

國立臺灣大學理學院化學所

碩士論文

Department of Chemistry


College of Science

National Taiwan University

Master Thesis

d 位置胺基酸的側鏈結構對於雙螺旋穩定度的影響

Effect of Side Chain Structure of *d*-Position Residue  
on Coiled Coil Stability



翁乙壬

YI-JEN WENG

指導教授：陳平博士

Advisor: Richard Ping Cheng, Ph.D.

中華民國 102 年 1 月

January, 2013

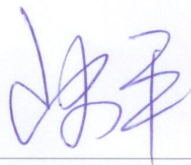
國立臺灣大學碩士學位論文  
口試委員會審定書

d 位置胺基酸的側鏈結構對於雙螺旋穩定度的影響

Effect of Side Chain Structure of *d*-Position Residue  
on Coiled Coil Stability

本論文係 翁乙壬 君 (學號 R99223123) 在國立臺灣大學化學系完成之碩士學位論文，於民國 101 年 7 月 11 日承下列考試委員審查通過及口試及格，特此證明。

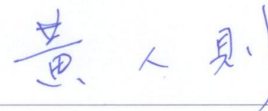
口試委員：



(簽名)

(指導教授)







系主任、所長

(簽章)

## 誌謝

四年的實驗室生涯經歷了許許多多：國科會大專生計畫、碩士班推甄、大學部壁報、化學年會、可怕的高等有機、儀器殺手、淬鍊心智的有機合成、培養做事習慣與效率的各種實驗組合 (5 timers at a time)、paper 海、每學期絞盡腦汁的大報告、研究所壁報、多少血淚多少肝的論文，以及最後的口試舞台。過五關斬六將，終於來到了尾聲。當年那個天真浪漫但什麼都不會的小朋友，如今，在能力上與心智上，某種程度而言，也能獨當一面了吧。而這一切--這四年來我的改變與成長--都必須歸功於陳平老師的教誨。謝謝老師四年來給予的教導、機會與容忍，而塑造出現在的我。此外，也要感謝口試委員陳佩燁老師和黃人則老師對於這份論文的指導，與老師們的討論過程中知道了自己思考上的不周延之處。

對於每天一起生活、一起經歷各種歡笑與苦難的每一位實驗室夥伴，我必須獻上我最誠摯的感謝。真心謝謝各位的陪伴、安慰與教導，支持我度過這段歲月。在此之中，想特別感謝以下三位：吳政勳學長、郭秀婷學姊，和蔡幸芸小妹。政勳學長，我已經不知道該怎麼表達對你的感謝了。實驗上、學術上，你總是提供了非常寶貴的建議與幫助，是令人安心的後援。當我遭遇各種不順遂時，你總是耐心聆聽我雜亂無章的敘述，然後幫我理出頭緒、提出解決辦法。從你身上我真的學到了很多，不論是專業上或是待人處事上，謝謝你。秀婷學姊，如同親姊的存在，四年來也真是辛苦妳了，忍受我的聒噪、我的扭動、我的粗俗野蠻、我的無厘頭。謝謝妳總是扮演著實驗室潤滑劑的角色，也常很乾脆的幫大家擔下實驗，讓我們能夠順利度過研究所生涯。幸芸妹，妳的存在讓我灰暗的碩二生涯多了很多樂趣。妳真的是個很有趣又很可愛的人，和妳聊天、了解妳腦中的「蔡幸芸世界」真的是件非常紓解壓力的事情。謝謝妳當我的小福夥伴、常常幫我買午晚餐、也常常陪我這個膽小學姊一起待到最後、一起關實驗室。謝謝林珽蔚，在這段歲月中的幫助、關懷、陪伴、容忍與體諒。從大學到研究所，人生中重要階段之一，謝謝你和我一起走過，讓我有許多體悟與成長。謝謝隔壁楊吉水老師實驗室的學長與同學。在那段做有機合成的日子中，感謝你們很慷慨的出借藥品與儀器、並提供了許多寶貴的經驗。

最後，必須謝謝我的家人。這段日子中，實驗室生活占了絕大部分，全家人一起吃個飯都是件難得的事。感謝老爸，當我實驗做到半夜才結束時，你會打著哈欠騎著車載我回家。感謝老媽總是切好水果，讓我回家可以開了冰箱就開始吃；偶爾還會燉雞湯或煮魚湯讓我補補身體。感謝老弟偶爾讓我欺負，讓我得以調劑身心。如今我的學生生涯終於結束了，老爸老媽心頭上的壓力可以解除了吧。謝謝你們十幾年來的栽培，以後你們就開心到處遊山玩水吧。

## 中文摘要

雙螺旋由 2 到 5 個 $\alpha$ 螺旋互相纏繞而成，是個生物體內常見的結構單元。雙螺旋在於轉錄因子、肌肉收縮系統、細胞骨架等系統之中，而這些生理功能來自於雙螺旋明確且穩定的結構。此研究探討了位在雙螺旋交界面的 d 位置殘基的側鏈結構對於雙螺旋穩定度的影響。研究所使用的雙螺旋系統為 GCN4，將其第二個 d 位置殘基 (Leu12) 替換成帶有各式結構與極性的胺基酸。IaLd 雙螺旋則用來測定胺基酸本身對於雙螺旋結構的偏好程度。實驗利用固相勝肽合成法來合成出各個雙螺旋，藉由圓二色光譜儀來監控由鹽酸胍促成的雙螺旋變性過程，並計算出雙螺旋變性的自由能。除此之外，也利用薄層色層分析法量測了胺基酸的疏水性。至於胺基酸的側鏈結構則以各種結構參數去描述，這些參數包含  $E_s$ , MR, [L, B1, B5] 和側鏈體積。數據分析顯示，當雙螺旋的 d 位置殘基為烷基時，側鏈結構和胺基酸疏水性兩者相比之下，側鏈結構對於雙螺旋穩定度的影響較大。

關鍵字：雙螺旋、GCN4、 $\alpha$ 螺旋偏好程度、胺基酸疏水性、側鏈結構

## Abstract

Coiled coil is a superhelical twist formed by two to five wrapping  $\alpha$ -helices. It is a common structural motif that can be found in transcription factors, cytoskeletal systems, contractile systems and etc. These biochemical roles rely on well-defined and stable structures of coiled coils. Accordingly, the effect of side chain structure of *d*-position residue, which is buried in the coiled coil interface, on coiled coil stability was investigated. GCN4 coiled coil was employed, of which the 2<sup>nd</sup> *d*-residue (Leu12) was substituted with various amino acids. IaLd coiled coils were used to obtain the coiled coil propensities of these amino acids. Guanidinium denaturation of the coiled coils was monitored by circular dichroism spectroscopy. Free energy of unfolding was derived from the guanidinium denaturation data. The hydrophobicities of the amino acids were measured by thin layer chromatography. Structural parameters  $E_s$ , MR, [L, B1, B5], and side chain volume were employed in the analysis. Results show that the shape and size of the residue side chain contribute more than hydrophobicity to the coiled coil stability in coiled coils with the *d*-position residue bearing aliphatic side chains.

*Keywords:* coiled coil, GCN4, helix propensity, amino acid hydrophobicity, side chain structure

# Table of Contents

口試委員審定書 .....	I
誌謝 .....	II
中文摘要 .....	III
Abstract .....	IV
Table of Contents .....	V
List of Figures .....	VII
List of Tables .....	IX
Abbreviations .....	X
<b>Chapter 1 Introduction</b> .....	<b>1</b>
1.1 Proteins .....	2
1.2 Protein Structures .....	4
1.3 Forces in Protein Folding .....	6
1.4 $\alpha$ -Helix .....	8
1.5 Coiled coil .....	10
1.6 Thesis Overview .....	11
1.7 References .....	12
<b>Chapter 2 Effect of Side Chain Structure of <i>d-Position</i> Residue on Coiled Coil Stability</b> .....	<b>15</b>
2.1 Introduction .....	16
2.1.1 Coiled Coils in Biological Systems .....	16
2.1.2 GCN4 Coiled Coil Motif .....	17
2.1.3 Coiled Coil Sequence and Structure .....	18
2.1.4 Knobs-Into-Holes Interactions of a Dimeric Coiled Coil .....	21
2.1.5 Coiled Coil Stability .....	21
2.1.6 Chapter Overview .....	26
2.2 Results and Discussions .....	27
2.2.1 Design of GCN4-Derived Peptides .....	27
2.2.2 Peptide Synthesis of GCN4-Derived Peptides .....	30
2.2.3 UV-Visible Spectroscopy (UV-vis) of GCN4-Derived Peptides .....	32
2.2.4 Circular Dichroism (CD) Spectroscopy of GCN4-Derived Peptides .....	33
2.2.5 Guanidinium Denaturation of GCN4-Derived Peptides .....	37
2.2.6 $\Delta G_{\text{unfold, H}_2\text{O}}$ of GCN4-Derived Peptides .....	43

2.2.7 Design of IaLd-Derived Peptides .....	47
2.2.8 Peptide Synthesis of IaLd-Derived Peptides .....	49
2.2.9 UV-Visible Spectroscopy (UV-vis) of IaLd-Derived Peptides.....	50
2.2.10 Circular Dichroism (CD) Spectroscopy of IaLd-Derived Peptides.....	50
2.2.11 Guanidinium Denaturation of IaLd-Derived Peptides.....	53
2.2.12 $\Delta G_{\text{unfold, H}_2\text{O}}$ of IaLd-Derived Peptides .....	57
2.2.13 Measuring Hydrophobicities of the Amino Acids.....	58
2.2.14 Size and Shape Parameters .....	59
2.2.15 Discussion.....	60
2.3 Conclusion.....	65
2.4 Acknowledgement .....	65
2.5 Experimental Section.....	66
2.5.1 General Materials and Methods.....	66
2.5.2 Peptide Synthesis.....	70
2.5.3 UV-Visible Spectroscopy .....	88
2.5.4 Circular Dichroism Spectroscopy.....	88
2.5.5 Guanidinium Denaturation .....	89
2.5.6 Derivation of $\Delta G_{\text{unfold, H}_2\text{O}}$ .....	90
2.5.7 Measurement of Hydrophobicity by Thin Layer Chromatography.....	91
2.5.8 Calculation of Side Chain Volume .....	91
2.6 References .....	92
2.7 Appendix .....	97
2.7.1 Guanidinium Denaturation Curves of GCN4-Xaa .....	97
2.7.2 Guanidinium Denaturation Curves of IaLd-Xaa .....	100
2.7.3 NMR Spectrum of Fmoc-Allo Ile-OH.....	102
2.7.4 NMR Spectrum of Fmoc-Cpa-OH .....	103
2.7.5 NMR Spectrum of Fmoc-Nle-OH .....	104
2.7.6 ESI-MS Spectrum of Fmoc-Allo Ile-OH.....	105
2.7.7 ESI-MS Spectrum of Fmoc-Cpa-OH .....	106
2.7.8 ESI-MS Spectrum of Fmoc-Nle-OH .....	107
2.7.9 Elsevier License for Reprint of Figure .....	108

## List of Figures

<b>Figure 1-1.</b> Dihedral angles $\phi$ and $\psi$ .....	3
<b>Figure 1-2.</b> Ramachandran plot. ....	3
<b>Figure 1-3.</b> Four levels of protein structure. ....	4
<b>Figure 1-4.</b> A hydrogen bond.....	7
<b>Figure 1-5.</b> The $\alpha$ -helix backbone hydrogen bond. ....	9
<b>Figure 1-6.</b> Illustration of a coiled coil. ....	11
<b>Figure 2-1.</b> GCN4-DNA complex. ....	18
<b>Figure 2-2.</b> Illustration of knobs-into-hole interactions, parallel packing, and perpendicular packing. ....	22
<b>Figure 2-3.</b> Structures of leucine, trifluoroleucine, and hexafluoroleucine.....	25
<b>Figure 2-4.</b> The sequence and the helical wheel diagram of GCN4.....	29
<b>Figure 2-5.</b> CD spectrum of GCN4-Phe <sub>2</sub> .....	29
<b>Figure 2-6.</b> Chemical structures of the amino acids. ....	30
<b>Figure 2-7.</b> CD spectra of GCN4-Xaa peptides at 30 $\mu$ M peptide concentration in 50 mM phosphate and 150 mM NaCl buffer at pH 7 and 4 $^{\circ}$ C.....	34
<b>Figure 2-8.</b> A typical guanidinium titration curve with the proper folded and unfolded baselines depicted. ....	38
<b>Figure 2-9.</b> Guanidinium denaturation curves for GCN4-Xaa peptides at 30 $\mu$ M peptide in 50 mM phosphate, 150 mM NaCl, and 0 M to 6 M (with 0.1 M intervals) guanidinium chloride at pH 7 and 4 $^{\circ}$ C as monitored by CD at 222 nm reported in mean residue ellipticity. ....	39
<b>Figure 2-10.</b> Fraction unfolded as a function of guanidinium concentration for GCN4-Xaa peptides as derived from the guanidinium denaturation curves. .....	40
<b>Figure 2-11.</b> The plot of $\Delta G_{\text{unfold}}$ against guanidinium concentration for GCN4-Tba ..	44
<b>Figure 2-12.</b> The bar graph of $\Delta G_{\text{unfold, H}_2\text{O}}$ of GCN4 peptides. ....	44
<b>Figure 2-13.</b> $\Delta G_{\text{unfold}}$ of GCN4-Leu and GCN4-Tba. ....	46
<b>Figure 2-14.</b> Sequence and helical wheel of IaLd. ....	48
<b>Figure 2-15.</b> CD spectra of IaLd-Xaa at 20 $\mu$ M peptide in 10 mM MOPS at pH 7.5 and 25 $^{\circ}$ C.....	51
<b>Figure 2-16.</b> Guanidinium denaturation curves for IaLd-Xaa peptides at 20 $\mu$ M peptide concentration in 10 mM MOPS with 0 to 6 M (at 0.1 M intervals)	

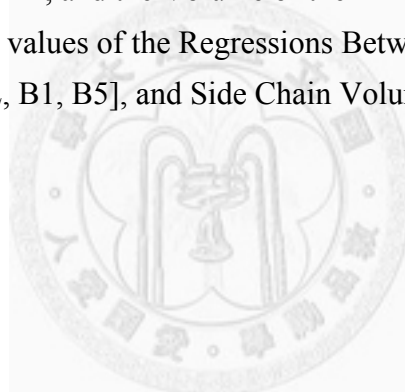


guanidinium chloride at pH 7.5 and 25 °C as monitored by CD at 222 nm reported in mean residue ellipticity .....	55
<b>Figure 2-17.</b> Fraction unfolded as a function of guanidinium concentration for IaLd-Xaa peptides as derived from guanidinium denaturation curves.....	56
<b>Figure 2-18.</b> The bar graph of $\Delta G_{\text{unfold, H}_2\text{O}}$ of IaLd peptides. ....	57
<b>Figure 2-19.</b> Regressions between the residual energy and log $P_{\text{ow}}$ , $E_s$ , MR, and side chain volume. ....	62



## List of Tables

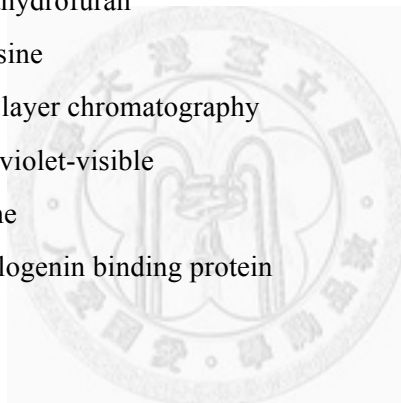
<b>Table 2-1.</b> Crude Yield of Peptide Synthesis, Molecular Formula, Calculated $[MH^+]$ , and Observed $m/z$ of GCN4 Peptides .....	32
<b>Table 2-2.</b> Concentrations and Regression Coefficients of GCN4 Peptides .....	33
<b>Table 2-3.</b> $[\theta]_{222}$ at 0 M Guanidinium, Melting Concentrations ( $[C]_m$ ), $m$ values, and $\Delta G_{\text{unfold, H}_2\text{O}}$ of GCN4 Peptides .....	35
<b>Table 2-4.</b> Crude Yield of Peptide Synthesis, Molecular Formula, Calculated $[MH^+]$ , and Observed $m/z$ of IaLd Peptides .....	49
<b>Table 2-5.</b> Concentrations and Regression Coefficients of IaLd Peptides .....	50
<b>Table 2-6.</b> $[\theta]_{222}$ at 0 M Guanidinium, Melting Concentrations ( $[C]_m$ ), $m$ values, and $\Delta G_{\text{unfold, H}_2\text{O}}$ of IaLd Peptides .....	52
<b>Table 2-7.</b> $\log P_{\text{ow}}$ of the Amino Acids .....	58
<b>Table 2-8.</b> $E_s$ , [L, B1, B5], MR, and the Volume of the Amino Acids Side Chains .....	60
<b>Table 2-9.</b> Equations and R values of the Regressions Between Residual Energy, $\log P_{\text{ow}}$ , $E_s$ , MR, [L, B1, B5], and Side Chain Volume .....	62



## Abbreviations

Abu	Aminobutyric acid
Allo Ile	Alloisoleucine
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
CaM	Calmodulin
CD	Circular dichroism
Cha	Cyclohexylalanine
Cpa	Cyclopentylalanine
DIEA	Diisopropylethylamine
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
ESI-MS	Electrospray ionization mass spectrometry
Fmoc	N-9-Fluorenylmethoxycarbonyl
Fmoc-Osu	N-(9-Fluorenylmethoxycarbonyloxy) succinimide
Gln	Glutamine
Glu	Glutamate
Gly	Glycine
HBTU	O-1H-benzotriazol-1-yl-1,1,3,3-tetramethyluronium hexafluoro phosphate
Hfl	Hexafluoroleucine
His	Histidine
HOBt	1-Hydroxybenzotriazole
Ile	Isoleucine
KIH	Knobs-into-holes
Leu	Leucine
Lys	Lysine
MALDI-TOF MS	Matrix-assisted laser desorption ionization time-of-flight
Met	Methionine
MOPS	3-(N-morpholino)propanesulfonic acid
MR	Molecular refractivity
MRE	Mean residue ellipticity
MW	Molecular weight

Nle	Norleucine
NMR	Nuclear magnetic resonance
Nva	Norvaline
Pff	Pentafluorophenylalanine
Phe	Phenylalanine
QSAR	Quantitative structure–activity relationship
RNA	Ribonucleic acid
RP-HPLC	Reverse phase-high performance liquid chromatography
Ser	Serine
SPPS	Solid phase peptide synthesis
Tba	<i>tert</i> -Butylalanine
TFA	Trifluoroacetic acid
Tfl	Trifluoroleucine
THF	Tetrahydrofuran
Tyr	Tyrosine
TLC	Thin layer chromatography
UV-vis	Ultraviolet-visible
Val	Valine
VBP	Vitellogenin binding protein



## Chapter 1 Introduction

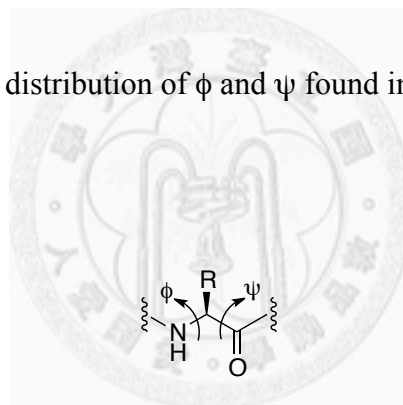


## Proteins

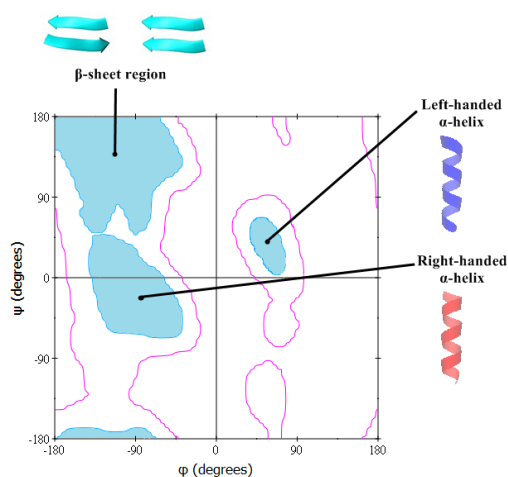
Proteins are one of the most versatile and most abundant macromolecules in living systems.<sup>1,2</sup> Proteins serve as the downstream terminus of the central dogma of molecular biology, which describes the direction of the flow of genetic information.<sup>3</sup> The genetic information stored in deoxyribonucleic acid (DNA) is passed down to ribonucleic acid (RNA) by transcription.<sup>4</sup> RNA then serves as the template in the translation process, leading to the synthesis of proteins.<sup>5</sup> These proteins then participate in a variety of physiological processes, such as catalysis,<sup>6</sup> immune response,<sup>7</sup> signal transduction,<sup>8</sup> and construction and support of cellular structures.<sup>9</sup> Proteins are important to living systems, and studies to enhance fundamental knowledge on proteins should facilitate our understanding of how Nature functions, and also lead to unprecedented technological advancements.

A protein is a biopolymer composed of amino acids.<sup>2</sup> There are twenty naturally occurring amino acids, nineteen of which are L- $\alpha$  amino acids bearing different side chains. These side chains have various functional groups and properties, and play different roles in protein functions. The remaining amino acid is glycine, which has no side chain and is therefore achiral. Amino acids are connected by an amide bond

between the  $\alpha$ -amino group of one amino acid and the  $\alpha$ -carbonyl group of the other, yielding a polypeptide chain. Due to resonance, the amide bond has considerable double bond character, limiting bond rotation and therefore restricting the conformation of the peptide backbone. Dihedral angles are defined along the peptide backbone. The dihedral angle between  $C_{\text{carbonyl}} - N$  plane and  $C_{\alpha} - C_{\text{carbonyl}}$  plane is defined as phi ( $\phi$ ), and the dihedral angle between  $N - C_{\alpha}$  plane and  $C_{\text{carbonyl}} - N$  plane is defined as psi ( $\psi$ ) (Figure 1-1).<sup>10</sup> A plot of  $\psi$  against  $\phi$  is called Ramachandran plot (Figure 1-2).<sup>10</sup> The shaded region in the plot shows the distribution of  $\phi$  and  $\psi$  found in several structures (Figure 1-2).<sup>10</sup>



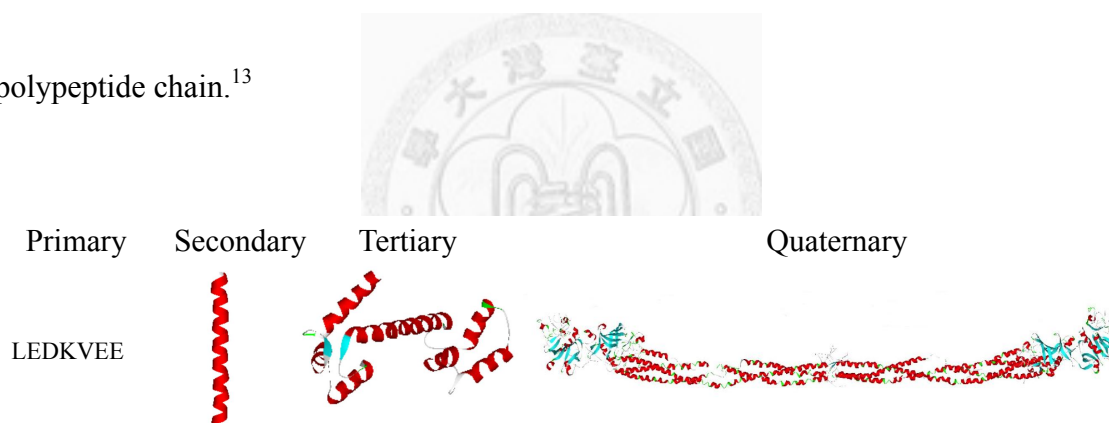
**Figure 1-1.** Dihedral angles  $\phi$  and  $\psi$ .



**Figure 1-2.** Ramachandran plot.<sup>10</sup>

## Protein Structures

The structure of a protein is critical for its biological function. A protein performs its function only when it is correctly folded.<sup>11</sup> Upon losing its structure, for instance, being denatured by heat or chemical denaturant, a protein becomes inactive.<sup>12</sup> The structure of a protein is categorized into four hierarchical levels: primary, secondary, tertiary, and quaternary structures (Figure 1-3). Primary structure refers solely to the amino acid sequence of the protein, without consideration of the three dimensional structures of the polypeptide chain.<sup>13</sup>



**Figure 1-3.** Four levels of protein structure. (Tertiary structure: calmodulin, protein data bank code 1CCL. Quaternary structure: fibrinogen, protein data bank code 3GHG.)

Secondary structures are local structural motifs composed of regularly repeating spatial arrangement of the residues. Secondary structures can be defined by the pattern of hydrogen bonds or backbone dihedral angles of the peptide chain.<sup>14, 15</sup> For example,  $\alpha$ -helix, a common secondary structure, is characterized by the main-chain hydrogen bond between the carbonyl oxygen and the amide hydrogen located four residues away



in the sequence.<sup>16</sup> The dihedral angles ( $\phi$ ,  $\psi$ ) of an  $\alpha$ -helix is defined to be  $(-57^\circ, -47^\circ)$  (Figure 1-2).<sup>10</sup> Another common secondary structure,  $\beta$ -sheet, is made of two or more  $\beta$  strands that are connected by hydrogen bonds between backbone carbonyl oxygen and the amide hydrogen.<sup>17</sup> The dihedral angles ( $\phi$ ,  $\psi$ ) of a  $\beta$ -sheet are  $(-119^\circ, +113^\circ)$  for parallel sheet and  $(-139^\circ, +135^\circ)$  for antiparallel sheet.<sup>10</sup>

Tertiary structure refers to the overall structure of one polypeptide chain, containing a combination of several secondary structures. A typical tertiary structure has nonpolar residues buried in the interior, making the hydrophobic core.<sup>18</sup> Polar and charged residues are usually found on the surface, where the protein is in contact with the aqueous surrounding.<sup>18</sup> Calmodulin (CaM) is an example of tertiary structure, which is a protein that binds calcium ion and evokes downstream biochemical reactions. CaM is a 148-amino-acid polypeptide chain made up of several  $\alpha$ -helices and turns.<sup>19</sup>

Quaternary structure is the spatial arrangement of several polypeptide chains. One example is the fibrinogen, a protein found in blood and is involved in blood coagulation.

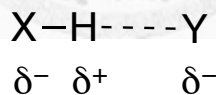
Fibrinogen consists of six chains,  $(\alpha\beta\gamma)_2$ , in which  $\alpha$ ,  $\beta$ , and  $\gamma$  are the different subunits.<sup>20</sup>

## Forces in Protein Folding

The forces that dictate protein folding include electrostatic interactions, hydrogen bond, hydrophobic effect, and van der Waal's interactions.<sup>21</sup> The first two occur in charged and polar groups, whereas the last two are important for nonpolar entities. Electrostatics include charge to charge, charge to dipole, and dipole to dipole interactions. The charges originate from the following amino acids: aspartic acid (Asp), glutamic acid (Glu), lysine (Lys), arginine (Arg), and to a much lesser content, histidine (His). At physiological pH (pH 7.4), side chains of Asp and Glu are deprotonated and are negatively charged. On the other hand, Lys, Arg, and His are protonated and bear a positive charge. Opposite charges attract one another, resulting charge-charge interactions between these amino acids. A dipole is the asymmetric distribution of electron density along a chemical bond caused by the difference in electronegativity of atoms. Dipoles can be attracted by ions or other dipoles on the protein, leading to dipole-charge and dipole-dipole interactions. Electrostatic interactions may provide extra stability to protein structures, as in the case of hyperthermophilic proteins.<sup>22</sup> These proteins are able to withstand high temperatures, and this stability originates from the numerous electrostatic interactions between the charged residues on the protein

surface.<sup>22</sup>

A hydrogen bond is the consequence of the interaction between a hydrogen bond donor (HX) and a hydrogen bond receptor (Y) (Figure 1-4). The atoms X and Y could be nitrogen, oxygen, or fluorine atoms, which are all highly electronegative atoms. The electronegativity of X causes uneven distribution of electron density along the H-X bond, resulting in a partial positive charge on H. This charge is attracted to the nonbonding electrons of Y, therefore forming a hydrogen bond. Hydrogen bond is the dominant force for the formation of secondary structures. The deprivation of these hydrogen bonds (e.g., using solvents like trifluoroacetic acid or formic acid, which protonate the peptide backbone) diminishes peptide structure.<sup>23, 24</sup>



**Figure 1-4.** A hydrogen bond.

Hydrophobic effect is the release of water molecules around a hydrophobic group when two or more such groups approach one another.<sup>25</sup> Water molecules that are in contact with nonpolar solutes are, to some extent, orderly arranged.<sup>25</sup> When the nonpolar groups come into contact with one another, these ordered water molecules are excluded, contributing to the increase in entropy and decrease in free energy.<sup>25</sup> An

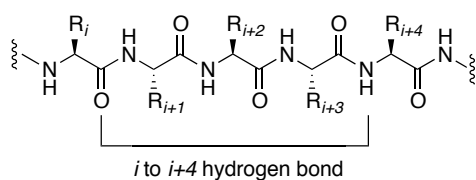
example of the importance of the hydrophobic effect in protein folding is the coiled coil motif. A coiled coil is composed of intertwining  $\alpha$ -helices.<sup>26</sup> The residues in helices interface are mainly hydrophobic.<sup>26, 27</sup> Substituting the hydrophobic residues with polar residues destabilizes the coiled coil.<sup>27</sup>

van der Waal's interaction (or London dispersion force) is the attractive force between apolar molecules.<sup>21</sup> The electron density distribution of the atoms in apolar group may fluctuate, generating an instantaneous dipole. This dipole may induce an opposite dipole in a nearby molecular entity. The attraction between these dipoles is termed London dispersion force. This is a weak interaction and is significant only when the apolar groups are close in distance. The huge number of London dispersion forces in proteins make them an important driving force in protein folding.

