0.552 ± 0.030

[a] Since larger structuring shifts provide f_F measures with higher precision, the T4α and I3H_N CSDs were doubly weighted to obtain the 1st hairpin f_F and the I18H_N CSD was double weighted in obtaining the 2nd hairpin f_F.

Four-State Analysis

The four-state model used in conjunction with this study is presented in Figure 7-1 and includes states in which either the N-terminal (N) or C-terminal (C) hairpin is formed, in addition to U and F. Because these systems apparently conform to such a folding model, and CSDs are equilibrium measures of folding, we cannot directly measure the population of any of the states with the wild type peptide or the NG-NG mutant. This is because the determination of population at either hairpin 1 or 2 corresponds to the sum of both the folded state (χ_F) and the partially folded state of the corresponding hairpin (χ_N or χ_C) at which the measurement is made [$\chi_{(N+F)}$ and $\chi_{(C+F)}$]. Therefore, an algebraic solution providing the equilibrium folded fractions populating each state is only possible when it is assumed that the XG-PG and PG-XG (where X=p or N) turn mutants given in Table 1 provide both the unfolded reference CSD values about the destabilized

turn *and* equilibrium hairpin populations about the undisturbed turn that correspond to the N and C states for both the SG and SG NG-NG peptides. We must also assume that no direct inter-conversion between the unfolded and folded states (U ↔ F) or the N and C states (N ↔ C) can occur in order to find an algebraic solution to this problem.

The quantitation of $\chi_{(N+F)}$ and $\chi_{(C+F)}$ from the SG and SG NG-NG peptides and the determination of the fractions populating N and C from the turn destabilized mutants provide the necessary data for calculating the fractions populating all states. From these mole fractions, the cooperativity of the system can be defined as in Equation 7-1 and the equilibrium constants required for its calculation can be derived from the following procedure.

Equation 7-1: Definition of the cooperativity index.

$$K_{NF} = \mathbb{C}\, K_{UC}$$

or

$$K_{CF} = \mathbb{C}\, K_{UN}$$

Turning to the algebraic solution, we know that stoichiometry provides that:

Equation 7-2: Stoichiometric description of a four-state system.

$$\chi_U + \chi_N + \chi_C + \chi_F = 1.00$$

and since we can measure $\chi_{(N+F)}$ and $\chi_{(C+F)}$, as f_F(N-term) and f_F(C-term), we have a way of eliminating one of the unknowns, χ_F, from equation (7-2). Substituting values for $\chi_{[N+F]}$ and $\chi_{[C+F]}$ into equation 7-2 and simplifying yields the following equality:

Equation 7-3: Parameter X for incorporation of measurement limits.

$$X = 1 - [\chi_{(N+F)} + \chi_{(C+F)}] = (\chi_U - \chi_F)$$

The stoichiometry condition can then be restated as:

$$2\,\chi_U + \chi_U K_{UN} + \chi_U K_{UC} - X = 1.00$$

Which rearranges to give: $\chi_U = (1 + X) / (2 + K_{UN} + K_{UC})$

At this point it is possible to calculate χ_U from only observables which can be entered as arbitrary variables: X a function of f_F(N-term) and f_F(C-term), and K_{UN} and K_{UC}, for which estimates are available from f_F(N-term) of SG(XG-PG) and f_F(C-term) of SG(PG-XG), respectively. For each calculated χ_U value, we can calculate χ_N and χ_C because we have the correlated values of K_{UN} and K_{UC} :

$$\chi_N = \chi_U K_{UN}$$

and

$$\chi_C = \chi_U K_{UC}$$

The corresponding values of χ_F are obtained by substituting χ_N, χ_C, and χ_U into equation 7-2. At this point, the two remaining equilibrium constants K_{NF} and K_{CF} can be calculated and the cooperativity index $\mathbb{C}$ can then be calculated from equation 7-1. The cyclic nature of a four-state system requires the two methods for calculating the cooperativity index to produce the same solution, giving one cooperativity index per system.

Results

The key experimental factor for the calculation of $\mathbb{C}$ turns out to be the difference in the folded populations of the two hairpins in each system. Since this difference is larger for the NG-NG peptide than SG, and previous 2- and 3-state estimates predict a larger effect about NG turns, the NG-NG analysis is presented first. As was previous stated, the PG-NG and NG-PG peptides provide the K_{UC} and K_{UN} values as derived from the f_F of the C

and N terminal hairpin, χ_C and χ_N respectively, and the NG-NG peptide provided the $\chi_{(C+F)}$ and $\chi_{(N+F)}$ values. Because of the uncertainty inherent in experimentally determined values, wide variability in $\mathbb{C}$ values was initially noted. These values were sorted and fit to all measured values resulting in the elimination of many erroneous $\mathbb{C}$ values. Ultimately, the calculations that fell within the measured tolerances provided mole fraction limits for the four states: χ_U=0.52-0.56, χ_N=0.055-0.067, χ_C=0.199-0.233 and χ_F=0.165-0.202 and the values for $\mathbb{C}$ that were determined from these were limited to the range 5.7-9.8. Using the widest experimental ranges that appear in Table 7-2, the full range of acceptable solutions found was χ_U = 0.48 – 0.57, χ_N = 0.042 – 0.077, χ_C = 0.193 – 0.26, and χ_F = 0.142 – 0.234; with the calculated cooperativity index also having an increased range – $\mathbb{C}$ = 4 – 13.6.

Turning to the cooperative boost to folding observed in the SG system with pG turns, the error limits were determined by comparison of the calculated values of χ_N, χ_C and χ_F to the measured values of $\chi_{(N+F)}$ and $\chi_{(C+F)}$, as determined in the native peptide. Additionally, solutions were deemed erroneous when the quantity, [$\chi_{(N+F)}$ - $\chi_{(C+F)}$], was less than 0.05. This yielded mathematical solutions for $\mathbb{C}$ of 2.7±0.7. As in the case of the NG-NG system, the SG system populates four states but, in this case, the system is dominated by the folded state with: χ_U =0.09-0.13, χ_N =0.09-0.15, χ_C =0.15-0.23 and χ_F =0.48-0.66. This system provides a less satisfactory solution to the equations relative to the NG-NG system because the calculated quantity [$\chi_{(N+F)}$ - $\chi_{(C+F)}$] was often significantly less than the experimental value.

Additional examinations of the robustness of the calculations were provided by a change in the assumption of the folded population of the reference peptide, SG in 8% HFIP. This reference value was reduced from 0.92 to 0.84 for both hairpins in the system and these values propagated through the calculations, resulting in values for $\mathbb{C}$ between 5.9 and 10.2 for the SG NG-NG peptide. The effect of differential folding of the hairpins

in the reference peptide was also examined. In this case, the reference was assumed to populate the N-terminal hairpin to 0.84 while the C-terminal hairpin was assumed to have f_F=0.92. This also produced $\mathbb{C}$ values that are within the tolerances previously reported and not outside the range: 5.3-10.2. The use of these different folding reference values allowed for the calculation of a "best fit" $\mathbb{C}$ value of 7.6 ±1.2 for the NG-NG peptide, independent of folding reference calibration.

Because the chemical shift probes from the center strand that indicate hairpin populations can receive some contribution to their structuring CSD due to the presence of the other hairpin, "f_F"-values obtained from individual 1st and 3rd strand probes of hairpin populations are viewed as more dependable. Therefore, the calculations were repeated for the four-state model using these to provide the limits for K_{UN}, K_{UC}, $\chi_{(N+F)}$ and $\chi_{(C+F)}$. Relying exclusively on 1st and 3rd strand measures of the folding for the folded fractions of the N- and C-terminal hairpins afford the following range of cooperativity ($\mathbb{C}$) values: 5.7 – 10.3 (with values between 6.3 and 8.4 providing the best fits) for hairpin formation about NG turns, and 2.1 – 3.8 about pG turns. Throughout, the calculated populations of all four states are significant; however, the fits for the SG pGpG were not as good, with the difference in $\chi_{(N+F)}$ and $\chi_{(C+F)}$ often being significantly less than the experimental value.

The cooperativity index, $\mathbb{C}$, can also be used to calculate the free-energy stabilization of a hairpin upon going to the three-stranded sheet context. Since $\mathbb{C}$ is a ratio of equilibrium constants, the relationship $\Delta\Delta G = -RT \ln \mathbb{C}$ holds. By this equation, the folding increment for SG NG-NG, corresponding to a $\mathbb{C}$ value of 7.6 ±1.2, gives an increase in free energy upon going to the three-stranded context of 4.9 ±0.4 kJ/mol. Since prior reports provided values based upon a two-state assumption, an equivalent analysis is presented here for comparison. The largest cooperativity increment is observed for the N-terminal hairpin of the SG NG-pG and SG NG-PG peptides and is 3.9 kJ/mol. While the four-state model

provides a larger estimate of the cooperativity of sheet formation, the measures are nonetheless smaller than one would estimate based purely upon the entropic advantage of prestructuring four residues (TYTE) in a β-strand suitable for docking in an antiparallel sheet. Values determined from protein folding suggest a TΔS value on the order of 4-6 kJ/residue (Matthews et al. 1987). Estimates generated in this lab (Barua et al. 2005) indicate 4.2 kJ/residue, based upon a Pro→Ala mutation in a Trp-cage miniprotein, and would suggest a 16 kJ advantage in the case of the SG systems. The lack of such a dramatic cooperativity increment is likely due to the significant β-propensity expected for this sequence. Since β-sheets do not typically have such large intrinsic β-propensities in their strands, one would expect to observe a larger cooperativity increment in that environment.

The analogs made for this study also provide the basis for comparing the ΔG_f increments for the different turn loci. These values have been previously presented for X→p mutations where X=N and G, earlier in this work. The mutations cause a $\Delta\Delta G_U$ of 3.7 kJ/mol in SG (XG-pG) and 3 kJ/mol in SG (pG-XG); values that can be directly compared to previously determined ΔΔG values for NG→pG mutations that have been reported for four distinct sets of β strands. These studies indicate such a turn mutation causes changes of 2.7 kJ/mol (Cochran et al. 2001a; Espinosa et al. 2002), 2.8 kJ/mol, and 3.1 kJ/mol (Fesinmeyer et al. 2004). These values are smaller than I observed when comparing the differences between SG pG-pG and SG NG-NG. Here, 4.9 and 3.9 kJ/mol of destabilization is observed for the 1st and 2nd hairpin, respectively. An even larger difference is observed when looking at the first hairpin of SG NG-PG versus SG pG-PG. Here the populations are 0.104 and 0.557, respectively, which equates to a ΔΔG of 5.7 kJ/mol for the NG→pG mutation. These observations suggest that turn propensities and strand effects are not completely additive: turn stabilization appears to be more effective when the enthalpic advantage of strand alignment is diminished.

The two-state analyses of hairpin formation about the first and second turns of the SG and SG NGNG peptides indicate a greater propensity for hairpin formation about turn 2. These differences in propensity ranged from 1.5-2.3 kJ/mol and are substantially less than the 12kJ difference determined by molecular dynamics simulations (Roe et al. 2005). The present four-state analysis provides energetic differences corresponding to the difference between K_{UN} and K_{UC}. While the most accurate determination involves that for SG NG-NG, $\Delta\Delta G$ = 3.15±0.25 kJ/mol, it is not clear whether this difference is sufficient to provide a folding pathway exclusively through the C-terminal hairpin. At this point I turned my attention to the kinetics of three-stranded-sheet folding.

Dynamics of SG Peptide Mutants

Dynamic ^{1}HNMR has recently been used to measure folding and unfolding rates in peptides. This linewidth measurement technique was, for example, implemented in the measurement of rates of GB1p and mutants (Olsen et al. 2005a). In the course of study of the SG sequence and the various turn mutants prepared, I determined that these sequences did not display the large CSDs necessary for an analysis of this sort. As a result, a sequence needed to be designed with the interactions required to make such an analysis possible. Prior work with Trp-zip-like peptides showed that a W/W cross-strand interaction would provide a chemical shift probe for dynamic analysis.

Determination of Rates by ^{1}HNMR

The use of internally-referenced dynamic NMR for relaxation rate measurements requires the presence of two specific resonances in the proton spectrum: one which possesses a large difference in chemical shift ($\Delta\upsilon$) between the unfolded and the folded state and another which remains near random coil upon folding. The Trp-zip peptides and the HP7 peptides display this spectroscopic signature due to the edge-to-face orientation of the two indole rings. The Hε3 proton of one indole ring is thrust into the

face of the other upon folding, while the Hε3 proton of the other is directed into solvent in both the folded and unfolded state. This leads to a large shielding effect on one of these resonances, resulting in a large $\Delta\upsilon$ upon folding, while the other remains essentially random coil, even in a well-folded peptide. The Hε3 resonance with the large upfield $\Delta\upsilon$ is observed broadened to an extent that is proportional to the folding and unfolding rates experienced by that proton, a property known as exchange broadening (Δ_{ex}), the observed coupling constant (J), and its intrinsic linewidth (Δ°). The other Hε3 proton, because it experiences only a small change in chemical shift upon folding, does not exhibit appreciable exchange broadening so is observed as very sharp, with all broadening a function of the observed coupling constant and the intrinsic linewidth. Because both resonances used in this technique are chemically equivalent and are in the same molecule, the same coupling constant is observed and they feel the same amount of magnetic field inhomogeneity. Th is allows for the assumption that the intrinsic linewidth of both resonances is the same and is the basis for the technique that is used here to measure the folding and unfolding rates in SG-like systems.

Since the Bloch equations (Bloch 1946) provided the mathematical description for the appearance of NMR spectral lines, the relationship between lineshape and exchange rates has been known. The solution, or approximate solution, to these equations has since provided a function [g(v)] for the calculation of dynamics based upon these equations and is given in Equation 7-4 (Saika 1953).

Equation 7-4: Equations Describing NMR Lineshape.

$$g(\nu) = [(1+\tau\pi\Delta)P + QR]/(4\pi^2P^2 + R^2)$$
$$where:$$
$$P = (0.25\Delta^2 - \nu^2 + 0.25\delta\nu^2)\tau + \Delta/4\pi$$
$$Q = [-\nu - 0.5(\chi_A - \chi_B)\delta\nu]\tau$$
$$R = 0.5(\chi_A - \chi_B)\delta\nu - \nu(1+2\pi\tau\Delta)$$

The symbols used in these equations have the following definitions:

$\tau = \tau_F\tau_U/(\tau_F + \tau_U)$, where τ_F and τ_U are the average life times of the nuclei in the folded and unfolded states (s)

χ_A and χ_B are the populations of the folded and unfolded states.

$\delta\nu = \nu_F - \nu_U$, where ν_F and ν_U are the resonance frequencies of the nuclei in the folded and unfolded states (Hz)

Δ is the width at half-height (Hz) of the signal in the absence of exchange ($\tau \rightarrow \infty$) in which case Δ_F would be set equal to Δ_U for simplicity

ν is a variable frequency (Hz).

Due to the relative complexity of these equations, a simpler approximation has been derived that allows for the calculation of rates in the fast exchange limit (Sandstrom 1982; Gunther 1995) and those are presented as Equation 7-5.

