Design of New Bio-Materials: Fluorous Peptides and Metal-Peptides Frameworks

by

Hyang-Yeol Lee

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in The University of Michigan
2008

Doctoral Committee:

Professor E. Neil. G. Marsh, Chair
Professor Vincent L. Pecoraro
Professor Ronald W. Woodard
Associate Professor Anna Mapp
Associate Professor Nils G. Walter

DEDICATION

This thesis is dedicated to my wife Young-Ran, my daughter Sungyun Kim Lee and my parents for their encouragement during my years here at the University of Michigan, to my advisor Dr. E. Neil G. Marsh and our lab members, Lindsey Gottler, Dustin Patterson, Chris Fox, Miri Yoon, Eric Rodrigues and my best friend, Lei Li.

ACKNOWLEDGMENTS

I would like to thank Dr. Vincent L. Pecoraro, Dr. Nils G. Walter, Dr. Anna Mapp, and Dr. Ronald W. Woodard for their kind advice and hospitality and also thank the Horace H. Rackam School for financial support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure

LIST OF APPENDICES

Appendix

Chapter 1 Modification of Natural Proteins and *De Novo* Designed Coiled Coils.

1.1. Introduction

Proteins play an essential role in biological systems, functioning as enzymes, which catalyze chemical reactions, as regulators of reactions, and as signal transducers. A protein's function can often be understood based on the three dimensional structure of the folded protein. Structural levels of a protein can be broken down into four categories:[1] Primary, secondary, tertiary and quaternary structures. Primary structure is simply the amino acid sequence of the polypeptide chain. Secondary structure is the local spatial arrangement of a polypeptide, such as helices (2.2_7 ribbon, 3_{10} helix, α-helix, π-helix), beta pleated sheets and loop conformation. Secondary structure is repetitive and common in proteins. Tertiary structure refers to the three dimensional structure of a polypeptide, however often it is ambiguous to distinguish between secondary and tertiary structure. Lastly, quaternary structure refers to the spatial arrangement of protein subunits.

Protein folding is one of the most intriguing and important unsolved problems in biology. As such there have been extensive studies on characterizing and predicting protein structures. However, predicting protein structure from its linear sequence still remains challenging because protein folding involves complicated weak interactions, such as electrostatic interactions including ionic interaction and dipole-dipole interactions, hydrogen bonding and hydrophobic interactions, which easily extend a secondary structure to a complex tertiary structure.

Such complexity makes the prediction of protein structure exclusively difficult. Extensive studies over the past several decades have shed light on the enigmatic problem and there has been progress. Among them, three general

approaches have been well established. Computational simulation has been used to take a protein structure and predict ligand binding sites. This approach together with docking routes to determine the position of two molecules relative to each other, has been used extensively in pharmaceutical design.[2,3]

Another approach is the re-design of natural proteins. Redesign of natural proteins provides a stringent test of our understanding of the basic principles underlying the structure and function of natural proteins and the opportunity to design properties and functions not present in natural proteins. One way to modify natural proteins is by performing mutations on active site residues in order to change substrate recognition. Such mutational studies of natural proteins provide useful information such as various folding characteristics, structures and catalytic reaction mechanisms. By using systematic mutation, the effects of repacking the hydrophobic cores of the natural proteins have been investigated.[4-10]

The third approach is *de novo* design. *De novo* design implies that the structure is designed from first principles. The sequence is not based on any naturally occurring protein, although the tertiary structure may closely mimic a natural protein. *De novo* design approaches to constructing new protein sequences have employed two global strategies, which are random sequence libraries and rational design.[11-14] Recent advances in genomic sequencing have shed light on new area of proteomics to investigate protein functions and structures. The construction and characterization of large libraries of *de novo* proteins can now be easily achieved, so that studies on protein functions and structures need not be limited to investigating existing proteins. For even very small proteins the hypothetical library of 20 naturally occurring amino acid sequences is too enormous for every permutation to be synthesized. Besides the vast quantity of possible sequences, it is also worth considering the quality of the sequences. Sequences capable of well folded and water soluble protein-like structures are extremely rare.

Rational design has shown great promise in creating novel proteins. Protein sequences have been designed residue by residue to construct the desired structure. Although rational design strategies have proved successful in designing some desired protein structures, rational design does not explore the extreme range of structures. Combinatorial libraries and rational design are often blended such that *de novo* proteins derived from large combinatorial libraries of sequences have been guided by rational design.