### **$\alpha$ -Helix**

$\alpha$ -Helix is a common secondary structure that constitutes about 30% of all protein structures known to date.<sup>15, 28</sup> The naturally occurring  $\alpha$ -helix is a right-handed, rod-like structure characterized by consecutive, main-chain,  $i \leftarrow i + 4$  hydrogen bonds between each carbonyl oxygen ( $i$ ) and an amide hydrogen ( $i+4$ ) on the adjacent helical turn

(Figure 1-5).<sup>16</sup> A helix has backbone dihedral angles ( $\phi$ ,  $\psi$ ) around  $(-57^\circ, -47^\circ)$  and has 3.6 residues per turn.<sup>10, 16</sup> Many factors determine the stability of a helix, one of which is the helix propensity of the constituting amino acids. Chou and Fasman analyzed the frequency of appearance of each amino acid in helices.<sup>15</sup> The statistics revealed that different amino acids have different probabilities to be found in a helix. The thermodynamic tendency to form a helix can be described in a statistical mechanical manner in Lifson-Roig theory, in which helix propensity ( $w$ ) of an amino acid is employed to describe the equilibrium between  $\alpha$ -helix and random coil.<sup>29</sup> The greater the helix propensity, the greater the probability to find an amino acid in the helical state. Helices are often combined with other secondary structures to form motifs, which usually exhibit certain functions.<sup>1</sup> For example, the helix-turn-helix motif is a DNA binding motif,<sup>30</sup> and the helix-loop-helix ("EF hand") is a calcium ion binding motif.<sup>19</sup>

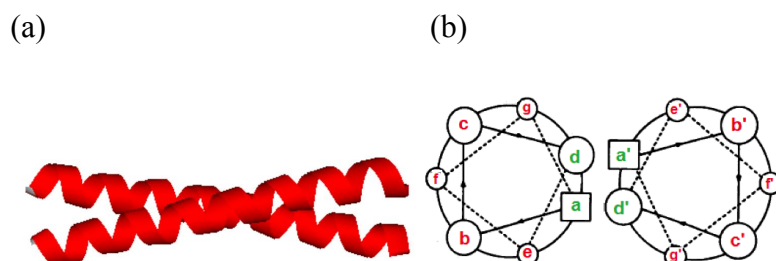


**Figure 1-5.** The  $\alpha$ -helix backbone hydrogen bond.

## Coiled coil

Coiled coils are formed by 3-5% of all amino acids in proteins,<sup>31</sup> and can be found in transcription factors,<sup>32</sup> cytoskeletal<sup>33</sup> and contractile<sup>34</sup> systems, viral envelope proteins,<sup>35</sup> and other systems.<sup>36</sup> A coiled coil is formed by two to five  $\alpha$ -helices wrapping around each other to form a slight left-handed superhelical twist (Figure 1-6).<sup>26</sup> The sequence of a coiled coil is characterized by the “heptad repeat”, *abcdefg*, in which positions *a* and *d* are predominantly hydrophobic residues, such as leucine (Leu) and isoleucine (Ile), leading to a hydrophobic interface between two wrapping helices. In contrast, the *e* and *g* positions are frequently charged residues that form interhelical ion pairs.<sup>26</sup> The three dimensional architecture was first proposed by Crick in 1953 as the “knobs into holes” model, which states that a coiled coil is stabilized by the packing of hydrophobic side chains “knobs” into the “holes” that are the space between hydrophobic side chains of the neighboring helix.<sup>37</sup> According to the type and orientation of the monomers, coiled coils can be further categorized into homodimer/heterodimer and parallel/antiparallel coiled coils.<sup>38</sup> A homodimer is made up of two identical monomers, whereas a heterodimer consists of two different monomers. A parallel coiled coil has the N-termini of the two helices located on the

same end, whereas an antiparallel coiled coil has the two N-termini located at opposite ends.



**Figure 1-6.** Illustration of a coiled coil. (a) A coiled coil. (GCN4 leucine zipper. Protein data bank code 2ZTA.) (b) The helical wheel of a coiled coil.

## Thesis Overview

Coiled coil is a common structure in proteins. The factors contributing to coiled coil stability include hydrophobicity of the coiled coil interface, coiled coil propensity of the constituting amino acids, and the packing (or the van der Waal's interactions) of the residues at the interface.<sup>39</sup> Studies involving natural amino acids have shown that geometric properties of the buried amino acids can influence the stability of coiled coils.<sup>39, 40</sup> To further elucidate the relationship between structure, hydrophobicity, and coiled coil stability, natural and non-natural amino acids with various side chain structures and properties were introduced into different coiled coil systems. Also, the hydrophobicities of the amino acids were measured. Several structural parameters were included in the analysis (Chapter 2).

## Reference

1. Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemistry* 6th ed., W.H. Freeman: New York, **2006**.
2. Nelson, D.; Cox, M. *Lehninger Principles of Biochemistry* 5th ed, W. H. Freeman: New York, **2009**.
3. Crick, F. Central Dogma of Molecular Biology. *Nature* **1970**, *227*, 561-563.
4. Conaway, J. W.; Shilatifard, A.; Dvir, A.; Conaway, R. C. Control of elongation by RNA polymerase II. *Trends Biochem. Sci.* **2000**, *25*, 375-380.
5. Ramakrishnan, V. Ribosome structure and the mechanism of translation. *Cell* **2002**, *108*, 557-572.
6. Radzicka, A.; Wolfenden, R. A proficient enzyme. *Science* **1995**, *267*, 90-93.
7. Aderem, A.; Ulevitch, R. J. Toll-like receptors in the induction of the innate immune response. *Nature* **2000**, *406*, 782-787.
8. Nishizuka, Y. The role of protein kinase C in cell surface signal transduction and tumour promotion. *Nature* **1984**, *308*, 693-698.
9. Hall, A. Rho GTPases and the actin cytoskeleton. *Science* **1998**, *279*, 509-514.
10. Edsall, J. T.; Flory, P. J.; Kendrew, J. C.; Liquori, A. M.; Nemethy, G.; Ramachandran, G. N.; Scheraga, H. A. A proposal of standard conventions and nomenclature for the description of polypeptide conformations. *J. Mol. Biol.* **1966**, *15*, 399-407.
11. Wright, P. E.; Dyson, H. J. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999**, *293*, 321-331.
12. Tanford, C. Protein denaturation. *Adv. Protein Chem.* **1968**, *23*, 121-282.
13. Sanger, F. The arrangement of amino acids in proteins. *Adv. Protein Chem.* **1952**, *7*, 1-67.
14. Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **1963**, *7*, 95-99.
15. Chou, P. Y.; Fasman, G. D. Conformational parameters for amino acids in helical, beta-sheet, and random coil regions calculated from proteins. *Biochemistry* **1974**, *13*, 211-222.
16. Pauling, L.; Corey, R. B. The structure of synthetic polypeptides. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37*, 241-250.
17. Pauling, L.; Corey, R. B. The pleated sheet, a new layer configuration of



- polypeptide chains. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37*, 251-256.
18. Pace, C. N.; Shirley, B. A.; McNutt, M.; Gajiwala, K. Forces contributing to the conformational stability of proteins. *FASEB J.* **1996**, *10*, 75-83.
  19. Chattopadhyaya, R.; Meador, W. E.; Means, A. R.; Quioco, F. A. Calmodulin structure refined at 1.7 Å resolution. *J. Mol. Biol.* **1992**, *228*, 1177-1192.
  20. Kollman, J. M.; Pandi, L.; Sawaya, M. R.; Riley, M.; Doolittle, R. F. Crystal structure of human fibrinogen. *Biochemistry* **2009**, *48*, 3877-3886.
  21. Dill, K. A. Dominant forces in protein folding. *Biochemistry* **1990**, *29*, 7133-7155.
  22. Xiao, L.; Honig, B. Electrostatic contributions to the stability of hyperthermophilic proteins. *J. Mol. Biol.* **1999**, *289*, 1435-1444.
  23. Hanlon, S.; Klotz, I. M. Protonation of polypeptides in "helix-breaking" solvents: spectral and optical-rotatory properties in solutions containing strong organic acids. *Biochemistry* **1965**, *4*, 37-48.
  24. Lotan, N.; Bixon, M.; Berger, A. Alpha-helix formation by solvent-solvent interaction. *Biopolymers* **1967**, *5*, 69-77.
  25. Southhall, N. T.; Dill, K. A.; Haymet, A. D. A View of the Hydrophobic Effect. *J. Phys. Chem. B* **2002**, *106*, 521-533.
  26. Mason, J. M.; Arndt, K. M. Coiled coil domains: Stability, specificity, and biological implications. *ChemBioChem* **2004**, *5*, 170-176.
  27. Gonzalez, L., Jr.; Woolfson, D. N.; Alber, T. Buried polar residues and structural specificity in the GCN4 leucine zipper. *Nat. Struct. Biol.* **1996**, *3*, 1011-1018.
  28. Cheng, R. P.; Girinath, P.; Suzuki, Y.; Kuo, H. T.; Hsu, H. C.; Wang, W. R.; Yang, P. A.; Gullickson, D.; Wu, C. H.; Koyack, M. J.; Chiu, H. P.; Weng, Y. J.; Hart, P.; Kokona, B.; Fairman, R.; Lin, T. E.; Barrett, O. Positional effects on helical Ala-based peptides. *Biochemistry* **2010**, *49*, 9372-9384.
  29. Lifson, S.; Roig, A. On the theory of helix - coil transition in polypeptides. *J. Chem. Phys.* **1961**, *34*, 1963-1974.
  30. Brennan, R. G.; Matthews, B. W. The helix-turn-helix DNA-binding motif. *J. Biol. Chem.* **1989**, *264*, 1903-1906.
  31. Wolf, E.; Kim, P. S.; Berger, B. MultiCoil: A program for predicting two- and three-stranded coiled coils. *Protein Sci.* **1997**, *6*, 1179-1189.
  32. O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **1991**, *254*, 539-544.

33. Pauling, L.; Corey, R. B. Compound helical configurations of polypeptide chains: structure of proteins of the alpha-keratin type. *Nature* **1953**, *171*, 59-61.
34. Sodek, J.; Hodges, R. S.; Smillie, L. B.; Jurasek, L. Amino-acid sequence of rabbit skeletal tropomyosin and its coiled-coil structure. *Proc. Natl. Acad. Sci. U. S. A.* **1972**, *69*, 3800-3804.
35. Wilson, I. A.; Skehel, J. J.; Wiley, D. C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **1981**, *289*, 366-373.
36. Cohen, C.; Parry, D. A. Alpha-helical coiled coils: more facts and better predictions. *Science* **1994**, *263*, 488-489.
37. Crick, F. H. C. The Packing of Alpha-Helices: Simple Coiled-Coils. *Acta Cryst.* **1953**, *6*, 689-697.
38. Grigoryan, G.; Keating, A. E. Structural specificity in coiled-coil interactions. *Curr. Opin. Struct. Biol.* **2008**, *18*, 477-483.
39. Moitra, J.; Szilak, L.; Krylov, D.; Vinson, C. Leucine is the most stabilizing aliphatic amino acid in the d position of a dimeric leucine zipper coiled coil. *Biochemistry* **1997**, *36*, 12567-12573.
40. Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **1993**, *262*, 1401-1407.

## **Chapter 2 Effect of Side Chain Structure of *d*-Position**

### **Residue on Coiled Coil Stability**



## Introduction

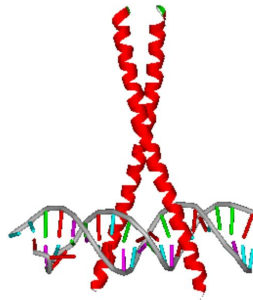
### *Coiled Coils in Biological Systems*

Coiled coils are formed by 3-5% of all amino acids in proteins,<sup>1</sup> and can be found in transcription factors<sup>2</sup>, cytoskeletal<sup>3</sup> and contractile<sup>4</sup> systems, viral envelope proteins,<sup>5</sup> and other systems.<sup>6</sup> In particular, transcription factors are an excellent example to illustrate the importance of coiled coils. Transcription factors are proteins that bind specific DNA sequences and thereby regulate gene expression.<sup>2</sup> Among the numerous transcription factors, bZIP transcription factors constitute an important class of DNA binding proteins.<sup>7,8</sup> bZIP transcription factors achieve specific DNA binding through the dimerization of the coiled coil motifs.<sup>7,8</sup> Such dimerization results in the formation of various homodimers and heterodimers, which enhance specific binding between the proteins and DNA, generate a large number of variations for binding of different DNA sequences, and therefore form a complex regulatory network in a variety of organisms.<sup>7,8</sup> This suggests an important role for coiled coils in living systems, and studies of this structural motif should facilitate the understanding of protein functions and the design of novel proteins.

## *GCN4 Coiled Coil Motif*

GCN4 is a bZIP transcription factor in yeast. GCN4 contains a basic region and a coiled coil region.<sup>9</sup> The former is responsible for DNA recognition, and the latter promotes dimerization (Figure 2-1).<sup>9</sup> The crystal structure of the coiled coil region of GCN4 was first solved in 1991.<sup>2</sup> This achievement has turned coiled coils, which was mainly considered to be structural for decades because coiled coils were first recognized in fibrous proteins, into a popular research topic that has blossomed in the past twenty years.<sup>6, 10, 11</sup> There were two significances. First, it showed that coiled coils participate in crucial interactions such as transcription. Second, the knowledge obtained from the physical details of the coiled coil structure provided a means to recognize tertiary structures by the inspection of primary structures, leading to *de novo* coiled coil design.<sup>6, 10, 11</sup> The coiled coil region of GCN4 is a 33-residue, 4-heptad repeat *leucine zipper* (see Chapter 1 for the general introduction of coiled coils). Leucine zipper bears its name due to the conserved leucine at the *d* positions. The *a* positions are taken up by valine, so that both *a* and *d* positions are occupied by hydrophobic residues. One exception occurs at the *a* position in the third heptad, where there is an asparagine.<sup>2</sup> This buried polar residue determines the oligomeric state of the coiled coil, making GCN4 a

homodimeric parallel coiled coil.<sup>12</sup>



**Figure 2-1.** GCN4-DNA complex (protein data bank code 1YSA).

### *Coiled Coil Sequence and Structure*

The sequence of a coiled coil is characterized by the “heptad repeat”, usually designated *abcdefg*. A regular  $\alpha$ -helical turn is composed of 3.6 residues.<sup>10</sup> However, this value is lowered to 3.5 for helices in a coiled coil, leading to the heptad repeat every two turns of a helix.<sup>10</sup> Inspection of coiled coil sequences led to the concept of "peptide Velcro", which points out three elements for the formation of a specific coiled coil.<sup>10, 13</sup>

First, the residues at *a* and *d* positions, which are located at the interface between a coiled coil, need to be hydrophobic residues like leucine, valine, or isoleucine, so that the coiled coil is stabilized through hydrophobics and van der Waals interactions.

Second, the *e* and *g* positions, which are located adjacent to the coiled coil interface, have to be occupied by charge residues such as glutamate or lysine, which form interhelical electrostatic interactions that also contribute to coiled coil stability. Last, the

remaining *b*, *c*, and *f* positions have to be hydrophilic residues, since these positions are located at the solvent-accessing face of a coiled coil.

These features are insufficient to conclude the variety of coiled coils observed in Nature and laboratory. More precisely, the features are insufficient to predict the oligomerization (dimer, trimer, or tetramer, etc.), pairing specificity (homodimer or heterodimer), and orientation (parallel or antiparallel). 20% of the residues at the *a* and *d* positions are polar or charged.<sup>12</sup> Buried polar residues affect the oligomeric state of a coiled coil.<sup>10, 12, 14-16</sup> For instance, the substitution of the two buried asparagines of GCN4 coiled coil to valine resulted in change in oligomeric state from dimers to a mixture of dimers and trimers.<sup>17, 18</sup> These asparagines confer the dimer state through interhelical hydrogen bonds between the Asn side chains.<sup>2, 10, 17</sup> Besides, the structure of the *a* and *d* residues also have affect on oligomerization.<sup>18</sup> GCN4 coiled coil mutants, with the 4 *a* residues replaced with isoleucine, valine, or leucine yielded dimers, mixture of dimers and trimers, and trimers, respectively.<sup>18</sup> This switch between oligomeric states is a consequence of the packing interactions at the coiled coil interface.<sup>18</sup>

Orientation and pairing specificity are greatly influenced by the electrostatic

interactions between the *e* and *g* residues.<sup>10, 19</sup> For instance, placing only negatively charged residue at the *e* and *g* positions in one strand of a coiled coil and only positively charged residues at the *e'* and *g'* positions in another strand, favors the formation of a heterodimer due to attractive interhelical Coulombic interactions.<sup>13</sup> Also, deliberate positioning of oppositely charged residues at *g* and *g'* led to the preference for the antiparallel orientation as a result of the repulsive Coulombic interactions that would be encountered if the coiled coil was parallel.<sup>20</sup>

There are still many undiscovered underlying rules for the structures of coiled coils. Amazing examples include a GCN4 variant with all *e* positions mutated to valine gave a parallel tetramer,<sup>21</sup> whereas another GCN4 variant with all *g* residues substituted with either valine or alanine resulted in an antiparallel tetramer.<sup>21</sup> Still, when all *e* and *g* residues of GCN4 were replaced with alanine, the coiled coil turned out to be a parallel heptamer.<sup>21</sup> Predicting the oligomeric state, pairing specificity, and orientation remains an intriguing challenge.



### *Knobs-Into-Holes Interactions of a Dimeric Coiled Coil*

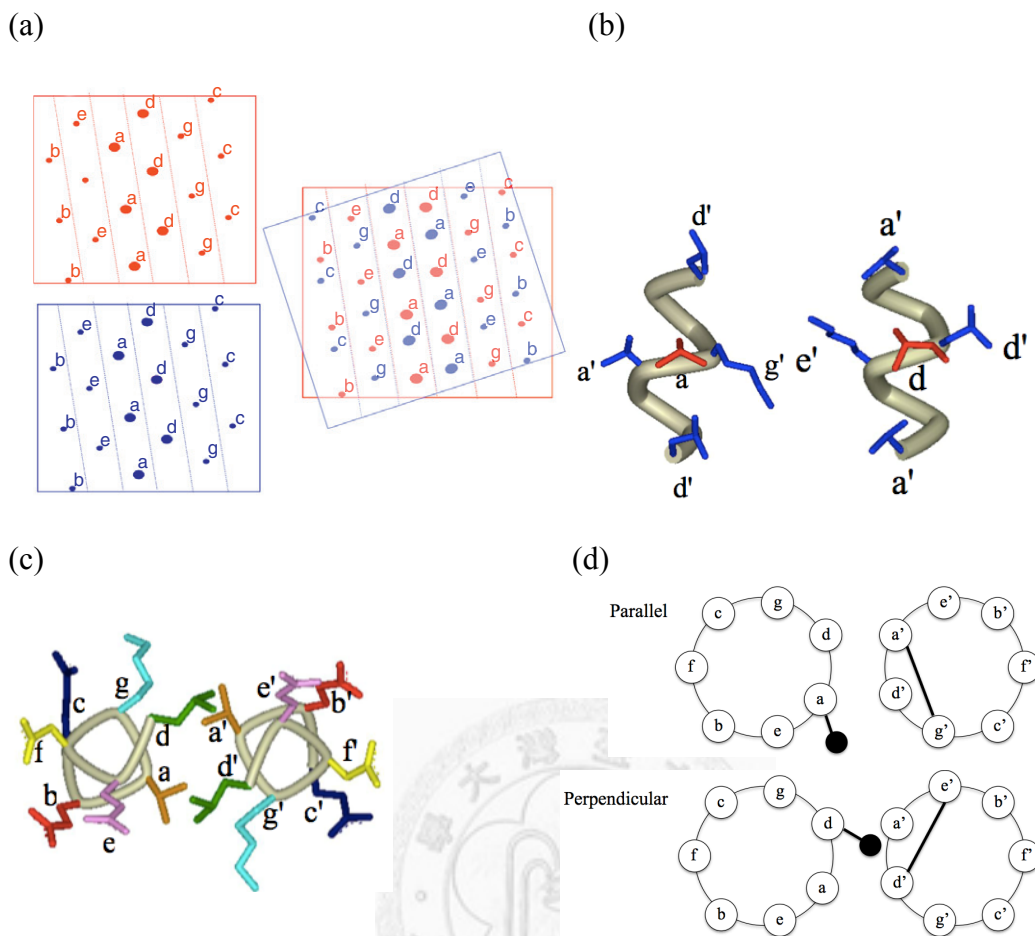
Knobs-into-holes (KIH) interactions of a dimeric coiled coil was first described by Crick in 1953.<sup>22</sup> KIH states that a side chain from one helix, referred to as the *knob*, fits into the *hole* surrounded by the side chains of the partner helix.<sup>22</sup> For instance, in a parallel, dimeric coiled coil, the *a* knob fits into a hole made by *d'*, *a'*, *d'*, and *g'* of the neighboring helix, whereas the *d* knob fits into a hole made by *a'*, *e'*, *a'*, and *d'* residues (Figure 2-2a and b).<sup>2, 11, 18</sup> Furthermore, the *a* knob projects *out of the interface*; the C $\alpha$ -C $\beta$  vector of the knob is parallel to the C $\alpha$ -C $\alpha$  vector at the bottom of the recipient hole.<sup>2, 11, 18</sup> The *d* knob projects directly into the interface; the C $\alpha$ -C $\beta$  vector of the knob is perpendicular to the C $\alpha$ -C $\alpha$  vector at the bottom of the hole (Figure 2-2c and d).<sup>2, 11,</sup>

18

### *Coiled Coil Stability*

Many factors contribute to the stability of a coiled coil, such as the coiled coil propensities of the constituting amino acids,<sup>23</sup> hydrophobic interactions,<sup>10, 12, 18, 24, 25</sup> packing of the core residues,<sup>18, 26-30</sup> and electrostatic interactions.<sup>10, 31, 32</sup> Herein, coiled coil propensities, hydrophobic interactions, and side chain packing will be discussed.

Amino acids have different conformational preferences that influence the stability of



**Figure 2-2.** Illustration of knobs-into-hole interactions, parallel packing and perpendicular packing. (a) Cartoons showing knobs-into-holes interactions. On the left hand side are the projections of each helix strand in a coiled coil onto flat surfaces. On the right is the overlap of the two sheets which reflects the interdigitation of side chains in coiled coil. The angle between the two sheets is the angle between the two strands of helices, which is  $\sim 20^\circ$ . (Reprinted from Woolfson *et al.* New currency for old rope: from coiled-coil assemblies to  $\alpha$ -helical barrels. *Curr. Opin. Struct. Biol.* 2012, doi:10.1016/j.sbi.2012.03.002, with permission from Elsevier.) (b) KIH interactions in GCN4 coiled coil. On the left shows the *a* knob and the corresponding *d'a'd'g'* hole, the other is the *d* knob and the *a'e'a'd'* hole. (c) One heptad of GCN4 coiled coil. The *a* knob projects out of the interface, whereas the *d* knob projects into the interface. (d) Cartoons showing the parallel and perpendicular packing of the *a* and *d* knobs. Open circles represent C $\alpha$  atoms, and solid circles represent C $\beta$  atoms. (GCN4 PDB ID: 2ZTA.)

protein secondary and tertiary structure.<sup>23, 33</sup> "Helix propensity" describes the probability of an amino acid to be in a helical state.<sup>34</sup> Many studies use monomeric model helices to measure helix propensities.<sup>23</sup> O'Neil *et al.* employed a parallel, homodimeric coiled coil instead.<sup>23</sup> They claimed that a monodimeric helix may not be representative of helices found in a protein because a helix in a protein experiences non-uniform solvent accessibility and dielectric constant.<sup>23</sup> The guest site, which was substituted with the 20 natural amino acids, was situated at the *f* position of the middle heptad. The *f* position is distant from the dimerization surface, therefore interhelical packing interactions are avoided. Residues near the guest site were designated to be small and neutral in order to exclude possible interactions between the guest residue and nearby residues. As such, differences in coiled coil stabilities solely depended on the different preferences for the coiled coil conformation of each guest residue. In other words, this system measures the coiled coil propensity. The result showed an overall good correlation with data obtained from various monomeric helix models.<sup>23</sup>

Hydrophobic interactions arise from the exclusion of surrounding orderly-arranged water molecules when hydrophobic moieties approach one another.<sup>35</sup> The release of

water molecules leads to the increase in entropy and decrease in free energy.<sup>35</sup>

Hydrophobicity is quantified by measuring the partition coefficient of a solute between water and a non-polar solvent, which is further transformed into transfer free energy.<sup>35</sup>

Hydrophobic interaction is one of the main determinants for coiled coil dimerization.<sup>10</sup>

A great deal of work has shown that decrease in the hydrophobicity of the coiled coil

interface decreases coiled coil stability. For instance, the substitution of the buried

asparagine in GCN4 to valine increases the melting temperature (the temperature at

which 50% of the coiled coils are unfolded) by 40°C.<sup>18</sup> The gain in stability is at the

expense of the specificity of oligomeric state (*vide supra*).<sup>10, 18</sup> In another study, the

*α*-position valines and *d*-position leucines in GCN4 were replaced with asparagine,

threonine, serine, and glutamine, and all of the mutants exhibited lower melting

temperatures.<sup>12</sup> Instead of substituting with polar entities, residues with higher

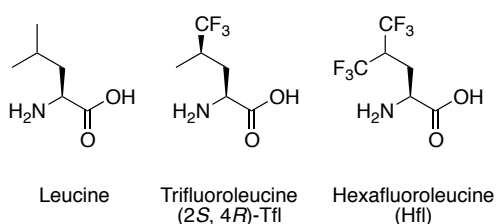
hydrophobicity were also employed, and coiled coil stability was enhanced. When the

*d*-position leucines in coiled coil A1 were replaced with trifluoroleucine or

hexafluoroleucine (Figure 2-3), the melting temperatures increased by 10 °C and 22 °C

respectively. The increase in thermal stability was a consequence of the superior

hydrophobicity of trifluoromethyl over methyl groups.<sup>24, 25</sup>



**Figure 2-3.** Structures of leucine, trifluoroleucine and hexafluoroleucine.

The hydrophobic effect is generally agreed to be one of the major determinant factors in protein folding and stability,<sup>26, 35</sup> but packing of the residue side chains also play a significant role in protein stability.<sup>26, 30</sup> In a typical experiment, a hydrophobic residue in the core of a protein was substituted with a smaller hydrophobic residue (for example, leucine to alanine), and the difference between the folding free energies of both the mutant and wild-type protein was measured.<sup>26</sup> This energy difference reflected the different stabilities of the mutant and wild-type protein.<sup>26</sup> Studies revealed that the energy difference was *larger* than what would be expected for the transfer free energy, which is a measure of hydrophobicity.<sup>26</sup> The change in protein stability, in this case, was not caused by hydrophobicity alone, and the discrepancy can be explained by the difference in packing efficiency.<sup>26, 27</sup> The formation of cavities within a protein is destabilizing.<sup>26, 30</sup> In coiled coils, the geometric properties of buried amino acids influenced the overall structure and stability.<sup>18, 27</sup> For example, substituting the 5<sup>th</sup> *d*-position leucine of the bZIP domain of VBP (vitellogenin binding protein) with

isoleucine and valine lowers the thermal stability, even though these amino acids are of similar sizes.<sup>27</sup> Leucine is the most favored amino acid at the  $d$  position, and is nearly 3 kcal/mole more stabilizing than the similarly sized amino acid isoleucine.<sup>27</sup>  $\beta$ -Branched amino acids are conformationally constrained in  $\alpha$ -helices.<sup>27</sup> Molecular modeling indicates that at  $d$  positions, the favored rotamers of isoleucine and valine, which are both  $\beta$ -branched, produce a steric clash between the two amino acids on opposite helices.<sup>27</sup> Recent studies from Gellman's group shows that  $a'$ - $a'$  vertical interactions are more prominent than  $d'$ - $d'$  in antiparallel dimeric coiled coil.<sup>28, 29</sup> Coiled coil database mining reveals that the  $C\alpha$ - $C\beta$  vectors of  $a$  and  $a'$  residues points together and their trajectories appear to cross, leading to steric repulsion between side chains.<sup>29</sup> On the other hand,  $C\alpha$ - $C\beta$  vectors of  $d$  and  $d'$  residues are nearly parallel but offset, and the steric repulsion is avoided.<sup>29</sup>

### *Chapter Overview*

Coiled coil stability has been intensively investigated over the past two decades, and the hydrophobicity at the coiled coil interface is the most recognized factor that determines the stability of this structural motif. Although packing interactions are

known to affect the stability of proteins and coiled coils, little has been carried out to probe the relationship between packing and stability. Furthermore, most of these studies only employed natural amino acids. To further elucidate the relationship between hydrophobicity, side chain packing, and stability, both natural and non-natural amino acids with various side chain structures and hydrophobicities were introduced into two coiled coil systems.