Equation 7-5: Folding and Unfolding Rates Equations.

$$k_F = (4\pi\chi_U\chi_F{}^2\Delta\upsilon^2)/\Delta^{ex}{}_A$$
$$k_U = (4\pi\chi_F\chi_U{}^2\Delta\upsilon^2)/\Delta^{ex}{}_A$$

The rates are a function of fraction folded (χ_F), the amount of exchange broadening (Δ^{ex}) exhibited by the peak with the large $\Delta\upsilon$, and the change in chemical shift of this

resonance ($\Delta\upsilon$) upon going from 0% to 100% folded. These properties are all determined from a 1D proton experiment.

Determination of Fraction Folded

The timescale of β-hairpin and sheet folding requires that chemical shift deviations are population weighted averages of the folded and unfolded state. Systems containing an edge-to-face W/W interaction provide a greatly shifted Hε3 resonance whose chemical shift is related to the fraction folded of the peptide and can be used in its determination. Using this resonance provides a simple and accurate determination of fraction folded as long as the fully folded and unfolded reference values are known. The random coil shift is used to approximate the chemical shift value observed in the unfolded state, and is 7.62 ppm (Wüthrich 1986). The determination of a 100% value has proven to be more problematic because the chemical shift of the Hε3 proton can vary in fully folded peptides. Prior studies (Cochran et al. 2001a; Andersen et al. 2006) suggest values around 5.3 ppm are reasonable in fully folded systems with edge-to-face W/W interactions.

Determination of Exchange Broadening and Rates

The determination of exchange broadening requires the measurement of the half-height linewidth of both the broadened resonance and the reference resonance. The half-height linewidths are used with the coupling constant (J=7.6 Hz for Hε3) and the modeling equations (Equation 7-4) to yield Δ_{total}, the broadening that is both intrinsic and due to exchange. Δ_{total} for the unshifted reference resonance is assumed to have a negligible contribution from Δ^{ex} and is, therefore the intrinsic linewidth. It must also be assumed that the two signals have the same intrinsic linewidth for Δ^{ex} to be calculated. Δ° is then subtracted from Δ_{total} to give Δ^{ex}. Δ^{ex} is then used in Equation 7-5 to calculate the folding and unfolding rates of the peptide.

Design of SG Systems

To apply the line broadening rate analysis to the three-stranded sheet systems, a cross-strand Trp/Trp interaction was engineered into the sequence. The design of these three-stranded sheet systems was based upon what was learned from the study of the HP5 and 7 series peptides previously reported (Chapter 3) and the SG systems (Chapter 5). The HP7 system, of which a mutant had been made with the two-residue INGK loop (Table 7-3), was shown to provide for both a significant hairpin population and the desired edge-to-face interaction between the tryptophan residues flanking the loop. Incorporation of the W/W interaction into the SG sequence could be done at one of two possible sites. Looking at the NGNG species, it has an SNGK turn at the T1 locus and a VNGO turn about T2 (Table 7-3). Strictly looking at turn sequence similarity, T2 is a slightly closer match with the INGK sequence of the HP7 INGK peptide (I is more like V than S). This is not, however, the only factor that needs to be considered: the nature of the W/W interaction is highly stabilizing and the purpose of the analysis is to detect changes in rates of a sequence due to remote mutations. A sequence with a very large hairpin propensity will be less sensitive to perturbations of this sort, as was demonstrated in the Cooperativity section. Because fold population estimates are most accurate, and the sensitivity of the folded state to remote perturbation is greatest, at intermediate fold populations, it is desirable to design a sequence with this characteristic. Since the hairpin propensity was previously shown to be lower in the first hairpin of the SG NGNG sequence, the T1 locus was indicated as the mutation site more likely to yield intermediate hairpin populations.

Table 7-3 Mutants of the SG Series for Rate Analysis

HP7 INGK	KTWINGKWTE
SG NGNG	Ac-VFITSNGKTYTEVNGOKILQ-NH_2
SG NGNGW2	Ac-VFIWSNGKWYTEVNGOKILQ-NH_2
SG NGPGW2	Ac-VFIWSNGKWYTEVPGOKILQ-NH_2
SG NGpGW2	Ac-VFIWSNGKWYTEVpGOKILQ-NH_2

The mutations made to the sequence are shown in Table 7-3, which shows that the two tryptophan residues were mutated into the S±2 positions of hairpin 1 of the SG NGNG sequence while the second turn was modified to the turn prohibiting PG sequence in one case and the highly favorable pG turn in the other. Another peptide contained the NG turn at this locus. These peptides, if the interactions prove successful, will provide the change in the folding and unfolding rates of a hairpin upon going from the isolated, to the three-stranded sheet context. While CSD plots of the backbone proton resonances do not prove the existence of the appropriate geometry, they are illustrative of the folding behavior of these systems. In Figure 7-2, the CSDs of the SG W2 mutants are plotted along with the SG NG-NG peptide for reference. This shows that the W/W cross-strand pair has dramatically increased the folding propensity about T1 relative to the NG-NG peptide; even complete destabilization of T2 results in a well folded hairpin about T1, though it is measurably destabilized versus the NG-pG variant. While the large increase in fold propensity is disadvantageous for the purposes of this study, the incremental increase in the population of hairpin 1 upon three-stranded sheet formation is a positive attribute. Turning to 1D ^{1}HNMR data, the desired chemical shift of the Hε3 proton falls in an otherwise vacant region of the spectrum, far upfield of the random coil value for both W2 sequences. Figure 7-3 shows that the edge-to-face interaction between the tryptophan residues is indeed present in the mutants, as evidenced by the Hε3 resonances at about 3100 Hz (~6 ppm at 500 MHz) at 300K. Also, the reference Hε3 resonances appear around 3750 Hz (~7.5 ppm), very near to random coil expectation. These spectral qualities make the systems amenable to dynamic ^{1}HNMR rate analysis.

Application of Linewidth Analysis to SG W2 Peptides

Data of sufficient quality for rate analysis was not gathered for the NGNGW2 mutant and it will not be discussed in this context. Subsequent figures may still contain data from this mutant when available to enrich the discussion. The 1D spectra show that both SG NGpGW2 and NGPGW2 mutants had W4 Hε3 proton chemical shifts that were virtually identical at all temperatures (Figure 7-3 for 300K data). Since these chemical shifts are used to estimate fold population, the lack of observed difference in chemical shift for these two resonances would imply that the f_F of each hairpin is the same if the same 100% fold shift applied, which is problematic. The CSD plots (Figure 7-2) indicate measurable differences in fraction folded between the two peptides which make the coincidental chemical shifts for the two Hε3 protons puzzling. A possible explanation is that the association of the third strand causes a slight change in the geometry of the indole/indole association leading to less shielding and a 100% reference chemical shift that is less upfield. The different shift for *downfield* Hε3 resonances also suggests geometry or folding extent differences in the hairpin forming peptide. Measurement of the fraction folded from key strand and turn CSDs (F2α, W9α, I3N and Y10N for strands and G7α, G7N and K8N for turns) demonstrates that there is a measurable difference in the hairpin populations present in the sequences (Figure 7-4) but panel (b) shows that the chemical shifts of the Hε3 resonances do not reflect this, particularly at low temperatures. The difference in the fraction folded between hairpin 1 of the two W2 mutants was conservatively estimated to be no less than a few percent, due to these backbone resonance CSDs. This allowed for the use of a different 100% folded chemical shift reference value of about 5.5ppm for the SG NGpGW2 mutant, in contrast to the value of 5.3 ppm from Andersen et al. used for the SG NGPGW2 peptide.

With hairpin populations estimated, the folding and unfolding rates were calculated for both peptides and these results are presented in Figure 7-5. The NGPGW2 data could not be reproducibly analyzed at the lowest temperatures (<290K) due to overlap of the

reference peak and neighboring aromatic resonances. Looking at the data for the isolated hairpin, the temperature dependence of both the folding and unfolding rates is essentially linear, indicating that the activation energy, approximately equal to $\Delta H^{\ddagger}$, is not highly temperature dependent. The slope of the unfolding rates is much steeper than that for folding, a condition generally observed since the barrier to folding is smaller than that for unfolding. Turning to the three-stranded sheet, the temperature dependence of the folding and unfolding rates are not linear hints at either a substantial ΔC_p component or non two-state behavior. If we consider a four-state model, the curvature could indicate the cooperativity effect of the three-stranded sheet. In this scenario, the folding and unfolding rates of hairpin 1 (which is being monitored by the W/W intereaction) is not insensitive to the presence or absence of hairpin 2. The population of hairpin 2 in the "unfolded state" is temperature dependent; that would produce curvature for both the k_F and k_U plots. The curvature of the folding rate data could also indicate a ΔCp effect but the slower rate at lower temperature is puzzling.

Comparison of the two peptides shows that both hairpins have similar folding and unfolding rates at ambient temperature. This is unusual because stabilization of a sequence generally increases the folding rate constant (Olsen et al. 2005a). However, the similarity in these rates could indicate a four-state model is appropriate. The two-state folding and unfolding rates of a three-stranded sheet could be slower than the corresponding rates for a hairpin that is a truncation of the sheet sequence. Conversely, if the two-state model is accurate, the implication is that hairpin 1 is already formed in the transition state though the central strand may require some reorganization to reach the fully folded state.

Another striking difference between the rates of the sheet and the hairpin is that the temperature dependence of the folding rate is much greater in the sheet construct, suggesting a larger barrier to folding vs. the isolated hairpin. Turning to the differences

in unfolding rates between the two peptides, the temperature dependence of the rates is very similar at ambient temperatures but the curvature of the unfolding data for the three-stranded sheet suggests an increase in the barrier to unfolding with decreasing temperature which is reflected in an increased population of hairpin 2. The similar rates at ambient temperature that diverge with decreasing temperature suggest the influence of four-state folding in the sheet construct. In this analysis, restructuring of the central strand, which may be required to get to and from the folded state from either intermediate, creates a barrier between the intermediates and the folded state (Figure 7-6). In such an analysis, two folding pathways must be considered: the formation of the hairpin containing the dynamics probe followed by the docking of the third strand, and the formation of the other hairpin followed by the docking of the strand with the probe. In the former case, the docking of the third strand to the probe containing hairpin would cause a change in the conformation of the central strand, only after this reorganization is the stable, fully folded state reached. The data suggests this possibility because the greater temperature dependence (increasing slope) of the folding rates with decreasing temperature may indicate that the restructuring of the preformed central strand is slower at these temperatures due to an energy barrier to this process. Also, in this scenario it is reasonable for the probe to report a similar folding rate for the three-stranded sheet and the hairpin construct at ambient temperature since the formation of the sheet may not be necessary for close proximity of the probes to be achieved. The significant slowing of the folding rate with decreasing temperature might then reflect the increased population of the C-terminal hairpin upon cooling: if reorganization of the central strand is required to dock the first strand, then the apparent folding rate would have to decrease with the increase in the population of hairpin 2, and an according amount of peptide folding through this intermediate were observed. The implications of this are that the folding mechanism at ambient temperature proceeds through formation of the N-terminal hairpin first while at lower temperatures there is increased competition from the initial formation of the C-terminal hairpin. Turning to unfolding

rates, the slowing of the unfolding rate with decreasing temperatures is also reasonable to undock the first strand could require a similar reorganization of the central strand, though the effect might be less pronounced due to the entropic gain realized for the unfolding process. When viewing the unfolding process as the undocking of the third strand prior to the unfolding of hairpin 1, the observed rate would probably not be much reduced relative to the hairpin because the unfolding process would correspond well with the rate of the unfolding of the isolated hairpin. Since the k_U data show only a slight curvature upon cooling, the implication is that the unfolding mechanism proceeds predominantly through the intermediate containing the N-terminal hairpin, even at low temperature.

Conclusion

The design of the SG W2 mutants, though steps were taken to ensure an intermediate fold population, led to peptides with very stable N-terminal hairpins. Additionally, the mutants prepared showed very small differences in the fraction folded of the probe bearing N-terminal hairpin, despite the fact that the mutants with the greatest difference at hairpin two were made. These factors cloud the analysis of the data which shows that both hairpins have very similar rates. The curvature of the temperature dependence of the rates of the hairpin of the three-stranded sheet is reproducible and all data was of very high quality, lending credibility to the observation. The data for the PG mutant also provided highly reproducible data and its rates are comparable to other W-loop-W hairpins (Olsen 2006). An interesting situation would arise if a probe containing hairpin could be made with a $f_F \approx 50\%$ because the rates of the hairpin would be more influenced by stabilization or prevention of third strand formation. One possible way to create such a system would be to place the Trp/Trp pair at the S±4 positions of the strand. This has been shown to be much less stabilizing to the folded state, though some doubt exists about whether the Hε3 resonance in such a system would be sufficiently shifted and not

be overlapped by other resonances. Another strategy might be to decrease the turn propensity of the loop sequence. [3:5] loop sequences have been shown to provide decreased hairpin populations but they might also preclude the observation of a sufficiently large CSD for the Hε3 protons.

The errors in the analysis came from the estimation of fraction folded as well as peak width measurements. The difference in 100% folding references is particularly problematic, both for the assignment of the fraction folded and for the determination of $\Delta\upsilon$ (0%-100%); both values are required for the rates calculations. At low temperatures, peak overlap of the reference with neighboring aromatic resonances was a problem in the NGPG mutant precluding workup for those points. The higher temperature data for this peptide gives no indication of curvature, as in the NGpG peptide, suggesting these low temperature points would not change the overall conclusions.

Looking at the kinetic data and the equilibrium cooperativity data, it seems apparent that the partially folded intermediate of these three-stranded-sheet systems is dominated by the N-terminal hairpin for the W2 mutants. The mutations introduced in these peptides caused gains in stability to be realized in the N-terminal hairpin, reducing the fraction unfolding through the C-terminal pathway compared to the native SG sequences. The gains in stability to this hairpin increased its fold propensity sufficiently to change the probable folding pathway from one dominated by the C-terminal hairpin as was found in the SG series to one dominated by the N-terminal hairpin in the SG W2 mutants. Despite this, the curvature of the folding rates data for the W2 mutants can likely be taken as evidence that both intermediates are present, particularly at low temperature. Therefore the conclusion here supports the use of a four-state model for the equilibrium analysis to describe these sequences.

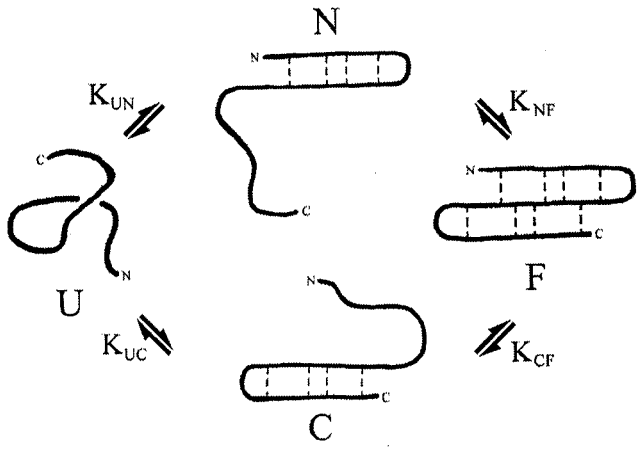


Figure 7-1. Four-state model of three-stranded sheet formation.