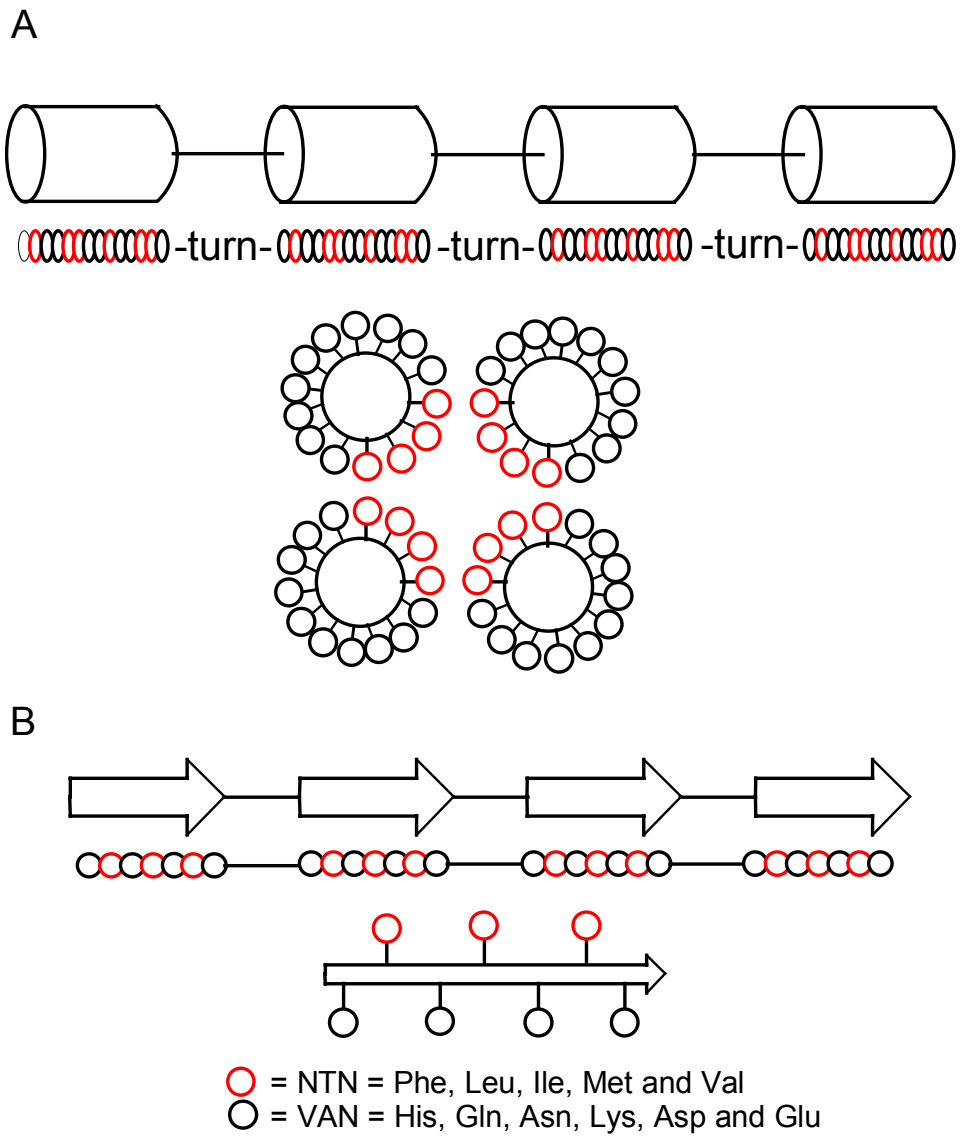


Fig. 1. Designed combinatorial libraries based on binary pattering. (A) The degenerate DNA codones, along with the nonplar (red circle) and polar (black circle) residues they encode in a library of α-helix. (B) Linear arrangement of polar and nonpolar residues in a library of β-sheet structures.[15]

Hecht and coworkers used a selection strategy to generate sequences that folding into their target protein structure. Their method generated a combinatorial library of four-helix bundles by relying on 'binary patterning' of polar and nonpolar amino acids.[16,17] About half of peptides sequence would be hydrophobic and half hydrophilic therefore these helices become amphiphilic. To design an amphipathic α-helix with on polar face and one nonpolar face, a binary patterns of P-N-P-P-N-N-P-P-N was used (where P = polar and N = nonpolar).[17] To design an amphipathic β-sheet with one polar face and one nonpolar face, an alternating pattern of P-N-P-N was used as shown in Fig. 1.

The library was prepared from a collection of synthetic genes expressed in bacteria. Nonpolar residues were encoded by the degenerate DNA codon NTN (N is a mixture of four nucleotide bases), which encodes Phe, Leu, Ile, Met and Val. Polar residues were encoded by the degenerate DNA codon VAN (V is a mixture of C, A and G). Among 4.7×10^{41} theoretical diversity of this library, over 50 proteins were isolated and characterized as water soluble and α-helical proteins. After the first round of collection, a second-generation library of binary-coded proteins was tested to encode collections of nativelike proteins.[18] The sequences identified were non-repetitive and not homologous to any known proteins.

Degrado and coworkers used a different strategy to design model peptide, α_4.[19] Their peptide is composed of all leucine residues on the hydrophobic core and all glutamic acid and lysine residues on the outer surface for complementary electrostatic interactions. The helical sequence is repetitive and capable of forming an amphiphilic α-helical structure that can form a tetramer. This rational strategy is more extensively discussed in next section with helical wheel diagrams for *de novo* designed multimeric coild coils.

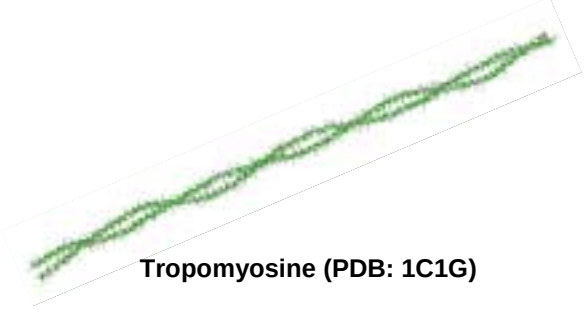
De novo designed proteins provide opportunity to test the hypotheses about protein folding and protein structure. The α-helix is the most common secondary structure and has proved to be the simplest and best characterized

motif to design, providing a useful model for structural and thermodynamic studies. Many *de novo* proteins have been designed that mimic the structures and/or properties of natural proteins.[14,20-24]

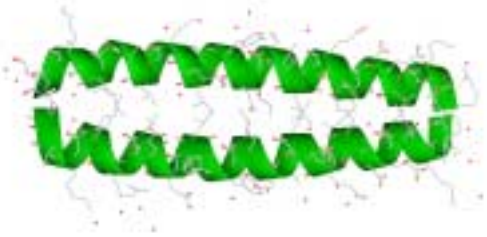
Numerous natural proteins that adopt helical domain motifs exist that can be used as models for re-designing proteins and for *de-novo* model studies.

1.2. Re-Designing Natural Proteins.

One of the simplest protein structures that has been used for re-design of a natural proteins is the dimerization domain of GCN4.[25] GCN4 is a yeast gene regulatory protein whose dimerization domain is comprised of a 2-stranded parallel helical coiled coil or 'leucine zipper' domain. Fig. 2. (A) shows some examples of proteins that adopt dimeric coiled coil motifs such as Tropomyosine,[26] GCN leucine zipper,[25] Sir4P[27] and Geminin[28] whereas (B) shows proteins that adopt trimeric motifs, Lectin,[29] α-Actinin 4[30], Coronine 1[31] and Influenza Heamagglutinin[32].



Tropomyosine (PDB: 1C1G)



GCN Leucine Zipper (PDB: 2ZTA)

Sir4p C-Terminal Coiled Coil (PDB: 1PL5)

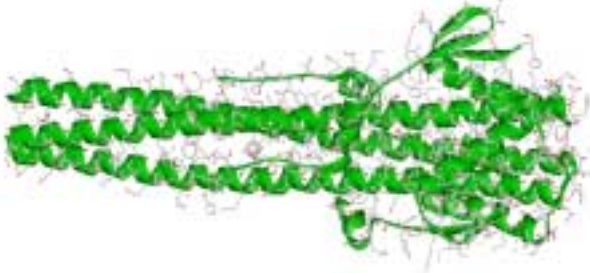
Coiled-Coil Dimerization Motif of Geminin (PDB: 1T6F)

A) Dimeric Coiled Coil Proteins.

Lectin (PDB:1RTM)

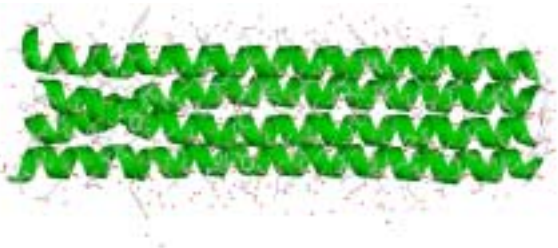
3rd Spectrin Repeat of α-Actinin-4 (PDB: 1WLX)

The Coiled-Coil Domain of Coronin 1 (PDB:2AKF)

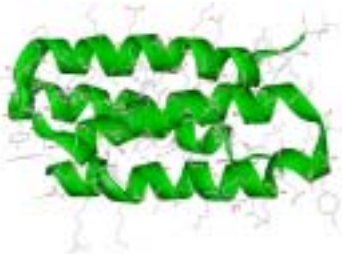


Influenza Haemagglutinin (PDB:1HTM)

B) Trimeric Coiled Coil Proteins.



Tetrabrachion (PDB: 1FE6)



ROP (PDB: 1ROP)

***E. Coli* cytochome b_{562} (PDB: 256B)**

Human Growth Hormone, Prolactin (PDB: 1N9D)

C) Tetrameric Coiled Coil Proteins.