## Results and Discussions

### *Design of GCN4-Derived Peptides*

The GCN4 coiled coil region is a well-studied model system for coiled coils, therefore we chose to study the effect of altering the residues at the *d* position in this particular system. There are 4 *a* positions and 4 *d* positions in a GCN4 coiled coil (Figure 2-4a). Initially, all 4 *d* residues were mutated simultaneously (previously performed by Hsien-Po Chiu in our lab), but some of these mutants exhibited no coiled coil conformation by circular dichroism (CD) spectroscopy. This was probably due to the profound changes in the interface constituents. Therefore, mutation sites were narrowed down to the central 2 *d* positions. Still, the second-generation mutant did not

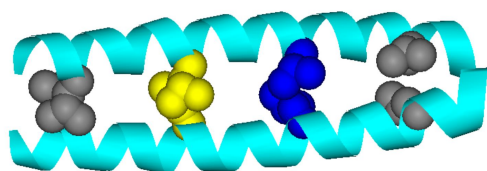
exhibit a coiled coil conformation (Figure 2-5). To enable significant formation of the coiled coil conformation, only one of the two central *d* sites (positions 12 and 19) could be mutated. The GCN4 coiled coil crystal structure showed that both sites were positioned at the same distance from the center of the coiled coil (Figure 2-4b). In addition, the side chain of a residue in an  $\alpha$ -helix points towards the N-terminus. For position 12 (Leu12), the side chain points towards position 5, which is a valine at an  $\alpha$  position. Position 19 (Leu19) points towards position 16, which is an asparagine at an  $\alpha$  position. Therefore, position 12 was chosen as the mutation site to exclude the unfavorable interactions between aliphatic Leu19 and polar Asn16 (Figure 2-4c). Sixteen different amino acids were incorporated individually at the mutation site (Figure 2-6). These amino acids can be divided into linear (Abu, aminobutyric acid; Nva, norvaline; and Nle, norleucine),  $\beta$ -branched (Ile, isoleucine; Allo Ile, alloisoleucine; and Val, valine),  $\gamma$ -branched (Leu, leucine; Tba, *tert*-butylalanine; Cpa, cyclopentylalanine; Cha, cyclohexylalanine; Phe, phenylalanine; and Pff, pentafluorophenylalanine), and polar (Asp, aspartic acid; Asn, asparagine; Glu, glutamic acid; Gln, glutamine).



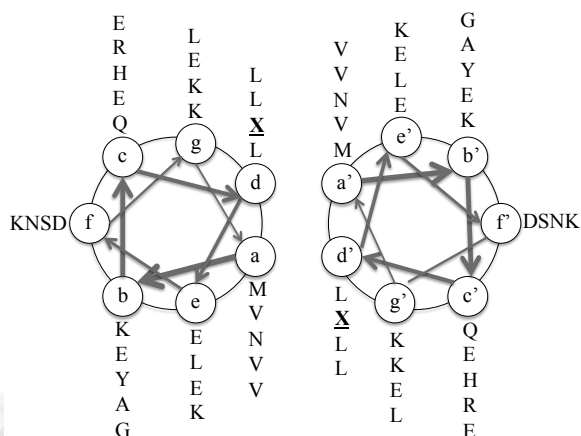
(a)

Ac-RMKQ      5                      12                      19                      26  
LEDKVEE LLSKNYH LENEVAR LKKLVGE R - OH  
*d*    *a*    *d*    *a*    *d*    *a*    *d*    *a*

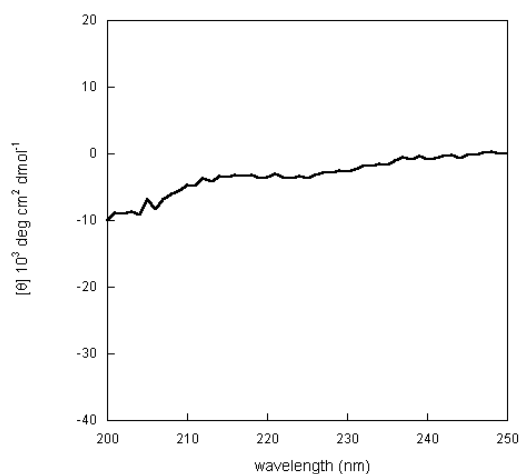
(b)



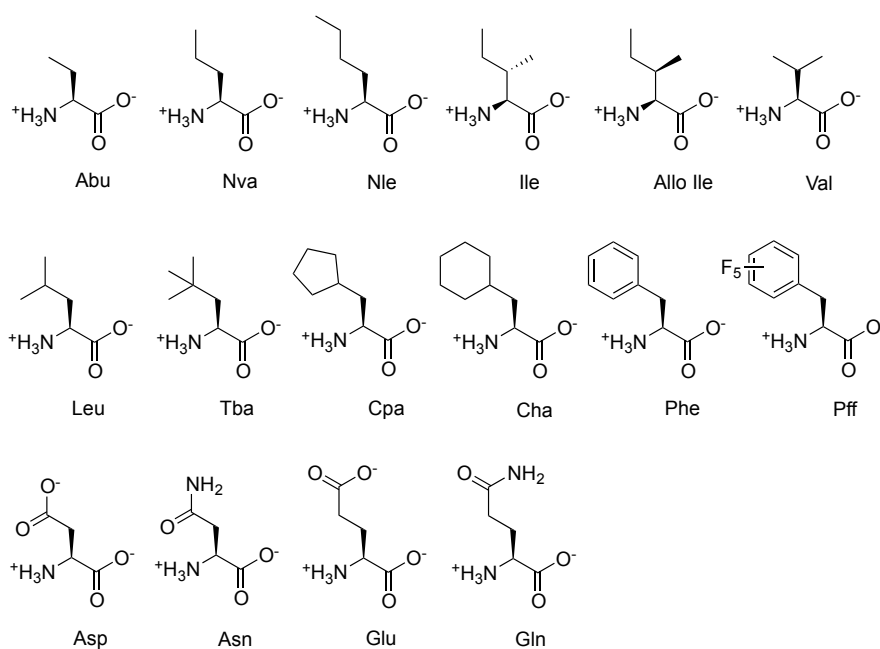
(c)



**Figure 2-4.** The sequence and the helical wheel diagram of GCN4. (a) Sequence of GCN4. The *a* and *d* positions are labeled. The *d* positions are numbered. (b) Ribbon representation of GCN4 with *d*-position leucines in space filling. Yellow: Leu12, blue: Leu19. (c) Helical wheel diagram of GCN4. X denotes the mutation site.



**Figure 2-5.** CD spectrum of GCN4-Phe<sub>2</sub>. The central 2 *d* positions of GCN4-Phe<sub>2</sub> were substituted with Phe.



**Figure 2-6.** Chemical structures of the amino acids.

### *Peptide Synthesis of GCN4-Derived Peptides*

Peptides were synthesized using commercially available reagents by Fmoc-based solid peptide synthesis (SPPS). However, Allo Ile, Nle, and Cpa were purchased in the unprotected form, and the Fmoc group had to be incorporated for SPPS (Scheme 1).

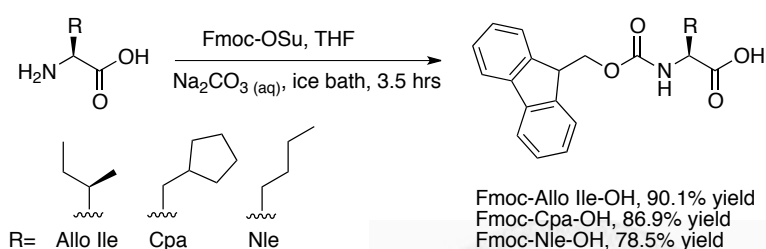
The amino acid was dissolved in  $Na_2CO_3$  solution. This aqueous solution was kept at pH 10 over ice-water bath. A solution of Fmoc-OSu in tetrahydrofuran (THF) was added dropwise into the amino acid solution, and the mixture was stirred for 3.5 hours.

For work up, the THF in the mixture was first removed under reduced pressure. The remaining basic aqueous solution was extracted with diethyl ether. The aqueous fraction

was then acidified with 1 M HCl until pH 2, followed by extraction with ethyl acetate.

The organic fraction was dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated in vacuo. The crude was purified by flash chromatography.

### Scheme 1. Incorporation of Fmoc Onto Amino Acid



The identity of the peptides was confirmed by matrix-assisted laser desorption ionization – time of flight (MALDI-TOF) mass spectrometry. Upon confirmation of the peptides, purification was carried out on reverse phase high performance liquid chromatography (RP-HPLC) to greater than 95% purity. The crude yield of peptide synthesis, molecular formula, calculated [MH<sup>+</sup>], and observed *m/z* of GCN4 peptides are shown in Table 2-1.

**Table 2-1.** Crude Yield of Peptide Synthesis, Molecular Formula, Calculated [MH<sup>+</sup>], and Observed *m/z* of GCN4 Peptides

Peptide	Crude Yield (%)	Molecular Formula	Calculated [MH <sup>+</sup> ]	Observed <i>m/z</i> in MALDI-TOF MS
GCN4-Abu	44.8	C <sub>174</sub> H <sub>294</sub> N <sub>52</sub> O <sub>54</sub> S	4009.17	4008.9
GCN4-Allo Ile	57.7	C <sub>176</sub> H <sub>298</sub> N <sub>52</sub> O <sub>54</sub> S	4037.20	4036.5
GCN4-Asn	58.2	C <sub>174</sub> H <sub>293</sub> N <sub>53</sub> O <sub>55</sub> S	4038.16	4037.6
GCN4-Asp	51.3	C <sub>174</sub> H <sub>292</sub> N <sub>52</sub> O <sub>56</sub> S	4039.14	4038.5
GCN4-Cha	49.2	C <sub>179</sub> H <sub>302</sub> N <sub>52</sub> O <sub>54</sub> S	4077.23	4077.1
GCN4-Cpa	57.1	C <sub>178</sub> H <sub>300</sub> N <sub>52</sub> O <sub>54</sub> S	4063.21	4062.8
GCN4-Gln	58.8	C <sub>175</sub> H <sub>295</sub> N <sub>53</sub> O <sub>55</sub> S	4052.17	4051.4
GCN4-Glu	61.3	C <sub>175</sub> H <sub>294</sub> N <sub>52</sub> O <sub>56</sub> S	4053.16	4052.3
GCN4-Ile	47.0	C <sub>176</sub> H <sub>298</sub> N <sub>52</sub> O <sub>54</sub> S	4037.20	4036.9
GCN4-Leu	54.8	C <sub>176</sub> H <sub>298</sub> N <sub>52</sub> O <sub>54</sub> S	4040.64 <sup>a</sup>	4039.3 <sup>a</sup>
GCN4-Nle	58.4	C <sub>176</sub> H <sub>298</sub> N <sub>52</sub> O <sub>54</sub> S	4037.20	4036.5
GCN4-Nva	50.7	C <sub>175</sub> H <sub>296</sub> N <sub>52</sub> O <sub>54</sub> S	4023.18	4022.6
GCN4-Pff	57.7	C <sub>179</sub> H <sub>291</sub> F <sub>5</sub> N <sub>52</sub> O <sub>54</sub> S	4161.14	4160.8
GCN4-Phe	63.9	C <sub>179</sub> H <sub>296</sub> N <sub>52</sub> O <sub>54</sub> S	4073.66 <sup>a</sup>	4073.1 <sup>a</sup>
GCN4-Tba	57.3	C <sub>177</sub> H <sub>300</sub> N <sub>52</sub> O <sub>54</sub> S	4051.21	4051.2
GCN4-Val	43.0	C <sub>175</sub> H <sub>296</sub> N <sub>52</sub> O <sub>54</sub> S	4023.18	4023.1

<sup>a</sup>Observed *m/z* and calculated [MH<sup>+</sup>] of GCN4-Leu and GCN4-Phe were given as average molecular weight. For the rest of the peptides, exact mass was given.

#### *UV-Visible Spectroscopy (UV-vis) of GCN4-Derived Peptides*

The concentration of the peptide stock solutions was determined by the tyrosine absorbance in 6 M guanidinium chloride as described by Edelhoch.<sup>36, 37</sup> UV data were obtained using 1 mm pathlength cells. Absorbance at 276 nm, 278 nm, 280 nm and 282 nm were measured at different concentrations of peptide. Linear regression was

performed to obtain the concentration. Data with correlation coefficients (R) less than 0.99 were discarded. Concentrations and regression coefficients for the GCN4 peptides are shown in Table 2-2.

**Table 2-2.** Concentrations and Regression Coefficients of GCN4 Peptides

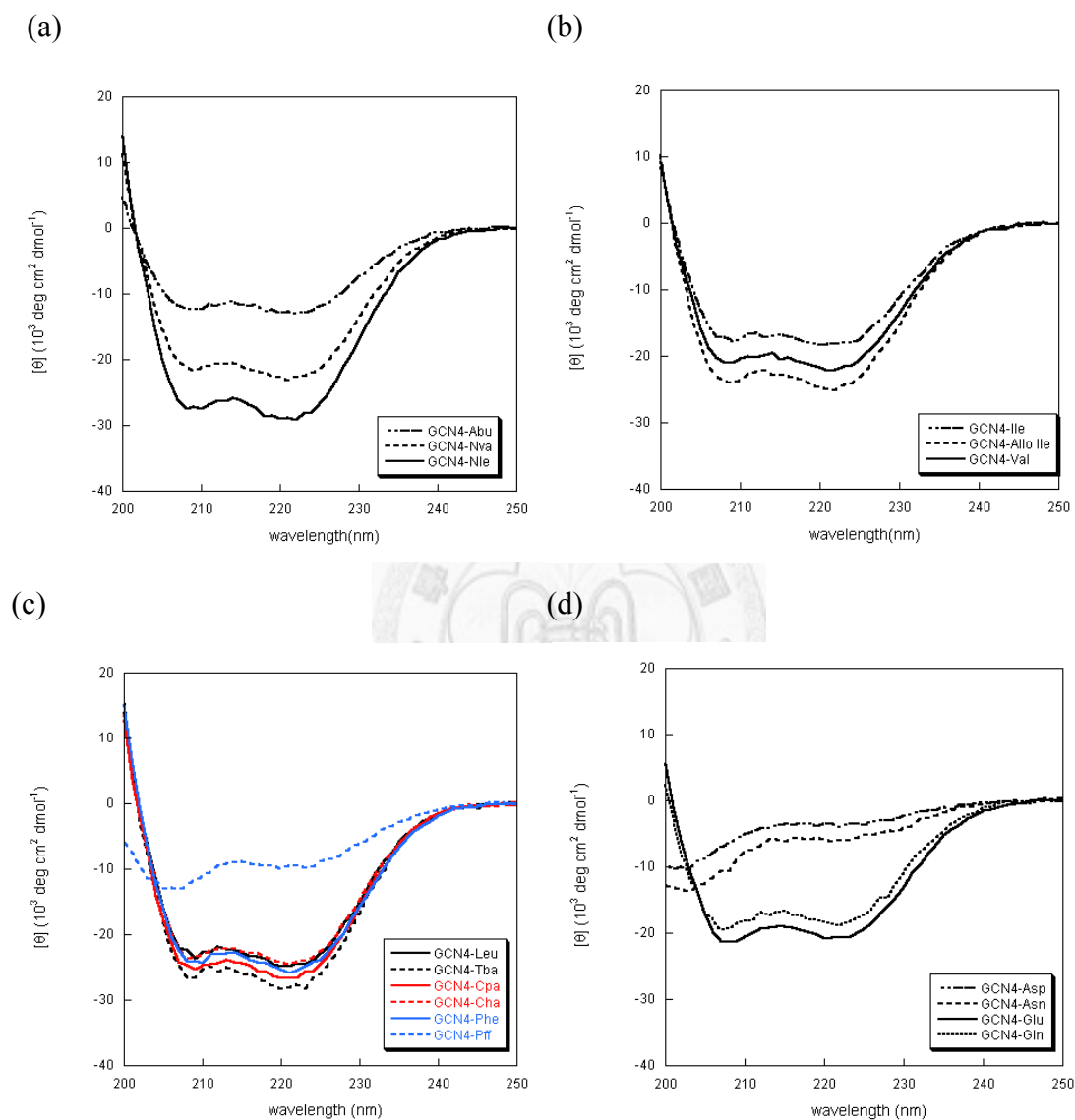
Peptide	Concentration (mM)	Regression Coefficient
GCN4-Abu	5.11±0.05	0.99914
GCN4-Allo Ile	4.72±0.03	0.99950
GCN4-Asn	5.01±0.03	0.99950
GCN4-Asp	5.53±0.06	0.99856
GCN4-Cha	4.9±0.1	0.99488
GCN4-Cpa	5.05±0.04	0.99921
GCN4-Gln	4.46±0.03	0.99940
GCN4-Glu	4.32±0.04	0.99911
GCN4-Ile	5.23±0.05	0.99914
GCN4-Leu	4.90±0.04	0.99908
GCN4-Nle	4.50±0.03	0.99954
GCN4-Nva	5.40±0.04	0.99926
GCN4-Pff	5.14±0.07	0.99781
GCN4-Phe	6.65±0.04	0.99959
GCN4-Tba	4.54±0.04	0.99920
GCN4-Val	3.73±0.03	0.99906

#### *Circular Dichroism (CD) Spectroscopy of GCN4-Derived Peptides*

CD data was acquired at 30  $\mu$ M peptide concentration in 50 mM phosphate and 150 mM NaCl buffer at pH 7 and 4  $^{\circ}$ C.<sup>18</sup> The magnitude of the CD signal at 222 nm reflects

the helical content of a coiled coil.<sup>38</sup> The CD spectra of peptides are shown in Figure

2-7, and the MRE values at 222 nm ( $[\theta]_{222}$ ) are listed in Table 2-3.



**Figure 2-7.** CD spectra of GCN4-Xaa peptides at 30  $\mu$ M peptide concentration in 50 mM phosphate and 150 mM NaCl buffer at pH 7 and 4  $^{\circ}$ C. (a) CD spectra of GCN4-Xaa peptides in which Xaa are residues with linear chains. (b) CD spectra of GCN4-Xaa peptides in which Xaa are residues with  $\beta$ -branched side chains. (c) CD spectra of GCN4-Xaa peptides in which Xaa are residues with  $\gamma$ -branched side chains. (d) CD spectra of GCN4-Xaa peptides in which Xaa are residues with polar side chains.

**Table 2-3.**  $[\theta]_{222}$  at 0 M Guanidinium, Melting Concentrations ( $[C]_m$ ), m values, and  $\Delta G_{\text{unfold, H}_2\text{O}}$  of GCN4 Peptides

Peptides	$[\theta]_{222}$ (deg cm <sup>2</sup> dmol <sup>-1</sup> )	$[C]_m$ (M) <sup>a</sup>	m Value <sup>b</sup>	$\Delta G_{\text{unfold, H}_2\text{O}}$ (kcal/mole) <sup>c</sup>
GCN4-Abu	-12800±300	0.45	-1.26±0.04	3.46±0.02
GCN4-Nva	-22900±400	1.35	-1.023±0.008	4.25±0.01
GCN4-Nle	-28900±500	1.15	-1.167±0.009	4.21±0.01
GCN4-Ile	-18500±400	0.85	-1.14±0.01	3.85±0.01
GCN4-Allo Ile	-24700±400	0.56	-1.241±0.005	3.557±0.003
GCN4-Val	-21900±400	0.66	-1.21±0.03	3.69±0.02
GCN4-Leu	-24800±400	2.44	-1.04±0.02	5.42±0.04
GCN4-Tba	-28100±600	2.67	-0.923±0.008	5.33±0.02
GCN4-Cpa	-26400±500	1.60	-1.15±0.02	4.71±0.03
GCN4-Cha	-24300±600	1.25	-1.033±0.008	4.15±0.01
GCN4-Phe	-25400±400	0.45	-1.34±0.02	3.478±0.009
GCN4-Pff	-9700±400	0.05	-1.4±0.1	2.92±0.03
GCN4-Asp	-3600±300	-- <sup>d</sup>	-- <sup>d</sup>	-- <sup>d</sup>
GCN4-Asn	-5900±400	-- <sup>d</sup>	-- <sup>d</sup>	-- <sup>d</sup>
GCN4-Glu	-20900±400	0.05	-1.13±0.01	2.922±0.004
GCN4-Gln	-18500±500	0.15	-1.25±0.02	3.053±0.009

<sup>a</sup> $[C]_m$  is the concentration of guanidinium chloride at which 50% of the total peptide is unfolded. <sup>b</sup>m value is the slope of the regression line for fitting of  $\Delta G_{\text{unfold, H}_2\text{O}}$ . <sup>c</sup> $\Delta G_{\text{unfold, H}_2\text{O}}$  is the free energy required for the coiled coil to unfold at 0 M guanidinium chloride. <sup>d</sup>These peptides barely fold even at 0 M of guanidinium chloride. Therefore,  $[C]_m$ , m value, and  $\Delta G_{\text{unfold, H}_2\text{O}}$  cannot be deduced.

The helical content of GCN4-Xaa with linear Xaa residues (Abu, Nva, and Nle) increased with increasing side chain length. This indicates that a longer linear side chain at the coiled coil interface reinforces the coiled coil helical structure. When side chains of Xaa are  $\beta$ -branched, the helical content of the peptides followed the trend: GCN4-Ile < GCN4-Val < GCN4-Allo Ile. There does not appear to be any structural reason for

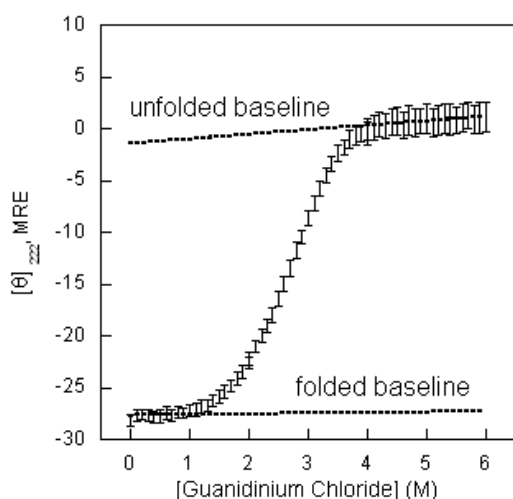
this trend. Regardless, Ile and Allo Ile are diastereomers with opposite chirality at the C $\beta$ . Surprisingly, the unnatural side chain structure yielded a higher helical content in the coiled coil. For  $\gamma$ -branched amino acids, the helical content followed the trend: GCN4-Pff < GCN4-Cha < GCN4-Leu < GCN4-Phe < GCN4-Cpa < GCN4-Tba. Still, no simple obvious structural explanation can be deduced from this trend. Regardless, the helical content of GCN4-Pff was significantly lower than the rest of the group, indicating that the bulky and very hydrophobic pentafluorophenyl group imposes adverse effects onto the coiled coil structure. For the polar amino acids, the helical content followed the trend: GCN4-Asp < GCN4-Asn < GCN4-Gln < GCN4-Glu. Apparently, the longer the side chain, the higher the helical content. However, the contribution of the carboxylate (-COO<sup>-</sup>) and amide (-CONH<sub>2</sub>) groups to the coiled coil structure remains unclear.

The  $\theta_{222}$  signal is a measure of the helical content of a peptide. It does not provide information on the *stability* of a coiled coil. In other words, a peptide with a higher helical content does not guarantee that the peptide is also more stable. Therefore, guanidinium titration was performed to assess the stability of the coiled coils.



### *Guanidinium Denaturation of GCN4-Derived Peptides*

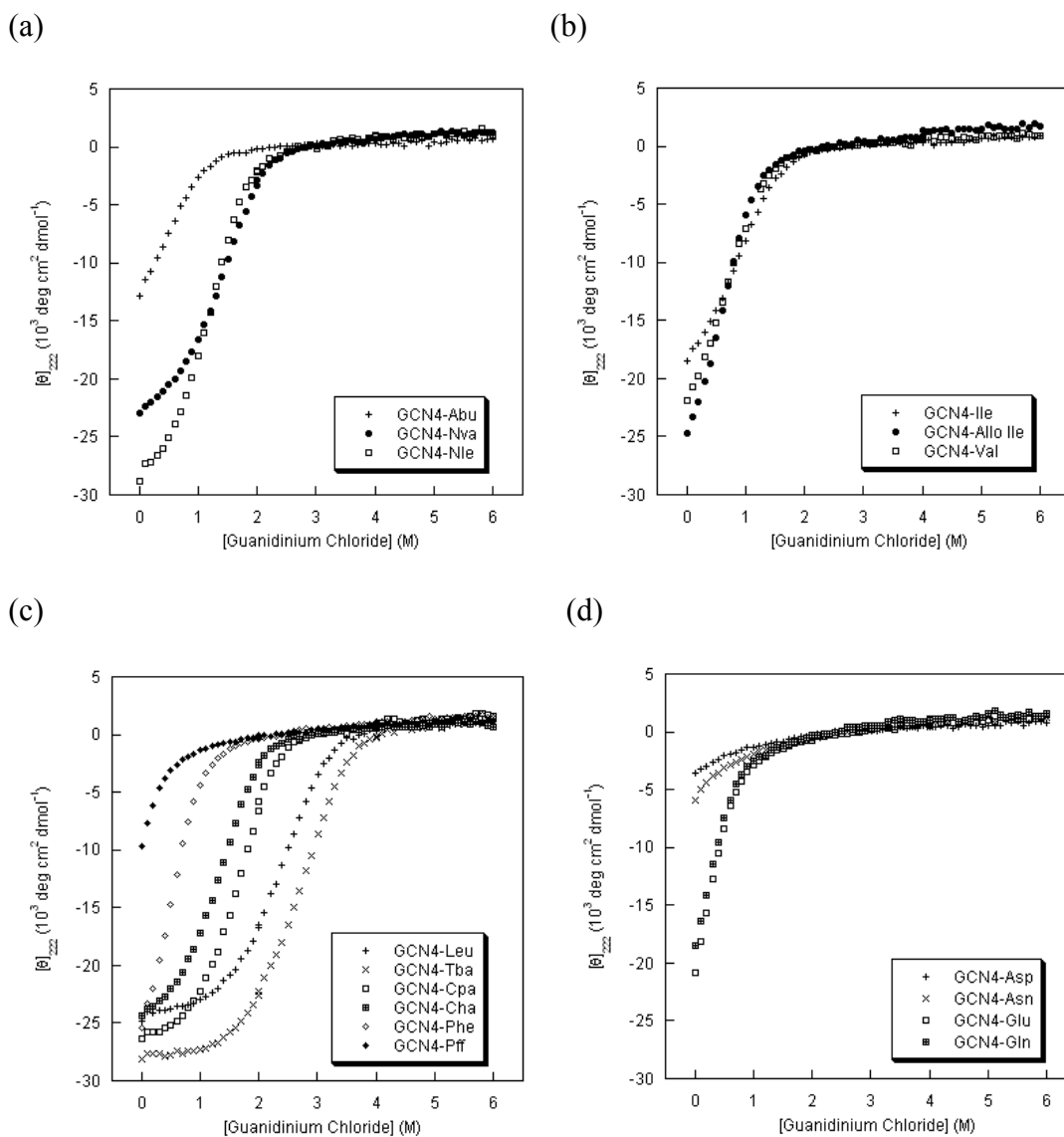
Guanidinium denaturation experiments were performed at 30  $\mu$ M peptide in 50 mM phosphate, 150 mM NaCl, and 0 M to 6 M (with 0.1 M intervals) guanidinium chloride at pH 7 and 4  $^{\circ}$ C.<sup>18</sup> CD was used to monitor the denaturation process and the signal at 222 nm was acquired at various guanidinium chloride concentrations. The coiled coil structure gradually unfolded upon addition of guanidinium chloride, leading to the decrease in CD signal. Suitable CD signals were chosen to derive the folded and unfolded baselines.<sup>39</sup> These baselines describe the expected CD signal at different guanidinium concentrations for a fully folded dimer and a fully unfolded monomer (Figure 2-8). The fraction unfolded of a given peptide at each guanidinium concentration can be derived using the unfolded and folded baselines.<sup>39</sup> The melting concentrations ( $[C]_m$ , concentration of guanidinium chloride at which 50% of the total peptides is unfolded) of the GCN4-based peptides are shown in Table 2-3. The denaturation curves and graphs showing the fraction unfolded are shown in Figures 2-9 and 2-10.