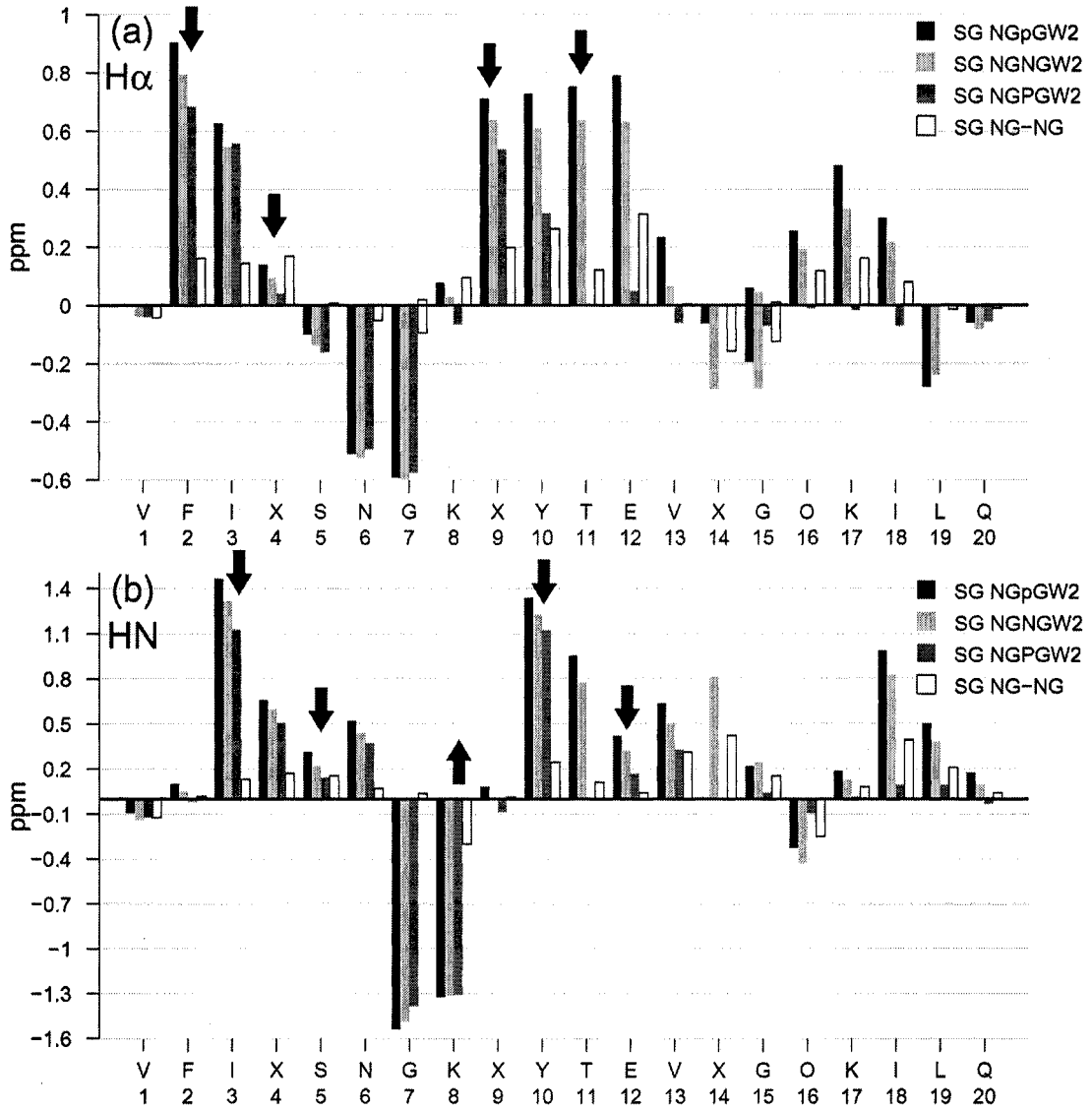


Figure 7-2. CSDs of W2 mutants and wild type sequence.
The NG-NG(W2) vs. NG-NG (no W) comparison at 290K provides another 2-state estimate of cooperativity; in this case, enhanced folding about T2. The arrows indicate cross-strand directed resonances which are indicative of hairpin 1 formation.

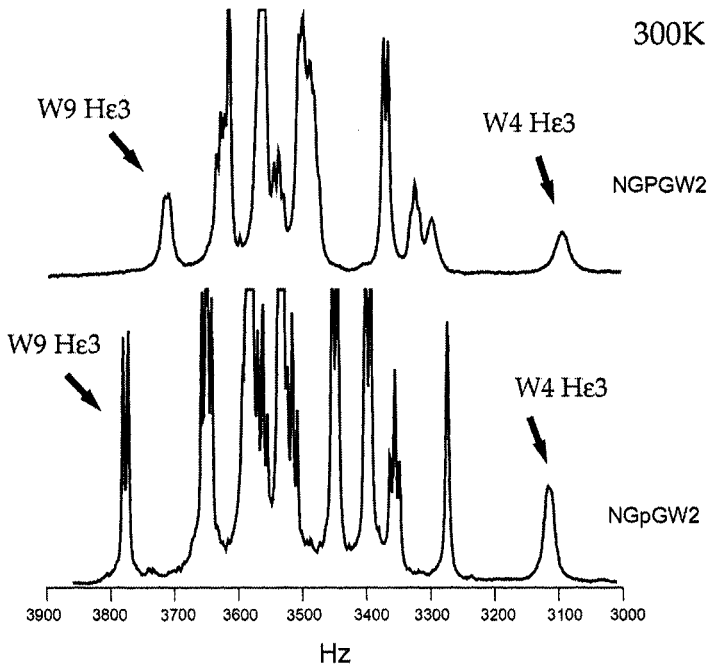


Figure 7-3. Aromatic regions of 1D ^{1}HNMR spectra of W2 mutants.
Both spectra acquired at 300K.

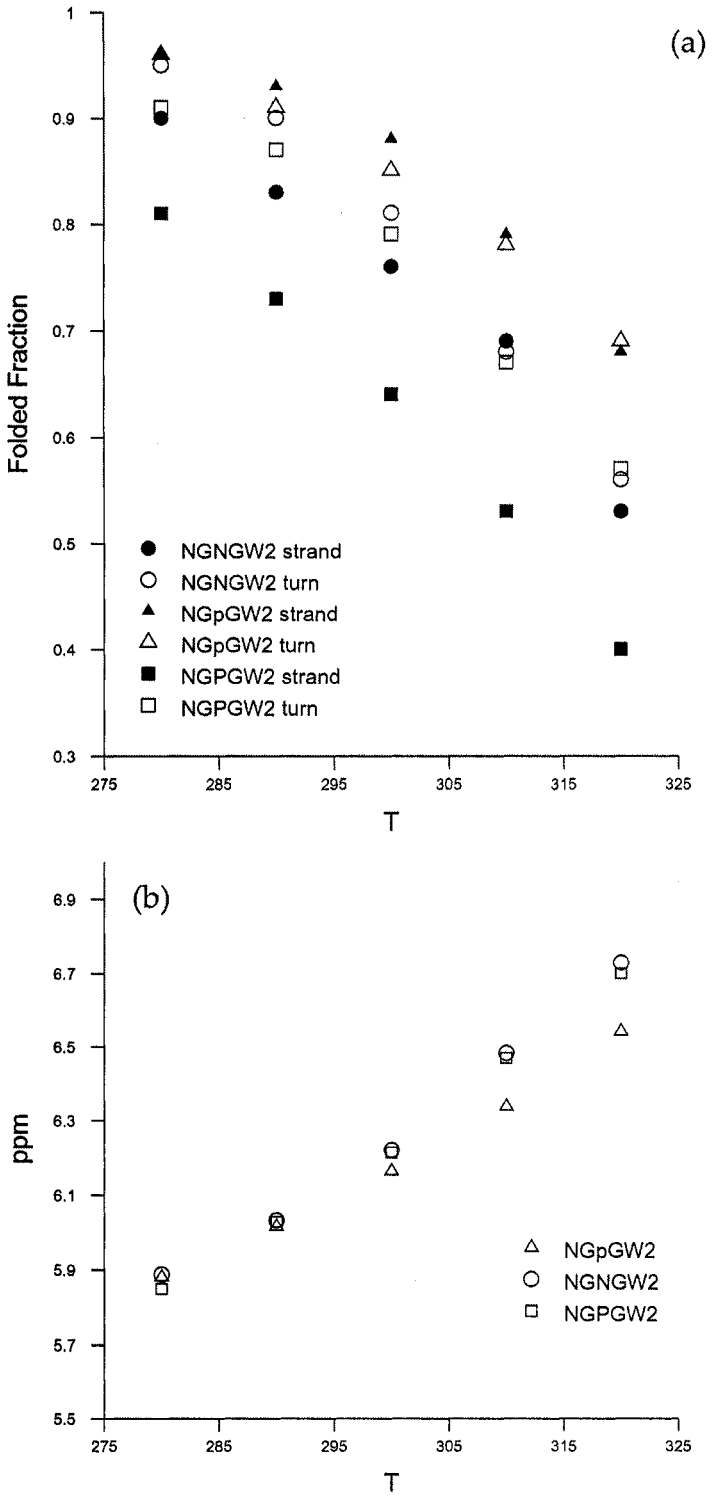


Figure 7-4. Correlation between Hε3 and % fold for W2 mutants.

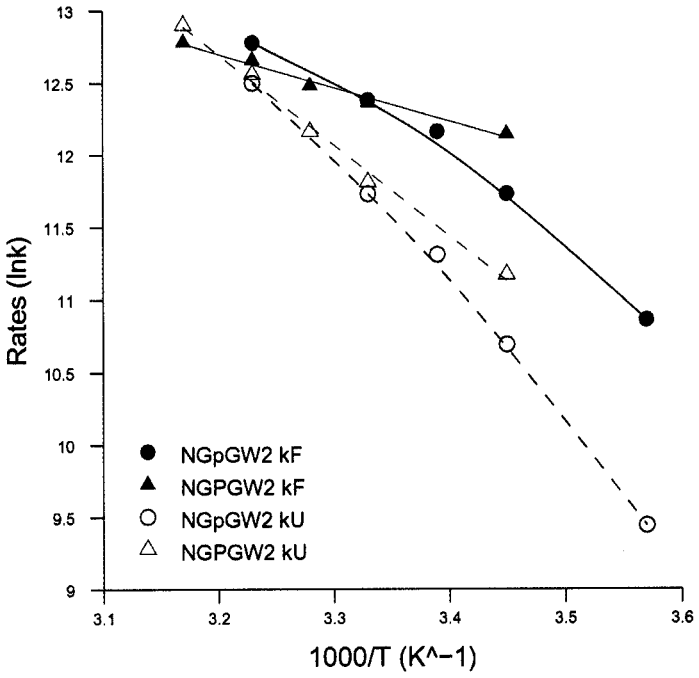


Figure 7-5. Arrhenius plots showing folding and unfolding limbs for W2 mutants.

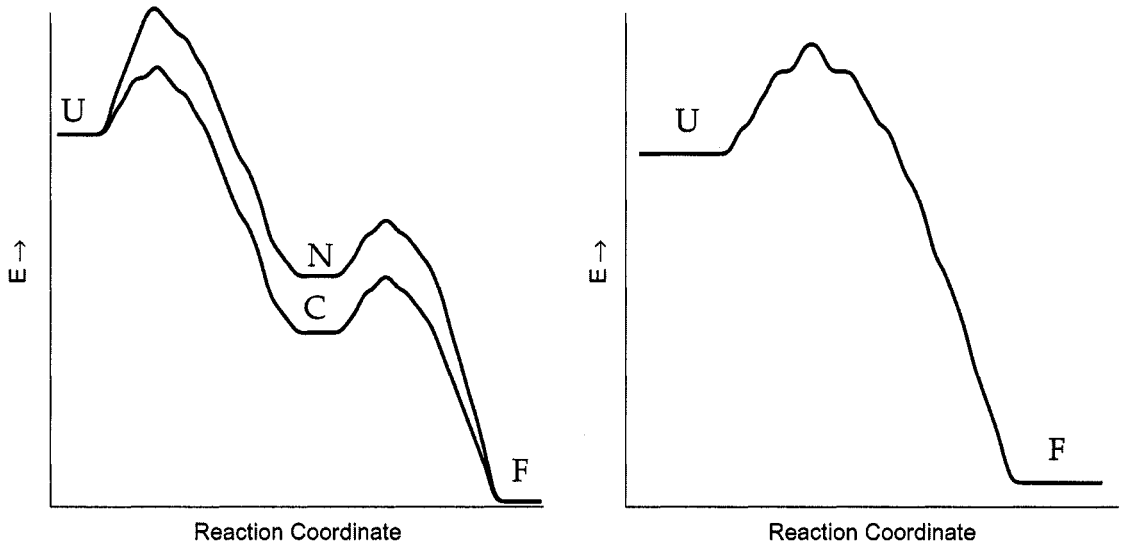


Figure 7-6. 4-state and 2-state reaction diagrams.
A 4-state folding mechanism (left) and a two-state mechanism (right) are compared.

Bibliography

Andersen, N.H., Barua, B., Fesinmeyer, R.M., Hudson, F.M., Lin, J., Euser, A., and White, G. 2002a. Chemical shifts, the ultimate test of polypeptide folding cooperativity. In *Peptides 2002: Proc. 27th Eur. Peptide Symp.*, pp. 824-825. Edizioni Ziino, Napoli, Italy, Sorrento, Italy.

Andersen, N.H., Brodsky, Y., Neidigh, J.W., and Prickett, K.S. 2002b. Medium-Dependence of the Secondary Structure of Exendin-4 and Glucagon-Like-Peptide-1. *Bioorganic and Medicinal Chemistry* 10: 79-85.

Andersen, N.H., Dyer, R.B., Fesinmeyer, R.M., Gai, F., Liu, Z., Neidigh, J.W., and Tong, H. 1999. Effect of Hexafluoroisopropanol on the Thermodynamics of Peptide Secondary Structure Formation. *Journal of the American Chemical Society* 121: 9879-9880.

Andersen, N.H., Dyer, R. Brian, Fesinmeyer, R. Matthew, Gai, Feng, Maness, Shelia, Werner, James H. 2001. Timescales for Peptide Secondary and Tertiary Structure Formation. In *Peptides 2000.* (ed. J. Martinez, Fehrentz, Jean-Alain), pp. 553-554. EDK, Paris, France.

Andersen, N.H., Fesinmeyer, R.M., and Hudson, F.M. 2004. Analysis of β Peptides Using Chemical Shift Deviations. In *Peptide Revolution: Genomics, Proteomics & Therapeutics.* (eds. M. Chorev, and T.K. Sawyer), pp. 462-463. American Chemical Society.

Andersen, N.H., Neidigh, J.W., Harris, S.M., Lee, G.M., Liu, Z., and Tong, H. 1997. Extracting Information from the Temperature Gradients of Polypeptide NH Chemical Shifts .1. The Importance of Conformational Averaging. *Journal of the American Chemical Society* 119: 8547-8561.

Andersen, N.H., Olsen, K.A., Fesinmeyer, R.M., Eidenschink, L.A., Tan, X., Hudson, F.M., and Farazi, S.R. 2006. Minimization and Optimization of Designed β Hairpin Folds. *Journal of the American Chemical Society* 128: 6101-6110.

Anfinsen, C.B. 1973. Principles that govern the folding of protein chains. *Science* 181: 223-230.

Anfinsen, C.B., Haber, E., Sela, M., and White, F.H., Jr. 1961. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc Natl Acad Sci U S A* 47: 1309-1314.