Fig. 2. Examples of natrual proteins that adopt multimeric coiled-coils.

Natural proteins such as Tetrabrachionin,[33] ROP,[34] *E.coli* cyt b_{562}[35] and Prolactin,[36,37] adopt four-helix arrangement as shown in Fig. 2. (C). Prolactin (191 residues) is a human growth hormone, which has up-up-down topology. ROP protein will be described in detail in Chapter 3.

Cytochrome b_{562} is a four-helix bundle protein containing a non-covalently bound b-type heme prosthetic group. Many groups work on reengineering this protein. For example, Gibney et al reported synthesis of ferredoxin-heme marquette where sequence is –Cys-X-X-X-Cys-X-X-Cys-X-X-Cys-. Successful assembly of the $[4Fe\text{-}4S]^{2+/1+}$ cluster within a heme-bound peptide derived from a motif of natural ferrodoxines was demonstrated in model studies.[22,38,39] Most recently, Salgado et al engineered a cyt cb_{562} variant (His^4-cb_{562}) near the N- and C- termini of which helix at the positions of 59/63 and 73/77 for controlling protein-protein interactions through metal coordination. As a result, 4 copies of cyt cb_{562} associated to form a 16 helix complex Zn-coordination.[40]

By simply borrowing motifs from the natural proteins, such structures and biological properties can be explored utilizing *de novo* design strategy.

1.3. *De novo* Design of Coiled-Coil Peptides.

De novo design is from first principles, not based on any naturally occurring protein, although structure may closely mimic a natural protein. The coiled coil was first described in 1953 by Pauling et al[41] and Crick[42] as the main structural element of a large class of fibrous proteins that included keratin, myosin and fibrinogen. A coiled coil is a bundle of α-helices that is wound into a super-helix. They have been observed to exist with two, three or four helices in the bundle and they can present either in a parallel or anti-parallel direction.[43] Coiled coils are well characterized for their structures and have been developed as tools for protein design.

Undistorted α-helices have approximately 3.6 residues per turn. Coiled-coils effectively reduce the number of residues per turn to 3.5 with respect to the

super-coil axis. The packing of core residues can determine the number of helices in a coiled coil. This is because residues in positions ‘a’ and ‘d’ have a different geometry with respect to the backbone of the facing helix in two-, three- and four-stranded coiled coils, and therefore also have different side chain preferences.[4,5] The parallel or anti-parallel orientation of helices in a coiled coil is primarily determined by polar and ionic interactions even though electrostatic interactions can contribute only about 0.5 kcal/mol or less.[44,45]

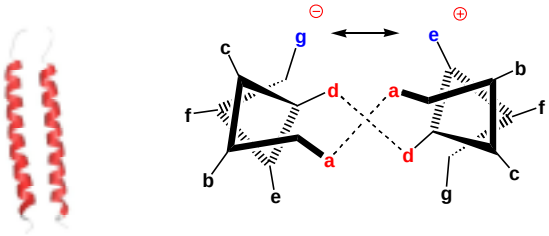
Designing and engineering *de novo* proteins has been attractive for many reasons including the understanding of protein structure and function. Today, design and synthesis of *de novo* designed metalloproteins, enzymes, and pharmaceutically relevant peptides is becoming more routine. Using these simple and well characterized coiled coils as a model along with *de novo* design as a strategy, it will be powerful tool to test our knowledge and hypothesis.

De novo design has been known for decades and its strategy is well established for the last decades for designing small numbers of coiled-coils. With the simple helical wheel diagrams, which the DeGrado group well established, it can be relatively easy to design multimeric coiled-coils.

1.3.1. Parallel Helical Bundles: Two-, Three- and Four Stranded Coiled Coils.

In the design of dimeric helical coiled coils, it is important to consider alternate folding topologies and to incorporate both positive and negative design elements to favor the desired fold and discourage unwanted folds. The geometry of parallel two-stranded coiled coil is shown in Fig. 3. The helices are represented as having a heptad repeat, and the positions are labeled ‘a’ to ‘g’. Folding topologies for parallel bundles are fundamentally simple. Hydrophobic residues are located at ‘a’ and ‘d’ positions of heptads and ‘g’ and ‘e’ residues have an electrostatic interaction (*i.e.* a salt bridge). Buried polar groups, such as Asn, in the hydrophobic core also have an important role for maintaining the dimeric and parallel orientation via hydrogen bonding since it destabilizes the alternative orientation.[25]

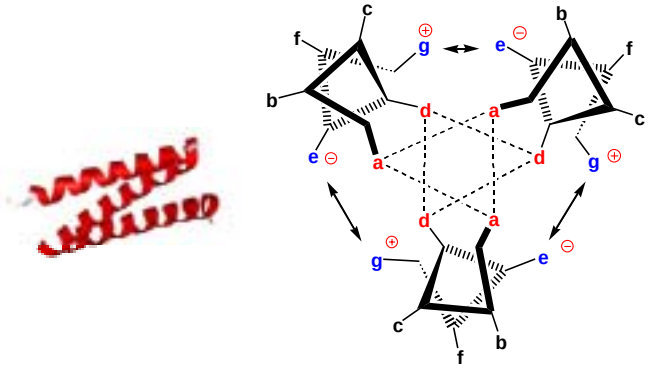
Fig. 3. Parallel Two Stranded Coiled-Coil Models.



For instance, in the *de novo* design of mercury-binding coiled coil peptides, Tri variants and Coil-V_aL_d variants have been reported to form either two or three helix bundles.[20,21,46,47] The trimeric model peptide, L16C, based on Coil-Ser, was designed by replacing Leu at position 16 with Cys. This peptide forms dimers and trimers depending on the pH. At low pH, it was observed as a dimer, but trimer was formed at pH 8.5.[20]

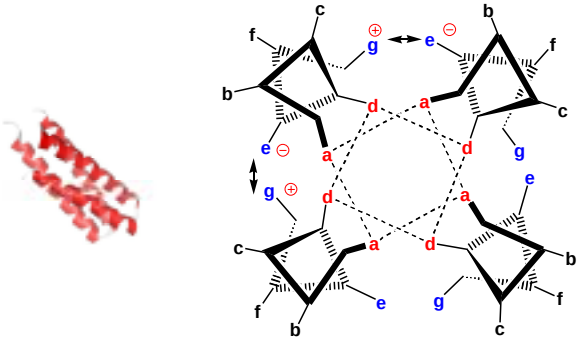
Many groups also work on incorporation of unnatural amino acid into the hydrophobic core of bundles to investigate the properties of unnatural proteins. The Tirrell group substituted 5,5,5-trifluoroleucine at the 'd' positions of the heptads of GCN4-pld and observed that the thermal stability of the peptide increases without changing oligomerization states.[7] Interestingly, the Kumar group found out that their highly fluorinated peptide, in which 5,5,5,5',5',5'-hexafluoroleucine was incorporated at 'a' and 'd' positions of their designed peptide, FF, did change the oligomerization state to tetramers whereas their controlled nonfluorinated peptide, HH, was a dimer.[48,49]

Fig. 4. Parallel Three Stranded Coiled-Coil Models



Parallel three-helix bundles such as coiled coil-V_aL_d contains valine in the 'a' heptad positions and leucine in the 'd' heptad positions.[50,51] The geometry of a parallel three-stranded coiled coil is shown in Fig. 4. Formation of three salt bridges between charged residues in the 'e' and 'g' positions in each layer stabilizes such bundles in a parallel orientation. However, this peptide also changes oligomeric states from trimer to dimer depending on the pH to minimize the electrostatic repulsion between the peptide chains.[14,21] This behavior illustrates the instability of the parallel arrangement of helices.