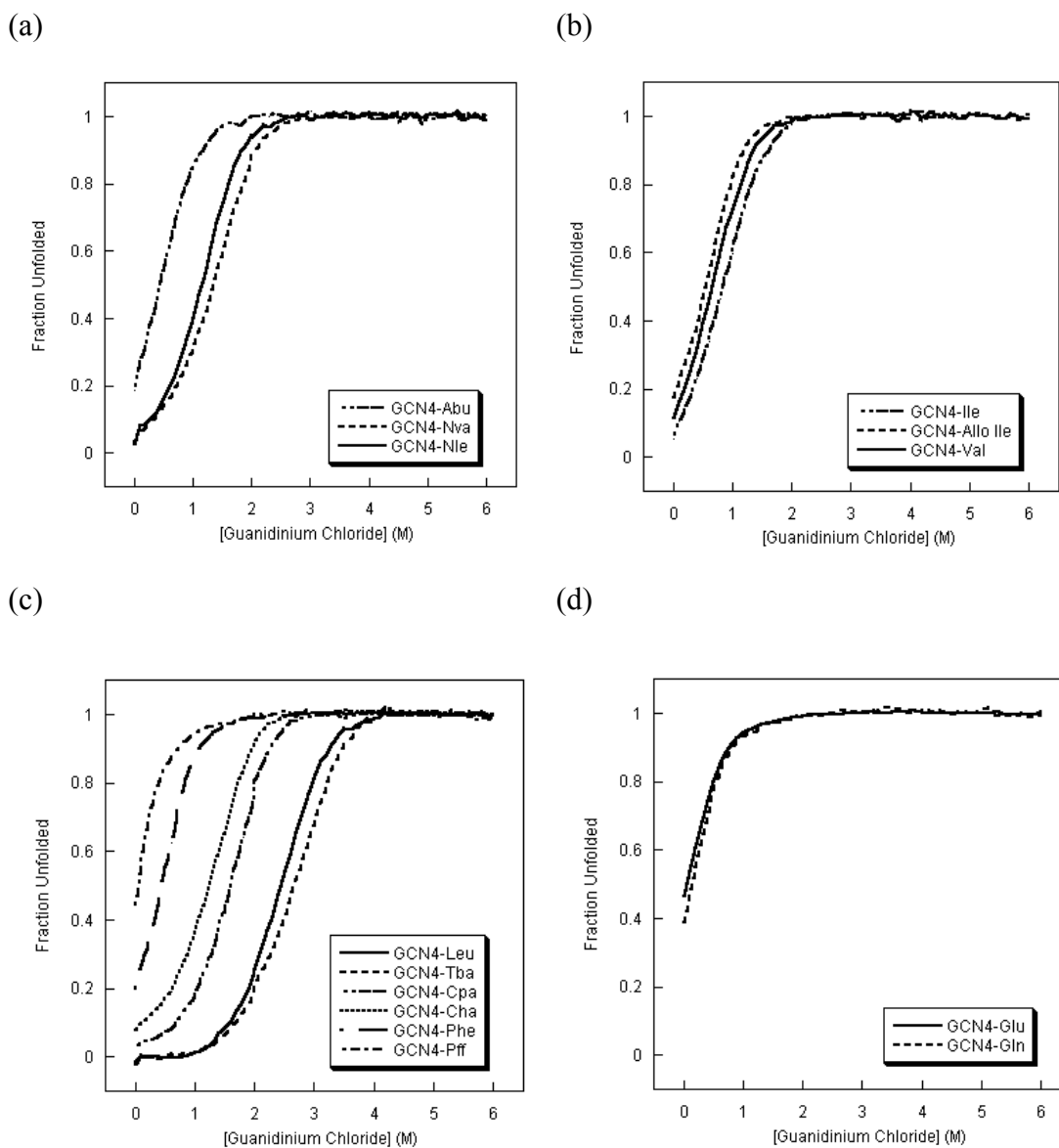
**Figure 2-8.** A typical guanidinium titration curve with the proper folded and unfolded baselines depicted. The denaturation curve for GCN4-Tba is shown.

Peptide GCN4-Nva and GCN4-Nle exhibited different  $\theta_{222}$  at 0 M guanidinium (Figure 2-9a). However, the denaturation curves gradually overlapped as the guanidinium concentration increased. This also happened for GCN4-Ile, GCN4-Allo Ile, and GCN4-Val (Figure 2-9b). On the other hand, GCN4-Cha and GCN4-Leu exhibited similar  $\theta_{222}$  at 0 M guanidinium, but the denaturation curves diverged as guanidinium increased in concentration (Figure 2-9c). As mentioned above,  $\theta_{222}$  represents the helical content of a peptide but provide no information on the stability of a coiled coil. Therefore, it is not surprising for some coiled coils to exhibit different starting  $\theta_{222}$  but gradually overlapping denaturation curves, or to exhibit similar  $\theta_{222}$  but gradually diverging denaturation curves.

A peptide with a higher  $[C]_m$  means that a higher concentration of guanidinium is



**Figure 2-9.** Guanidinium denaturation curves for GCN4-Xaa peptides at 30  $\mu$ M peptide in 50 mM phosphate, 150 mM NaCl, and 0 M to 6 M (with 0.1 M intervals) guanidinium chloride at pH 7 and 4  $^{\circ}$ C as monitored by CD at 222 nm reported in mean residue ellipticity. (a) Guanidinium denaturation curves of GCN4-Xaa peptides in which Xaa are residues with linear side chains. (b) Guanidinium denaturation curves of GCN4-Xaa peptides in which Xaa are residues with  $\beta$ -branched side chains. (c) Guanidinium denaturation curves of GCN4-Xaa peptides in which Xaa are residues with  $\gamma$ -branched side chains. (d) Guanidinium denaturation curves of GCN4-Xaa peptides in which Xaa are residues with polar side chains.



**Figure 2-10.** Fraction unfolded as a function of guanidinium concentration for GCN4-Xaa peptides as derived from the guanidinium denaturation curves. (a) Fraction unfolded plots for GCN4-Xaa peptides in which Xaa are residues with linear side chains. (b) Fraction unfolded plots for GCN4-Xaa peptides in which Xaa are residues with  $\beta$ -branched side chains. (c) Fraction unfolded plots for GCN4-Xaa peptides in which Xaa are residues with  $\gamma$ -branched side chains. (d) Fraction unfolded plots for GCN4-Xaa peptides in which Xaa are residues with polar side chains.

needed to denature half of the peptides. In other words, the peptide is more resilient towards guanidinium denaturation. For the linear residues,  $[C]_m$  followed the trend: GCN4-Abu < GCN4-Nle < GCN4-Nva (Table 2-3). There does not appear to be any structural reason for this trend. For  $\beta$ -branched amino acids,  $[C]_m$  followed the trend: GCN4-Allo Ile < GCN4-Val < GCN4-Ile (Table 2-3). Peptide GCN4-Ile is more stable than GCN4-Val, consistent with the hydrophobicity trend Ile > Val. Allo Ile and Ile share the same stereochemistry at the  $C\alpha$ , but have opposite chirality at the  $C\beta$ . This may impair the packing of the Allo Ile side chain at the coiled coil interface, making GCN4-Allo Ile more susceptible to guanidinium denaturation compared to GCN4-Ile. For the aliphatic  $\gamma$ -branched residues, the  $[C]_m$  followed the trend: GCN4-Cha < GCN4-Cpa < GCN4-Leu < GCN4-Tba (Table 2-3). The difference between GCN4-Leu and GCN4-Tba can be attributed to the hydrophobicity of Leu versus Tba. However, Cpa and Cha exhibit higher hydrophobicity, and are larger in size and higher in rigidity compared to Leu and Tba. The larger and more rigid side chains may introduce unfavorable steric clashes at the coiled coil interface, overwhelming the stabilization effects from the increase in hydrophobicity. The Cha, Phe, and Pff residues all bear a six-membered ring. The  $[C]_m$  followed the trend: GCN4-Pff < GCN4-Phe < GCN4-Cha

(Table 2-3). The side chain of Cha and Phe are similar in size. Nevertheless, the benzene ring of Phe possesses a quadrupole whereas Cha is simply hydrophobic. The quadrupole may interfere with the packing of side chains in the interface, making the coiled coil more prone to be denatured by guanidinium. The superb hydrophobicity of Pff may disrupt the packing at the coiled coil interface in such a way that the coiled coil is to some extent distorted to allow for the maximum fluorophilic effect,<sup>25</sup> which describes the superior affinity between fluorocarbons such that fluorocarbons form a fluorophilic phase distinct from organic phase,<sup>25</sup> between the pentafluorophenyl groups. This can lead to the very low  $[C]_m$  for GCN4-Pff. For the polar residues, GCN4-Asn and GCN4-Asp did not fold, whereas GCN4-Gln and GCN4-Glu were more well folded. The side chain of Asn and Asp are similar in structure compared to Leu. Therefore, Asn and Asp may pack in a similar pattern as Leu. While such a packing pattern can create a stabilizing hydrophobic contact for GCN4-Leu, the same packing pattern for GCN4-Asn and GCN4-Asp would place the polar entities at the hydrophobic interface, thereby destabilizing the coiled coil structure. On the other hand, one-methylene-longer side chains of Gln and Glu not only increased the side chain hydrophobicity, more side chain rotamers with the longer side chains can place the polar entities toward the

aqueous environment instead of at the interface, resulting in the relatively higher stability of the coiled coil structure.

#### *$\Delta G_{\text{unfold, H}_2\text{O}}$ of GCN4-Derived Peptides*

The equilibrium constant ( $K_{\text{eq}}$ ) of “folded dimer  $\rightleftharpoons$  2 unfolded monomer” was derived from the denaturation curves using the folded baseline and unfolded baseline.

The  $\Delta G_{\text{unfold}}$  for each guanidinium concentration was derived from  $\Delta G_{\text{unfold}} = -RT \ln K_{\text{eq}}$ .

The  $\Delta G_{\text{unfold}}$  was plotted as a function of the concentration of guanidinium chloride

(Figure 2-11 a). Data points near the melting concentration were fit linearly and

extrapolated to 0 M guanidinium to obtain the  $\Delta G_{\text{unfold, H}_2\text{O}}$  (Figure 2-11 b). The slope of

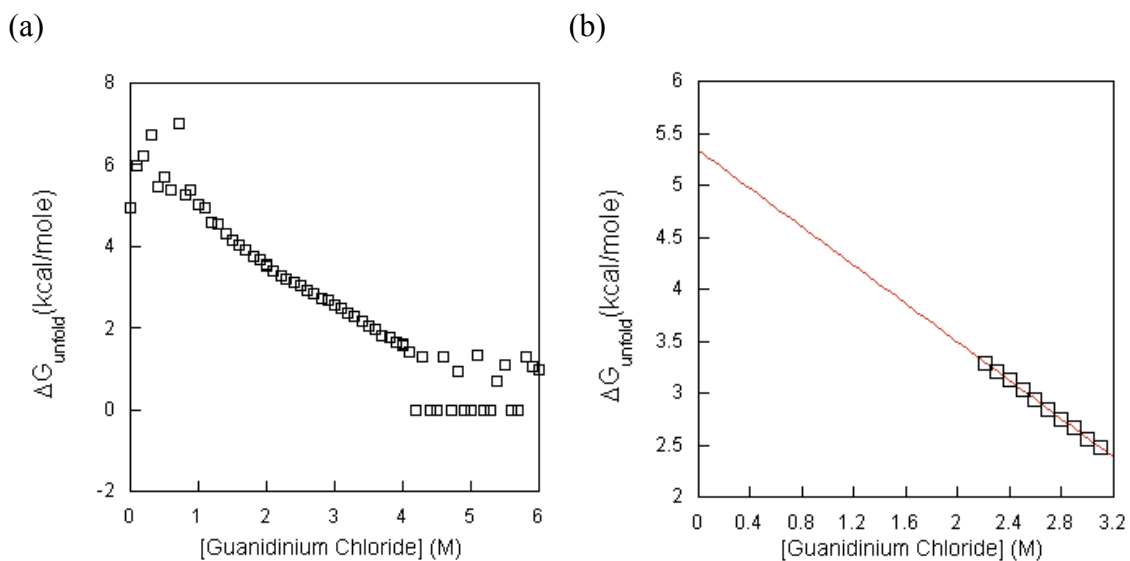
the line of the fit (m value) and the free energy of unfolding of GCN4 peptides are listed

in Table 2-3 and graphed in Figure 2-12.

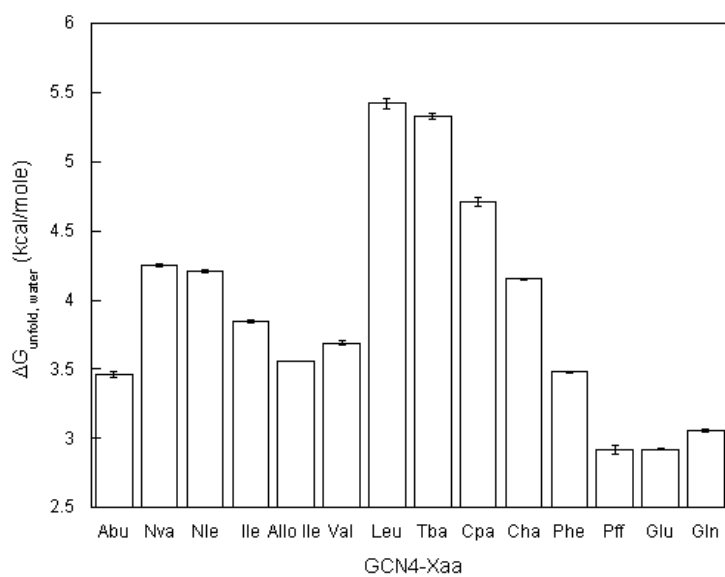
A higher  $\Delta G_{\text{unfold, H}_2\text{O}}$  indicates that it is more unfavorable to unfold the coiled coil,

and thus the coiled coil is more stable. The trend for  $\Delta G_{\text{unfold, H}_2\text{O}}$  is generally the same

as the trend for  $[C]_{\text{m}}$  except for peptides GCN4-Leu and GCN4-Tba. The  $[C]_{\text{m}}$  of



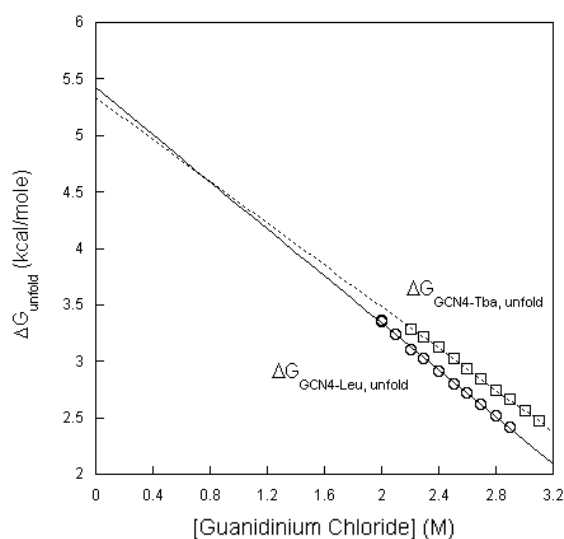
**Figure 2-11.** The plot of  $\Delta G_{\text{unfold}}$  against guanidinium concentration for GCN4-Tba. (a) The plot of  $\Delta G_{\text{unfold}}$  against guanidinium concentration using all  $\Delta G_{\text{unfold}}$  data points. (b) The plot of  $\Delta G_{\text{unfold}}$  against guanidinium concentration using data points near the melting concentration. Line of fit was also shown.



**Figure 2-12.** The bar graph of  $\Delta G_{\text{unfold, H}_2\text{O}}$  of GCN4 peptides.



GCN4-Tba is greater than that of GCN4-Leu, whereas the order is reversed for  $\Delta G_{\text{unfold}}$ ,  $\text{H}_2\text{O}$ . This originates from the different slopes (m values) in the  $\Delta G_{\text{unfold}}$  - guanidinium concentration plots for the two peptides (Figure 2-10c, Figure 2-13). The slope is related to the number of the denaturant binding sites of a peptide.<sup>40-42</sup> The steeper the slope, the more the denaturant binding sites, the more the surface area burial of a peptide.<sup>40-42</sup> A larger surface area burial means that it is more difficult to denature the peptide, because the denaturant binding sites are buried.<sup>40-42</sup> However, once the denaturation begins, the peptide unfolds readily because the numerous denaturant binding sites are now exposed.<sup>40-42</sup> In other words, the peptide exhibits a higher cooperativity when the slope is steeper.<sup>40-42</sup> A peptide with a higher  $\Delta G_{\text{unfold, H}_2\text{O}}$ , a lower  $[\text{C}]_m$  and a steeper slope (as in GCN4-Leu) suggests that it is more difficult to initiate the denaturation (a higher  $\Delta G_{\text{unfold, H}_2\text{O}}$ ), but denaturation occurs more readily (a lower  $[\text{C}]_m$  and a steeper slope) once the process is initiated. On the other hand, a peptide with lower  $\Delta G_{\text{unfold, H}_2\text{O}}$ , a higher  $[\text{C}]_m$  and a flatter slope (as in GCN4-Tba) suggests that it is easier to initiate the denaturation (a lower  $\Delta G_{\text{unfold, H}_2\text{O}}$ ), but is more difficult for the denaturation process to complete (a higher  $[\text{C}]_m$  and a flatter slope).



**Figure 2-13.**  $\Delta G_{\text{unfold}}$  of GCN4-Leu and GCN4-Tba.

The  $\Delta G_{\text{unfold, H}_2\text{O}}$  generally followed the trend  $\gamma$ -branched > linear >  $\beta$ -branched for the peptides with aliphatic Xaa. More specifically, the  $\Delta G_{\text{unfold, H}_2\text{O}}$  followed the trend:

GCN4-Leu > GCN4-Nle > GCN4-Ile > GCN4-Allo Ile (Figure 2-12). The

hydrophobicity of the four residues Leu, Nle, Ile, and Allo Ile are similar. As such, the side chain structures must play a role in coiled coil stability, leading to the significant differences in  $\Delta G_{\text{unfold, H}_2\text{O}}$ .

Although Cha bears a  $\gamma$ -branched side chain and is more hydrophobic compared to Nle and Nva, the  $\Delta G_{\text{unfold, H}_2\text{O}}$  is lower for GCN4-Cha compared to GCN4-Nle and GCN4-Nva. As mention earlier, Cha bears a large (and bulky) side chain, which may cause unfavorable steric clashes at the coiled coil interface. The stability gained from hydrophobicity and the  $\gamma$ -branched structure may not be sufficient to compensate for the

unfavorable steric clashes. This indicates that hydrophobicity, side chain shape, and side chain size all affect coiled coil stability. Therefore, as a first step to further elucidate the factors contributing to coiled coil stability, namely coiled coil propensity, hydrophobicity, and side chain size and shape, the coiled coil propensity for various amino acids were measured.

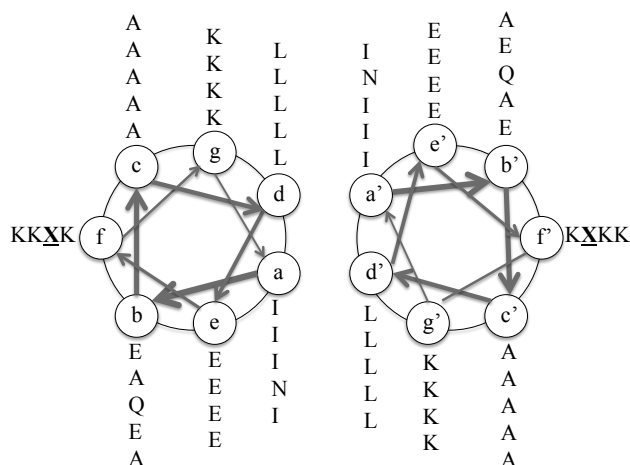
#### *Design of IaLd-Derived Peptides*

Coiled coil propensity of the residues at the guest site affects inherent coiled coil stability. To measure coiled coil propensities, homodimeric parallel coiled coils were used (Figure 2-14). The sequence was based on DeGrado's<sup>23</sup> and Hu's design.<sup>43</sup> Isoleucines were placed at the *a* positions, since Ile-Ile provides very stable a-a' interactions.<sup>43,44</sup> One asparagine was placed at the 4<sup>th</sup> *a* position to control the oligomeric state and orientation of the coiled coil.<sup>43</sup> The *d* positions were occupied by leucines, the most favored residue at this position.<sup>27</sup> Glutamic acid and lysine were placed at *e* and *g* positions respectively, providing interhelical Coulombic attractions and controlling coiled coil orientation. Many alanines were used because alanine has the highest coiled coil propensity.<sup>23</sup> Tyrosine (Tyr) was incorporated to facilitate

concentration determination by UV-vis,<sup>36</sup> and the Gly-Gly intervening sequence was included to minimize interference in the CD signal by the Tyr chromophore.<sup>45</sup> Studies on IaLd-Allo Ile, IaLd-Ile, IaLd-Leu, IaLd-Tba and IaLd-Val were previously performed by Hsien-Po Chiu in our lab. These peptides, along with IaLd-Abu, IaLd-Cha, IaLd-Cpa, IaLd-Nle, IaLd-Nva, IaLd-Pff and IaLd-Phe, are reported in this chapter. IaLd-Asp, IaLd-Asn, IaLd-Glu and IaLd Gln were not included due to the following reasons. First, GCN4-Asn and GCN4-Asp barely folded. It would be difficult to discuss factors affecting the stabilities of these peptides. Second, the packing of polar residues in the coiled coil interface often involves buried water molecules and the formation of species of higher oligomeric state,<sup>12, 17</sup> which complicates the discussion. Therefore, polar substituents were excluded in these experiments.

IaLd-Xaa:

Ac-YGGE IEALEKK IAALEXK IQALEKK NEALEKK IAAL - NH<sub>2</sub>



**Figure 2-14.** Sequence and helical wheel of IaLd. X denotes the mutation sites.

## Peptide Synthesis of IaLd-Derived Peptides

Peptides were synthesized using commercially available reagents by Fmoc-based solid peptide synthesis (SPPS). Nle and Cpa were Fmoc protected as described earlier (Scheme 1). The identity of the peptides was confirmed by MALDI-TOF mass spectrometry. Upon confirmation of the peptides, purification was carried out on reverse RP-HPLC to greater than 95% purity. The crude yield of peptide synthesis, molecular formula, calculated  $[MH^+]$ , and observed  $m/z$  of GCN4 peptides are shown in Table 2-4.

**Table 2-4.** Crude Yield of Peptide Synthesis, Molecular Formula, Calculated  $[MH^+]$ , and Observed  $m/z$  of IaLd Peptides

Peptide	Crude Yield (%)	Molecular Formula	Calculated $[MH^+]$	Observed $m/z$ in MALDI-TOF MS
IaLd-Abu	62.6 <sup>a</sup>	C <sub>180</sub> H <sub>308</sub> N <sub>46</sub> O <sub>54</sub>	3979.29 <sup>b</sup>	3978.6 <sup>b</sup>
IaLd-Cha	54.3	C <sub>185</sub> H <sub>316</sub> N <sub>46</sub> O <sub>54</sub>	4047.35 <sup>b</sup>	4047.2 <sup>b</sup>
IaLd-Cpa	55.7	C <sub>184</sub> H <sub>314</sub> N <sub>46</sub> O <sub>54</sub>	4033.33 <sup>b</sup>	4032.9 <sup>b</sup>
IaLd-Nle	55.9	C <sub>182</sub> H <sub>312</sub> N <sub>46</sub> O <sub>54</sub>	4009.71	4008.8
IaLd-Nva	56.7	C <sub>181</sub> H <sub>310</sub> N <sub>46</sub> O <sub>54</sub>	3995.68	3994.7
IaLd-Pff	69.2 <sup>a</sup>	C <sub>185</sub> H <sub>305</sub> F <sub>5</sub> N <sub>46</sub> O <sub>5</sub>	4133.68	4131.4
IaLd-Phe	43.7 <sup>a</sup>	C <sub>185</sub> H <sub>310</sub> N <sub>46</sub> O <sub>54</sub>	4043.73	4042.4

<sup>a</sup>Approximate yield is given here. These three peptides were initially synthesized together. Upon coupling of Xaa, the resin was divided into three portions, one portion for each peptide. The resin was not lyophilized, therefore, exact yield cannot be calculated. <sup>b</sup>Observed  $m/z$  and calculated  $[MH^+]$  were given as exact mass for IaLd-Abu, IaLd-Cha, and IaLd-Cpa. For other peptides, average molecular weights were given.

### *UV-Visible Spectroscopy (UV-vis) of IaLd-Derived Peptides*

The concentration of the peptide stock solutions was determined by the tyrosine absorbance in 6 M guanidinium chloride as described by Edelhoch.<sup>36, 37</sup> UV data were obtained using 1 mm pathlength cells. Absorbance at 276 nm, 278 nm, 280 nm and 282 nm were measured at different concentrations of peptide. Linear regression was performed to obtain the concentration. Data with correlation coefficients (R) less than 0.99 were discarded. Concentrations and regression coefficients of IaLd peptides are shown in Table 2-5.

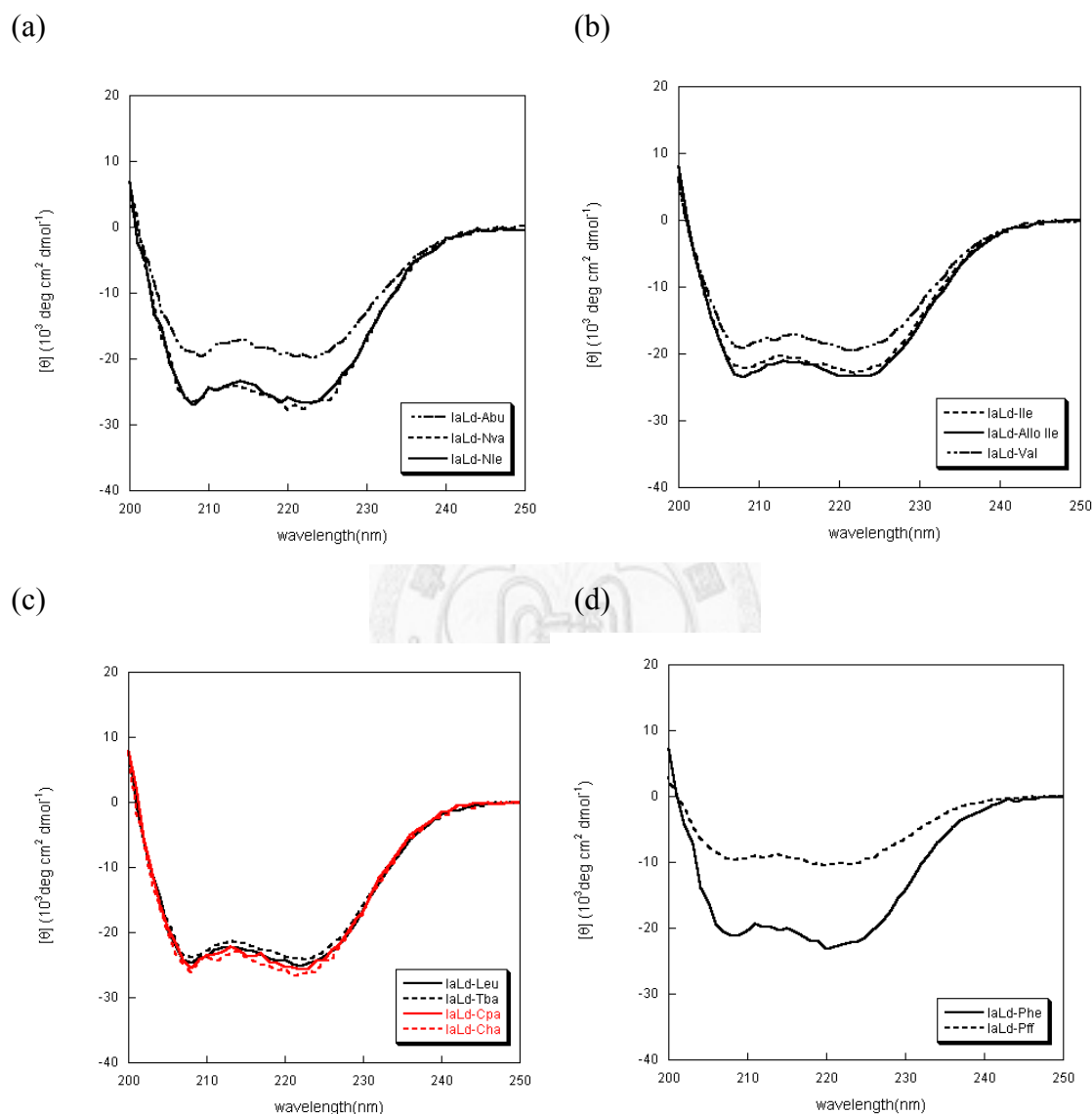
**Table 2-5.** Concentrations and Regression Coefficients of IaLd Peptides

Peptide	Concentration (mM)	Regression Coefficient
IaLd-Abu	2.97±0.06	0.99502
IaLd-Cha	3.57±0.05	0.998
IaLd-Cpa	5.52±0.09	0.99685
IaLd-Nle	2.73±0.04	0.99799
IaLd-Nva	2.06±0.06	0.99208
IaLd-Pff	8.7±0.1	0.99772
IaLd-Phe	5.14±0.04	0.99920

### *Circular Dichroism (CD) Spectroscopy of IaLd-Derived Peptides*

CD measurements were performed at 20  $\mu$ M peptide in 10 mM 3-(N-morpholino)-propanesulfonic acid (MOPS) at pH 7.5 and 25 °C.<sup>23</sup> The magnitude of the CD signal

at 222 nm reflects the helical content of a coiled coil.<sup>38</sup> The CD spectra of the peptides are shown in Figure 2-15, and the MRE values at 222 nm ( $[\theta]_{222}$ ) are listed in Table 2-6.



**Figure 2-15.** CD spectra of IaLd-Xaa at 20  $\mu$ M peptide in 10 mM MOPS at pH 7.5 and 25  $^{\circ}$ C. (a) CD spectra of IaLd-Xaa peptides in which Xaa are residues with linear side chains. (b) CD spectra of IaLd-Xaa peptides in which Xaa are residues with  $\beta$ -branched side chains. (c) CD spectra of IaLd-Xaa peptides in which Xaa are residues with  $\gamma$ -branched aliphatic side chains. (d) CD spectra of IaLd-Xaa peptides in which Xaa are residues with  $\gamma$ -branched aromatic side chains.