Anfinsen, C.B., Sela, M., and Cooke, J.P. 1962. The reversible reduction of disulfide bonds in polyalanyl ribonuclease. *J Biol Chem* 237: 1825-1831.

Aravinda, S.S., N.; Rajkishore, R.; Hosahudya, N.; Balaram, P.;. 2002. A Crystalline β-Hairpin Peptide Nucleated by a Type I' Aib-D-Ala β-Turn: Evidence for Cross-Strand Armonatic Interactions. *Angew. Chem. Int. Ed.* 41: 3863-3865.

Barua, B., and Andersen, N.H. 2002. Determinants of miniprotein stability: can anything replace a buried H-bonded Trp sidechain? *Letters in Peptide Science* 8: 221-226.

Barua, B., Lin, J.C., Kummler, P., Neidigh, J.W., and Andersen, N.H. 2005. The Trp-cage: Optimizing the Stability of a Globular Miniprotein. *Journal of Molecular Biology*: submitted.

Bax, A., and Davis, D.G. 1985. MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *Journal of Magnetic Resonance* 65: 355-360.

Blanco, F.J., Jiménez, M.A., Herranz, J., Rico, M., Santoro, J., and Nieto, J.L. 1993. NMR evidence of a short linear peptide that folds into a β-hairpin in aqueous solution. *Journal of the American Chemical Society* 115: 5887-5888.

Blanco, F.J., Jiménez, M.A., Pineda, A., Rico, M., Santoro, J., and Nieto, J.L. 1994a. NMR solution structure of the isolated N-terminal fragment of protein-G B1 domain. Evidence of trifluoroethanol induced native-like β-hairpin formation. *Biochemistry* 33: 6004-6014.

Blanco, F.J., Rivas, G., and Serrano, L. 1994b. A short linear peptide that folds into a native stable β-hairpin in aqueous solution. *Nature Structural Biology* 1: 584-590.

Blandl, T., Cochran, A.G., and Skelton, N.J. 2003. Turn stability in β-hairpin peptides: Investigation of peptides containing 3:5 type I G1 bulge turns. *Protein Science* 12: 237-247.

Bloch, F. 1946. Nuclear Induction. *Physical Review* 70: 460-474.

Bofill, R., Simpson, E.R., Platt, G.W., Crespo, M.D., and Searle, M.S. 2005. Extending the Folding Nucleus of Ubiquitin with an Independently Folding β-Hairpin Finger: Hurdles to Rapid Folding Arising from the Stabilization of Local Interactions. *Journal of Molecular Biology* 349: 205-221.

Boyden, M.N., and Asher, S.A. 2001. Raman Studies of Peptide Conformation Demonstrate that Betanova Does Not Cooperatively Unfold. *Biochemistry* 40: 13723-13727.

Bruch, M.D., Dhingra, M.M., and Gierasch, L.M. 1991. Side chain-backbone hydrogen bonding contributes to helix stability in peptides derived from an alpha-helical region of carboxypeptidase A. *Proteins* 10: 130-139.

Bursulaya, B.D., and Brooks III, C.L. 1999. Folding Free Energy Surface of a Three-dimensional Beta-sheet Protein. *Journal of the American Chemical Society* 121: 9947-9951.

Carulla, N., Woodward, C., and Barany, G. 2002. BetaCore, a designed water soluble four-stranded antiparallel β-sheet protein. *Protein Science* 11: 1539-1551.

Chen, H.I.a.S., M. 1995. The WW Domain of Yes-Associated Protein Binds a Proline-Rich Ligand that Differs from the Consensus Established for Src Homolohy 3-Binding Modules. *Proceedings of the National Academy of Sciences of the United States of America* 92: 7819-7823.

Chen, P.-Y., Lin, C.-K., Lee, C.-T., Jan, H., and Chan, S.I. 2001. Effects of turn residues in directing the formation of the β-sheet and in the stability of the β-sheet. *Protein Science* 10: 1794-1800.

Chen, R.P.-Y., Huang, Joseph, Chen, H., Jan, H., Velusamy, M., Lee, C., Fann, W., Larsen, R., Chan, S. 2004. Measuring the Refolding of β-Sheets with Different Turn Sequences on a Nanosecond Time Scale. *PNAS* 101: 7305-7310.

Ciani, B., Jourdan, M., and Searle, M.S. 2003. Stabilization of β-hairpin peptides by salt bridges: Role of preorganization in the energetic contribution of weak interactions. *Journal of the American Chemical Society* 125: 9038-9047.

Cochran, A.G., Skelton, N.J., and Starovasnik, M.A. 2001a. Tryptophan zippers: Stable, monomeric β-hairpins. *Proceedings of the National Academy of Sciences of the United States of America* 98: 5578-5583.

Cochran, A.G., Tong, R.T., Starovasnik, M.A., Park, E.J., McDowell, R.S., Theaker, J.E., and Skelton, N.J. 2001b. A Minimal Peptide Scaffold for β-Turn Display: Optimizing a Strand Position in Disulfide-Cyclized β-Hairpins. *Journal of the American Chemical Society* 123: 625-632.

Cox, J.P.H., Evans, P.A., Packman, L.C., Williams, D.H., and Woolfson, D.N. 1993. Dissecting the structure of a partially folded protein: circular dichroism and nuclear magnetic resonance studies of peptides from ubiquitin. *Journal of Molecular Biology* 234: 483-492.

Crane, J.C., Sabelko, J., and Gruebele, M. 1999. Laser temperature jump induced protein folding dynamics of betanova and protein. *Biophysical Journal* 76: A107-A107.

Creighton, T.E. 1993. *Proteins: Structures and Molecular Principles.*, 2nd Ed. ed. Freeman & Co., New York.

de Alba, E., Jiménez, M.A., and Rico, M. 1997. Turn Residue Sequence Determines β-Hairpin Conformation in Designed Peptides. *Journal of the American Chemical Society* 119: 175-183.

de Alba, E., Jiménez, M.A., Rico, M., and Nieto, J.L. 1996. Conformational investigation of designed short linear peptides able to fold into beta-hairpin structures in aqueous solution. *Folding & Design* 1: 133-144.

de Alba, E., Rico, M., and Jiménez, M.A. 1999a. The turn sequence directs β-strand alignment in designed β-hairpins. *Protein Science* 8: 2234-2244.

de Alba, E., Santoro, J., Rico, M., and Jiménez, M.A. 1999b. De novo design of a monomeric three-stranded antiparallel beta-sheet. *Protein Science* 8: 854-865.

Dill, K., and Chan, H. 1997. From Levinthal to pathways to funnels. *Nat Struct Biol* 4: 10-19.

Dinner, A.R., Lazaridis, T., and Karplus, M. 1999. Understanding beta-hairpin formation. *Proceedings of the National Academy of Sciences of the United States of America* 96: 9068-9073.

Dobson, C.M., and Karplus, M. 1999. The fundamentals of protein folding: bringing together theory and experiment. *Curr Opin Struct Biol* 9: 92-101.

Du, D., Zhu, Y., Huang, C.-Y., and Gai, F. 2004. Understanding the key factors that control the rate of β-hairpin folding. *Proceedings of the National Academy of Sciences of the United States of America* 101: 15915-15920.

Dyer, R.B., Manas, E.S., Peterson, E.S., Franzen, S., Fesinmeyer, R.M., and Andersen, N.H. 2004. The mechanism of β-hairpin formation. *Biochemistry* 43: 11560-11566.

Dyer, R.B., Maness, S.J., Franzen, S., Fesinmeyer, R.M., Olsen, K.A., and Andersen, N.H. 2005. Hairpin Folding Dynamics: The Cold-denatured State Is Predisposed for Rapid Refolding. *Biochemistry* 44: 10406-10415.

Eng, J., Kleinman, W.A., Singh, L., Singh, G., and Raufman, J.P. 1992. Isolation and characterization of exendin-4, an exendin-3 analogue, from Heloderma suspectum venom. Further evidence for an exendin receptor on dispersed acini from guinea pig pancreas. *Journal of Biological Chemistry* 267: 7402-7405.

Espinosa, J.F., and Gellman, S.H. 2000. A Designed β-Hairpin Containing a Natural Hydrophobic Cluster. *Angew. Chem. Int. Ed.* 39: 2330-2333.

Espinosa, J.F., Muñoz, V., and Gellman, S.H. 2001. Interplay Between Hydrophobic Cluster and Loop Propensity in β-Hairpin Formation. *Journal of Molecular Biology* 306: 397-402.

Espinosa, J.F., Syud, F.A., and Gellman, S.H. 2002. Analysis of the factors that stabilize a designed two-stranded antiparallel β-sheet. *Protein Science* 11: 1492-1505.

Ferguson, N., Johnson, C.M., Macias, M., Oschkinat, H., and Fersht, A. 2001a. Ultrafast folding of WW domains without structured aromatic clusters in the denatured

state. *Proceedings of the National Academy of Sciences of the United States of America* 98: 13002-13007.

Ferguson, N., Pires, J.R., Toepert, F., Johnson, C.M., Pan, Y.P., Volkmer-Engert, R., Schneider-Mergener, J., Daggett, V., and Oschkinat, H. 2001b. Using flexible loop mimetics to extend Φ-value analysis to secondary structure interactions. *Proceedings of the National Academy of Sciences of the United States of America* 98: 13008-13013.

Fersht, A.R. 1995. Characterizing transition states in protein folding: an essential step in the puzzle. *Curr Opin Struct Biol* 5: 79-84.

Fesinmeyer, R.M., Hudson, F.M., and Andersen, N.H. 2004. Enhanced Hairpin Stability through Loop Design: The Case of the Protein G B1 Hairpin. *Journal of the American Chemical Society* 126: 7238-7243.

Fesinmeyer, R.M., Hudson, F.M., White, G.W.N., Olsen, K.A., Euser, A., and Andersen, N.H. 2005. Chemical Shifts Provide Fold Populations and Register of β-Hairpins and β-Sheets. *Journal of Biomolecular NMR*: in press.

Finsinger, S., Serrano, L. and Lacroix, E. 2001. Computational estimation of specific side-chain interaction energies in a-helices. . *Protein Science* 10: 809-818.

Fischer, E. 1894. Einfluss der Configuration auf die Wirkung Derenzyme. *Ber. Dt. Chem. Ges.* 27: 2985-2993.

Gibbs, A.C., Bjorndahl, T.C., Hodges, R.S., and Wishart, D.S. 2002. Probing the Structural Determinants of Type II' β-Turn Formation in Peptides and Proteins. *Journal of the American Chemical Society* 124: 1203-1213.

Göke, R., Fehmann, H.C., Linn, T., Schmidt, H., Krause, M., Eng, J., and Göke, B. 1993. Exendin-4 is a high potency agonist and truncated exendin-(9-39)-amide an antagonist at the glucagon-like peptide 1-(7-36)-amide receptor of insulin-secreting β-cells. *Journal of Biological Chemistry* 268: 19650-19655.

Greenfield, N., and Fasman, G.D. 1969. Computed circular dichroism spectra for the evaluation of protein conformation. *Biochemistry* 8: 4108-4116.

Griffiths-Jones, S.R., and Searle, M.S. 2000. Structure, folding, and energetics of cooperative interactions between the beta-strands of a de novo designed three-stranded antiparallel beta-sheet peptide. *Journal of the American Chemical Society* 122: 8350-8356.

Gunther, H. 1995. *NMR Spectroscopy, 2nd Ed.* John Wiley & Sons, Chichester, pp. see pp. 339-345.

Haque, T.S., and Gellman, S. 1997. Insights on β-Hairpin Stability in Aqueous Solution from Peptides with Enforced Type I' and Type II' β-Turns. *Journal of the American Chemical Society* 119: 2303-2304.

Honda, S., Kobayashi, N., and Munekata, E. 2000. Thermodynamics of a β-Hairpin Structure: Evidence for Cooperative Formation of Folding Nucleus. *Journal of Molecular Biology* 295: 269-278.

Huang, G.S., and Oas, T.G. 1995. Submillisecond folding of monomeric λ Repressor (dynamic NMR). *Proceedings of the National Academy of Sciences of the United States of America* 92: 6878-6882.

Jager, M., Nguyen, H., Crane, J.C., Kelly, J.W., and Gruebele, M. 2001. The Folding Mechanism of a β-Sheet: The WW Domain. *Journal of Molecular Biology* 311: 373-393.

Karle, I.L., Awasthi, S.K., and Balaram, P. 1996. A designed β-hairpin peptide in crystals. *Proceedings of the National Academy of Sciences of the United States of America* 93: 8189-8193.

Kobayashi, N., Endo, S., and Munekata, E. 1993. Conformational study of the IgG binding domain of protein G. *Peptide Chem.*: 278-280.

Koepf, E.K., Petrassi, H.M., Sudol, M., and Kelly, J.W. 1999. WW: An isolated three-stranded antiparallel beta-sheet domain that unfolds and refolds reversibly; evidence for a structured hydrophobic cluster in urea and GdnHCl and a disordered thermal unfolded state. *Protein Science* 8: 841-853.

Kortemme, T., Ramírez-Alvarado, M., and Serrano, L. 1998. Design of a 20-Amino Acid, Three-Stranded β-Sheet Protein. *Science* 281: 253-256.

Kraemer-Pecore, C., Lecomte, J., and Desjarlais, J. 2003. A de novo redesign of the WW domain. *Protein Science* 12: 2194-2205.

Kuhlman, B., O'Neill, J.W., Kim, D., E., Zhang, K.Y.J., and Baker, D. 2002. Accurate Computer-based Design of a New Backbone Conformation in the Second Turn of Protein L. *Journal of Molecular Biology* 315: 471-477.

Kuznetsov, S.V., Hilario, J., Keiderling, T.A., and Ansari, A. 2003. Spectroscopic studies of structural changes in two β-sheet-forming peptides show an ensemble of structures that unfold noncooperatively. *Biochemistry* 42: 4321-4332.

López de la Paz, M., Lacroix, E., Ramírez-Alvarado, M., and Serrano, L. 2001. Computer-aided Design of β-sheet peptides. *Journal of Molecular Biology* 312: 229-246.

Macias, M.J., Gervais, V., Civera, C., and Oschkinat, H. 2000. Structural analysis of WW domains and design of a WW prototype. *Nature Structural Biology* 7: 375-379.

Macias, M.J., Hyvonen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M., and Oschkinat, H. 1996. Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature* 382: 646-649.

Matthews, B.W., Nicholson, H., and Becktel, W.J. 1987. Enhanced protein thermostability from site-directed mutations that decrease the entropy of unfolding. *Proceedings of the National Academy of Sciences of the United States of America* 84: 6663-6667.

Maynard, A.J., Sharman, G.J., and Searle, M.S. 1998. Origin of β-hairpin stability in solution: Structural and thermodynamic analysis of the folding of model peptide supports hydrophobic stabilization in water. *Journal of the American Chemical Society* 120: 1996-2007.

Merutka, G., Dyson, H.J., and Wright, P.E. 1995. 'Random coil' ^{1}H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. *Journal of Biomolecular NMR* 5: 14-24.

Muñoz, V., Henry, E.R., Hofrichter, J., and Eaton, W.A. 1998. A statistical mechanical model for beta-hairpin kinetics. *Proceedings of the National Academy of Sciences of the United States of America* 95: 5872-5879.