Fig. 5. Parallel four Stranded Coiled-Coil Models



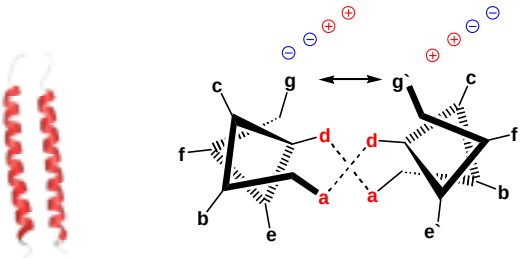
Hydrophobic residues are located at 'a' and 'd' positions of heptads and 'g' and 'e' residues have an electrostatic interaction, i.e., a salt bridge. In a tetramer, in addition to 'g' and 'e' interaction, there can be additional 'c' and 'b' interactions. Repacked GCN-L_aI_d with isoleucine at each 'd' and leucine at each 'a' is tetrameric in solution and in the solid state.[4,5] The geometry of parallel four-stranded coiled coil is shown in Fig. 5. This parallel four helical bundle system also provides the foundation for the design of hemoproteins that mimic the functions of a photosynthetic reaction center.[22,38,39,52]

1.3.2. Anti-Parallel Helical Bundles: Two-, Three- and Four-Stranded coiled coil.

Anti-parallel helical bundles have very different inter-helix interaction patterns from parallel coiled coils which are quite different depending upon

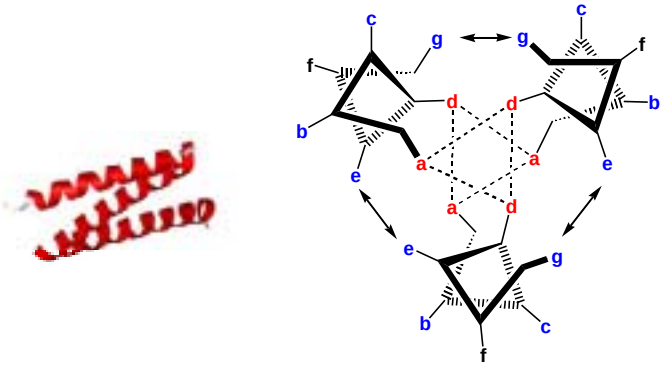
whether a two-, three-, or four-stranded coiled-coil is being considered. For a two-stranded coiled coil, the hydrophobic residues at ‘a’ and ‘d’ face each other as shown in Fig 6. One interface is formed between residues at ‘g-g`’ positions, and the other between residues at the ‘e-e`’ positions. ‘g’ and ‘e’ of heptads have different charges from those of the other end of the bundle so that it arranges in an anti-parallel pattern by electrostatic interaction.

Fig. 6. Anti-Parallel Two Stranded Coiled-Coil Models



An example of a naturally occurring antiparallel 2-stranded coiled coil is the repressor-activator MerR, which controls transcription of the mercury resistance (mer) operon has an anti-parallel dimeric coiled coil domain. The metal recognition domain of MerR has an anti-parallel orientation to form a trigonal Hg(II)-coordination site.[53-55] Dimeric helical bundles, which mimic MerR have been engineered for metal binding.[23,56]

Fig. 7. Anti-Parallel Three Stranded Coiled Coil Models

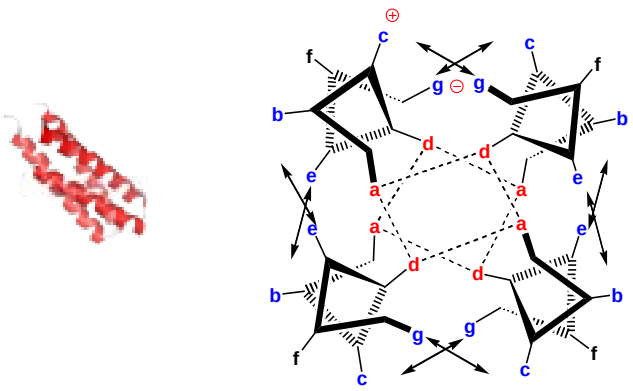


In a trimeric antiparallel coiled-coil, two helices are in parallel and one bundle is in anti-parallel orientation. Nonpolar residues at ‘a’ positions face two

nonpolor ‘d’ residues in addition to one ‘d-d’ interaction. The geometry of an anti-parallel three-stranded coiled coil is shown in Fig. 7. One example of an anti-parallel coiled coil, Coil-Ser in which all ‘a’ and ‘d’ positions of the heptads are Leu, was designed to be unfavorable at neutral pH, and it is reported to form parallel trimers in aqueous solution at neutral pH. However, the crystal structure of coil-Ser was found to be anti-parallel.[57]

De novo designed mercury-binding peptides have been reported. An anti-parallel trimeric bundle (L16C, up-up-down topology) based on CoilSer was designed for binding Hg(II). A three-coordinate site of the designed peptide for MerR was constructed by replacement of Leu with Cys.[20]

Fig. 8. Anti-Parallel Four Stranded Coiled-Coil Models



In the anti-parallel tetramer, apolar ‘a’ residues face ‘d’ residues in the core layer and there are two different electrostatic interactions between the interface residues ‘c’ and ‘g’ and ‘b and ‘e’ residues, which help to stabilize the helices in the anti-parallel orientation. For example, the DeGrado group is working on the hierarchic *de novo* design of α-helical bundles that adopt native-like structures such as the α_2 and RLP family of peptides. The α_2 series of the peptide form dimers of helix-loop-helix motifs, where their hydrophobic cores are packed with various hydrophobic residues such as Leu, Val and Ile.[58,59] RLP peptides were intended to fold into ROP-like, four-helix bundles.[60,61] The geometry of anti-parallel four-stranded coiled coil is shown in Fig. 8.

The antiparallel four-helix bundle fold is a structurally robust motif so that it is expected to tolerate the incorporation of unnatural amino acids, such as hFleu, which has been used in the model studies.

1.4. Goals

My research has three main aims. Firstly, to examine how these fluorous proteins differ from their natural counterparts with respect to their physicochemical properties such as thermal stability and unfolding by denaturants. Secondly, to both test and advance our understanding of how hydrophobic interactions contribute to the stability and correct folding of natural proteins by incorporating limited numbers of fluorous residues at specific positions within the hydrophobic core. Finally, I use the knowledge gained from these experiments to exploit fluorous interactions in the design of very small, stably folded proteins to explore whether the fluorous effect can be used to engineer specific proteins.