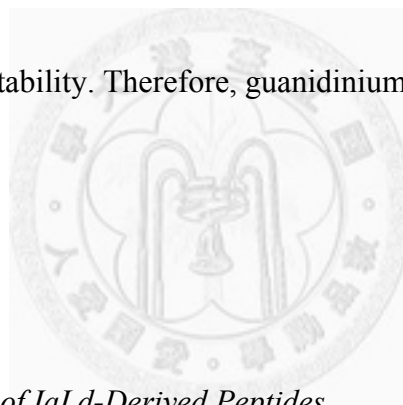
**Table 2-6.**  $[\theta]_{222}$  at 0 M Guanidinium, Melting Concentrations ( $[C]_m$ ), m values, and  $\Delta G_{\text{unfold, H}_2\text{O}}$  of IaLd Peptides

Peptides	$[\theta]_{222}$ (deg cm <sup>2</sup> dmol <sup>-1</sup> )	$[C]_m$ (M)	m Value	$\Delta G_{\text{unfold, H}_2\text{O}}$ (kcal/mole)
IaLd-Abu	-19300±500	3.49	-1.06±0.02	6.97±0.07
IaLd-Nva	-27500±900	3.64	-1.02±0.02	6.92±0.06
IaLd-Nle	-27100±800	3.58	-1.02±0.02	6.86±0.06
IaLd-Ile	-24100±300	3.31	-1.146±0.005	6.99±0.02
IaLd-Allo Ile	-25700±300	3.00	-1.08±0.01	6.46±0.03
IaLd-Val	-20800±400	3.15	-1.126±0.006	6.75±0.02
IaLd-Leu	-26700±300	3.73	-1.155±0.009	7.52±0.03
IaLd-Tba	-25700±200	3.53	-1.11±0.02	7.15±0.09
IaLd-Cpa	-25600±800	3.56	-1.08±0.02	7.05±0.08
IaLd-Cha	-26400±800	3.55	-1.05±0.02	6.96±0.08
IaLd-Phe	-22500±400	3.16	-1.08±0.02	6.61±0.06
IaLd-Pff	-10200±300	2.84	-0.94±0.02	5.88±0.06

The helical content of the the IaLd peptides with linear Xaa side chains followed the trend: IaLd-Abu < IaLd-Nva ≈ IaLd-Nle (Figure 2-15a). For β-branched amino acids, the helical content of the IaLd peptides followed the trend: IaLd-Val < IaLd-Ile ≈ IaLd-All Ile (Figure 2-15b). For these two groups, the longer side chain resulted in a higher helical content, as observed in monomeric helices.<sup>46</sup> Furthermore, the IaLd peptides with β-branched amino acids generally exhibited lower helical content compared to those with linear amino acids, which is also consistent with previous reports on monomeric helices.<sup>47, 48</sup> For γ-branched amino acids, the helical content of



the IaLd peptides followed the trend: IaLd-Pff < IaLd-Phe < IaLd-Tba < IaLd-Leu ≈ IaLd-Cpa < IaLd-Cha (Figure 2-15c and d). The helical content of all IaLd peptides with aliphatic  $\gamma$ -branched amino acids were similar, whereas the helical content of IaLd-Pff was significantly different from the others. The unique property of IaLd-Pff may originate from the superb hydrophobicity of the pentafluorophenyl group.<sup>49</sup> The coiled coil structure may be distorted to minimize the contact of Pff with surrounding water molecules. As described in GCN4 section, helical content alone does not provide information on coiled coil stability. Therefore, guanidinium denaturation experiments were performed.



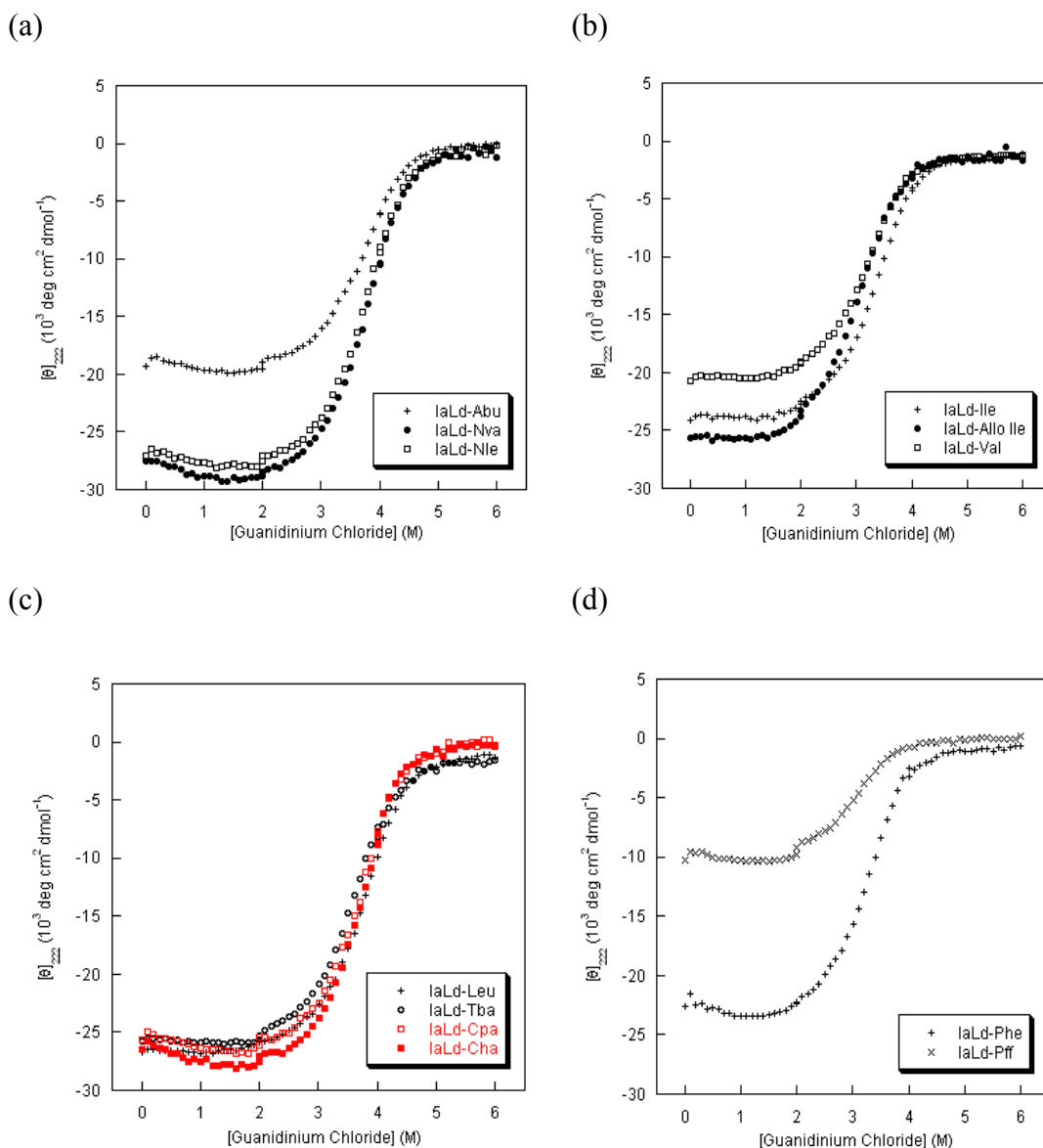
#### *Guanidinium Denaturation of IaLd-Derived Peptides*

Guanidinium denaturation experiments were performed with 20  $\mu$ M peptide concentration in 10 mM MOPS with 0 to 6 M (at 0.1 M intervals) guanidinium chloride at pH 7.5 and 25 °C.<sup>23</sup> The denaturation process was monitored by CD, and the CD signal at 222 nm was acquired. A coiled coil gradually unfolds upon addition of guanidinium chloride, leading to the decrease of CD signal. The folded and unfolded baselines for each peptide were determined using suitable CD signals by linear

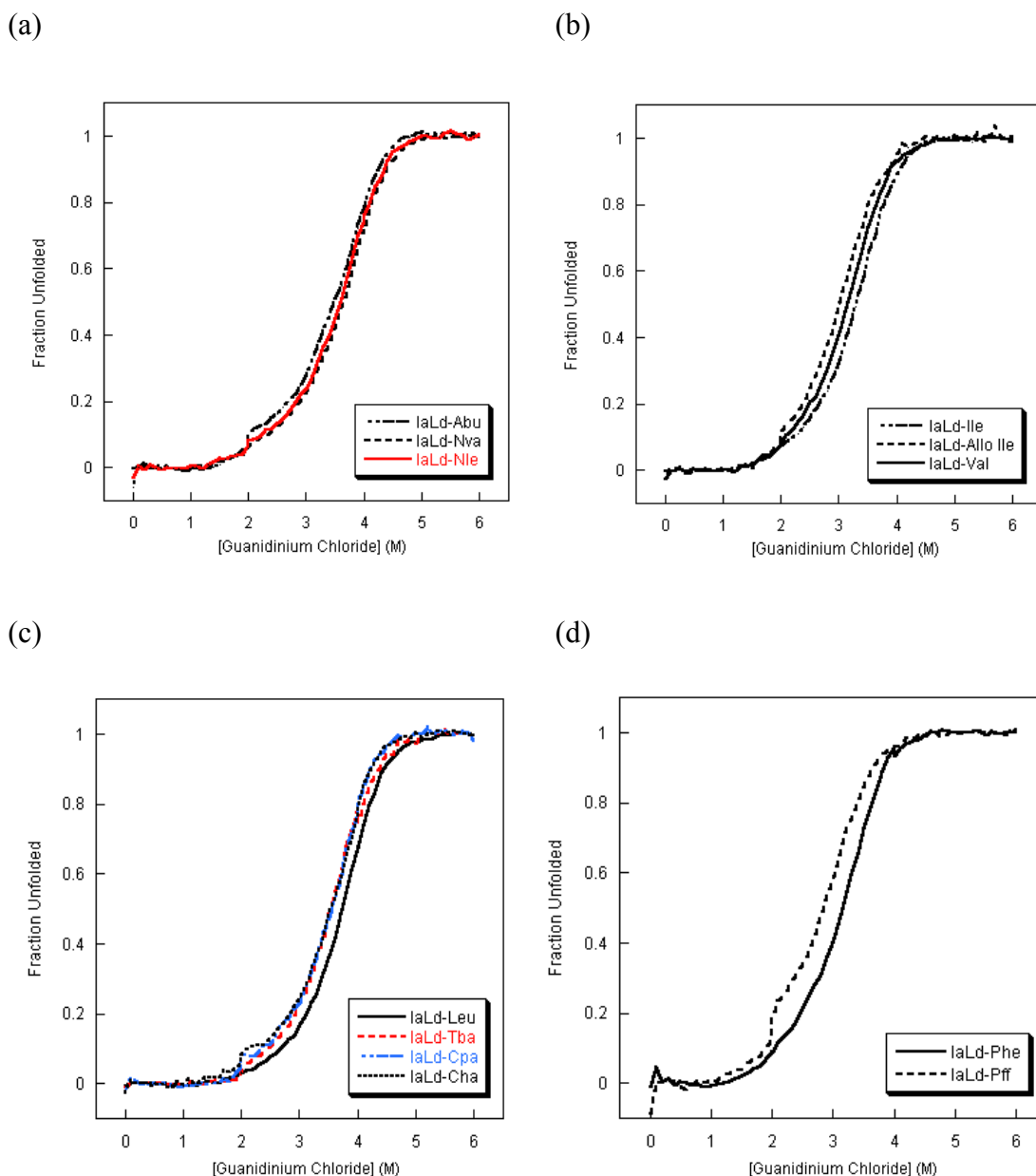
regression methods.<sup>39</sup> These baselines describe the expected CD signal at different guanidinium concentrations for the fully folded dimer and a fully unfolded monomer.

The fraction unfolded of a given peptide at each guanidinium concentration can be derived using the unfolded and folded baselines.<sup>39</sup> The melting concentrations ( $[C]_m$ ) of IaLd peptides are shown in Table 2-6. The denaturation curves and graphs showing fraction unfold are shown in Figure 2-16 and Figure 2-17.

The  $[C]_m$  of the IaLd peptides with linear amino acids were basically the same regardless of side chain length (Table 2-6). For the IaLd peptides with  $\beta$ -branched residues,  $[C]_m$  followed the trend: IaLd-Allo < IaLd-Val < IaLd-Ile. The  $[C]_m$  for the peptides with linear amino acids were higher than that for peptides with  $\beta$ -branched amino acids, indicating that the linear side chains were more favorable in coiled coil structures. For the aliphatic  $\gamma$ -branched amino acids, the  $[C]_m$  for IaLd-Leu was slightly higher, whereas the  $[C]_m$  of IaLd-Tba, IaLd-Cpa, and IaLd-Cha were the same. The difference in stability between IaLd-Phe and IaLd-Pff can be attributed to the difference in hydrophobicity.



**Figure 2-16.** Guanidinium denaturation curves for IaLd-Xaa peptides at 20  $\mu$ M peptide concentration in 10 mM MOPS with 0 to 6 M (at 0.1 M intervals) guanidinium chloride at pH 7.5 and 25  $^{\circ}$ C as monitored by CD at 222 nm reported in mean residue ellipticity. (a) Guanidinium denaturation curves of IaLd-Xaa peptides in which Xaa are residues with linear side chains. (b) Guanidinium denaturation curves of IaLd-Xaa peptides in which Xaa are residues with  $\beta$ -branched side chains. (c) Guanidinium denaturation curves of IaLd-Xaa peptides in which Xaa are residues with  $\gamma$ -branched aliphatic side chains. (d) Guanidinium denaturation curves of GCN4-Xaa peptides in which Xaa are residues with  $\gamma$ -branched aromatic side chains.



**Figure 2-17.** Fraction unfolded as a function of guanidinium concentration for IaLd-Xaa peptides as derived from guanidinium denaturation curves. (a) Fraction unfolded plots for IaLd-Xaa peptides in which Xaa are residues with linear side chains. (b) Fraction unfolded plots for IaLd-Xaa peptides in which Xaa are residues with  $\beta$ -branched side chains. (c) Fraction unfolded plots for IaLd-Xaa peptides in which Xaa are residues with  $\gamma$ -branched aliphatic side chains. (d) Fraction unfolded plots for IaLd-Xaa peptides in which Xaa are residues with  $\gamma$ -branched aromatic side chains.

### $\Delta G_{\text{unfold, H}_2\text{O}}$ of IaLd-Derived Peptides

The equilibrium constant ( $K_{\text{eq}}$ ) of “folded dimer  $\rightleftharpoons$  2 unfolded monomer” was derived from the denaturation curves using the folded baseline and unfolded baseline.

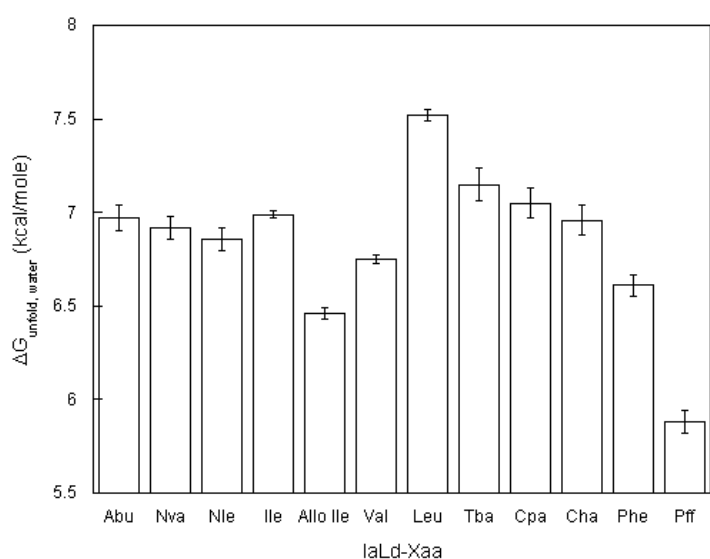
The  $\Delta G_{\text{unfold}}$  for each guanidinium concentration was derived from  $\Delta G_{\text{unfold}} = -RT \ln K_{\text{eq}}$ .

The  $\Delta G_{\text{unfold}}$  was plotted as a function of the concentration of guanidinium chloride.

Data points near the melting concentration were fit linearly and extrapolated to 0 M

guanidinium to obtain the  $\Delta G_{\text{unfold, H}_2\text{O}}$ . The slope of the linear fit (m value) and the free

energy of unfolding of IaLd peptides are listed in Table 2-6 and graphed in Figure 2-18.



**Figure 2-18.** The bar graph of  $\Delta G_{\text{unfold, H}_2\text{O}}$  of IaLd peptides.

The  $\Delta G_{\text{unfold, H}_2\text{O}}$  of the IaLd peptides with natural Xaa side chains followed the trend:

IaLd-Phe < IaLd-Val < IaLd-Ile < IaLd-Leu (Table 2-6, Figure 2-18), which is

generally the same as the trend of the helix propensity Val < Phe < Ile < Leu.<sup>23</sup> IaLd-Pff has the lowest  $\Delta G_{\text{unfold, H}_2\text{O}}$ , in coherence with previous reports that Pff has a low helix propensity.<sup>49</sup>

### *Measuring Hydrophobicities of the Amino Acids*

Hydrophobicity was measured by thin layer chromatography (TLC).<sup>50</sup> The TLC of the amino acids were performed with *n*-butanol - 0.05 M ammonium acetate (2:1) and *n*-butanol - 50% acetic acid (2:1) on cellulose plates. The spots were detected by ninhydrin.  $R_f$  value of each amino acids were measured and converted into  $\log P_{\text{ow}}$  values (ow: *n*-octanol - water) to conform with the Hansch definition of the hydrophobic parameter.<sup>50, 51</sup> The amino acids include: Abu, Nle, Nva, Ile, Allo Ile, Val, Leu, Tba, Cha, Phe, and Pff. The experimentally derived  $\log P_{\text{ow}}$  values are listed in Table 2-7. Data (except for Cpa) was obtained by Hsien-Po Chiu.

**Table 2-7.**  $\log P_{\text{ow}}$  of the Amino Acids

	Abu	Nva	Nle	Ile	Allo Ile	Val	Leu	Tba	Cpa	Cha	Phe	Pff
$\log P_{\text{ow}}$	-2.54	-1.87	-1.29	-1.57	-1.60	-2.11	-1.44	-1.14	-0.94	-0.61	-1.58	-0.65

### *Size and Shape Parameters*

Several parameters have been developed over the past decades to describe the size and shape of a molecule. Common parameters include Taft's steric parameter ( $E_s$ ),<sup>52</sup> sterimol parameters (L, B1, B5),<sup>53, 54</sup> and molecular refractivity (MR).<sup>55, 56</sup> Taft's steric parameter ( $E_s$ ) is defined as  $E_s = \log (k_R/k_{CH_3})_A$ , in which  $k_R$  is the rate constant of the hydrolysis of RCOOMe,  $k_{CH_3}$  is the rate constant of the hydrolysis of CH<sub>3</sub>COOMe, and A denotes that it is an acid-catalyzed hydrolysis.<sup>52, 57</sup>  $E_s$  can be expressed in terms of the van der Waals radius of the molecule.<sup>52, 57</sup>  $E_s$  is a parameter that considers both the size and shape of the molecule.<sup>57</sup> Sterimol parameters describe the shape of a molecule with L, B1, and B5.<sup>53, 54, 57</sup> L represents the length of a substituent along the axis of the bond between the first atom of the substituent and the parent molecule.<sup>54, 57</sup> B1 is the parameter for the minimum width that is perpendicular to L, and B5 is for the maximum width.<sup>54, 57</sup> Sterimol parameters consider both the size, and especially the shape, of a molecule.<sup>57</sup> Molecular refractivity (MR) is defined as  $MR = [(n^2-1)/(n^2+2)] (MW/d)$ , in which  $n$  is the refractive index of the molecule, MW is the molecular weight, and  $d$  is the density of the molecule.<sup>57</sup> For organic compounds,  $n$  differs in a small range. Therefore, MR is actually a small correction on the volume (MW/ $d$ ) of the molecule,

and it contains no information on the shape of the molecule.<sup>57</sup> Apart from these parameters, the volumes of the molecules were calculated using Discovery Studio. The  $E_s$ , [L, B1, B5], MR, and the volumes of the side chain of Abu, Nva, Nle, Ile, Allo Ile, Val, Leu, Tba, and Cha are listed in Table 2-8.

**Table 2-8.**  $E_s$ , [L, B1, B5], MR, and the Volume of the Amino Acids Side Chains

	$E_s^a$	$L^a$	$B1^a$	$B5^a$	$MR^a$	Volume ( $\text{\AA}^3$ ) <sup>b</sup>
Abu	-1.31	4.11	1.52	3.17	1.03	38.452
Nva	-1.6	4.92	1.52	3.49	1.5	55.810
Nle	-1.63	6.17	1.52	4.54	1.96	79.112
Ile	-2.37	4.92	1.9	3.49	1.96	79.112
Allo Ile	-2.37	4.92	1.9	3.49	1.96	79.112
Val	-1.71	4.11	1.9	3.17	1.5	55.810
Leu	-2.17	4.92	1.52	4.45	1.96	77.843
Tba	-2.98	4.89	1.52	4.18	2.42	91.012
Cha	-2.22	6.09	1.52	5.42	3.13	114.385
Phe	-1.62	4.62	1.52	6.02	3	93.091

<sup>a</sup>Values from Hansch, C.; Leo, A.; Heller, S. R. Exploring QSAR: Hydrophobic, Electronic, and Steric Constants. ACS, 1995. <sup>b</sup>Volume calculated by Discovery Studio 2.1 using CFF forcefield and dielectric constant 80, and was probed with 1.4  $\text{\AA}$  radius.

### Discussion

$\Delta G_{\text{unfold, H}_2\text{O}}$  of GCN4 peptides reflect the overall outcome originating from the different contribution of coiled coil propensity, hydrophobicity, and side chain size and shape to coiled coil stability. The coiled coil propensity, hydrophobicity, and side chain



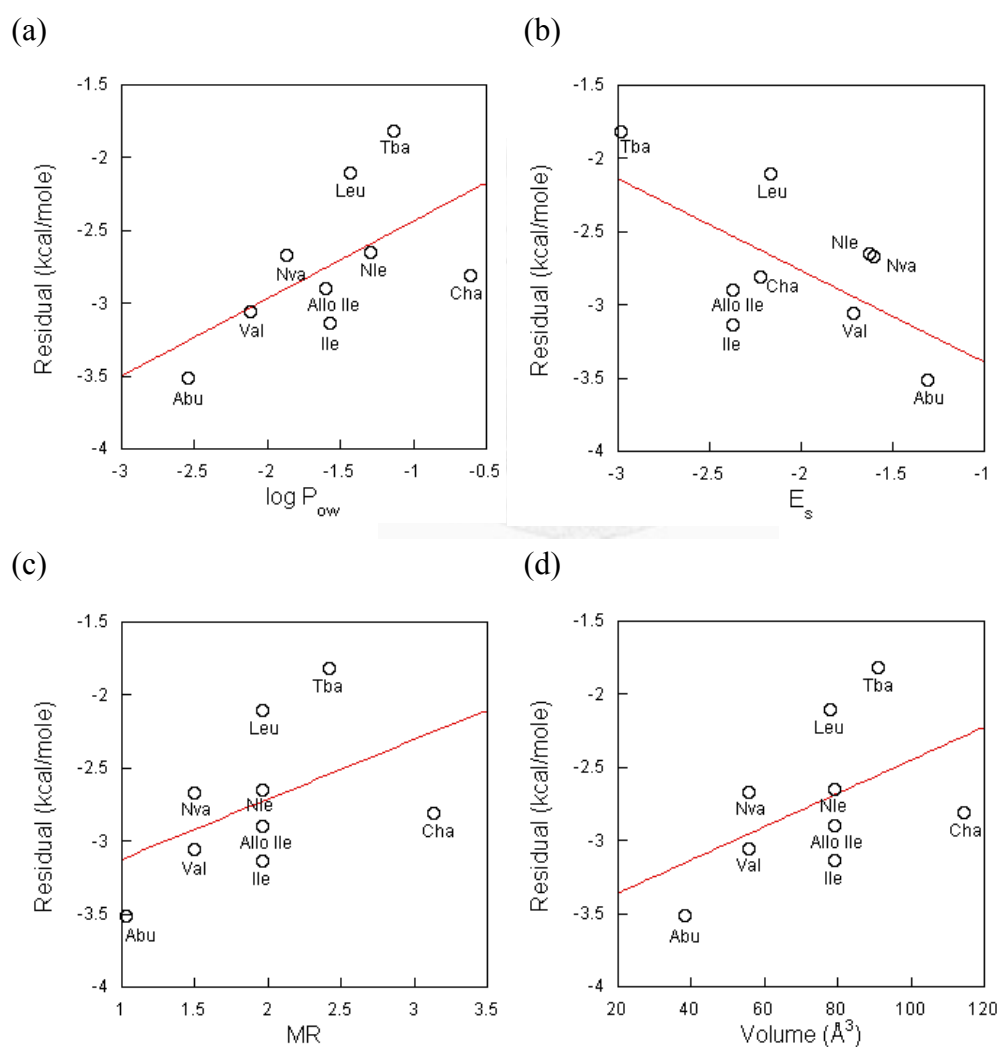
size and shape of the substituents were obtained to deduce the detailed contribution of these factors to GCN4 coiled coil stability. To remove the intrinsic preference of an amino acid to adopt a coiled coil structure from the overall GCN4 stability,  $\Delta G_{\text{unfold, H}_2\text{O}}$  of IaLd peptides was subtracted from  $\Delta G_{\text{unfold, H}_2\text{O}}$  of GCN4 peptides. The "residual energy" should represent the interhelical interaction in the GCN4-based coiled coils. The residual energies were then fitted to  $\log P_{\text{ow}}$ ,  $E_s$ , MR, [L, B1, B5], and side chain volume respectively. However, data related to Phe and Pff were not included in the fitting. Unlike other amino acids with aliphatic side chains, Phe and Pff bear aromatic and fluoro groups, complicating the hydrophobicity and interface packing. Also, GCN4-Pff folded poorly, making accurate determination of  $\Delta G_{\text{unfold, H}_2\text{O}}$  difficult. Furthermore, preliminary results showed that the exclusion of Phe and Pff in the process of fitting yielded higher correlations between the parameters and energies. Therefore, only residues with aliphatic side chain were included in the analysis. The graph, equation, and R value of each regression are shown in Figure 2-19 and Table 2-9. The regression between the residual energy and [L, B1, B5] can not be shown graphically because it is in a 4-dimensional space.

The higher the R value, the higher the correlation, the greater the contribution of the

**Table 2-9.** Equations and R values of the Regressions Between Residual Energy<sup>a</sup>, log P<sub>ow</sub>, E<sub>s</sub>, MR, [L, B1, B5], and Side Chain Volume

Parameter	Equation	R value
log P <sub>ow</sub>	residual energy = 0.5317 log P <sub>ow</sub> - 1.9038	0.57668
E <sub>s</sub>	residual energy = -0.6265 E <sub>s</sub> - 4.0184	0.62627
L, B1, B5	residual energy = -0.2612 L - 0.4911 B1 + 0.4772 B5 - 2.5013	0.56036
MR	residual energy = 0.4094 MR - 3.5328	0.47187
Side chain volume	residual energy = 0.0113 side chain volume - 3.5881	0.48617

<sup>a</sup>residual energy =  $\Delta G_{\text{unfold, H}_2\text{O}}(\text{GCN4-Xaa}) - \Delta G_{\text{unfold, H}_2\text{O}}(\text{IaLd-Xaa})$



**Figure 2-19.** Regressions between the residual energy and log P<sub>ow</sub>, E<sub>s</sub>, MR, and side chain volume. (a) Residual energy - log P<sub>ow</sub>. (b) Residual energy - E<sub>s</sub>. (c) Residual energy - MR. (d) Residual energy - side chain volume.

parameter to interhelical interaction for coiled coil stability. The R values for the various parameters followed the trend:  $E_s > \log P_{ow} \geq [L, B1, B5] > \text{side chain volume} \geq \text{MR}$ . As described above, MR and side chain volume provide information on the size of a molecule, whereas  $E_s$  and  $[L, B1, B5]$  consider both the size and shape of a molecule. Within these four parameters, the R values for  $E_s$  and  $[L, B1, B5]$  are higher than those of MR and side chain volume, meaning that  $E_s$  and  $[L, B1, B5]$  are more correlated to the residual energy than MR or side chain volume. This indicates that the size of the residue side chain in the coiled coil interface alone is insufficient to explain the interhelical interaction, and the shape of a molecule should also be taken into account. The R value of  $\log P_{ow}$  is slightly lower than that of  $E_s$ , meaning that  $\log P_{ow}$  is slightly less correlated to the residual energy than  $E_s$ . This indicates that hydrophobicity participates slightly lesser than the structural properties of the residue side chain to the interhelical interaction in a coiled coil. To build up the relationship between the interhelical interaction of a coiled coil, hydrophobicity, and the structural property of the residues in the coiled coil interface, the residual energy was fitted with  $\log P_{ow}$  and  $E_s$ .  $E_s$  was chosen because it is the parameter that correlates best to the residual energy among the four structural parameters. The regression yielded as follows:

$$\text{residual energy} = -0.4359 E_s + 0.2755 \log P_{ow} - 3.1961, R = 0.66723.$$

The coefficient of  $E_s$  is -0.6265 when the residual is fitted with  $E_s$  alone, and changes to -0.4359 when the residual energy is fitted with both  $E_s$  and  $\log P_{ow}$ . The latter is 69.9% of the former. For  $\log P_{ow}$ , the coefficient changes from 0.5317 to 0.2755, and the latter is 51.8% of the former. The percentage shows "to what extent is the relationship between the residual energy and  $E_s$  (or  $\log P_{ow}$ ) remained when more parameters are included in the fitting". For instance, 69.9% of  $E_s$  means that the dependence of the residual energy on  $E_s$  dropped to 69.9% of the original when  $\log P_{ow}$  was taken into the fitting. Therefore, the greater percentage of  $E_s$  (69.9%, in comparison to 51.8% of  $\log P_{ow}$ ) means that the dependence of the residual energy on  $E_s$  remains more than  $\log P_{ow}$ , indicating that  $E_s$  contributes more to the residual energy. Therefore, in the system discussed here (coiled coils with residues bearing aliphatic side chains in the interface), the structural properties of the residue in the coiled coil interface participate more than the hydrophobicity to the interhelical interaction.