Muñoz, V., and Serrano, L. 1995. Elucidating the Folding Problem of Helical Peptides using Empirical Parameters. II. Helix Macrodipole Effects and Rationale Modification of the Helical Content of Natural Peptides. *Journal of Molecular Biology* 245: 275-296.

Muñoz, V., and Serrano, L. 1997. Development of the multiple sequence approximation within the AGADIR model of alpha-helix formation: comparison with Zimm-Bragg and Lifson- Roig formalisms. *Biopolymers* 41: 495-509.

Muñoz, V., Thompson, P.A., Hofrichter, J., and Eaton, W.A. 1997. Folding dynamics and mechanism of β-hairpin formation. *Nature* 390: 196-199.

Myers, J.K., and Oas, T.G. 2001. Preorganized secondary structure as an important determinant of fast protein folding. *Nature Structural Biology* 8: 552-558.

Nauli, S., Kuhlman, B., and Baker, D. 2001. Computer-based redesign of a protein folding pathway. *Nature Structural Biology* 8: 602-605.

Neidigh, J.W., and Andersen, N.H. 2002. Peptide Conformational Changes Induced by Tryptophan-Phosphocholine Interactions in a Micelle. *Biopolymers* 65: 354-361.

Neidigh, J.W., Fesinmeyer, R.M., and Andersen, N.H. 2002. Designing a 20-residue protein. *Nature Structural Biology* 9: 425-430.

Neidigh, J.W., Fesinmeyer, R.M., Prickett, K.S., and Andersen, N.H. 2001. Exendin-4 and Glucagon-Like-Peptide-1: NMR Structural Comparisons in the Solution and Micelle-Associated State. *Biochemistry* 40: 13188-13200.

Olsen, K. 2006. NMR Studies Reveal Kinetics and Thermodynamics of Hairpin Formation. In *Chemistry*, pp. 189. University of Washington, Seattle.

Olsen, K.A., Fesinmeyer, R.M., Stewart, J., and Andersen, N.H. 2005a. Elucidating polypeptide folding dynamics by ^{1}H NMR. I. GB1 hairpin folding rates reflect mutations within and remote from the turn. *Proceedings of the National Academy of Sciences of the United States of America*: submitted.

Olsen, K.A., Scian, M., and Andersen, N.H. 2005b. Elucidating polypeptide folding dynamics by ^{1}H NMR. III. Loop sequence effects of hairpin dynamics in minimized models. *Biochemistry*: in preparation.

Osapay, K., and Case, D.A. 1991. A new analysis of proton chemical shifts in proteins. *Journal of the American Chemical Society* 113: 9436-9444.

Osterman, D.G.a.K., E.T. 1985. Context-Dependent Secondary Structure Formation of a Designed Protein Sequence. *Nature* 380: 730-734.

Otte, L.W., U.; Schlegel, B.; Pires, J.; Beyermann, M.; Schmieder, P.; Krause, G.; Volkmer-Engert, R.; Schneider-Mergener, J. and Oschkinat, H. 2003. WW Domain Sequence Activity Relationships Identified Using Ligand Recognition Propensities of 42 WW Domains. *Protein Science* 12: 491-500.

Pauling, L.a.D., M. 1940. Gradient-tailored excitation for single quantum NMR spectroscopy of aqueous solutions. *Journal of Biomolecular NMR* 2: 661-665.

Piotto, M., Saudek, V., and Sklenar, V. 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *Journal of Biomolecular NMR* 2: 661-665.

Platt, G.W., Simpson, S.A., Layfield, R., and Searle, M.S. 2003. Stability and Folding Kinetics of a Ubiquitin Mutant with a Strong Propensity for Nonnative β-Hairpin Conformation in the Unfolded State. *Biochemistry* 42: 13762-13771.

Ramírez-Alvarado, M., Blanco, F.J., Niemann, H., and Serrano, L. 1997. Role of β-turn residues in β-hairpin formation and stability in designed peptides. *Journal of Molecular Biology* 273: 898-912.

Ramírez-Alvarado, M., Blanco, F.J., and Serrano, L. 1996. De novo design and structural analysis of a model β-hairpin peptide system. *Nature Structural Biology* 3: 604-612.

Roe, D.R., Hornack, V., and Simmerling, C. 2005. Folding Cooperativity in a Three-stranded β-Sheet Model. *Journal of Molecular Biology* 352: 370-381.

Saika, H.S.G.a.A. 1953. Dissociation, Chemical Exchange, and teh Proton Magnetic Resonance in Some Aqueous Electrolytes. *Journal of Chemical Physics* 21: 1688-1694.

Sandstrom, J. 1982. *Dynamic NMR Spectroscopy*. Academic Press, London.

Santiveri, C.M., Rico, M., and Jiménez, M.A. 2001. 13C(alpha) and 13C(beta) chemical shifts as a tool to delineate beta-hairpin structures in peptides. *Journal of Biomolecular NMR* 19: 331-345.

Santiveri, C.M., Rico, M., Jiménez, M.A., Pastor, M.T., and Pérez-Payá, E. 2003. Insights into the determinants of β-sheet stability: ^{1}H and ^{13}C NMR conformational investigation of three-stranded antiparallel β-sheet-forming peptides. *Journal of Peptide Research* 61: 177-188.

Santiveri, C.M., Santoro, J., Rico, M., and Jiménez, M.A. 2004. Factors involved in the stability of isolated β-sheets: Turn sequence, β-sheet twisting and hydrophobic surface burial. *Protein Science* 13: 1134-1147.

Schenck, H., and Gellman, S. 1998. Use of a Designed Triple-Standard Antiparallel β-Sheet to Probe β-Sheet Cooperativity in Aqueous Solution. *Journal of the American Chemical Society* 120: 4869-4870.

Searle, M.S., and Ciani, B. 2004. Design of β-sheet systems for understanding the thermodynamics of kinetics of protein folding. *Current Opinion in Structural Biology* 14: 1-7.

Searle, M.S., Williams, D.H., and Packman, L.C. 1995. A short linear peptide derived from the N-terminal sequence of ubiquitin folds into a water-stable non-native β-hairpin. *Nature Structural Biology* 2: 999-1006.

Sharman, G.J., Griffiths-Jones, S.R., Jourdan, M., and Searle, M.S. 2001. Effects of Amino Acid ϕ,ψ Propensities and Secondary Structure Interactions in Modulating Hα Chemical Shifts in Peptide and Protein β-Sheets. *Journal of the American Chemical Society* 123: 12318-12324.

Sharman, G.J., and Searle, M.S. 1997. Dissecting the effects of cooperativity on the stabilisation of a *de novo* designed three stranded anti-parallel β-sheet. *Chemical Communications*: 1955-1956.

Sharman, G.J., and Searle, M.S. 1998. Cooperative Interaction between the Three Strands of a Designed Antiparallel β-Sheet. *Journal of the American Chemical Society* 120: 5291-5300.

Sibanda, B.a.T., J. 1984. β-Hairpin Families in Globular Proteins. *Nature* 316: 170-174.

Sibanda, B.L., and Thornton, J.M. 1991. Conformation of β Hairpins in Protein Structures: Classification and Diversity in Homologous Structures. *Methods in Enzymology* 202: 59-82.

Snow, C.D., Qiu, L., Du, D., Gai, F., Hagen, S.J., and Pande, V.S. 2004. Trp zipper folding kinetics by molecular dynamics and temperature-jump spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* 101: 4077-4082.

Stanger, H.E., Syud, F.A., Espinosa, J.F., Giriat, I., Muir, T., and Gellman, S.H. 2001. Length-dependent stability and strand length limits in antiparallel β-sheet secondary structure. *Proceedings of the National Academy of Sciences of the United States of America* 98: 12015-12020.

Sudol, M. 1996. Structure and function of the WW domain. *Prog Biophys Mol Biol* 65: 113-132.

Sudol, M.B., P.; Einbond, A.; Kastury, K.; Druck, T.; Negrini, M.; Huebner, K.; Lehman, D. 1995a. Characterization of the Mammalian YAP (Yes-associated Protein) Gene and Its Role in Defining a Novel Protein Module, the WW Domain. *Journal of Biological Chemistry* 270: 14733-14741.

Sudol, M.C., H.; Bougeret, C.; Einbond, A. and Bork, P. 1995b. Characterization of a Novel Protein-Binding Module-the WW Domain. *FEBS Letters* 369: 67-71.

Syud, F.A., Espinosa, J.F., and Gellman, S.H. 1999. NMR-Based Quantification of β-Sheet Populations in Aqueous Solution through Use of Reference Peptides for the Folded and Unfolded States. *Journal of the American Chemical Society* 121: 11577-11578.

Syud, F.A., Stanger, H.E., and Gellman, S.H. 2001. Interstrand Side Chain-Side Chain Interactions in a Designed β-Hairpin: Significance of Both Lateral and Diagonal Pairings. *Journal of the American Chemical Society* 123: 8667-8677.

Syud, F.A., Stanger, H.E., Mortell, H.S., Espinosa, J.F., Fisk, J.D., Fry, C.G., and Gellman, S.H. 2003. Influence of Strand Number on Antiparallel β-Sheet Stability in Designed Three- and Four-stranded β-Sheets. *Journal of Molecular Biology* 326: 553-568.

Tatko, C.D., and Waters, M.L. 2002. Selective Aromatic Interactions in β-Hairpin Peptides. *Journal of the American Chemical Society* 124: 9372-9373.

Tatko, C.D., and Waters, M.L. 2004. Comparison of C-H $\cdots$ π and Hydrophobic Interactions in a β-Hairpin Peptide: Impact on Stability and Specificity. *Journal of the American Chemical Society* 126: 2028-2034.

Wang, M., Tang, Y., Sato, S., Vugmeyster, L., McKnight, C.J., and Raleigh, D.P. 2003. Dynamic NMR Line-Shape Analysis Demonstrates that the Villin Headpiece Subdomain Folds on the Microsecond Time Scale. *Journal of the American Chemical Society* 125: 6032-6033.

Williamson, M.P. 1990. Secondary Structure Dependent Chemical Shifts in Proteins. *Biopolymers* 29: 1423-1431.

Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S., and Sykes, B.D. 1995. ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *Journal of Biomolecular NMR* 5: 67-81.

Wishart, D.S., Sykes, B.D., and Richards, F.M. 1991. Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. *Journal of Molecular Biology* 222: 311-333.

Wolynes, P.G., Onuchic, J.N., and Thirumalai, D. 1995. Navigating the folding routes. *Science* 267: 1619-1620.

Wüthrich, K. 1986. *NMR of Proteins and Nucleic Acids*. J. Wiley, New York.

Xu, X.P., and Case, D.A. 2001. Automated prediction of ^{15}N, $^{13}C\alpha$ $^{13}C\beta$ and $^{13}C'$ chemical shifts in proteins using a density functional database. *J Biomol NMR* 21: 321-333.

Xu, Y., Oyola, R., and Gai, F. 2003. Infrared Studies of the Stability and Folding Kinetics of a 15-Residue β-Hairpin. *Journal of the American Chemical Society* 125: 15388-15394.

Yang, J.T., Wu, C.S., and Martinez, H.M. 1986. Calculation of protein conformation from circular dichroism. *Methods in Enzymology* 130: 208-269.

Appendix A: Chemical Shift Assignments

Chemical shift assignments for the peptides characterized personally and mentioned in this volume follow at the temperature and solvent conditions specified.

MrH2a pH 5.90 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Arg	7.085	4.032	1.876	Hγ:1.536 Hδ: 3.169
2	Lys	8.675	4.539	1.709	Hγ:1.400,1.310 Hδ:1.654 Hε:2.946
3	Tyr	8.525	4.744	3.044,2.936	Hδ:7.087 Hε:6.769
4	Thr	8.266	4.574	4.064	Hγ:1.131
5	Val	8.321	4.284	1.872	Hγ:0.897
6	Ser	8.435	4.717	3.799,3.796	
7	Ile	8.459	4.252	1.84	Hγ: Hδ:0.892
8	Asn	8.71	4.642	2.849	Hδ:6.935,7.592
9	Gly	8.368	3.997,3.812		
10	Lys	7.974	4.389	1.768	Hγ:1.385 Hδ:1.666 Hε:2.983
11	Lys	8.362	4.487	1.735	Hγ:1.307 Hδ:1.637 Hε:2.942
12	Ile	8.493	4.336	1.856	Hδ:0.875
13	Thr	8.37	4.514	4.122	Hγ:1.138
14	Val	8.413	4.217	1.979	Hγ:0.891
15	Ser	8.408	4.564	3.844,3.841	
16	Glu	8.385	4.399	2.158,1.971	Hγ:2.460, 2.448

MrH2b pH 5.90 290K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Arg		4.028	1.857,1.854	Hγ:1.495,1.466 Hδ:3.038
2	Lys	8.731	4.598	1.716	Hγ:1.399,1.331 Hδ: Hε:
3	Trp	8.614	4.845	3.312,3.195	Hδ:7.196 Hε:10.154,7.549 Hζ:7.453,7.073 Hη:7.187
4	Thr	8.278	4.596	4.054	Hγ:1.123
5	Val	8.416	4.267	2.035	Hγ:0.932
6	Ser	8.515	4.728	3.791,3.792	
7	Ile	8.538	4.238	1.835	Hγ:1.407, Hδ:0.884
8	Asn	8.776	4.636	2.924,2.752	Hδ:6.985,7.657
9	Gly	8.421	4.004,3.803		
10	Lys	8.002	4.39	1.817,1.753	Hγ:1.425,1.375 Hδ:1.680 Hε:2.985
11	Lys	8.455	4.495	1.723,1.722	Hγ:1.427,1.308 Hδ:1.603 Hε:2.913
12	Ile	8.617	4.355	2.397	Hγ:1.423, 0.878 Hδ:0.825
13	Thr	8.46	4.517	4.088	Hγ:1.140
14	Val	8.488	4.188	1.896	Hγ:0.860,0.817
15	Ser	8.434	4.557	3.814,3.812	
16	Glu	8.45	4.357	1.976,1.850	Hγ:2.128,2.126

MrH2b' pH 6.00 280K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Arg		4.014	1.850	Hγ:1.467 Hδ:3.025 Hε:6.983
2	Lys	8.801	4.569	1.730	Hγ:1.405,1.337 Hδ:1.668 Hε:2.938
3	Trp	8.689	4.826	3.307,3.201	Hδ:8.696 Hε:7.198 Hζ:7.456,7.081 Hη:7.556
4	Thr	8.303	4.55	4.049	Hγ:1.122
5	Val	8.458	4.23	2.034	Hγ:0.934
6	Ser	8.589	4.69	3.797,3.796	
7	Ile	8.579	4.227	1.837	Hγ:1.146, 1.444 Hδ:0.891
8	Asn	8.812	4.643	2.916,2.744	Hδ:7.028, 7.719
9	Gly	8.478	3.90, 3.813		
10	Lys	8.07	4.377	1.817,1.754	Hγ:1.431,1.379 Hδ:1.683 Hε:2.985
11	Lys	8.523	4.463	1.722	Hγ:1.427,1.321 Hδ:1.601 Hε:2.911
12	Ile	8.662	4.33	1.852	Hγ:1.180 1.441 Hδ:0.854
13	Thr	8.532	4.492	4.095	Hγ:1.142
14	Val	8.542	4.175	1.923	Hγ:0.847
15	Ser	8.526	4.534	3.820,3.818	
16	Glu	8.519	4.36	2.137,1.952	Hγ:2.430