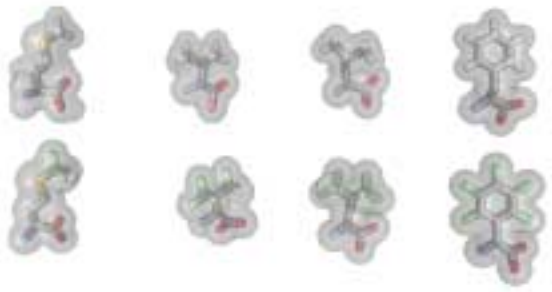
Chapter 2 Synthesis and Characterization of Fluorinated Antiparallel 4-Helix (α_4) Peptides

2.1. Exploiting the fluorous effect in protein design

This chapter 2 is based on our publication in *Biochemistry*[24] and *JACS*[62]. Fluorocarbons are thermally and chemically inert, e.g. Teflon, and a typical C-F bond is 10 – 15 kcal/mol more stable than C-H bond. Fluorocarbons are extremely hydrophobic and exhibit interesting self-segregating properties. By incorporating these unique fluorous properties into proteins, it is predicted that fluorous proteins should be thermally more stable than natural proteins, resistant to denaturation by organic solvents and "non-sticky" to other proteins.

Fig. 9. Typical bond-length of C-H (1.1Å) and C-F (1.4Å).

As shown in Fig. 10, fluorinated analogs of hydrophobic amino acids such as trifluoro-methionine, hexafluoro-valine, hexafluoro-leucine[63,64] and pentafluoro-phenylalanine are sterically similar to the nonfluorinated amino acids and can be used to substitute in the place of hydrocarbon counterparts. They are either commercially available, or can be synthesized from known synthetic procedures.



$_L$-methionine $_L$-valine $_L$-leucine $_L$-phenylalanine

$_L$-trifluoromethionine $_L$-hexafluorovaline $_L$-hexafluoroleucine $_L$-pentafluorophenylalanine

Fig. 10. Fluorocarbon analogs (below) of hydrocarbon amino acids (upper) that have similar shapes.

2.2. Fluorinated α-Helical Coiled-Coil Peptides

Fluorocarbons possess unusual and useful physicochemical properties. They are extremely hydrophobic and chemically inert; properties that have been successfully exploited to develop inert materials such as Teflon, fire retardants and anesthetics. Fluorocarbons also exhibit unusual phase segregation behavior. Extensively fluorinated molecules preferentially partition into fluorocarbon solvents, rather than hydrocarbon solvents – a phenomenon referred to as the "fluorous effect". Recently, the fluorous effect has found increasing use in organic synthesis to facilitate the purification of compounds tagged with long perfluorocarbon "tails".[65-68] Fluorocarbons are quintessentially man-made compounds because, with one or two interesting exceptions,[69,70] fluorine is essentially absent from biological molecules. This raises the question of whether similarly novel and useful properties can be engineered into proteins that incorporate extensively fluorinated analogues of hydrophobic amino acids into their structures.[71,72] Therefore, recently there has been considerable interest in the properties of peptides incorporating fluorinated analogues of

hydrophobic amino acids, in particular leucine and valine, within their hydrophobic cores.

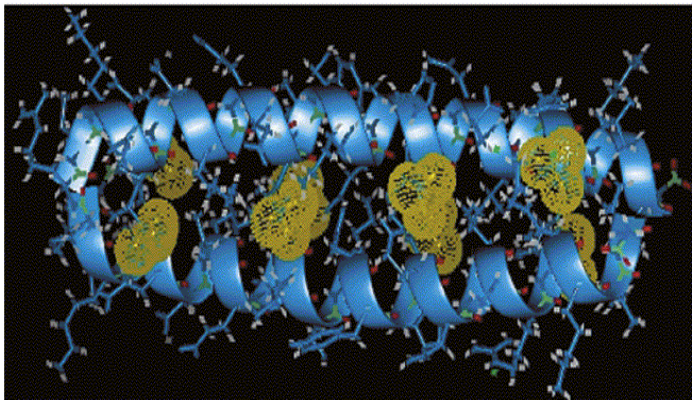


Fig. 11. Fluorous GCN-4[7]

Several studies have focused on synthesizing fluorous analogues of peptides designed to adopt dimeric α-helical coiled-coil structures typified by the "leucine zipper" region of the yeast transcription factor GCN4 dimerization domain. Incorporation of fluorinated analogues of leucine and valine have been shown to significantly increase the stability of coiled-coil peptides toward unfolding by heat and chemical denaturants.[6-8,48,73] Tang et al[7] observed that thermal stability of GCN4-pld increases when the unnatural amino acid, 5,5,5-trifluoroleucine, was substituted at the 'd' positions of the heptads and Bilgicer et al[6] designed a highly fluorinated peptide, in which all four leucine residues at the 'a' positions and three valine residues at the 'd' positions were replaced by 5,5,5-trifluoroleucine and 4,4,4-trifluorovaline.

Specific self-association of two α-helical peptides incorporating hexafluoroleucine (hFLeu) at the hydrophobic interface has also been accomplished.[6] The fluorous effect has also been used to control the self-association of membrane-spanning α-helical peptides.[74] Helix propensity of highly fluorinated amino acids, such as *S*-hexafluoroleucine (Hfl), *S*-2-amino-4-

fluorobutyric acid (Atb) and *S*-pentafluorophenylalanine (Pff), was investigated. Their studies showed that, when fluorinated amino acids were incorporated at the same position of the model peptides, nonfluorinated *S*-2-aminobutyric acid (Abu) showed the second highest helical content in the model pepetides whereas its fluorinated analog, Atb, had the lowest (helical propensity: Pff > Hfl > Atb).[64]

My research has focused on a different structural motif, the antiparallel 4-α-helix bundle, in my investigations into the effect of incorporating fluorous amino acids on the physical properties of proteins.[24] The 4-α-helix bundle fold is a structurally robust motif that is widely distributed in Nature, being found in proteins such as the RNA-binding protein ROP, transcription factors, cytokines and in the cores of many dinuclear metalloenzymes.[14,75-77] On the basis of *de novo* design principles established by DeGrado and co-workers, I have synthesized peptides that adopt an antiparallel 4-helix bundle fold in which the hydrophobic core is exclusively packed with leucine residues. This allows me to investigate the effect on protein structure and stability of systematically replacing leucine with the fluorous analogue hFleu.

Here I describe the synthesis and characterization of a series of peptides, designated α_4-F_n (where n = 2, 4, 6), in which 2, 4, or 6 layers of the hydrophobic core have been repacked with hFLeu, as shown in Figure 11. This introduces progressively more trifluoromethyl groups into the hydrophobic core of the protein: thus α_4-F_2 contains 16 trifluoromethyl groups packed in the central 2 layers of the 4-helix bundle; α_4-F_4 contains 32 trifluoromethyl groups packed in the central 4 layers; and finally in α_4-F_6 a total of 48 trifluoromethyl groups pack all 6 layers. I find that increasing the number of hFLeu residues increases G_{unfold} for the peptides in an almost linear fashion. And, unexpectedly, that increasing the number of hFLeu residues appears to result in better structured peptides as judged by 1-D ^{1}H, ^{19}F NMR spectra and 2-D ^{1}H-^{15}N HSQC NMR spectra.