## Conclusion

Hydrophobicity of the residues at the  $d$  positions in a coiled coil has long been recognized as the determining factor for coiled coil stability.<sup>10</sup> Although previous studies have shown that side chain structure also affects coiled coil stability, no systematic investigation has been performed to probe the relationship between side chain structure and coiled coil stability. In this study, unnatural amino acids with various side chain structures and properties were incorporated into coiled coils.

Examining coiled coil propensities, hydrophobicities, and side chain size and shape, conclusions can be drawn from coiled coil bearing substituent with aliphatic side chain.

First, side chain structural preference follows the trend  $\gamma$ -branched > linear >  $\beta$ -branched. Second, the size and shape of the residue side chain contribute slightly more than hydrophobicity to the coiled coil stability in coiled coils with aliphatic residues in the interface.

## Acknowledgement

This work was supported by the National Science Council (NSC-97-2113-M-002-019-MY2 and NSC-98-2119-M-002-025) and National Taiwan University. I would like to

thank Professor R. P. Cheng for discussing and revising this thesis, and the members in the Cheng lab, especially Hsien-Po Chiu for contributing parts of the data, and Cheng-Hsun Wu for helpful discussions. I would also like to thank Yi-Che Li and Dr. Rofeamor Obena from Professor Yu-Ju Chen's lab at Academia Sinica for MALDI-TOF MS.

## **Experimental Section**

### *General Materials and Methods*

Reagents and solvents were used without further purification. All of the chemical reagents except those indicated otherwise were purchased from Sigma Aldrich.

Diisopropylethylamine (DIEA), piperidine, trifluoroacetic acid (TFA), and acetic anhydride were from Acros. Guanidine hydrochloride was from Fluka.

Dimethylformamide (DMF) was from Mallinckrodt. Methanol, acetonitrile, and

Fmoc-Arg(Pbf)-Wang resin (100-200 mesh) were from Merck. Amino acids,

1-hydroxybenzotriazole (HOBt), and *O*-1*H*-benzotriazol-1-yl-

1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were from NovaBiochem.

Analytical reverse phase (RP)-HPLC was performed on an Agilent 1200 series

chromatography system using a Vydac C18 column (4.6 mm diameter, 250 mm length).

Preparative RP-HPLC was performed on a Waters Breeze chromatography system

using Vydac C4 and C18 column (22 mm diameter, 250 mm length). Mass

spectrometry of the peptides was performed on a matrix-assisted laser desorption

ionization time-of-flight (MALDI-TOF) spectrometer (Bruker BIFLEX) using

$\alpha$ -cyano-4-hydroxycinnamic acid as the matrix. Determination of peptide concentration

was performed on a UV-Vis spectrophotometer (Jasco V-650). Circular dichroism (CD)

spectra were collected on a Jasco J-815 spectrometer using 1 mm pathlength cells.

**Fmoc-Allo-Ile-OH.** H-Allo-Ile-OH (0.08 g, 0.61 mmol, 1 equivalent) was dissolved in

$\text{Na}_2\text{CO}_3$  solution (4 mL). This solution was kept at pH 10 and over an ice-water bath.

Fmoc-OSu (0.2160 g, 0.64 mmol, 1.05 equivalents) was dissolved in tetrahydrofuran

(1.6 mL). The Fmoc-OSu/THF solution was added dropwise into the amino acid

solution. The mixture was stirred for 3 hours. The pH was kept around 10 throughout

the course of the reaction. For work up, THF was first removed under reduced pressure.

The remaining basic aqueous solution was extracted with diethyl ether (3x10 mL). The

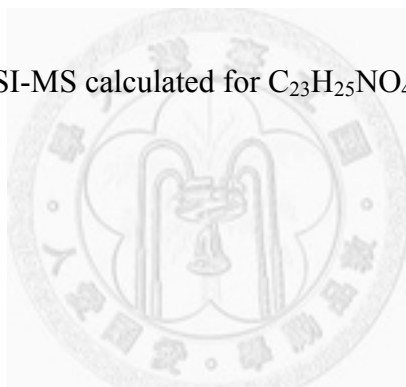
aqueous fraction was then acidified with 1 M HCl until pH 2, followed by extraction

with ethyl acetate (3x10 mL). The organic fraction was dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated in vacuo. The crude was purified by flash chromatography (silica gel, 95% dichloromethane, 5% methanol) to yield a white solid (0.1953 g, 90.1% yield). TLC (silica gel, dichloromethane: methanol = 95: 5) R<sub>f</sub> = 0.30; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) (*J* in hertz) δ 7.75 (d, *J* = 7.6 Hz, 2H), 7.58 (d, *J* = 6.0 Hz, 2H), 7.38 (t, *J* = 7.6 Hz, 2H), 7.29 (t, *J* = 7.2 Hz, 2H), 5.19 (d, *J* = 9.2 Hz, 1H), 4.49 (dd, *J* = 9.2 Hz, 3.6 Hz, 1H), 4.41 (d, *J* = 7.2 Hz, 2H), 4.22 (t, *J* = 6.8 Hz, 1H), 2.02-1.98 (m, 1H), 1.44 (quin, *J* = 6.8 Hz, 1H), 1.24 (quin, *J* = 7.6 Hz, 1H), 0.95 (t, *J* = 7.2 Hz, 3H), 0.89 (d, *J* = 6.8 Hz, 3H). ESI-MS calculated for C<sub>21</sub>H<sub>23</sub>NO<sub>4</sub> [M+Na<sup>+</sup>]: 377.40, observed: 376.17.

**Fmoc-Cpa-OH.** H-Cpa-OH (0.0805 g, 0.51 mmol, 1 equivalent) was dissolved in Na<sub>2</sub>CO<sub>3</sub> solution (4 mL). This solution was kept at pH 10 and over an ice-water bath. Fmoc-OSu (0.1754 g, 0.52 mmol, 1.02 equivalents) was dissolved in tetrahydrofuran (1.5 mL). The Fmoc-OSu/THF solution was added dropwise into the amino acid solution. The mixture was stirred for 4 hours. The pH was kept at around 10 throughout the course of reaction. For work up, THF was first removed under reduced pressure. The remaining basic aqueous solution was extracted with diethyl ether (3x10 mL). The



aqueous fraction was then acidified with 1 M HCl until pH 2, followed by extraction with ethyl acetate (3x10 mL). The organic fraction was dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated in vacuo. The crude was purified by flash chromatography (silica gel, methanol:dichloromethane = 1:15) to yield white solid (0.1679 g, 86.9% yield). TLC (silica gel, methanol:dichloromethane = 1:15) R<sub>f</sub> = 0.29; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ 7.74 (d, *J* = 7.6 Hz, 2H), 7.58-7.56 (m, 2H), 7.38 (t, *J* = 7.2 Hz, 2H), 7.29 (t, *J* = 7.2 Hz, 2H), 5.15 (d, *J* = 8.4 Hz, 1H), 4.48-4.32 (m, 3H), 4.21 (t, *J* = 6.8 Hz, 1H), 1.94-1.42 (m, 10H), 1.12 (bs, 2H) ppm. ESI-MS calculated for C<sub>23</sub>H<sub>25</sub>NO<sub>4</sub> [M+Na<sup>+</sup>]: 402.44, observed: 402.21.



**Fmoc-Nle-OH.** H-Nle-OH (0.0810 g, 0.61 mmol, 1 equivalent) was dissolved in Na<sub>2</sub>CO<sub>3</sub> solution (4 mL). This solution was kept at pH 10 and over an ice-water bath. The Fmoc-OSu (0.2163 g, 0.64 mmol, 1.05 equivalents) was dissolved in tetrahydrofuran (3.2 mL). Fmoc-OSu/THF solution was added dropwise into the amino acid solution. The mixture was stirred for 3 hours. The pH was kept at around 10 throughout the course of the reaction. For work up, THF was first removed under reduced pressure. The remaining basic aqueous solution was extracted with diethyl

ether (3x10 mL). The aqueous fraction was then acidified with 1 M HCl until pH 2, followed by extraction with ethyl acetate (3x10 mL). The organic fraction was dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated in vacuo. The crude was purified by flash chromatography (silica gel, 95% dichloromethane, 5% methanol) to yield a white solid (0.1693 g, 78.5% yield). TLC (silica gel, dichloromethane: methanol = 95: 5) R<sub>f</sub> = 0.30; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) (*J* in hertz) δ 7.74 (d, *J* = 7.6 Hz, 2H), 7.58 (dd, *J* = 6.8 Hz, 2.4 Hz, 2H), 7.38 (t, *J* = 7.6 Hz, 2H), 7.29 (t, *J* = 7.2 Hz, 2H), 5.21 (d, *J* = 8 Hz, 1H), 4.42-4.37 (m, 3H), 4.21 (t, *J* = 6.8 Hz, 1H), 1.90-1.86 (m, 1H), 1.71-1.67 (m, 1H), 1.35 (m, 4H), 0.90 (t, *J* = 6 Hz, 3H). ESI-MS calculated for C<sub>21</sub>H<sub>23</sub>NO<sub>4</sub> [MH<sup>+</sup>]: 354.41, observed: 354.21.

### *Peptide Synthesis*

Fmoc-PAL-PEG-PS (0.05 mmole) or Fmoc-Arg(Pbf)-Wang resin (100-200 mesh) was swollen in *N,N*-dimethylformamide (DMF, 5 mL) for 30 minutes. The resin was then deprotected by 25% piperidine/DMF (3 mL, 5x8 min) and rinsed with DMF (3 mL, 7x1 min). A mixture of 3 equivalents of the appropriately protected Fmoc amino acid, HOBt and HBTU was dissolved in DMF (1 mL). Diisopropylethylamine (DIEA, 8 equivalents)

was then added to the solution and mixed thoroughly. The solution was then applied to the resin. The vial that contained the solution was rinsed with DMF (2x1 mL) and added to the reaction. The first coupling was carried out for 8 hours. Other residues were coupled for 1.5 hours (IaLd peptides) or 3 hours (GCN4 peptides). After each coupling, the resin was washed with DMF (7x1 min). For capping with acetic anhydride ( $\text{Ac}_2\text{O}$ ), a solution of  $\text{Ac}_2\text{O}$  (20 equivalents), DIEA (20 equivalents), DMF (1 mL) was added to the resin. The reaction was shaken for 3 hours. The resin was subsequently washed with DMF (3 mL, 7x1 min) and methanol (15 mL), and was lyophilized. The peptide was deprotected and cleaved off the resin by treating the resin with 95:5 trifluoroacetic acid (950  $\mu\text{L}$ )/triisopropylsilane (50  $\mu\text{L}$ ) and shaken for 3 hours. The solution was then filtered through glass wool and the resin was washed with TFA (3x1 mL). The combined filtrate was evaporated by a gentle stream of  $\text{N}_2$ . The resulting material was washed with hexanes (1 mL), dissolved in water, and lyophilized. The peptide (1 mg/mL aqueous solution) was analyzed using analytical RP-HPLC on a C18 column with flow rate 1 mL/min, temperature  $25^\circ\text{C}$ , linear 1 %/min gradient from 100% A to 0% A (solvent A: 99.9% water, 0.1% TFA; solvent B: 90% acetonitrile, 10% water, 0.1% TFA). Appropriate linear solvent A/solvent B gradient was used for purification

on preparative RP-HPLC on C4 and C18 columns. The identity of the peptide was confirmed by MALDI-TOF.

#### **GCN4-Abu**

(Ac-RMKQ LEDKVEE AbuLSKNYH LENEVAR LKKLVEG R-OH)

The peptide was synthesized using 79.5 mg (0.05 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 216.0 mg resin (44.8% yield). The cleavage yielded 84.1 mg of crude peptide (96.2% yield, 78.4% yield if included trifluoroacetic acid). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG28\_40 and PLG33\_43, respectively. GCN4-Abu was purified to 95.2% purity. Retention time on analytical RP-HPLC was 42.7 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{174}H_{294}N_{52}O_{54}S$   $[MH^+]$ : 4009.17; observed: 4008.9.

#### **GCN4-Allo Ile**

(Ac-RMKQ LEDKVEE Allo IleLSKNYH LENEVAR LKKLVEG R-OH)

The peptide was synthesized using 159.5 mg (0.1 mmol) of Fmoc-Arg(Pbf)-Wang resin

(100-200 mesh). The synthesis gave 512.9 mg of resin (57.7% yield). The cleavage yielded 262.8 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG29\_39 and PLG34\_43, respectively. GCN4-Allo Ile was purified to 96.5% purity. Retention time on analytical RP-HPLC was 44.0 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{176}H_{298}N_{52}O_{54}S$   $[MH^+]$ : 4037.20; observed: 4036.5.

#### **GCN4-Asn**

(Ac-RMKQ LEDKVEE NLSKNYH LENEVAR LKKLVGE R-OH)

The peptide was synthesized using 160.1 mg (0.1 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 530.7 mg of resin (58.2% yield). The cleavage yielded 403.3 mg of crude peptide (quantitative yield). The peptide is being purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG24\_34 and PLG29\_39, respectively. GCN4-Asn was purified to 95.6% purity. Retention time on analytical RP-HPLC was 40.4 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{174}H_{293}N_{53}O_{55}S$   $[MH^+]$ : 4038.16;

observed: 4037.6.

#### **GCN4-Asp**

(Ac-RMKQ LEDKVEE DLSKNYH LENEVAR LKKLVGE R-OH)

The peptide was synthesized using 79.8 mg (0.05 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 238.2 mg of resin (51.3% yield). The cleavage yielded 135.1 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG26\_36 and PLG31\_41, respectively. GCN4-Asp was purified to 95.9% purity. Retention time on analytical RP-HPLC was 41.4 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for C<sub>174</sub>H<sub>292</sub>N<sub>52</sub>O<sub>56</sub>S [MH<sup>+</sup>]: 4039.14; observed: 4038.5.

#### **GCN4-Cha**

(Ac-RMKQ LEDKVEE ChaLSKNYH LENEVAR LKKLVGE R-OH)

The peptide was synthesized using 79.7 mg (0.05 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 231.4 mg of resin (49.2% yield). The cleavage

yielded 161 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG30\_40 and PLG35\_45, respectively. GCN4-Cha was purified to 96.8% purity. Retention time on analytical RP-HPLC was 45.5 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{179}H_{302}N_{52}O_{54}S$   $[MH^+]$ : 4077.23; observed: 4077.1.

#### **GCN4-Cpa**

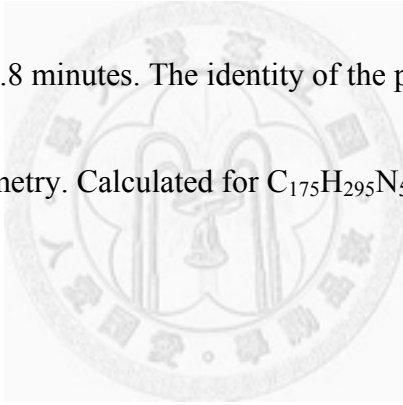
(Ac-RMKQ LEDKVEE **Cpa**LSKNYH LENEVAR LKKLVGE R-OH)

The peptide was synthesized using 159.5 mg (0.1 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 510.8 mg of resin (57.1% yield). The cleavage yielded 321.8 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG30\_40 and PLG35\_44, respectively. GCN4-Cpa was purified to 95.7% purity. Retention time on analytical RP-HPLC was 44.6 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{178}H_{300}N_{52}O_{54}S$   $[MH^+]$ : 4063.21; observed: 4062.8.

### GCN4-Gln

(Ac-RMKQ LEDKVEE QLSKNYH LENEVAR LKCLVGE R-OH)

The peptide was synthesized using 160.5 mg (0.1 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 535.3 mg of resin (58.8% yield). The cleavage yielded 404.5 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG26\_36 and PLG31\_40, respectively. GCN4-Gln was purified to 98.2% purity. Retention time on analytical RP-HPLC was 40.8 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{175}H_{295}N_{53}O_{55}S$   $[MH^+]$ : 4052.17  
observed: 4051.4.



### GCN4-Glu

(Ac-RMKQ LEDKVEE ELSKNYH LENEVAR LKCLVGE R-OH)

The peptide was synthesized using 81.0 mg (0.05 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 274.5 mg of resin (61.3% yield). The cleavage yielded 166.9 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG26\_36 and



PLG31\_41, respectively. GCN4-Glu was purified to 95.8% purity. Retention time on analytical RP-HPLC was 41.4 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{175}H_{294}N_{52}O_{56}S$   $[MH^+]$ : 4053.16; observed: 4052.3.

### **GCN4-Ile**

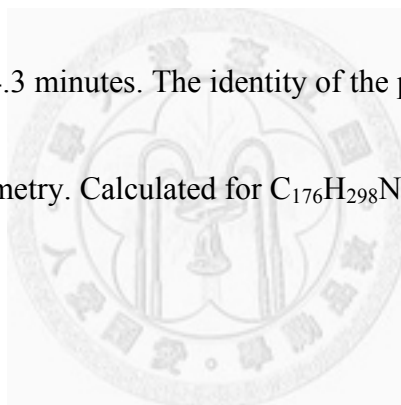
(Ac-RMKQ LEDKVEE ILSKNYH LENEVAR LKKLVGE R-OH)

The peptide was synthesized using 79.5 mg (0.05 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 223.4 mg of resin (47.0% yield). The cleavage yielded 133.4 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG29\_39 and PLG33\_43, respectively. GCN4-Ile was purified to 95.6% purity. Retention time on analytical RP-HPLC was 43.9 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{176}H_{298}N_{52}O_{54}S$   $[MH^+]$ : 4037.20; observed: 4036.9.

### GCN4-Leu

(Ac-RMKQ LEDKVEE LLSKNYH LENEVAR LKKLVGE R-OH)

The peptide was synthesized using 79.7 mg (0.05 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 255.6 mg of resin (54.8% yield). The cleavage yielded 187.7 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG29\_39 and PLG33\_43, respectively. GCN4-Leu was purified to 95.4% purity. Retention time on analytical RP-HPLC was 44.3 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for C<sub>176</sub>H<sub>298</sub>N<sub>52</sub>O<sub>54</sub>S [MH<sup>+</sup>]: 4040.64; observed: 4039.3.



### GCN4-Nle

(Ac-RMKQ LEDKVEE NleLSKNYH LENEVAR LKKLVGE R-OH)

The peptide was synthesized using 160.2 mg (0.1 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 517.5 mg of resin (58.4% yield). The cleavage yielded 162.6 mg of crude peptide (56.5% yield, with TFA counter ion). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients

PLG29\_39 and PLG34\_43, respectively. GCN4-Nle was purified to 95.2% purity.

Retention time on analytical RP-HPLC was 43.8 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{176}H_{298}N_{52}O_{54}S$   $[MH^+]$ : 4037.20; observed: 4036.5.

#### **GCN4-Nva**

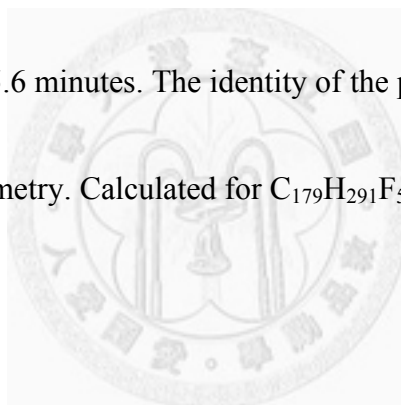
(Ac-RMKQ LEDKVEE NvaLSKNYH LENEVAR LKKLVGE R-OH)

The peptide was synthesized using 79.9 mg (0.05 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 234.6 mg of resin (50.7% yield). The cleavage yielded 152.5 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG28\_38 and PLG33\_43, respectively. GCN4-Nva was purified to 96.0% purity. Retention time on analytical RP-HPLC was 43.3 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{175}H_{296}N_{52}O_{54}S$   $[MH^+]$ : 4023.18; observed: 4022.6.

### GCN4-Pff

(Ac-RMKQ LEDKVEE PffLSKNYH LENEVAR LKKLVGE R-OH)

The peptide was synthesized using 160.1 mg (0.1 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 520.5 mg of resin (57.7% yield). The cleavage yielded 460.7 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG30\_40 and PLG35\_45, respectively. GCN4-Pff was purified to 96.5% purity. Retention time on analytical RP-HPLC was 45.6 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{179}H_{291}F_5N_{52}O_{54}S$   $[MH^+]$ : 4161.14; observed: 4160.8.



### GCN4-Phe

(Ac-RMKQ LEDKVEE FLSKNYH LENEVAR LKKLVGE R-OH)

The peptide was synthesized using 79.4 mg (0.05 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 276.1 mg of resin (63.9% yield). The cleavage yielded 161.9 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG28\_38 and

PLG33\_43, respectively. GCN4-Phe was purified to 97.3% purity. Retention time on analytical RP-HPLC was 44.5 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{179}H_{296}N_{52}O_{54}S$   $[MH^+]$ : 4073.66; observed: 4073.1.

### **GCN4-Tba**

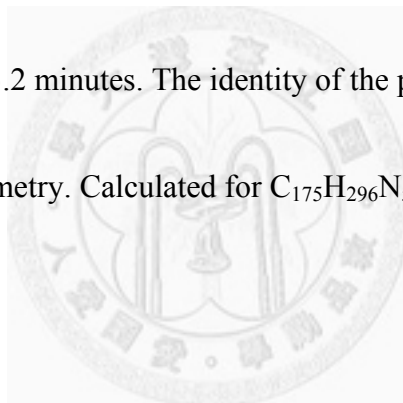
(Ac-RMKQ LEDKVEE TbaLSKNYH LENEVAR LKKLVGE R-OH)

The peptide was synthesized using 160.2 mg (0.1 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 511.5 mg of resin (57.3% yield). The cleavage yielded 306.7 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG30\_40 and PLG34\_44, respectively. GCN4-Tba was purified to 96.5% purity. Retention time on analytical RP-HPLC was 43.0 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{177}H_{300}N_{52}O_{54}S$   $[MH^+]$ : 4051.21; observed: 4051.2.

### **GCN4-Val**

(Ac-RMKQ LEDKVEE VLSKNYH LENEVAR LKKLVGE R-OH)

The peptide was synthesized using 79.9 mg (0.05 mmol) of Fmoc-Arg(Pbf)-Wang resin (100-200 mesh). The synthesis gave 211.3 mg of resin (43.0% yield). The cleavage yielded 161.0 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradients PLG28\_38 and PLG33\_43, respectively. GCN4-Val was purified to 97.5% purity. Retention time on analytical RP-HPLC was 43.2 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{175}H_{296}N_{52}O_{54}S$   $[MH^+]$ : 4023.18; observed: 4023.1.



### **IaLd-half**

(Fmoc- KIQALEKKNEALEKKIAAL-NH<sub>2</sub>)

The peptide was synthesized using 715.2 mg (0.12 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 1106.9 mg of resin (57.3% yield). Retention time on analytical RP-HPLC was 43.3 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{111}H_{183}N_{27}O_{29}$   $[MH^+]$ : 2360.81;

observed: 2358.5. IaLd-half was further divided into 3 portions for synthesis of

IaLd-Abu, IaLd-Phe and IaLf-Pff.

### **IaLd-Abu**

(Ac-YGG EIEALEKK IAALEAbuK IQALEKK NEQLEKK IAAL-NH<sub>2</sub>)

The peptide was synthesized using approximately 238.4 mg (0.04 mmol) of

Fmoc-PAL-PEG-PS resin. The synthesis gave 376.3 mg of resin (approximately 62.6%

yield). The cleavage yielded 116.4 mg of crude peptide (quantitative yield). The peptide

was purified by preparative RP-HPLC using C4 and C18 column with linear gradient

PLG32\_43 and PLG37\_47, respectively. IaLd-Abu was purified to 98.9% purity.

Retention time on analytical RP-HPLC was 47.9 minutes. The identity of the peptide

was confirmed by MALDI-TOF mass spectrometry. Calculated for C<sub>180</sub>H<sub>308</sub>N<sub>46</sub>O<sub>54</sub>

[MH<sup>+</sup>]: 3979.29; observed: 3978.6.

### **IaLd-Cha**

(Ac-YGG EIEALEKK IAALEChaK IQALEKK NEQLEKK IAAL-NH<sub>2</sub>)

The peptide was synthesized using 501.0 mg (0.1 mmol) of Fmoc-PAL-PEG-PS resin.

The synthesis gave 796.4 mg of resin (approximately 54.3% yield). Half of the synthesis product was cleaved, and the cleavage yielded 133.8 mg of crude peptide

(quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and

C18 column with linear gradient PLG34\_44 and PLG39\_49, respectively. IaLd-Cha

was purified to 96.1% purity. Retention time on analytical RP-HPLC was 48.0 minutes.

The identity of the peptide was confirmed by MALDI-TOF mass spectrometry.

Calculated for C<sub>185</sub>H<sub>316</sub>N<sub>46</sub>O<sub>54</sub> [MH<sup>+</sup>]: 4047.35; observed: 4047.2.

### **IaLd-Cpa**

(Ac-YGG EIEALEKK IAALECpaK IQALEKK NEQLEKK IAAL-NH<sub>2</sub>)

The peptide was synthesized using 503.2 mg (0.1 mmol) of Fmoc-PAL-PEG-PS resin.

The synthesis gave 805.4 mg of resin (approximately 55.7% yield). Half of the total

synthesis product was cleaved, and the cleavage yielded 108.6 mg of crude peptide

(80.6 % yield). The peptide was purified by preparative RP-HPLC using C4 and C18



column with linear gradient PLG34\_44 and PLG39\_49, respectively. IaLd-Cpa was purified to 96.5% purity. Retention time on analytical RP-HPLC was 48.6 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{184}H_{314}N_{46}O_{54}$   $[MH^+]$ : 4033.33; observed: 4032.9.

### **IaLd-Nle**

(Ac-YGG EIEALEKK IAALENleK IQALEKK NEQLEKK IAAL-NH<sub>2</sub>)

The peptide was synthesized using 502.7 mg (0.1 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 804.8 mg of resin (approximately 55.9% yield). Half of the total synthesis product was cleaved, and the cleavage yielded 79.1 mg of crude peptide (58.8% yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradient PLG34\_44 and PLG39\_49, respectively. IaLd-Nle was purified to 98.0% purity. Retention time on analytical RP-HPLC was 48.4 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{182}H_{312}N_{46}O_{54}$   $[MH^+]$ : 4009.71; observed: 4008.8.

### **IaLd-Nva**

(Ac-YGG EIEALEKK IAALENvaK IQALEKK NEQLEKK IAAL-NH<sub>2</sub>)

The peptide was synthesized using 502.9 mg (0.1 mmol) of Fmoc-PAL-PEG-PS resin.

The synthesis gave 808.4 mg of resin (approximately 56.7% yield). Half of the synthesis product was cleaved, and the cleavage yielded 135.8 mg of crude peptide

(quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and

C18 column with linear gradient PLG34\_44 and PLG39\_49, respectively. IaLd-Nva

was purified to 96.3% purity. Retention time on analytical RP-HPLC was 47.5 minutes.

The identity of the peptide was confirmed by MALDI-TOF mass spectrometry.

Calculated for C<sub>181</sub>H<sub>310</sub>N<sub>46</sub>O<sub>54</sub> [MH<sup>+</sup>]: 3995.68; observed: 3994.7.

### **IaLd-Pff**

(Ac-YGG EIEALEKK IAALEPffK IQALEKK NEQLEKK IAAL-NH<sub>2</sub>)

The peptide was synthesized using approximately 238.4 mg (0.04 mmol) of

Fmoc-PAL-PEG-PS resin. The synthesis gave 394.9 mg of resin (approximately 69.2% yield). The cleavage yielded 137.0 mg of crude peptide (quantitative yield). The peptide

was purified by preparative RP-HPLC using C4 and C18 column with linear gradient

PLG32\_43 and PLG37\_47, respectively. IaLd-Pff was purified to 98.9% purity.

Retention time on analytical RP-HPLC was 47.8 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{185}H_{305}F_5N_{46}O_{54}$   $[MH^+]$ : 4133.68; observed: 4131.4.

### **IaLd-Phe**

(Ac-YGG EIEALEKK IAALEPheK IQALEKK NEQLEKK IAAL-NH<sub>2</sub>)

The peptide was synthesized using approximately 238.4 mg (0.04 mmol) of Fmoc-PAL-PEG-PS resin. The synthesis gave 335.7 mg of resin (approximately 43.7% yield). The cleavage yielded 103.3 mg of crude peptide (quantitative yield). The peptide was purified by preparative RP-HPLC using C4 and C18 column with linear gradient PLG32\_44 and PLG37\_47 respectively. IaLd-Phe was purified to 97.6% purity.

Retention time on analytical RP-HPLC was 48.0 minutes. The identity of the peptide was confirmed by MALDI-TOF mass spectrometry. Calculated for  $C_{185}H_{310}N_{46}O_{54}$   $[MH^+]$ : 4043.73; observed: 4042.4.