HP5A pH 6.00 280K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Lys		4.003	1.818	Hγ:1.360 Hδ:1.818,1.670 Hε:2.963
2	Lys	8.75	4.421	1.755	Hγ:1.393,1.339 Hδ:1.755,1.679 Hε:2.959
3	Tyr	8.702	4.744	2.845,2.844	Hδ:7.044 Hε:6.774
4	Thr	8.404	4.391	4.062	Hγ:1.148
5	Trp	8.61	4.635	3.176,3.176	Hδ:7.169 Hε:10.232,7.513 Hζ:7.431,7.185 Hη:7.103
6	Asn	8.053	4.9	2.984,2.542	Hδ:7.154,7.585
7	Pro		4.065	2.305	Hγ:1.963 Hδ:3.718,3.578
8	Ala	8.139	4.279	1.419	
9	Thr	7.617	4.338	4.242	Hγ:1.144
10	Gly	8.354	3.970,3.795		
11	Lys	7.84	4.39	1.807	Hγ:1.375 Hδ:1.791,1.656 Hε:2.966
12	Ala	8.607	4.559	1.295	
13	Thr	8.463	4.388	4.122	Hγ:1.178
14	Val	8.419	4.133	1.991	Hγ:0.914,0.874
15	Gln	8.615	4.365	2.049,1.936	Hγ:2.325 Hε:6.935,7.656
16	Glu	8.405	4.21	2.026	Hγ:2.398

HP5R pH 6.00 280K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Lys		3.98	1.796	Hγ:1.341 Hδ:1.659 Hε:
2	Lys	8.726	4.411	1.762	Hγ:1.385 Hδ:1.667 Hε:2.958
3	Tyr	8.657	4.816	2.817,2.816	Hδ:7.021 Hε:6.763
4	Thr	8.472	4.411	4.065	Hγ:1.137
5	Trp	8.604	4.727	3.158,3.158	Hδ:7.175 Hε:10.236,7.501 Hζ:7.450,7.099 Hη: 7.187
6	Asn	8.137	4.886	2.958,2.556	Hδ:7.169,7.608
7	Pro		4.021	2.277	Hγ:1.938 Hδ:3.696,3.556
8	Ala	8.122	4.283	1.416	
9	Thr	7.626	4.342	4.16	Hγ:1.144
10	Gly	8.357	3.984,3.814		
11	Lys	7.895	4.404	1.765	Hγ:1.384 Hδ:1.667 Hε:2.957
12	Arg	8.687	4.558	1.717,1.610	Hγ:1.455 Hδ:2.954
13	Thr	8.591	4.448	4.136	Hγ:1.170
14	Val	8.454	4.123	1.992	Hγ:0.908,0.871
15	Gln	8.606	4.344	2.050,1.898	Hγ:2.325 Hε:6.936,7.631
16	Glu	8.449	4.244	2.137,1.915	Hγ:2.420

HP7 INGK pH 6.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Lys		4.268	1.879,1.819	Hγ:1.324 Hδ:1.384 Hε:2.503,2.445
2	Thr	8.879	4.742	3.999	Hγ:1.220
3	Trp	8.799	4.65	2.961,2.176	Hδ:6.792 Hε:9.887,6.088 Hζ:7.269,6.674 Hη:7.012
4	Ile	8.4	4.116	1.643	Hγ:1.326,1.102,0.995 Hδ:0.772
5	Asn	8.854	4.221	2.906,2.591	Hδ:6.857,7.517
6	Gly	7.362	3.800,3.301		
7	Lys	6.984	4.184	1.669	Hγ:1.323 Hδ:1.576
8	Trp	8.364	5.025	3.133,3.058	Hδ:7.384 Hε:10.170,7.083 Hζ:7.288 Hη:7.089
9	Thr	8.931	4.534	4.187	Hγ:1.190
10	Glu	8.255	4.199	2.138,2.003	Hγ:2.450

HP7 NPDGT pH 6.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Lys	8.228	4.192	1.829	Hγ:1.386,1.281 Hδ:2.536 Hε: Hζ:7.236
2	Thr	8.805	4.534	4.075	Hγ:1.206
3	Trp	8.681	4.48	3.094,2.866	Hδ:6.956 Hε:9.991,6.589 Hζ:7.129
4	Asn	8.177	4.838	2.643,2.499	Hδ:7.337,6.817
5	Pro		4.10	2.254	Hγ:1.915 Hδ:3.593,3.476
6	Asp	8.003	4.554	2.912,2.754	
7	Gly	8.148	4.044,3.76		
8	Thr	7.581	4.388	4.155	Hγ:1.083
9	Trp	8.251	5.021	3.064,3.063	Hδ:7.221 Hε:10.132,7.430 Hζ:7.090,7.195 Hη:7.095
10	Thr	8.463	4.386	4.149	Hγ:1.144
11	Glu	8.228	4.195	2.125,1.967	Hγ:2.442

HP7 EPDGK pH 6.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Lys		4.12	1.506	Hγ:1.314 Hδ:1.838 Hε:2.699
2	Thr	8.708	4.457	4.095	Hγ:1.215
3	Trp	8.604	4.474	3.146,3.006	Hδ:7.099 Hε:10.024,7.021 Hζ:7.412,7.281 Hη:7.169
4	Glu	7.831	4.519	1.948,1.673	Hγ:2.272,2.272
5	Pro		4.113	2.237	Hγ:1.919 Hδ:3.567,3.430
6	Asp	8.155	4.553	2.861,2.782	
7	Gly	8.196	3.892,3.761		
8	Lys	7.632	4.291	1.672,1.579	Hγ:1.214 Hδ:1.595 Hε:2.901 Hζ:7.466
9	Trp	8.148	4.879	3.219,3.142	Hδ:7.236 Hε:10.111,7.531 Hζ:7.456,7.119 Hη:7.219
10	Thr	8.146	4.344	4.155	Hγ:1.117
11	Glu	8.007	4.162	2.095,1.929	Hγ:2.401

MrH5a pH 6.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Lys				
2	Lys	8.603	4.46	1.710	Hγ:1.415 Hδ:1.658 Hε:2.957
3	Tyr	8.525	4.686	3.046,2.938	Hδ:7.111 Hε:6.791
4	Thr	8.187	4.501	4.09	Hγ:1.143
5	Val	8.323	4.199	2.086	Hγ:0.920
6	Ser	8.357	4.505	3.769,3.769	
7	Asp	8.469	4.934	2.899,2.636	
8	Pro		4.385	2.299,1.958	Hγ:2.010 Hδ:3.846,3.801
9	Ala	8.317	4.31	1.429	
10	Thr	7.834	4.337	4.264	Hγ:1.213
11	Gly	8.357	4.002,3.923		
12	Arg	7.996	4.316	1.800,1.735	Hγ:1.599 Hδ:
13	Lys	8.407	4.425	1.749	Hγ:1.432,1.369 Hδ:1.652 Hε:2.975
14	Ile	8.407	4.282	1.844	Hγ:1.170l,0.884 Hδ:0.840
15	Thr	8.329	4.468	4.123	Hγ:1.162
16	Val	8.359	4.216	2.019	Hγ:0.894
17	Ser	8.373	4.524	3.839,3.804	
18	Ala	8.197	4.25	1.374	

MrH5b pH 7.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Lys		3.952	1.825	Hγ:1.337 Hδ:1.675 Hε:2.957
2	Lys				
3	Tyr	8.474	4.662	3.048,2.934	Hδ:6.806 Hε:7.124
4	Thr	8.148	4.368	4.10	Hγ:1.149
5	Val	8.199	4.126	2.044	Hγ:0.924
6	Ser	8.34	4.424	3.776,3.756	
7	His	8.443	4.967	3.202,3.067	Hδ:7.233 Hε:8.342
8	Pro		4.429	2.296,1.932	Hγ:1.975 Hδ:3.539,3.731
9	Ala	8.642	4.402	1.425	
10	Thr	8.114	4.352	4.263	Hγ:1.223
11	Gly	8.394	3.984,3.985		
12	Arg	8.129	4.329	1.776	Hγ:1.603 Hδ:3.192,3.183
13	Lys	8.431	4.359	1.789,1.735	Hγ:1.417,1.323 Hδ:1.672 Hε:2.977
14	Ile	8.318	4.258	1.868	Hγ:1.474,1.187,0.885 Hδ:0.845
15	Thr	8.3	4.416	4.16	Hγ:1.179
16	Val	8.257	4.219	2.097	Hγ:0.943
17	Ser	8.387	4.476	3.858,3.837	
18	Ala	8.024	4.154	1.337	

MrH6a pH 6.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Lys		4.379		Hγ:1.384,1.281 Hδ:1.810 Hε:2.974
2	Lys	8.379	4.522	1.712	Hγ:1.407 Hδ:1.651 Hε:2.961
3	Tyr	8.542	4.724	3.048,2.943	Hδ:6.811 Hε:7.116
4	Thr	8.222	4.553	4.08	Hγ:1.144
5	Val	8.239	4.203	2.088	Hγ:0.939
6	Ser	8.379	4.604	3.808,3.788	
7	Asn	8.588	5.004	2.832,2.739	Hδ:7.570,6.741
8	Pro		4.394	2.315	Hγ:2.017 Hδ:4.152,3.768
9	Asp	8.081	4.569	2.783,2.687	
10	Gly	8.236	4.120,3.883		
11	Thr	8.095	4.294	4.212	Hγ:1.206
12	Lys	8.337	4.582	1.748	Hγ:1.430 Hδ:1.644 Hε:2.972
13	Ile	8.229	4.236	1.838	Hγ:1.457,0.883Hδ:0.839
14	Thr	8.381	4.509	4.093	Hγ:1.155
15	Val	8.341	4.31	1.832	Hγ:0.879
16	Ser	8.394	4.466	3.845	
17	Ala	8.017	4.145		

UTH3 pH 5.80 290K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Ser				
2	Glu	8.786	4.467	1.979,1.866	Hγ:2.238
3	Ile	8.331	4.403	2.066	Hγ:1.720,1.387,0.961 Hδ:6.792
4	Tyr	8.566	4.747	3.025,2.900	Hδ:6.790 Hε:7.060
5	Ser	8.251	4.33	3.542,3.510	
6	Asn	8.497	4.958	2.807,2.663	Hδ:6.956,7.508
7	Pro		4.366	2.305,1.973	Hγ:2.035 Hδ:3.851,3.760
8	Asp	7.98	4.556	2.843,2.663	
9	Gly	8.175	4.083,3.808		
10	Thr	7.932	4.35	4.235	Hγ:1.122
11	Trp	8.37	5.005	3.242,3.137	Hδ:7.187 Hε:10.111,7.217 Hζ:7.474,7.099 Hη:7.435
12	Thr	8.512	4.398	3.935	Hγ:1.126
13	Val	8.38	4.394	2.282	Hγ:0.956,0.795
14	Thr	8.551	4.449	4.238	Hγ:1.208
15	Glu	8.192	4.134	2.065,1.893	

EX4 30% TFE/pH 4.50 298K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	His		4.733		
2	Gly	8.869	4.114,3.898		
3	Glu	9.006	4.287	2.076	Hγ:2.498,2.367
4	Gly	8.562	4.019,4.018		
5	Thr	7.906	4.365	4.237	Hγ:1.150
6	Phe	8.236	4.676	3.194,3.197	
7	Thr	7.943	4.299	4.227	Hγ:1.246
8	Ser	8.128	4.411	4.036,3.966	
9	Asp	8.37	4.607	2.783,2.786	
10	Leu	8.177	4.212	1.655	Hγ:1.611 Hδ:0.933,0.901
11	Ser	7.67	4.427	3.975,3.927	
12	Lys	7.857	4.171	1.974	Hγ:1.486 Hδ:1.756 Hε:3.014
13	Gln	8.028	4.135	2.246	Hγ:2.466,2.437
14	Met	8.388	4.224		
15	Glu	8.411	4.008	2.262	Hγ:2.509,2.305
16	Glu	8.135	4.197		
17	Glu	7.849	4.300		
18	Ala	8.149	4.092	1.578	
19	Val	8.104	3.785	2.253	
20	Arg	8.201	4.019	1.983,1.900	Hγ:1.745 Hδ:3.262 Hε:7.494
21	Leu	8.424	4.196	2.072	Hγ: Hδ:0.944
22	Phe	8.328	4.303	3.459,3.319	
23	Ile	8.818	3.716	2.158	
24	Glu	8.346	4.007		
25	Trp	8.412	4.226	3.817,3.240	
26	Leu	8.671	3.375	1.671,1.794	Hγ:1.617 Hδ:0.827,0.798
27	Lys	8.946	3.927	1.971,1.917	Hγ:1.472 Hδ:1.661 Hε:3.384
28	Asn	7.593	4.751	2.992,2.768	
29	Gly	7.671	4.199,3.529		
30	Gly	8.288	2.710,1.331		
31	Pro		4.519	2.458,2.013	Hγ:2.057 Hδ:3.646,3.031
32	Ser	7.775	4.361	3.858,3.771	
33	Ser	8.056	4.31	4.011,3.768	
34	Gly	7.851	4.313,3.797		
35	Ala	8.106	4.924	1.48	
36	Pro		4.636	2.205,1.871	Hγ:2.005 Hδ:3.840
37	Pro		3.024	1.973,1.477	Hγ:1.764,1.619 Hδ:3.380
38	Pro		4.337	2.201	Hγ:1.943 Hδ:3.206,3.049
39	Ser	7.673	4.359	3.845,3.775	

SG pH 3.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	7.971	4.045	1.978	Hγ:0.875
2	Phe	8.298	5.172	2.986,2.905	Hδ:7.284,Hε:7.167
3	Ile	8.871	4.454	1.679	Hγ:1.422,1.092 Hδ:0.844
4	Thr	8.495	5.251	4.034	Hγ:1.150
5	Ser	8.856	4.984	3.816,3.773	
6	dPro		4.414	2.417,2.023	Hγ:2.128,2.064,Hδ:3.898
7	Gly	8.706	4.052,3.750		
8	Lys	7.938	4.595	1.812	Hγ:1.574,1.433,Hδ:1.701 Hε:3.017
9	Thr	8.369	4.95	3.861	Hγ:1.083,1.087
10	Tyr	9.108	5.217	2.868,2.735	Hδ:6.790 Hε:6.876
11	Thr	8.86	4.861	3.971	Hγ:1.052
12	Glu	8.629	5.129	1.988,1.823	Hγ:2.266,2.220
13	Val	8.636	4.616	1.979	Hγ:0.930
14	dPro		4.375	2.358,1.990	Hγ:2.121,2.053,Hδ:3.869
15	Gly	8.549	4.010,3.770		
16	Orn	7.688	4.566	1.852	Hγ:1.694 Hδ:3.019
17	Lys	8.475	4.701	1.580,1.486	Hγ:1.263 Hδ:1.484,1.075 Hε:2.969 Hζ:7.362
18	Ile	8.974	4.415	1.937	Hγ:1.420,1.186,0.897 Hδ:0.892
19	Leu	8.708	4.163	1.657,1.539	Hγ:1.545 Hδ:0.807,0.722
20	Gln	8.541	4.323	1.999,1.819	Hγ:2.264,2.226 Hε:7.063,7.651