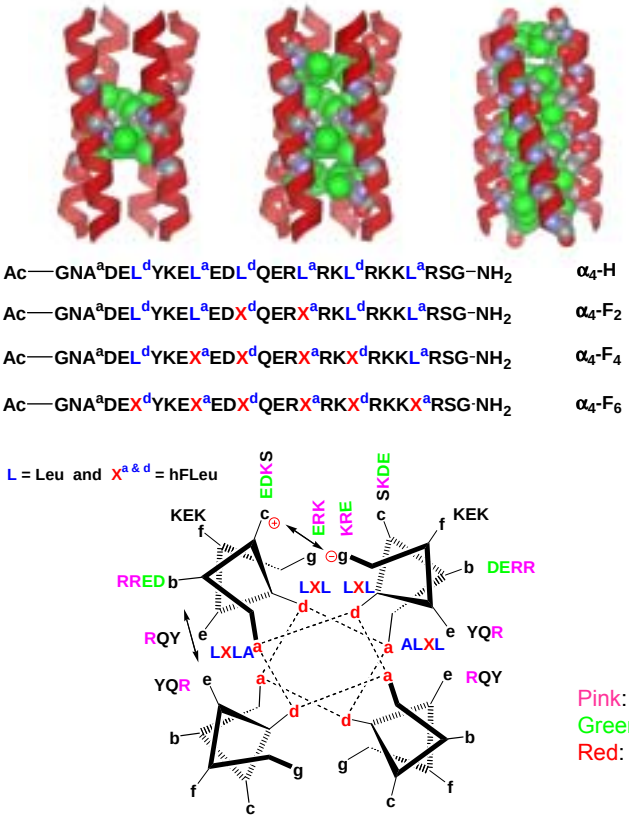


Fig. 12 (Top) Models showing the packing of hFLeu in the hydrophobic cores of α_4-F_2, α_4-F_4, and α_4-F_6 peptides. The trifluoromethyl groups are represented by green spheres). (Middle) Sequences of α_4-H, α_4-F_2, α_4-F_4, and α_4-F_6 peptides (X = hFLeu). (Bottom) Helical wheel diagram illustrating the side-chain interactions between helices for the α_4-H sequence.

Materials and Methods

Materials. Rink Amide resin, Fmoc-protected and *t*-Boc-protected amino acids, *N*-hydroxybenzotriazole (HOBt), and 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetra-methyluronium hexafluorophosphate (HBTU) were purchased from NovaBiochem. Peptide synthesis grade *N*-methylpyrrolidinone (NMP), *N*,*N*-dimethylformamide (DMF), and trifluoroacetic acid (TFA) and ACS grade *N*,*N*-diisopropylethylamine (DIEA) and piperidine were purchased from Fisher. GuHCl, 99%+ purity, was obtained from Gibco BRL. L-5,5,5,5',5',5'-Hexafluoroleucine

was synthesized as described previously and converted to Fmoc- or *t*-Boc-protected derivatives by standard procedures.

Peptide Synthesis. Peptide α_4-H was synthesized using Fmoc-protected amino acids by standard protocols on an ABI 433A automated synthesizer. Peptides were cleaved from the resin by stirring for 2 h at room temperature with 10 mL of a mixture of 90% TFA, 3% ethanedithiol, 5% thioanisole, and 2% anisole. The resin beads were filtered off and rinsed with an additional 4 mL of TFA. TFA was evaporated from the filtrate under a stream of nitrogen, and 50 mL of cold ethyl ether was added to precipitate the peptide. The crude peptide was collected by filtration on a fritted funnel, dissolved in 10% aqueous acetic acid, and lyophilized. Peptides α_4-F_2, α_4-F_4, and α_4-F_6 were synthesized using *t*-Boc-protected amino acids for Merrifield manual solid-phase synthesis on MBHA resin; couplings were performed using the in situ neutralization/HBTU protocol described by Schnolzer et al.,[78] typically on a 0.25 mM scale. The peptide was cleaved from the resin using "high" HF conditions.

Peptide Purification. Peptides were redissolved at ~10 mg/mL in 10% aqueous acetic acid and purified by reverse-phase HPLC on a Waters semipreparative C_{18} column equilibrated in 0.1% TFA and eluted with a linear gradient of 0-90% acetonitrile containing 0.1% TFA. The peptides were determined to be pure by analytical HPLC and MALDI-TOF mass spectrometry: expected mass for peptide α_4-H ($C_{142}H_{244}N_{46}O_{44}$) = 3299.7 amu, detected mass = 3300.1 amu; expected mass for peptide α_4-F_2 ($C_{142}H_{232}N_{46}O_{44}F_{12}$) = 3515.6 amu, detected mass = 3515.8 amu; α_4-F_4 ($C_{142}H_{220}N_{46}O_{44}F_{24}$) = 3731.5 amu, detected mass = 3732.9 amu; α_4-F_6 ($C_{142}H_{208}N_{46}O_{44}F_{36}$) = 3947.4 amu, detected mass = 3946.6 amu. The concentration of the peptides was determined by their absorbance at 275 nm due to the single tyrosine residue, using an extinction coefficient of 1420 cm^{-1} M^{-1}.

Circular Dichroism. Circular dichroism (CD) spectra of peptides were recorded with an Aviv 62DS spectropolarimeter at 25 °C. Mean residue ellipticities, [θ], were calculated using

$$[\theta] = \theta_{obsd}/10lcn,$$

where θ_{obsd} is the ellipticity measured in millidegrees, c is the molar concentration, l is the cell path length in centimeters, and n is the number of residues in the protein. To examine the unfolding of the peptide by GuHCl, two stock solutions were prepared containing 200 μM peptide (concentration of monomer) in 100 mM potassium phosphate buffer, pH 7.0, one with 8.0 M GuHCl and one without. The solutions were mixed in various proportions to obtain samples at different concentrations of GuHCl, and after equilibration for several minutes the ellipticity at 222 nm was measured.

Fluorescence Measurements. Fluorescence spectra of the ANS dye in the presence of peptides were measured using a Cary Eclipse fluorescence spectrophotometer with a cuvette with a path length of 5 mm. Peptides were titrated against a constant concentration of ANS (5 μM) in 100 mM potassium phosphate buffer (pH 7.0). The excitation wavelength was 370 nm.

Gel Filtration Chromatography. Peptides were subjected to chromatography using an FPLC system equipped with a Pharmacia "Peptide" Superose column equilibrated in 100 mM potassium phosphate buffer (pH 7.0). The flow rate was 0.25 mL/min, and the initial peptide loading concentration was 500 μM. Peptides were detected by their absorbance at 280 nm.