### *UV-Visible Spectroscopy*

UV data were obtained using 1 mm pathlength cells. Peptide stock solutions were prepared at a concentration of 5 mM. 3  $\mu$ L, 6  $\mu$ L, 9  $\mu$ L, 12  $\mu$ L and 15  $\mu$ L of the peptide solutions were added to 400  $\mu$ L of 6 M guanidinium chloride solution. At each concentration of peptide, tyrosine absorbance at 276 nm, 278 nm, 280 nm and 282 nm were measured. Linear regression was performed to determine the precise concentrations of the stock solutions using known extinction coefficients.<sup>58</sup> Data with correlation coefficients (R) less than 0.99 was discarded.

### *Circular Dichroism Spectroscopy*

CD data were obtained using 1 mm pathlength cells. CD data was acquired under two sets of conditions. CD spectrum of GCN4 peptides were acquired at 30  $\mu$ M peptide in 50 mM phosphate and 150 mM NaCl at pH 7 and 4 °C.<sup>18</sup> CD spectrum of IaLd peptides were acquired at 20  $\mu$ M peptide, 10 mM 3-(N-morpholino)-propanesulfonic acid (MOPS), and 0 M to 6 M (at 0.1 M interval) guanidinium chloride at pH 7.5 and 25 °C.<sup>23</sup> Each CD value was determined at least three times to obtain experimental standard deviations. The data were analyzed using Microsoft Excel 2007 and

Kaleidagraph (Synergy Software, CA). Data were expressed in terms of mean residue ellipticity ( $\text{deg cm}^2 \text{dmol}^{-1}$ ).

### *Guanidinium Denaturation*

Guanidinium denaturation was performed under two sets of conditions. For GCN4 peptides, guanidinium denaturation was performed at 30  $\mu\text{M}$  peptide in 50 mM phosphate, 150 mM NaCl, and 0 M to 6 M (at 0.1 M interval) guanidinium chloride at pH 7 and 4  $^\circ\text{C}$ .<sup>18</sup> For IaLd peptides, guanidinium denaturation was performed at 20  $\mu\text{M}$  peptide, 10 mM 3-(N-morpholino)-propanesulfonic acid (MOPS), and 0 M to 6 M (at 0.1 M interval) guanidinium chloride at pH 7.5 and 25  $^\circ\text{C}$ .<sup>23</sup> Denaturation process was monitored by CD using 1 mm pathlength cell. Two solutions, one with peptide, buffer (as described above), 0 M guanidinium chloride and the other one with peptide, buffer, 6 M guanidinium chloride, were prepared. Denaturation process began at 0 M, 2 M, and 6 M of guanidinium chloride. Suitable amounts of the two solutions were added to the cell to adjust the concentration of guanidinium chloride. Each CD value was determined at least three times to obtain experimental standard deviations. The data were analyzed using Kaleidagraph (Synergy Software, CA) and Microsoft Office Excel 2007.

*Derivation of  $\Delta G_{\text{unfold, H}_2\text{O}}$*

$\Delta G_{\text{unfold, H}_2\text{O}}$  was determined using following equations:

dimer  $\rightleftharpoons$  2 unfolded monomer

$$[\text{unfolded monomer}] + 2 [\text{dimer}] = [\text{Peptide}]_{\text{total}}$$

$$K_{\text{eq}} = \frac{[\text{unfolded monomer}]^2}{[\text{dimer}]}$$

$$K_{\text{eq}} = \frac{\{[\text{Peptide}]_{\text{total}} \left( \frac{\text{Signal at } [C] - \text{fold}_{100,[C]}}{\text{unfold}_{100,[C]} - \text{fold}_{100,[C]}} \right)\}^2}{\{[\text{Peptide}]_{\text{total}} \left( \frac{\text{unfold}_{100,[C]} - \text{Signal at } [C]}{\text{unfold}_{100,[C]} - \text{fold}_{100,[C]}} \right)\} / 2}$$

$$\Delta G_{\text{unfold}} = -RT \ln K_{\text{eq}},$$

in which [C] is the concentration of guanidinium chloride,  $\text{fold}_{100}$  is the expected CD signal of the peptide totally folded at guanidinium chloride concentration [C], and  $\text{unfold}_{100}$  is the expected CD signal of the peptide totally unfolded at guanidinium chloride concentration [C].  $\text{fold}_{100}$  and  $\text{unfold}_{100}$  were obtained from linear regression fits derived from the CD data. A graph of  $\Delta G_{\text{unfold}}$  plotted against the concentration of guanidinium chloride was plotted. Data points near the midpoint transition were chosen and used for extrapolation. The value at y-intercept is the  $\Delta G_{\text{unfold, H}_2\text{O}}$ .

### *Measurement of Hydrophobicity by Thin Layer Chromatography*

The TLC were performed with *n*-butanol - 0.05 M ammonium acetate (2:1) and *n*-butanol - 50% acetic acid (2:1) on cellulose plates. The alcohol and aqueous solution were shaken for 5 hours. The saturated alcoholic layer was separated by centrifugation and used as the mobile phase. The spots were detected by ninhydrin. The TLC of each amino acid were repeated at least three time to obtain average  $R_f$  values and experimental standard deviations. Average  $R_f$  values were converted into calculated log  $P_{ow}$  values following literature procedures.<sup>50, 51</sup>

### *Calculation of Side Chain Volume*

Calculation was performed on Discovery Studio 2.1 module (Accelrys, CA). Built-in molecules were used. The molecule was minimized using CFF forcefield. Dielectric constant was set as 80. The surface area and volume were probed with solvent molecule of 1.4 Å radius.

## References

1. Wolf, E.; Kim, P. S.; Berger, B. MultiCoil: A program for predicting two- and three-stranded coiled coils. *Protein Sci.* **1997**, *6*, 1179-1189.
2. O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **1991**, *254*, 539-544.
3. Pauling, L.; Corey, R. B. Compound helical configurations of polypeptide chains: structure of proteins of the alpha-keratin type. *Nature* **1953**, *171*, 59-61.
4. Sodek, J.; Hodges, R. S.; Smillie, L. B.; Jurasek, L. Amino-acid sequence of rabbit skeletal tropomyosin and its coiled-coil structure. *Proc. Natl. Acad. Sci. U. S. A.* **1972**, *69*, 3800-3804.
5. Wilson, I. A.; Skehel, J. J.; Wiley, D. C. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature* **1981**, *289*, 366-373.
6. Cohen, C.; Parry, D. A.  $\alpha$ -helical coiled coils: more facts and better predictions. *Science* **1994**, *263*, 488-489.
7. Newman, J.; Keating, A. Comprehensive identification of human bZIP interactions with coiled-coil arrays. *Science* **2003**, *300*, 2097-2198.
8. Vinson, C.; Myakishev, M.; Acharya, A.; Mir, A.; Moll, J.; Bonovich, M. Classification of human B-ZIP proteins based on dimerization properties. *Mol. Cell. Biol.* **2002**, *22*, 6321-6356.
9. Ellenberger, T.; Brandl, C.; Struhl, K.; Harrison, S. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. *Cell* **1992**, *71*, 1223-1260.
10. Mason, J. M.; Arndt, K. M. Coiled coil domains: Stability, specificity, and biological implications. *ChemBioChem* **2004**, *5*, 170-176.
11. Woolfson, D.; Bartlett, G.; Bruning, M.; Thomson, A. New currency for old rope: from coiled-coil assemblies to  $\alpha$ -helical barrels. *Curr. Opin. Struct. Biol.* **2012**. doi:10.1016/j.sbi.2012.03.002
12. Akey, D.; Malashkevich, V.; Kim, P. Buried polar residues in coiled-coil interfaces. *Biochemistry* **2001**, *40*, 6352-6412.
13. O'Shea, E.; Lumb, K.; Kim, P. Peptide 'Velcro': design of a heterodimeric coiled coil. *Curr. Biol.* **1993**, *3*, 658-725.
14. Lumb, K. A buried polar interaction imparts structural uniqueness in a designed

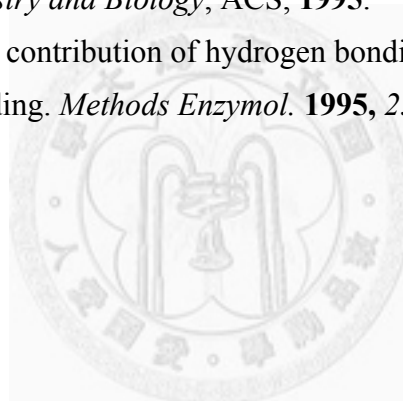


- heterodimeric coiled coil. *Biochemistry* **1995**, *34*, 8642-8648.
15. Wagschal, K.; Tripet, B.; Lavigne, P.; Mant, C.; Hodges, R. The role of position a in determining the stability and oligomerization state of alpha-helical coiled coils: 20 amino acid stability coefficients in the hydrophobic core of proteins. *Protein Sci.* **1999**, *8*, 2312-2341.
  16. Tripet, B.; Wagschal, K.; Lavigne, P.; Mant, C.; Hodges, R. Effects of side-chain characteristics on stability and oligomerization state of a de novo-designed model coiled-coil: 20 amino acid substitutions in position "d". *J. Mol. Biol.* **2000**, *300*, 377-779.
  17. Gonzalez, L., Jr.; Woolfson, D. N.; Alber, T. Buried polar residues and structural specificity in the GCN4 leucine zipper. *Nat. Struct. Biol.* **1996**, *3*, 1011-1018.
  18. Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **1993**, *262*, 1401-1407.
  19. Oakley, M.; Hollenbeck, J. The design of antiparallel coiled coils. *Curr. Opin. Struct. Biol.* **2001**, *11*, 450-457.
  20. McClain, D.; Woods, H.; Oakley, M. Design and characterization of a heterodimeric coiled coil that forms exclusively with an antiparallel relative helix orientation. *J. Am. Chem. Soc.* **2001**, *123*, 3151-3153.
  21. Grigoryan, G.; Keating, A. Structural specificity in coiled-coil interactions. *Curr. Opin. Struct. Biol.* **2008**, *18*, 477-560.
  22. Crick, F. H. C. The packing of helices: simple coiled-coils. *Acta Crystallogr.* **1953**, *6*, 689-1386.
  23. O'Neil, K.; DeGrado, W. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science* **1990**, *250*, 646-697.
  24. Van Deventer, J.; Fisk, J.; Tirrell, D. Homoioleucine: a translationally active leucine surrogate of expanded hydrophobic surface area. *ChemBioChem* **2011**, *12*, 700-702.
  25. Yoder, N.; Yuksel, D.; Dafik, L.; Kumar, K. Bioorthogonal noncovalent chemistry: Fluorous phases in chemical biology. *Curr. Opin. Chem. Biol.* **2006**, *10*, 576-659.
  26. Eriksson, A.; Baase, W.; Zhang, X.; Heinz, D.; Blaber, M.; Baldwin, E.; Matthews, B. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* **1992**, *255*, 178-261.

27. Moitra, J.; Szilak, L.; Krylov, D.; Vinson, C. Leucine is the most stabilizing aliphatic amino acid in the d position of a dimeric leucine zipper coiled coil. *Biochemistry* **1997**, *36*, 12567-12573.
28. Hadley, E.; Testa, O.; Woolfson, D.; Gellman, S. Preferred side-chain constellations at antiparallel coiled-coil interfaces. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 530-535.
29. Steinkruger, J.; Bartlett, G.; Hadley, E.; Fay, L.; Woolfson, D.; Gellman, S. The d'-d--d' vertical triad is less discriminating than the a'--a--a' vertical triad in the antiparallel coiled-coil dimer motif. *J. Am. Chem. Soc.* **2012**, *134*, 2626-2659.
30. Buer, B.; Meagher, J.; Stuckey, J.; Marsh, E. Structural basis for the enhanced stability of highly fluorinated proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 4810-4815.
31. Ryan, S.; Kennan, A. Variable stability heterodimeric coiled-coils from manipulation of electrostatic interface residue chain length. *J. Am. Chem. Soc.* **2007**, *129*, 10255-10315.
32. Matousek, W. M.; Ciani, B.; Fitch, C. A.; Garcia-Moreno, B.; Kammerer, R. A.; Alexandrescu, A. T. Electrostatic contributions to the stability of the GCN4 leucine zipper structure. *J. Mol. Biol.* **2007**, *374*, 206-219.
33. Chou, P. Y.; Fasman, G. D. Conformational parameters for amino acids in helical,  $\beta$ -sheet, and random coil regions calculated from proteins. *Biochemistry* **1974**, *13*, 211-222.
34. Lifson, S.; Roig, A. On the Theory of Helix - Coil Transition in Polypeptides. *J. Chem. Phys.* **1961**, *34*, 1963-1974.
35. Southhall, N. T.; Dill, K. A.; Haymet, A. D. A View of the Hydrophobic Effect. *J. Phys. Chem. B* **2002**, *106*, 521-533.
36. Edelhoch, H. Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* **1967**, *6*, 1948-1954.
37. Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein Sci.* **1995**, *4*, 2411-2423.
38. Chang, C.; Wu, C.; Yang, J. Circular dichroic analysis of protein conformation: inclusion of the  $\beta$ -turns. *Anal. Biochem.* **1978**, *91*, 13-44.
39. Pace, C. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **1986**, *131*, 266-346.

40. Myers, J.; Nick, C. Denaturant  $m$  values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* **1995**.
41. Courtenay, E.; Capp, M.; Saecker, R.; Record, M. Thermodynamic analysis of interactions between denaturants and protein surface exposed on unfolding: interpretation of urea and guanidinium chloride  $m$ -values and their correlation with changes in accessible surface area (ASA) using preferential interaction coefficients and the local-bulk domain model. *Proteins* **2000**, *Suppl 4*, 72-157.
42. Courtenay, E.; Capp, M.; Record, M. Thermodynamics of interactions of urea and guanidinium salts with protein surface: relationship between solute effects on protein processes and changes in water-accessible surface area. *Protein Sci.* **2001**, *10*, 2485-2582.
43. Zhu, H.; Celinski, S.; Scholtz, J.; Hu, J. The contribution of buried polar groups to the conformational stability of the GCN4 coiled coil. *J. Mol. Biol.* **2000**, *300*, 1377-1464.
44. Acharya, A.; Ruvinov, S.; Gal, J.; Moll, J.; Vinson, C. A heterodimerizing leucine zipper coiled coil system for examining the specificity of a position interactions: amino acids I, V, L, N, A, and K. *Biochemistry* **2002**, *41*, 14122-14153.
45. Chakrabarty, A.; Kortemme, T.; Padmanabhan, S.; Baldwin, R. L. Aromatic side-chain contribution to far-ultraviolet circular dichroism of helical peptides and its effect on measurement of helix propensities. *Biochemistry* **1993**, *32*, 5560-5565.
46. Padmanabhan, S.; Baldwin, R. Straight-chain non-polar amino acids are good helix-formers in water. *J. Mol. Biol.* **1991**, *219*, 135-142.
47. Padmanabhan, S.; Marqusee, S.; Ridgeway, T.; Laue, T.; Baldwin, R. Relative helix-forming tendencies of nonpolar amino acids. *Nature* **1990**, *344*, 268-338.
48. Chakrabarty, A.; Kortemme, T.; Baldwin, R. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci.* **1994**, *3*, 843-895.
49. Chiu, H.-P.; Suzuki, Y.; Gullickson, D.; Ahmad, R.; Kokona, B.; Fairman, R.; Cheng, R. Helix propensity of highly fluorinated amino acids. *J. Am. Chem. Soc.* **2006**, *128*, 15556-15563.
50. Pliska, V.; Schmidt, M.; Fauchere, J. L. Partition coefficients of amino acids and hydrophobic parameters  $\pi$  of their side-chains as measured by thin-layer chromatography. *J. Chromatogr.* **1981**, *216*, 79-171.
51. Fujita, T.; Iwasa, J.; Hansch, C. A new substituent constant,  $\pi$ , derived from

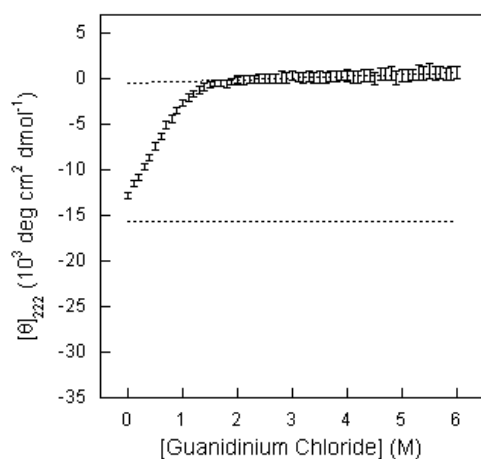
- partition coefficients. *J. Am. Chem. Soc.* **1964**, *86*, 5175-10355.
52. Taft Jr, R. W. Polar and steric substituent constants for aliphatic and o-Benzoate groups from rates of esterification and hydrolysis of esters1. *J. Am. Chem. Soc.* **1952**, *74*, 3120-6248.
53. Verloop, A. *Drug Design*, Academic Press: New York, **1976**.
54. Tipker, J.; Verloop, A. *The chemistry of excitation at interfaces*, ACS, **1984**.
55. Pauling, L.; Pressman, D. The serological properties of simple substances. IX. Hapten inhibition of precipitation of antisera homologous to the o-, m-, and p-azophenylarsonic acid groups. *J. Am. Chem. Soc.* **1945**, *67*, 1003-2015.
56. Agin, D.; Hersh, L.; Holtzman, D. The action of anesthetics on excitable membranes: a quantum-chemical analysis. *Proc. Natl. Acad. Sci. U. S. A.* **1965**, *53*, 952-960.
57. Hansch, C.; Leo, A.; Heller, S. R. *Exploring QSAR: Fundamentals and Applications in Chemistry and Biology*, ACS, **1995**.
58. Pace, C. N. Evaluating contribution of hydrogen bonding and hydrophobic bonding to protein folding. *Methods Enzymol.* **1995**, *259*, 538-554.