SG NG-pG pH 3.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	8.093	4.048	1.943	Hγ:0.858
2	Phe	8.501	4.906	2.995,2.836	Hδ:7.284 Hε:7.172
3	Ile	8.589	4.429	1.724	Hγ:1.438,1.123,0.874 Hδ:0.816
4	Thr	8.608	4.781	4.205	Hγ:1.198
5	Ser	8.82	4.524	3.850,3.833	
6	Asn	8.915	4.646	2.945,2.841	Hδ:7.029,7.685
7	Gly	8.563	4.026,3.791		
8	Lys	8.019	4.508	1.814	Hγ:1.429,1.367 Hδ:1.655 Hε:2.999
9	Thr	8.431	4.768	3.992	Hγ:1.118
10	Tyr	8.863	5.18	2.878,2.778	Hδ:6.770 Hε:6.915
11	Thr	8.834	4.698	4.042	Hγ:1.114
12	Glu	8.694	5.069	1.985,1.874	Hγ:2.292
13	Val	8.794	4.602	1.978	Hγ:0.939
14	dPro		4.383	2.388,2.026	Hγ:2.131,2.066 Hδ:3.882
15	Gly	8.658	3.994,3.766		
16	Orn	7.974	4.575	1.845	Hγ:1.705 Hδ:3.010
17	Lys	8.6	4.633	1.613,1.515	Hγ:1.112 Hδ:1.306 Hε:2.527
18	Ile	9.013	4.399	1.951	Hγ:1.411,1.208,0.878 Hδ:0.797
19	Leu	8.804	4.186	1.636,1.538	Hγ:1.536 Hδ:0.795,0.734
20	Gln	8.616	4.303	2.016,1.857	Hγ:2.277 Hε:7.194,7.794

SG DG-pG pH 3.00 280K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	8.101	4.056	1.953	Hγ:0.866
2	Phe	8.511	4.944	2.989,2.819	Hδ:7.170 Hε:7.277
3	Ile	8.65	4.454	1.718	Hγ:1.149,0.877 Hδ:0.819
4	Thr	8.645	4.858	4.198	Hγ:1.206
5	Ser	8.865	4.53	3.857,3.815	
6	Asp	9.026	4.671	3.069,2.913	
7	Gly	8.607	4.063,3.773		
8	Lys	7.96	4.518	1.839	Hγ:1.449 Hδ:1.691,1.646 Hε:3.018
9	Thr	8.472	4.838	3.976	Hγ:1.125
10	Tyr	8.933	5.214	2.877,2.779	Hδ:6.910 Hε:6.765
11	Thr	8.854	4.741	4.04	Hγ:1.114
12	Glu	8.721	5.097	2.006,1.884	Hγ:2.291,2.291
13	Val	8.797	4.613	1.988	Hγ:0.946,0.948
14	dPro		4.388	2.400,2.002	Hγ:2.138,2.082 Hδ:3.891,3.888
15	Gly	8.677	4.019,3.771		
16	Orn	7.982	4.584	1.846,1.846	Hγ:1.722,1.693,Hδ:3.027
17	Lys	8.608	4.655	1.610,1.501	Hγ:1.105 Hδ:1.302 Hε:2.911
18	Ile	9.054	4.41	1.962	Hγ:1.420,1.214,0.892 Hδ:0.810
19	Leu	8.835	4.187	1.655,1.548	Hγ:1.206 Hδ:0.812,0.738
20	Gln	8.65	4.311	2.031,1.866	Hγ:2.275,2.275 Hε:6.913,7.387

SG GG-pG pH 3.00 280K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	8.089	4.035	1.943	Hγ:0.851
2	Phe	8.486	4.902	2.995,2.847	Hδ:7.178 Hε:7.278
3	Ile	8.596	4.399	1.724	Hγ:1.602,1.298,0.867 Hδ:0.808
4	Thr	8.6	4.794	4.199	Hγ:1.101
5	Ser	8.772	4.541	3.872,3.835	
6	Gly	8.918	3.986,3.986		
7	Gly	8.562	4.026,3.828		
8	Lys	8.009	4.498	1.774,1.655	Hγ: Hδ:1.441,1.357 Hε:2.993 Hζ:7.610
9	Thr	8.433	4.748	3.989	Hγ:1.104
10	Tyr	8.851	5.18	2.875,2.779	Hδ:6.908 Hε:6.771
11	Thr	8.839	4.692	4.033	Hγ:1.095
12	Glu	8.685	5.069	1.977,1.872	Hγ:2.286
13	Val	8.788	4.603	1.98	Hγ:0.935
14	dPro		4.384	2.387,1.971	Hγ:2.127,2.040 Hδ:3.879
15	Gly	8.658	3.992,3.760		
16	Orn	7.971	4.576	1.832,1.833	Hγ:1.681 Hδ:3.007 Hε:7.679
17	Lys	8.594	4.634	1.605,1.503	Hγ: Hδ:1.290,1.093 Hε: Hζ:7.416
18	Ile	9.015	4.402	1.937	Hγ:1.394,1.216,0.879 Hδ:0.798
19	Leu	8.802	4.186	1.626,1.545	Hγ:1.540 Hδ:0.788,0.731
20	Gln	8.613	4.307	2.002,1.854	Hγ:2.270 Hε:7.189,7.789

SG GP-pG pH 3.00 280K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	8.131	3.999	1.944	Hγ:0.859
2	Phe	8.462	4.722	3.111,2.981	Hδ:7.312 Hε:7.232
3	Ile	8.304	4.222	1.816	Hγ:0.885 Hδ:0.748
4	Thr	8.409	4.392	4.243	Hγ:1.221
5	Ser	8.435	4.528	3.904,3.877	
6	Gly	8.434	4.130,3.894		
7	Pro		4.431	2.278,1.916	Hγ:2.001 Hδ:3.627
8	Lys	8.591	4.353	1.668	Hγ:1.485,1.396 Hδ: Hε:2.991
9	Thr	8.256	4.41	4.114	Hγ:1.136
10	Tyr	8.484	5.08	2.815,2.917	Hδ:6.961 Hε:6.777
11	Thr	8.808	4.508	4.063	Hγ:1.146
12	Glu	8.581	4.966	1.994,1.913	Hγ:2.341
13	Val	8.827	4.59	1.994	Hγ:0.941,0.876
14	dPro		4.395	2.382,2.00	Hγ:2.098 Hδ:3.881
15	Gly	8.627	4.001,3.780		
16	Orn	7.975	4.572	1.834	Hγ:1.699 Hδ:3.012
17	Lys	8.591	4.558	1.649	Hγ:1.353,1.154 Hδ:1.606 Hε:2.978
18	Ile	8.9	4.394	1.909	Hγ:1.662,1.418,0.884 Hδ:0.798
19	Leu	8.72	4.234	1.609	Hγ:1.527 Hδ:0.889,0.764
20	Gln	8.577	4.294	2.028,1.896	Hγ:2.292 Hε:6.969,7.493

SG PG-pG pH 3.00 280K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	8.127	3.999	1.94	Hγ:0.922,0.856
2	Phe	8.456	4.697	3.087,2.962	Hδ:7.216 Hε:7.307
3	Ile	8.277	4.207	1.801	Hγ:1.156,0.872 Hδ:0.841
4	Thr	8.36	4.317	4.157	Hγ:1.199
5	Ser	8.497	4.779	3.859,3.835	
6	Pro		4.428	2.293,1.993	Hγ:2.041 Hδ:3.830,3.733
7	Gly	8.581	3.947,3.947		
8	Lys	8.235	4.409	1.924	Hγ:1.425,1.362 Hδ:1.682 Hε:2.972
9	Thr	8.302	4.429	4.112	Hγ:1.140
10	Tyr	8.509	5.066	2.899,2.820	Hδ:6.782 Hε:6.963
11	Thr	8.804	4.504	4.062	Hγ:1.148
12	Glu	8.574	4.957	1.990,1.906	Hγ:2.344
13	Val	8.826	4.584	1.988	Hγ:0.942
14	dPro		4.391	2.380,1.995	Hγ:2.123,2.055 Hδ:3.873,3.879
15	Gly	8.604	3.982,3.866		
16	Orn	7.973	4.568	1.835	Hγ:1.718 Hδ:3.014
17	Lys	8.58	4.553	1.637	Hγ:1.352 Hδ:1.589 Hε:2.604
18	Ile	8.887	4.391	1.902	Hγ:1.195,0.876 Hδ:0.797
19	Leu	8.713	4.232	1.599,1.524	Hγ:1.524 Hδ:0.757
20	Gln	8.572	4.291	2.029,1.900	Hγ:2.295 Hε:7.204,7.769

SG pG-GG pH 3.00 280K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	8.099	4.01	1.957	Hγ:0.873,0.844
2	Phe	8.444	5.115	2.981,2.899	Hδ:7.149,Hε:7.282
3	Ile	8.929	4.417	1.975	Hγ:1.649,1.080,0.840 Hδ:0.803
4	Thr	8.61	5.218	4.029	Hγ:1.140
5	Ser	8.918	4.967	3.837,3.775	
6	dPro		4.42	2.406,2.009	Hγ:2.102,2.050 Hδ:3.890,3.889
7	Gly	8.817	4.030,3.755		
8	Lys	8.004	4.6	1.820	Hγ:1.430 Hδ:1.674,1.632 Hε:2.998
9	Thr	8.486	4.955	3.87	Hγ:1.091
10	Tyr	9.137	5.14	2.921,2.778	Hδ:6.792 Hε:6.903
11	Thr	8.758	4.759	3.988	Hγ:1.057
12	Glu	8.711	4.944	1.997,1.845	Hγ:2.269
13	Val	8.669	4.23	1.987	Hγ:0.932,0.813
14	Gly	9.269	3.929,3.929		
15	Gly	8.855	4.036,3.854		
16	Orn	7.955	4.515	1.840	Hγ:1.708 Hδ:3.015
17	Lys	8.579	4.572	1.626	Hγ:1.337 Hδ:1.559 Hε:2.616
18	Ile	8.935	4.334	1.936	Hγ:1.590,1.208,0.891 Hδ:0.814
19	Leu	8.786	4.212	1.634,1.582	Hγ:1.576 Hδ:0.933,0.811
20	Gln	8.629	4.3	2.050,2.027	Hγ:2.302,2.265 Hε:7.195,7.784

SG pG-PG pH 3.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	8	4.044		
2	Phe	8.26	5.000	3.030,2.909	Hδ:7.291 Hε:7.172
3	Ile	8.571	4.364	1.883	Hγ:1.676,0.921 Hδ:0.817
4	Thr	8.315	4.95	4.062	Hγ:1.144
5	Ser	8.66	4.909	3.794,3.784	
6	dPro		4.418	2.363,1.996	Hγ:2.057 Hδ:3.847,3.847
7	Gly	8.551	3.997,3.812		
8	Lys	7.988	4.539	1.794	Hγ:1.415 Hδ:1.694 Hε:2.977
9	Thr	8.253	4.759	3.957	Hγ:1.110
10	Tyr	8.785	4.849	3.030,2.850	Hδ:6.774 Hε:7.002
11	Thr	8.375	4.587	4.047	Hγ:1.100
12	Glu	8.419	4.362	2.276,2.276	Hγ:2.056
13	Val	8.365	4.392	2.29	Hγ:0.859
14	Pro		4.398	2.302,1.950	Hγ:2.055 Hδ:3.862,3.691
15	Gly	8.441	3.946,3.946		
16	Orn	8.183	4.357	1.758	Hγ:1.715 Hδ:3.010
17	Lys	8.357	4.318	1.770	Hγ:1.433,1.377 Hδ:1.717 Hε:2.986
18	Ile	8.255	4.14	1.835	Hγ:1.477,1.187,0.881 Hδ:0.838
19	Leu	8.364	4.391	1.950	Hγ: Hδ:0.933
20	Gln	8.321	4.304	2.114,1.977	Hγ:2.360 Hε:6.855,7.512

SG NG-NG pH 3.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	8.47	4.137		
2	Phe	8.986	4.548	2.948,2.742	Hδ:7.184 Hε:7.273
3	Ile	8.28	4.32	1.75	Hγ:1.387 Hδ:0.854
4	Thr	8.323	4.584	4.191	Hγ:1.174
5	Ser	8.515	4.475	3.833	
6	Asn	8.601	4.66	2.887,2.807	Hδ:7.563,6.891
7	Gly	8.378	3.985,3.816		
8	Lys	7.981	4.435	1.723,1.603	Hγ:1.394,1.324 Hδ:1.651 Hε:2.977
9	Thr	8.162	4.55	4.045	Hγ:1.110
10	Tyr	8.525	4.904	2.946,2.845	Hδ:6.980 Hε:6.763
11	Thr	8.443	4.495	4.04	Hγ:1.106
12	Glu	8.412	4.725	1.983,1.898	Hγ:2.311
13	Val	7.933	4.031	1.915	Hγ:0.820
14	Asn	8.299	4.805	3.038,2.884	Hδ:6.814,7.391
15	Gly	8.471	3.971,3.794		
16	Orn	7.95	4.45	1.740	Hγ:1.658Hδ:2.995
17	Lys	8.377	4.441	1.640	Hγ:1.225 Hδ:1.379 Hε:2.983
18	Ile	8.551	4.268	1.858	Hγ:1.223 Hδ:0.865
19	Leu	8.465	4.277	1.957	Hγ:1.571 Hδ:0.901
20	Gln	8.38	4.288	1.963	Hγ:2.292 Hε:7.080,7.593

SG NG-PG pH 6.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	7.951	4.026	1.929	Hγ:0.836
2	Phe	8.289	4.733	3.084,2.945	Hδ:7.207 Hε:7.251 Hζ:7.267
3	Ile	8.166	4.278	1.785	Hγ:1.396 0.865 Hδ:0.831
4	Thr	8.25	4.468	4.221	Hγ:1.198
5	Ser	8.427	4.474	3.891,3.843	
6	Asn	8.489	4.728	2.842,2.840	Hδ:7.575,6.896
7	Gly	8.341	3.987,3.874		
8	Lys	8.04	4.411	1.742	Hγ:1.325 Hδ:1.376 Hε:2.962
9	Thr	8.158	4.424	4.114	Hγ:1.137
10	Tyr	8.362	4.703	3.033,2.937	Hδ:7.080 Hε:6.791
11	Thr	8.079	4.332	4.084	Hγ:1.135
12	Glu	8.31	4.309	1.982,1.925	Hγ:2.217
13	Val	8.312	4.297	1.954	Hγ:0.864
14	Pro		4.403	2.302,1.930	Hγ:2.050,1.980 Hδ:3.856,3.676
15	Gly	8.449	3.945,3.945		
16	Orn	8.163	4.349	1.768,1.725	Hγ:1.691 Hδ:3.012
17	Lys	8.354	4.372	1.700	Hγ: Hδ:1.625 Hε:3.023
18	Ile	8.245	4.134	1.837	Hγ:1.479 0.879 Hδ:0.841
19	Leu	8.315	4.355	1.702,1.664	Hγ:1.658 Hδ:0.940,0.855
20	Gln	8.322	4.303	1.982	Hγ:2.357 Hε:7.511,6.851