NMR Spectroscopy. Peptides were dissolved in 10 mM potassium phosphate buffer, pH 7.0 containing 10% D_2O at concentrations ranging between 1 and 10 mM. One-dimension proton and ^{19}F spectra were acquired on a Bruker 500 MHz spectrometer using standard solvent suppression pulse sequences where appropriate. Natural abundance ^{1}H-^{15}N HSQC spectra were acquired using standard pulse sequences on a Bruker 600 MHz spectrometer equipped with a cryogenic probe.

Analytical Ultracentrifugation. Sedimentation equilibrium experiments were performed using a Beckman XLA analytical ultracentrifugation equipped with scanning UV-visible optics.[79,80] Initial peptide concentrations ranged from 500 to 70 μM in 100 mM phosphate buffer, pH 7.0. The temperature was 298 K. The samples were centrifuged at 35 000, 38 000, 41 000, 44 000, and 47 000 rpm and were judged to have obtained equilibrium when successive radial scans were indistinguishable. The data were fitted to either monomer-*n*-mer equilibria or to a single species using the *Ultrascan* software package (B. Demeler, University of Texas Health Science Center at San Antonio; www.ultrascan.uthscsa.edu). Partial specific volumes were calculated using the method of Cohn and Edsall:[80] the partial specific volume of Leu and hFLeu were calculated as 0.778 and 0.454 $cm^3\ g^{-1}$; the partial specific volume of α_4-H was calculated as 0.74 $cm^3 g^{-1}$; the partial specific volumes of α_4-F_2, α_4-F_4, and α_4-F_6 were calculated as 0.71, 0.69, and 0.66 $cm^3\ g^{-1}$, respectively.

Curve Fitting. The denaturation profiles for the peptides were analyzed assuming a two-state equilibrium between unfolded monomeric peptide and folded, tetrameric bundle, with no significantly populated intermediates being present, as described previously.[24,50] Igor Pro software (Wavemetrics, Inc.) was used to fit the denaturation curves.

2.4. Results

Design and Synthesis of Model α_4 peptides. I have focused on the antiparallel 4-helix bundle structure as such proteins are widely distributed in Nature and have been the subject of previous *de novo* design efforts.[14,59,81] To better understand the effects of fluorination on the structure and stability of proteins, I decided to investigate the effect of progressively increasing the number of hFLeu residues in the peptide. This was accomplished by the design of two further peptides in the series, α_4-F_4 in which positions 10, 13, 17, and 20

are substituted with hFLeu and α_4-F_6 in which positions 6, 10, 13, 17, 20, and 24 are substituted with hFLeu. These peptides were synthesized manually using standard protocols and *t*-Boc-protected amino acids as described in Materials and Methods.

Comparison and Characterization of Peptides. The CD spectra obtained at neutral pH for all four peptides, α_4-H, α_4-F_2, α_4-F_4, and α_4-F_6, are shown in Figure 13. All the peptides exhibited spectra characteristic of extensively α-helical secondary structure with minima at 208 and 222 nm.

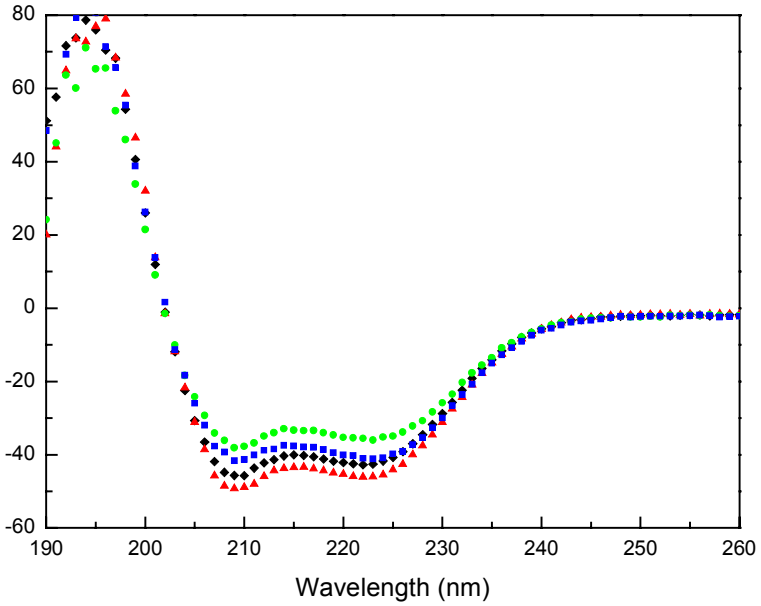
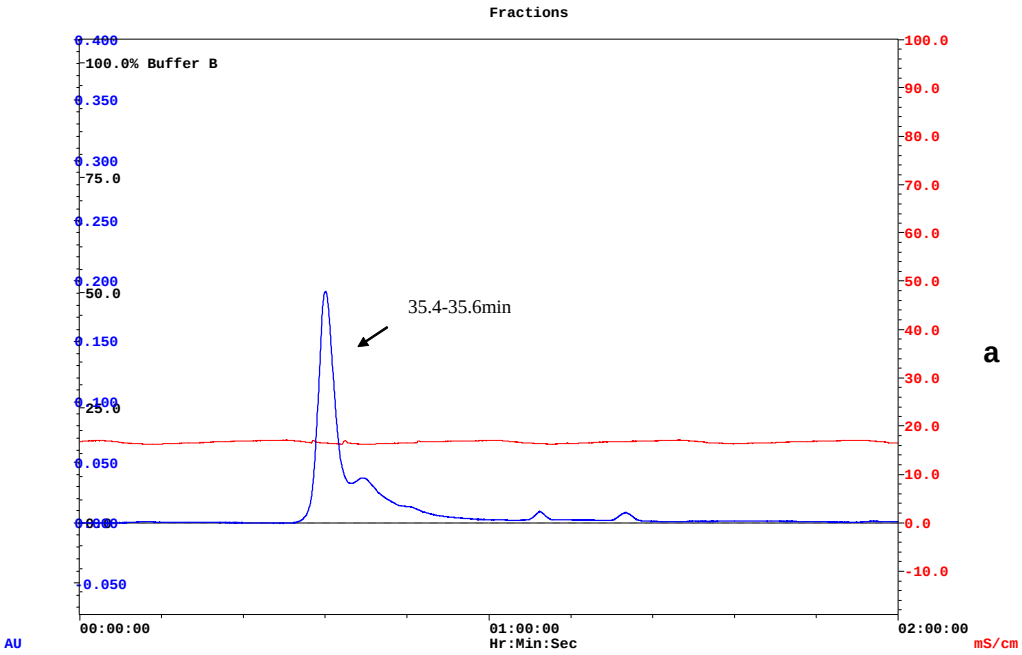


Figure 13 CD spectra of α_4-H (◆) α_4-F_2 (▲), α_4-F_4 (●), and α_4-F_6 (■) (peptide concentration, 30 μM) in 100 mM potassium phosphate buffer, pH 7.0.

Size Exclusion Chromatograhpy. The oligomerization state of the peptides was examined by gel filtration as shown in Fig. 14.