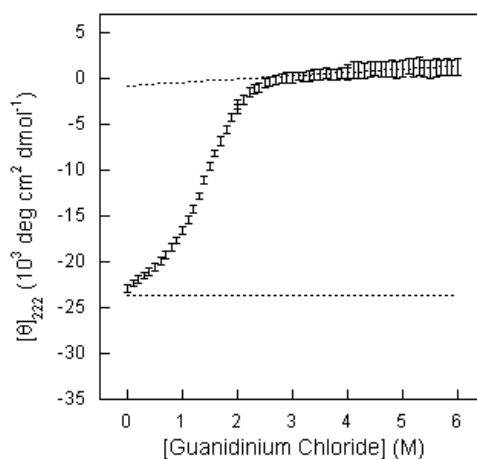
## Appendix

### *Guanidinium Denaturation Curves of GCN4-Xaa*

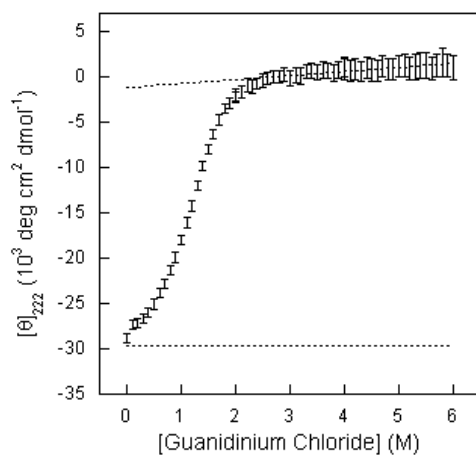
(a) GCN4-Abu



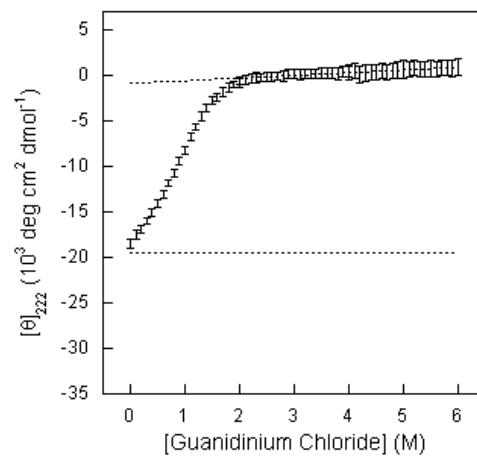
(b) GCN4-Nva



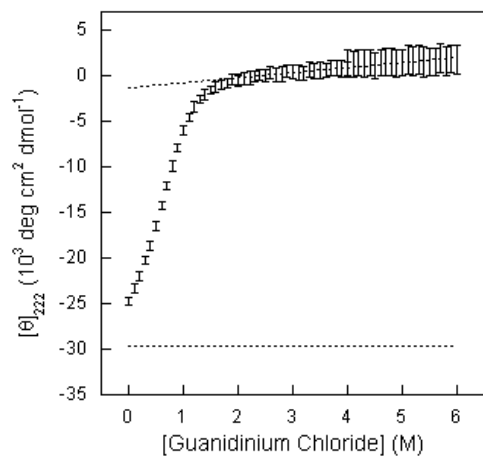
(c) GCN4-Nle



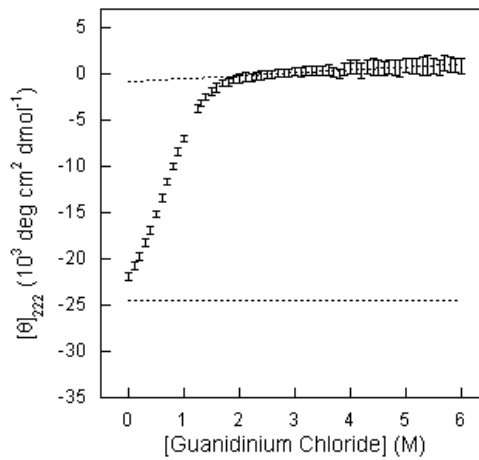
(d) GCN4-Ile



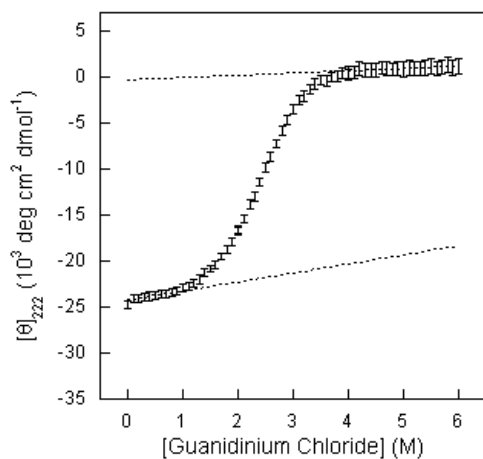
(e) GCN4-Allo Ile



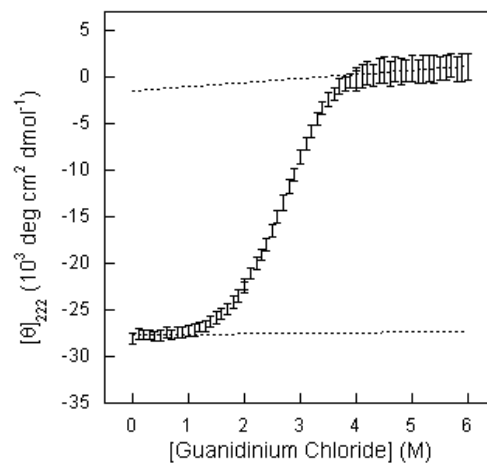
(f) GCN4-Val



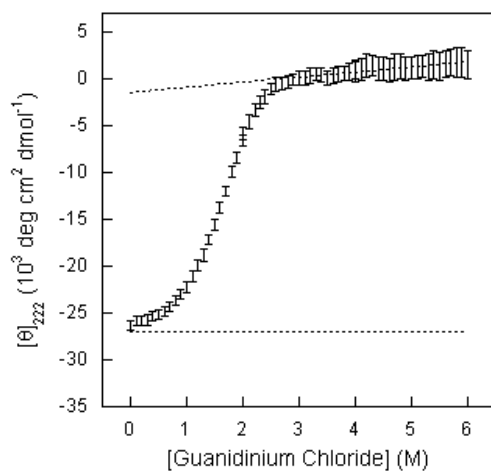
(g) GCN4-Leu



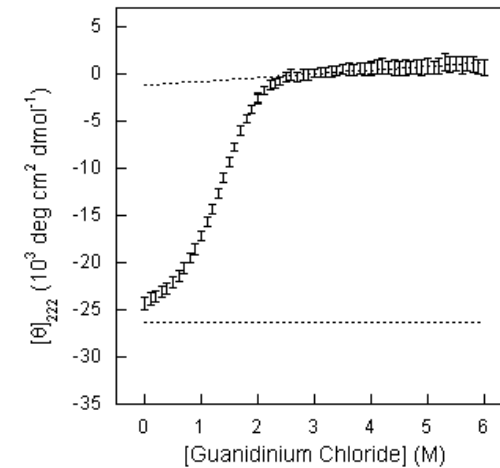
(h) GCN4-Tba



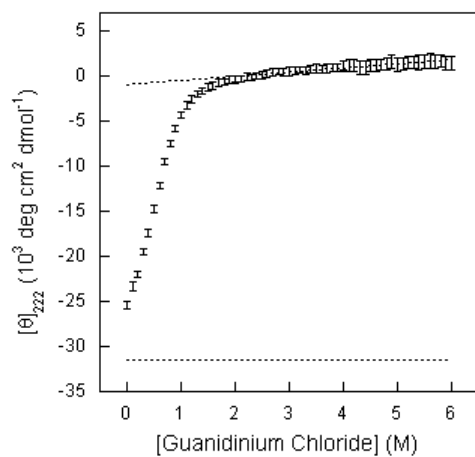
(i) GCN4-Cpa



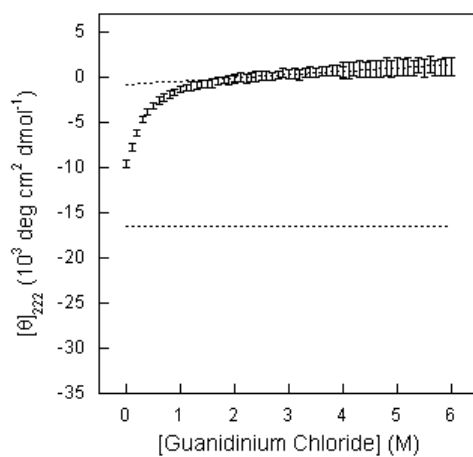
(j) GCN4-Cha



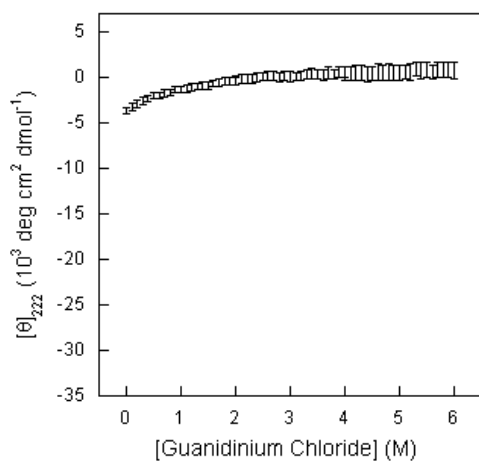
(k) GCN4-Phe



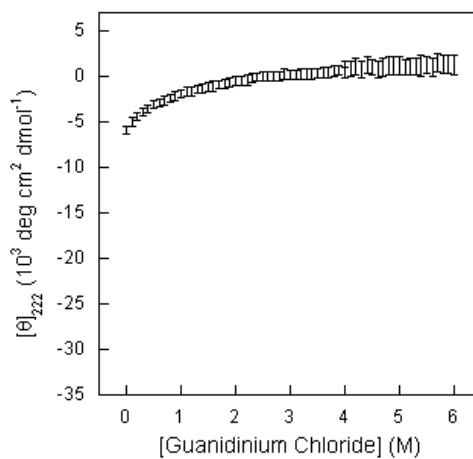
(l) GCN4-Pff



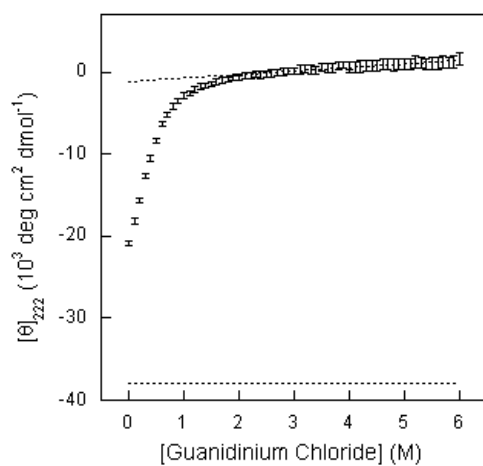
(m) GCN4-Asp



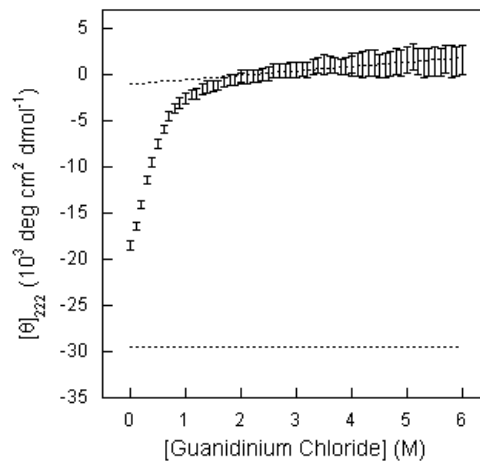
(n) GCN4-Asn



(o) GCN4-Glu

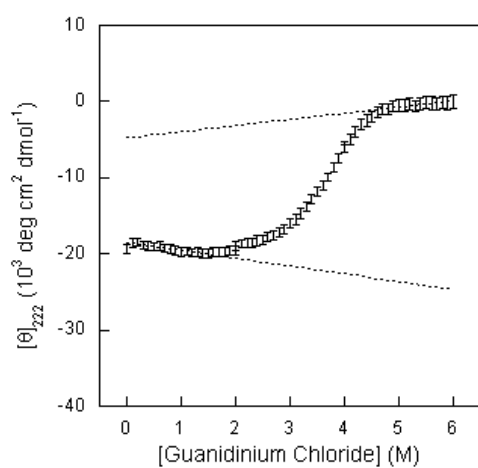


(p) GCN4-Gln

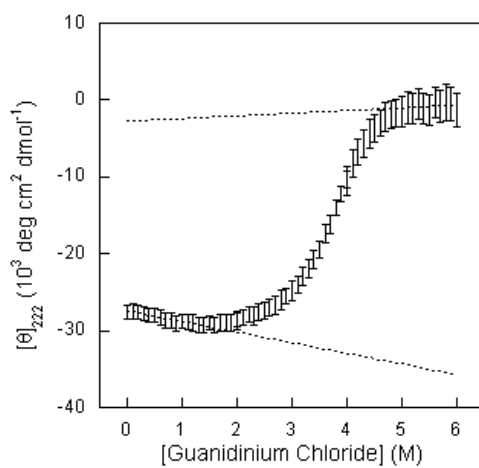


*Guanidinium Denaturation Curves of IaLd-Xaa*

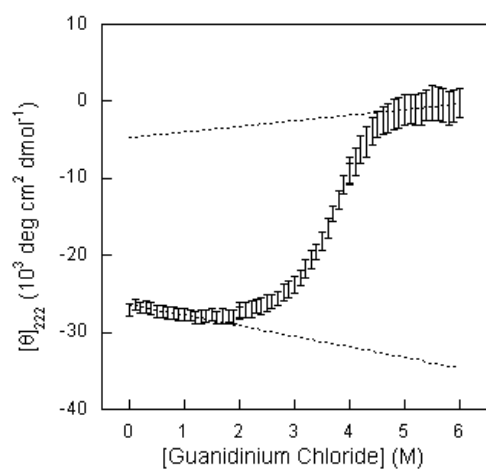
(a) IaLd-Abu



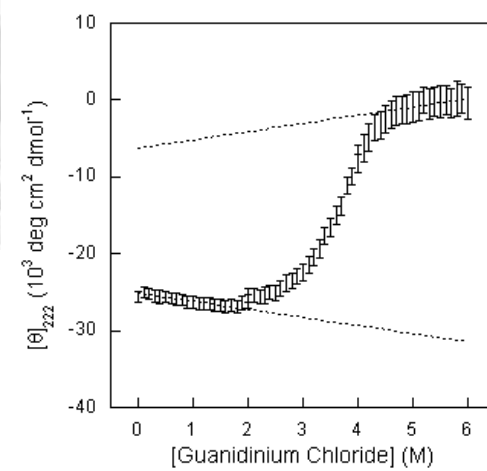
(b) IaLd-Nva



(c) IaLd-Nle

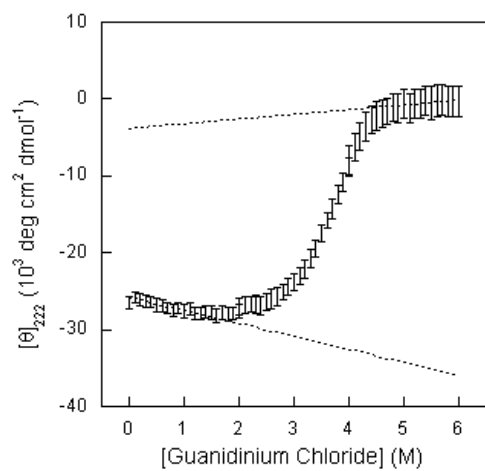


(d) IaLd-Cpa

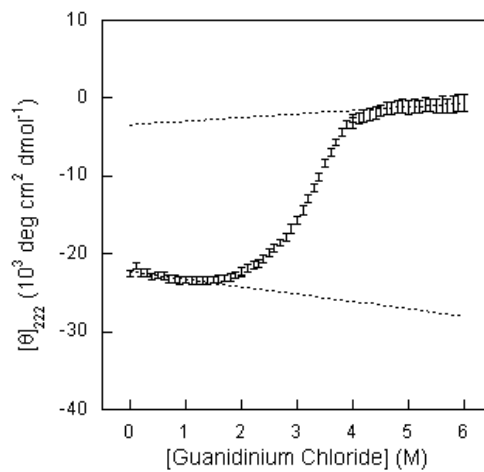




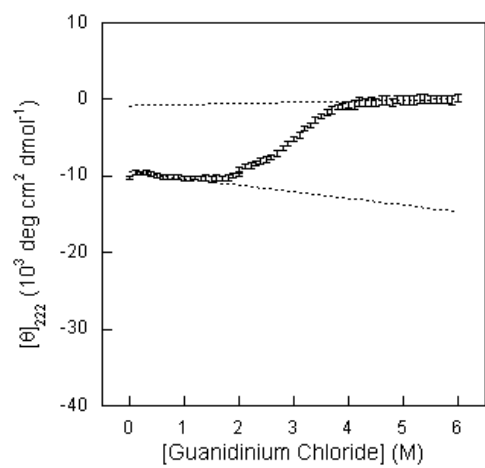
(e) IaLd-Cha



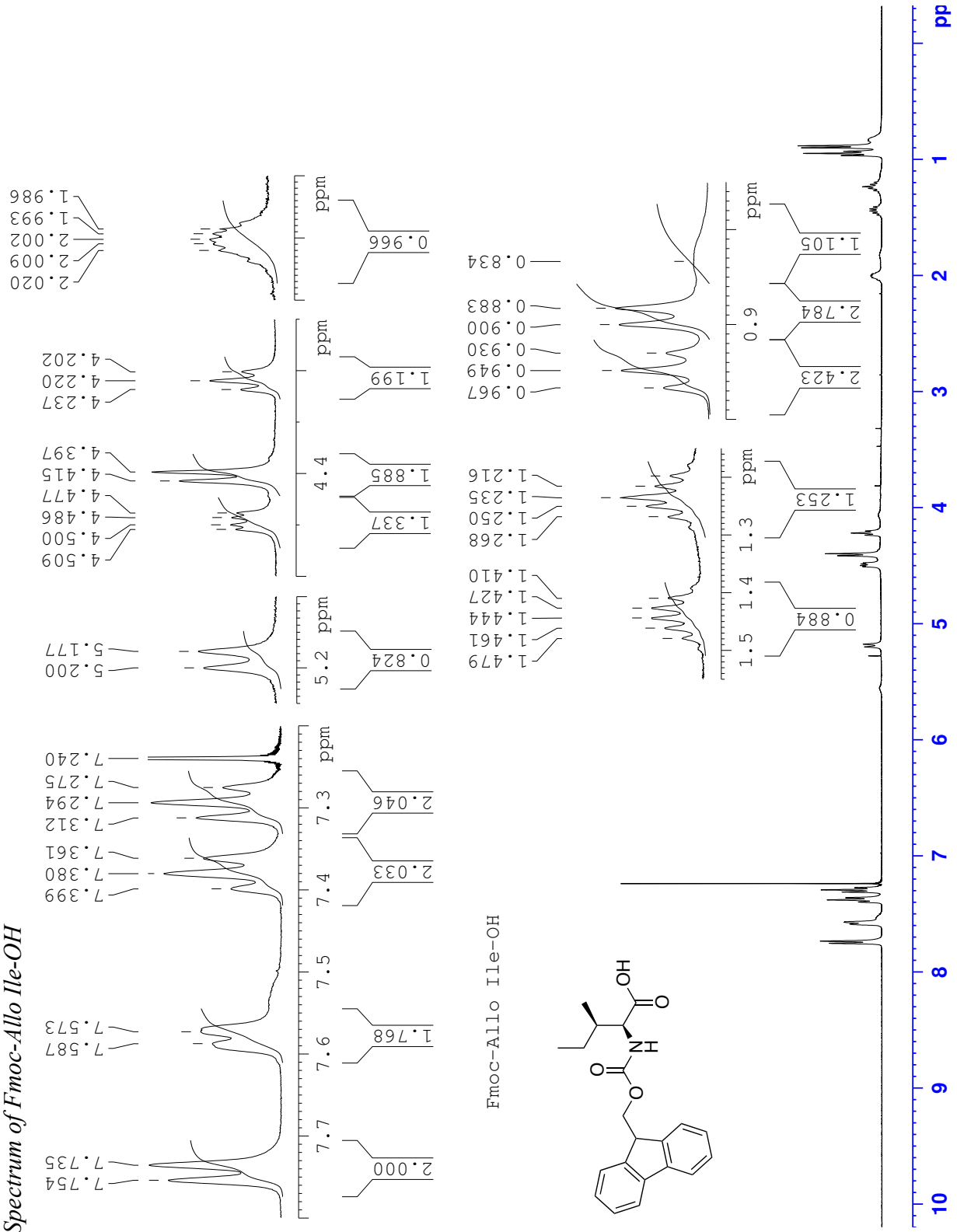
(f) IaLd-Phe



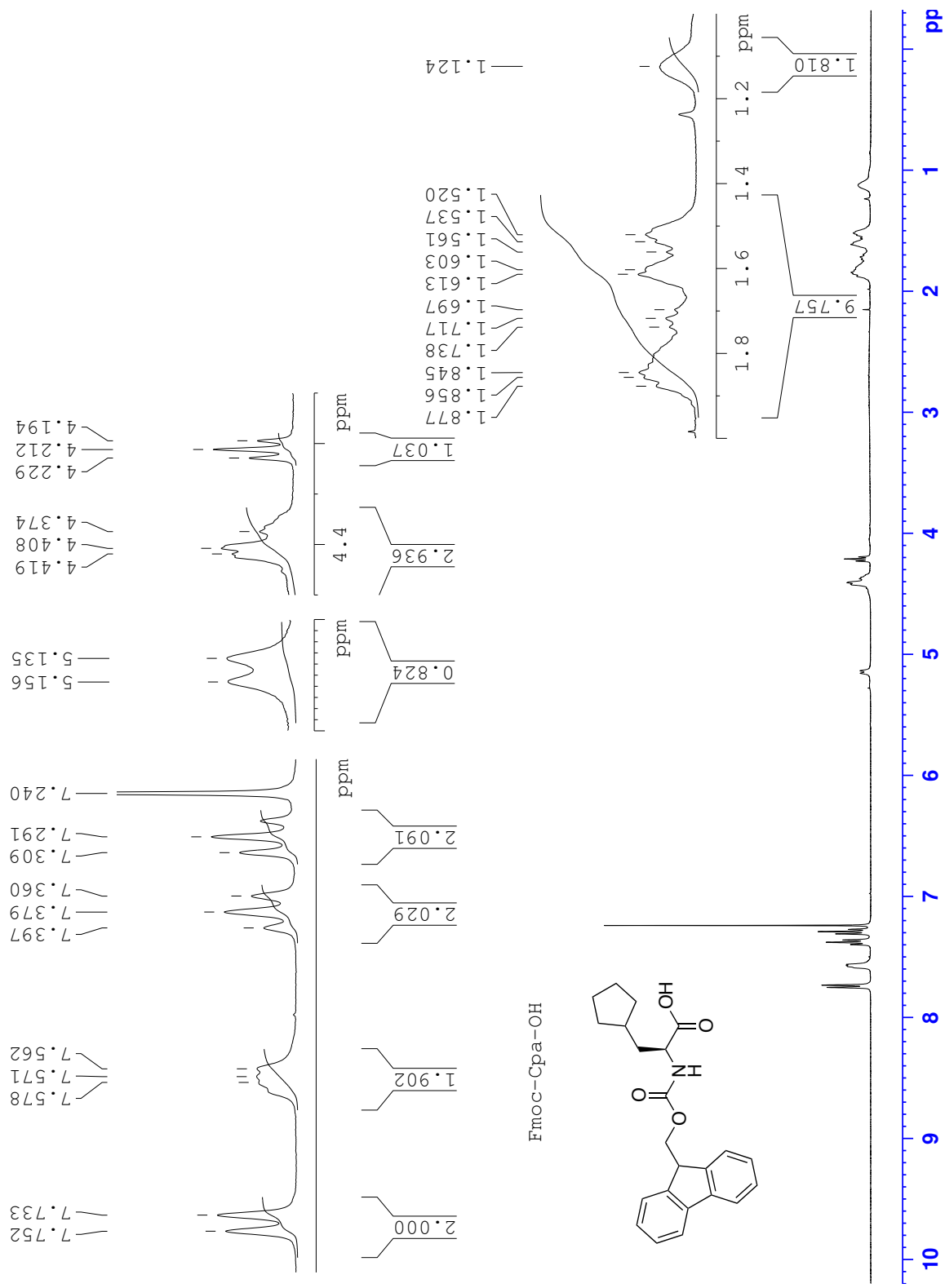
(g) IaLd-Pff



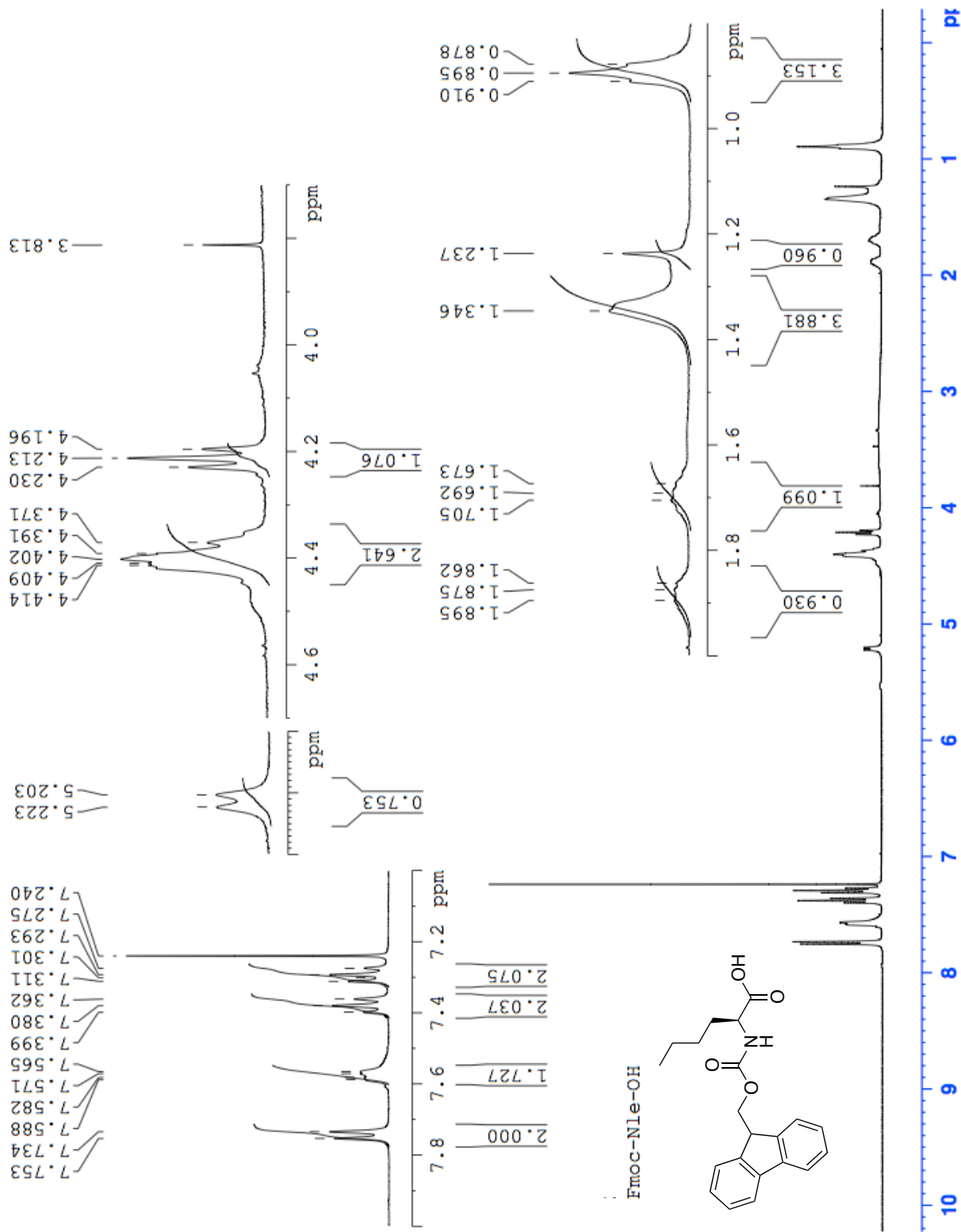
NMR Spectrum of Fmoc-Allo Ile-OH



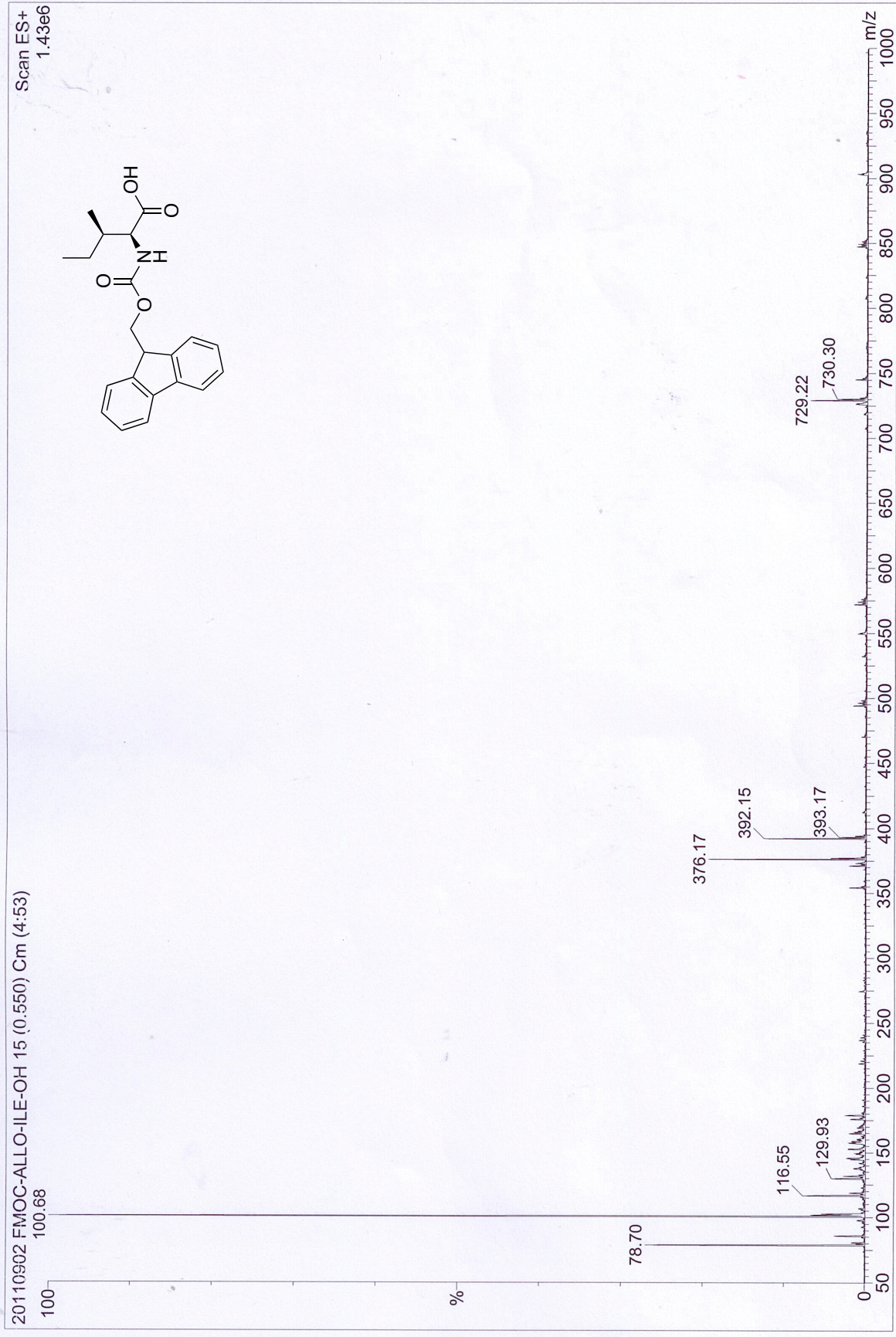
NMR Spectrum of Fmoc-Cpa-OH



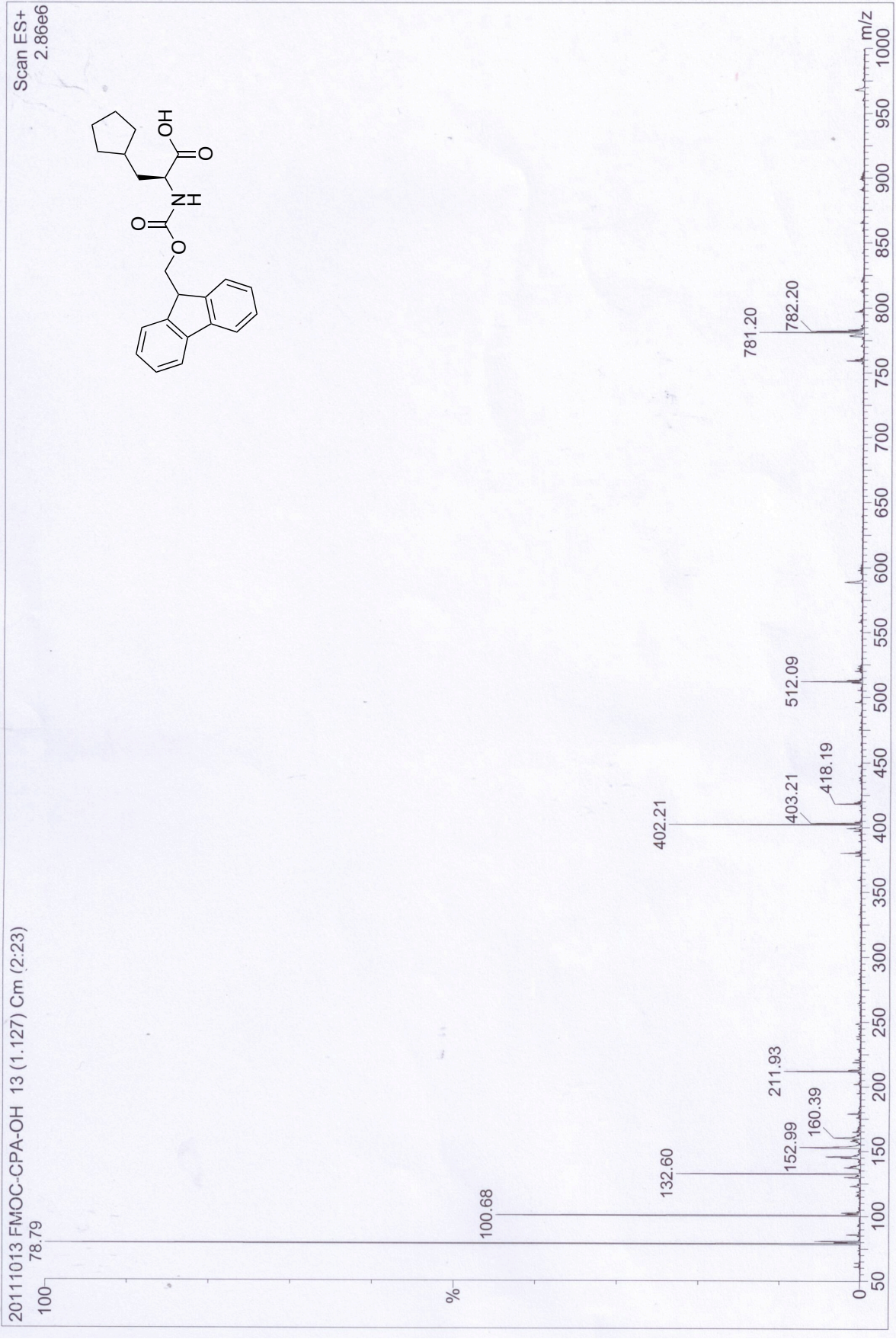
NMR Spectrum of Fmoc-Nle-OH



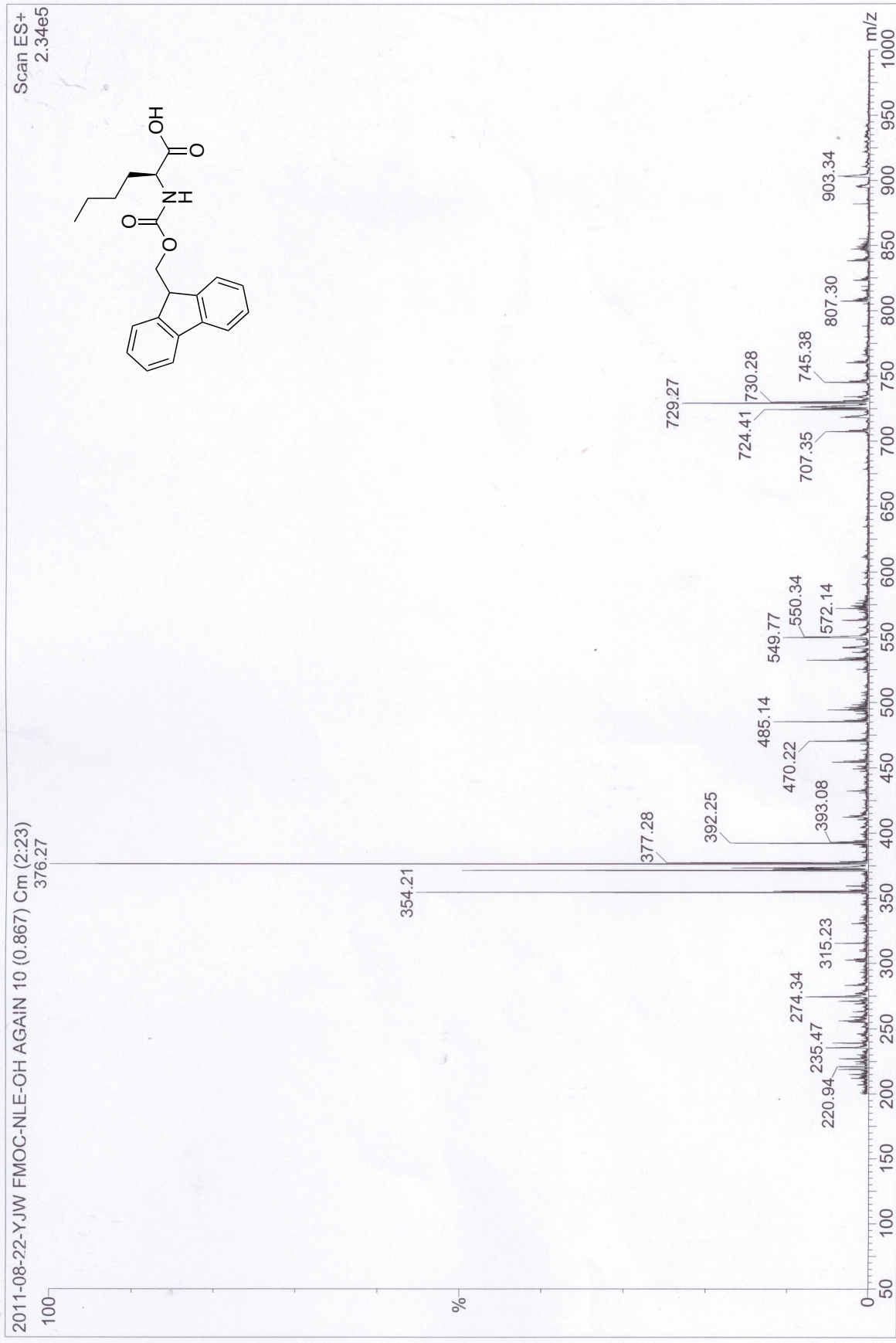
*ESI-MS Spectrum of Fmoc-Allo Ile-OH*



*ESI-MS Spectrum of Fmoc-Cpa-OH*



ESI-MS Spectrum of Fmoc-Nle-OH



## Elsevier License for Reprint of Figure

### ELSEVIER LICENSE TERMS AND CONDITIONS

Jul 17, 2012

This is a License Agreement between Yi-Jen Weng ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

**All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.**

Supplier	Elsevier Limited The Boulevard, Langford Lane Kidlington, Oxford, OX5 1GB, UK
Registered Company Number	1982084
Customer name	Yi-Jen Weng
Customer address	1F., No.3, Aly. 3, Ln. 77, Taipei, 116
License number	2887681314596
License date	Apr 14, 2012
Licensed content publisher	Elsevier
Licensed content publication	Current Opinion in Structural Biology
Licensed content title	New currency for old rope: from coiled-coil assemblies to $\alpha$ -helical barrels
Licensed content author	Derek N Woolfson, Gail J Bartlett, Marc Bruning, Andrew R Thomson
Licensed content date	22 March 2012
Licensed content volume number	
Licensed content issue number	
Number of pages	1
Start Page	0
End Page	0
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	Effect of Side Chain Structure of d Position Amino Acid on Coiled Coil Stability
Expected completion date	Jul 2012
Estimated size (number of pages)	50
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.0 USD / 0.0 GBP
Total	0.00 USD
Terms and Conditions	