WW Ncap1 pH 7.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Ser		4.322	3.983	
2	Pro		4.668	2.279,1.883	Hγ:1.975 Hδ:3.691,3.442
3	Pro		4.667	2.268,1.807	Hγ:1.961 Hδ:3.752,3.582
4	Pro		4.402	2.284,1.931	Hγ:2.015 Hδ:3.774,3.606
5	Gly	8.451	3.965,3.987		
6	Gly	8.274	3.987,3.965		
7	Trp	8.227	4.462	3.138,3.085	Hδ:7.025 Hε:10.086,7.164 Hζ:7.429,7.145 Hη:6.982
8	Asn	7.991	4.88	2.939,2.502	
9	Pro		3.992	2.283	Hγ:1.951 Hδ:3.685,3.530
10	Ala	8.018	4.244	1.398	
11	Thr	7.475	4.294	4.199	Hγ:1.121
12	Gly	8.16	3.873,3.664		
13	Arg	7.557	4.363	1.749	Hγ:1.573,1.441 Hδ:3.106 Hη:7.559
14	Tyr	8.347	4.707	3.146,2.861	Hδ:7.172 Hε:6.833
15	Thr	7.819	4.213	4.141	Hγ:1.140

WW Ncap2 pH 7.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Lys		4.276	1.895	Hγ:1.500 Hδ:1.735 Hε:3.016
2	Pro		4.49	2.303,1.869	Hγ:2.004 Hδ:3.744,3.580
3	Leu	8.414	4.591	1.753,1.473	Hγ:1.597 Hδ:0.925
4	Pro		4.672	2.336,1.784	Hγ:2.005 Hδ:3.859,3.555
5	Pro		4.383	2.291,1.885	Hγ:2.023 Hδ:3.819,3.600
6	Gly	8.515	4.103,3.901		
7	Thr	8.288	4.473	4.124	Hγ:1.278
8	Trp	8.909	4.22	3.096,3.030	Hδ:6.889 Hε:10.043,6.495 Hζ:7.375,6.800 Hη:7.056
9	Asn	7.734	4.893	3.114,2.385	Hδ:7.098,7.347
10	Pro		3.635	2.34,1.960	Hγ:2.026
11	Ala	7.839	4.19	1.397	
12	Thr	7.131	4.298	4.158	Hγ:1.045
13	Gly	8.084	3.817,3.434		
14	Arg	7.135	4.42	1.796,1.533	Hγ:1.424 Hδ:3.123
15	Tyr	8.42	4.74	2.991,2.871	Hδ:6.889 Hε:7.195
16	Thr	8.256	4.275	4.13	Hγ:1.175

WW Ncap3 pH 7.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Lys		4.277	1.900,	Hγ:1.505 Hδ:1.725 Hε:3.066
2	Pro		4.48	2.294,1.872	Hγ:1.995 Hδ:3.741,3.586
3	Leu	8.407	4.561	1.727	Hγ:1.549 Hδ:0.915
4	Pro		4.663	2.287,1.852	Hγ:1.984 Hδ:3.793,3.573
5	Pro		4.39	2.240,1.880	Hγ:1.997 Hδ:3.794,3.592
6	Gly	8.382	3.920,3.920		
7	Gly	8.181	4.028,3.991		
8	Thr	8.513	4.507	4.151	Hγ:1.203
9	Trp	8.732	4.303	3.090,3.038	Hδ:6.933 Hε:10.068,6.668 Hζ:7.397,6.847 Hη:7.082
10	Asn	7.776	4.902	2.992,2.487	Hδ:7.039,7.344
11	Pro		3.869	2.326	Hγ:1.982 Hδ:3.736,3.556
12	Ala	7.908	4.197	1.393	
13	Thr	7.25	4.288	4.16	
14	Gly	8.097	3.830,3.496		
15	Arg	7.238	4.421	1.781	Hδ:2.331,1.421
16	Tyr	8.432	4.789	3.029,2.860	Hδ:6.851 Hε:7.150
17	Thr	8.199	4.207	4.209	Hγ:1.125

WWm1b pH 7.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Lys		4.321		Hε:3.058
2	Pro		4.523	2.320,1.898	Hγ:2.021 Hδ:3.766,3.607
3	Leu	8.413	4.584	1.745,1.593	Hγ:1.525 Hδ:0.900
4	Pro		4.395	2.279,1.883	Hγ:1.958 Hδ:3.810,3.517
5	Gly	8.371	3.922,3.908		
6	Gly	8.285	3.967,3.964		
7	Trp	7.978	4.758	3.274,3.195	Hδ:7.160 Hε:10.014,7.512 Hζ:7.440,7.037 Hη:7.177
8	Thr	8.089	4.441	4.13	Hγ:1.104
9	Tyr	8.068	4.935	2.966,2.825	Hδ:6.789 Hε:7.022
10	Arg	8.421	4.512	1.748,1.449	Hγ:1.651 Hδ:2.883
11	Thr	8.399	4.607	4.135	Hγ:1.206
12	Val	8.206	4.473	2.182	Hγ:0.936
13	Pro		4.341	2.261,1.894	Hγ:2.000 Hδ:3.771,3.584
14	Asn	8.423	4.644	2.977,2.906	
15	Gly	8.406	4.053,4.034		
16	Lys	8.082	4.444	1.732	Hγ:1.460 Hδ:1.686 Hε:3.005
17	Thr	8.319	4.662	4.057	Hγ:1.109
18	Tyr	8.693	5.161	3.073,2.854	Hδ:6.705 Hε:7.061
19	Tyr	9.282	5.198	3.016,2.955	Hδ:7.182 Hε:7.343
20	Tyr	8.516	3.892		Hδ:6.406 Hε:5.842
21	Asn	7.663	4.768	3.112,3.108	
22	Pro		3.969	2.360,2.001	Hγ:2.032 Hδ:3.831
23	Ala	7.874	4.229	1.481	
24	Thr	6.972	4.425	4.343	Hγ:1.106
25	Gly	8.318	4.030,3.740		
26	Thr	7.159	4.72	4.22	Hγ:1.186
27	Trp	8.766	5.232	3.098,3.097	Hδ:7.344 Hε:10.084,7.507 Hζ:7.178,7.314 Hη:7.111
28	Thr	9.179	4.839	4.07	Hγ:1.227
29	Ser	8.629	4.89	3.684,3.551	
30	Lys	8.337	4.29	1.751	Hγ:1.417 Hδ:1.626 Hε:2.865

WWm2 pH 7.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Lys				
2	Leu	8.299	4.592	1.724,1.512	Hγ:1.649 Hδ:0.923,0.860
3	Pro		4.463		Hδ:3.727
4	Gly	8.422	4.794,3.924		
5	Gly	8.304	4.014,3.686		
6	Trp	7.885	4.897	3.292,3.284	
7	Thr	8.377	4.666	4.189	Hγ:1.169
8	Tyr	8.284	5.38	3.000,2.795	Hδ:6.995 Hε:6.623
9	Arg	7.1	4.726	1.529	Hγ:1.386 Hδ:3.054
10	Thr	8.765	4.886	4.008	Hγ:1.110
11	Val	8.21	4.565	1.89	Hγ:0.940
12	Ala				
13	Asp	8.287	4.497	2.839,2.800	
14	Gly	8.357	4.164,3.711		
15	Lys	7.938	4.421	1.899	Hγ:1.553,1.422 Hδ:1.698 Hε:3.027
16	Thr				
17	Tyr				
18	Tyr				
19	Tyr				
20	Asn				
21	Pro				
22	Ala				
23	Thr	7.055	4.693	4.195	Hγ:1.212
24	Gly				
25	Thr				
26	Trp	8.734	5.487	3.059,3.057	Hδ:7.205 Hε:10.110,7.516 Hζ:7.452,6.982 Hη:7.161
27	Thr	9.508	5.125	4.42	Hγ:1.374
28	Asn	8.85	5.094	2.876,2.722	Hδ:6.732,7.051
29	Glu	8.295	4.406	2.257,1.922	Hγ:2.218
30	Arg	8.321	3.494	1.523	Hγ:1.378 Hδ:3.057
31	Pro				
32	Ser				

SG NGNGW2 pH 3.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	7.943	4.045	1.953	Hγ:0.858
2	Phe	8.337	5.277	3.098,2.842	Hδ:7.051 Hε:7.146
3	Ile	9.189	4.564	1.678	Hγ:1.296,1.092,0.840 Hδ:0.837
4	Trp	8.603	4.718	2.864,2.082	Hδ:7.270 Hε:10.060,7.280 Hζ:7.496,7.501 Hη:7.145
5	Ser	8.513	4.27	3.502,3.488	
6	Asn	8.748	4.193	2.864,2.568	Hδ:6.820,7.350
7	Gly	6.95	3.877,3.364		
8	Lys	7.001	4.269	1.467	Hγ:1.173 Hδ:1.500 Hε:2.890 Hζ:5.177
9	Trp	8.105	5.197	2.980,2.980	Hδ:6.556 Hε:9.880,6.997 Hζ:6.220,6.218 Hη:6.225
10	Tyr	9.316	5.046	2.861,2.758	Hδ:6.910 Hε:6.775
11	Thr	8.803	4.83	3.939	Hγ:1.013
12	Glu	8.599	4.955	1.966,1.811	Hγ:2.206
13	Val	8.584	4.172	1.923	Hγ:0.885
14	Asn	9.226	4.469	2.997,2.735	Hδ:6.892,7.559
15	Gly	8.551	4.028,3.704		
16	Orn	7.868	4.498	1.810	Hγ:1.665 Hδ:2.994 Hε:7.589
17	Lys	8.422	4.585	1.568	Hγ:1.310 Hδ:1.597 Hε:2.886 Hζ:7.454
18	Ile	8.841	4.348	1.885	Hγ:1.401,1.147,0.871 Hδ:0.865
19	Leu	8.587	4.127	1.582,1.504	Hγ:1.165 Hδ:0.737
20	Gln	8.432	4.232	2.008,1.800	Hγ:2.226,2.175 Hε:7.046,7.592

SG NGPGW2 pH 3.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	7.957	4.03	1.896	Hγ:0.805
2	Phe	8.264	5.171	3.011,2.769	Hδ:7.135 Hε:7.009
3	Ile	8.992	4.545	1.682	Hγ:1.361
4	Trp	8.511	4.685	2.863,2.139	Hδ:6.615 Hε:9.862,6.216 Hζ:7.254,6.663 Hη:6.978
5	Ser	8.421	4.227	3.546,3.486	
6	Asn	8.665	4.206	2.831,2.557	
7	Gly	7.05	3.852,3.386		
8	Lys	7.006	4.183	1.458	Hγ:1.117,1.090 Hδ:1.451 Hε:2.949
9	Trp	8.001	5.087	2.957,2.960	Hδ:7.239 Hε:10.043,7.093 Hζ:7.277,7.091 Hη:7.081
10	Tyr	9.161	4.781	2.960,2.781	Hδ:6.989 Hε:6.741
11	Thr		4.563	3.973	Hγ:1.008
12	Glu	8.44	4.417	1.998,1.833	Hγ:2.301,2.301
13	Val	8.309	4.334	1.999	Hγ:0.880
14	Pro		4.33	2.223,1.869	Hγ:1.995 Hδ:3.612,3.788
15	Gly	8.39	3.896,3.894		
16	Orn	8.122	4.324	1.844,1.818	Hγ:1.564,1.527 Hδ:2.939 Hε:7.557
17	Lys	8.312	4.274	1.706,1.671	Hγ:1.352 Hδ:1.623 Hε:2.913 Hζ:7.485
18	Ile	8.19	4.09	1.766	Hγ:1.098 0.823 Hδ:0.801
19	Leu	8.31	4.32	1.693	Hγ: Hδ:0.797
20	Gln	8.279	4.249	2.045,1.927	Hγ:2.288

SG NGpGW2 pH 3.00 300K					
#	Res	HN	Hα	Hβ (Hβ')	Others
1	Val	7.986	4.081	2.019	Hγ:0.902
2	Phe	8.389	5.407	3.174,2.889	Hδ:7.094 Hε:7.177
3	Ile	9.368	4.649	1.715	Hγ:1.505 Hδ:0.877
4	Trp	8.68	4.784	2.864,2.004	Hδ:6.514 Hε:9.893,6.161 Hζ:7.035,6.708 Hη:7.173
5	Ser	8.614	4.291	3.528,3.504	
6	Asn	8.84	4.187	2.902,2.594	
7	Gly	6.851	3.907,3.355		
8	Lys	6.952	4.329	1.485	Hγ:1.317,1.090 Hδ: Hε:2.950
9	Trp	8.177	5.283	3.011,3.004	Hδ:7.315 Hε:10.078,7.127 Hζ:7.293,7.558 Hη:7.338
10	Tyr	9.461	5.188	2.867,2.761	Hδ:6.906 Hε:6.810
11	Thr	9.029	4.976	3.965	Hγ:1.040
12	Glu	8.702	5.159	1.995,1.832	Hγ:2.270,2.223
13	Val	8.665	4.621	1.989	Hγ:0.941
14	dPro		4.376	2.184,1.869	Hγ:2.122,2.059 Hδ:3.865,3.789
15	Gly	8.573	4.009,3.777		
16	Orn	7.962	4.581	1.859,1.818	Hγ:1.717,1.679 Hδ:3.023
17	Lys	8.49	4.747	1.706,1.671	Hγ:1.352 Hδ:1.623 Hε:2.913
18	Ile	9.051	4.444	1.952	Hγ:1.431,1.188,0.909 Hδ:0.807
19	Leu	8.716	4.089	1.632,1.520	Hγ:0.761 Hδ:0.714,0.797
20	Gln	8.513	4.255	2.017,1.800	Hγ:2.243,2.155

Vita

Frederick Michael Lewis Hudson

Education

Doctor of Philosophy, Chemistry; 2006
University of Washington
Seattle, WA

Masters of Science, Chemistry; 2002
University of Washington
Seattle, WA

Bachelors of Science, Chemistry; 1999
George Washington University
Washington, D.C.

Publications

Hudson, F.M. and Andersen, N. H. **2006**. Measuring Cooperativity in the Formation of a Three-stranded β Sheet (Double Hairpin). *Biopolymers* **83**: In press.

Andersen, N.H.; Olsen, K.A.; Fesinmeyer, R.M.; Tan, X.; Hudson, F.M.; Eidenschink, L.A.; Farazi, S.R. **2006.** Minimization and Optimization of Designed β-Hairpin Folds. *Journal of the American Chemical Society* **128:** 6101-6110.

Fesinmeyer, R.M.; Hudson, F.M.; Olsen, K.A.; White, G.W.N.; Euser, A.; and Andersen, N.H. **2005.** Chemical Shifts Provide Fold Populations and Register of β Hairpins and β Sheets. *Journal of Biomolecular NMR* **33:** 213-231.

Fesinmeyer, R.M.; Hudson, F.M.; and Andersen, N.H. **2004**. Enhanced Hairpin Stability Through Loop Design: the Case of the Protein G B1 Domain Hairpin. *Journal of the American Chemical Society* **126:** 7238-7243.

Hudson, F.M. and Andersen, N.H. **2004.** Exenatide: NMR/CD Evaluation of the Medium Dependence of Conformation and Aggregation State. *Biopolymers (Peptide Science)* **76:** 298-308.

Andersen, N.H.; Fesinmeyer, R.M.; Hudson, F.M.; and Lin, J.C. **2003**. Analysis of ß Peptides Using Chemical Shift Deviations. In *18th American Peptide Symposium on Peptide Revolution: Genomics, Proteomics and Therapeutics*. Biopolymers, Boston, MA.