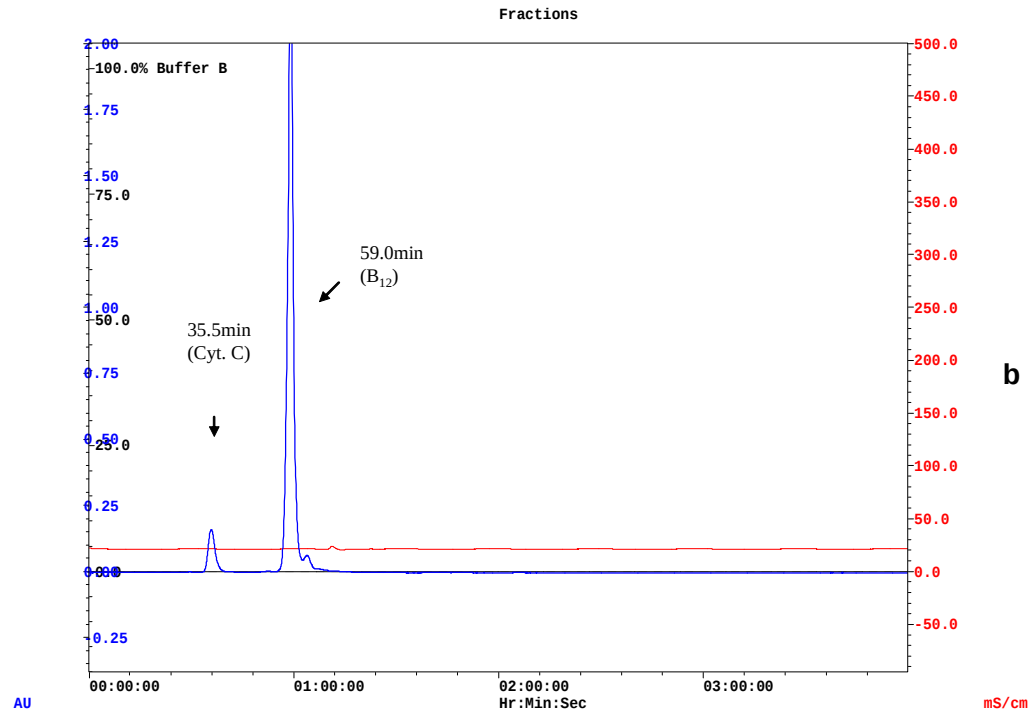


Fig. 14. FPLC spectra of α_4F_2 (**a**) and standard proteins (**b**). Standard proteins; Cytocrome C (12,500 D) and B12 (1352 D). Molecular weight of α_4F_2 is 3516 D as monomer and 14064 D as tetramer: 14064 D. Buffer solution is 20mM Mops, pH = 8.0, 100 mM KCl.

Both α_4-H (M_r = 3300) and α_4-F_2 (M_r = 3516) exhibited elution volumes very similar to that of cytochrome *c* (M_r = 12 500), indicating that they were adopting a predominantly tetrameric structure, as intended. Whereas cytochrome *c* eluted in a sharp symmetrical peak, characteristic of a monodisperse species, both peptides exhibited significant trailing of lower-molecular weight material, which is consistent with the peptides existing in equilibrium between tetrameric and monomeric species.

Analytical Ultracentrifugation. Analytical sedimentation equilibrium centrifugation was used to investigate the oligomerization states of the peptides as shown in Fig 15. The initial peptide concentrations varied between 70 μM and 500 μM. At high peptide concentrations the peptides sedimented as homogeneous species with apparent molecular weights that are close to those expected for a tetrameric structure. Fits of the curves assuming either a trimeric or pentameric structure for the peptides exhibited significant systematic deviation of the residuals from the data. At lower concentrations the sedimentation curves were better described by a monomer-tetramer equilibrium, however attempts to extract equilibrium constants from the fits yielded unreliable results.

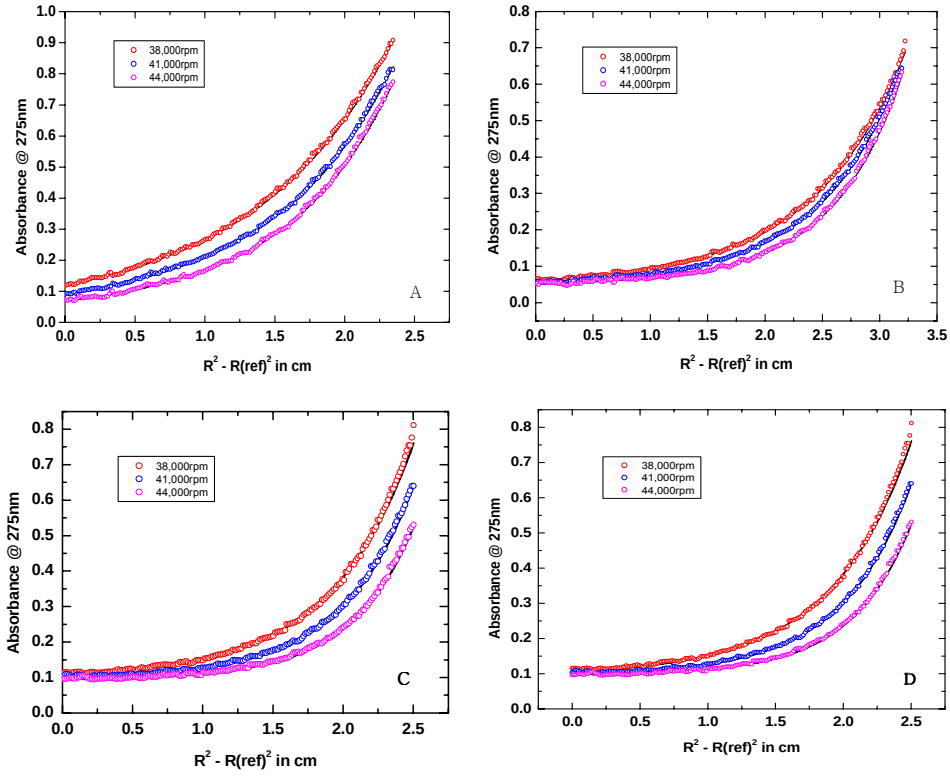


Fig. 15. Representative sedimentation equilibrium traces of α_4H (**A**), α_4F_2 (**B**), α_4F_4(**C**), and α_4F_6(**D**) for monomer-tetramer fitting. A 11,900 D (fitted), 13,200 D (cal. MW); B 13,390 D (fitted) 14.060 D (cal. MW); C 16,290 D (fitted) 14,926 D (cal. MW); D 17,190 D (fitted) 15,788 D (cal. MW).

Hydrophobic Dye Binding. To qualitatively assess how replacing two leucine residues in α_4-H with hFLeu in α_4-F_2 might affect the hydrophobic packing of the interior of the bundle, I investigated the binding of the hydrophobic dye 8-anilinophthalenesulfonic acid (ANS) to the peptides. ANS is commonly used as a probe for molten globule behavior in proteins because whereas well-folded proteins do not bind the dye strongly, those possessing molten hydrophobic cores bind ANS, typically with micromolar affinities, resulting in a large increase in fluorescence. When α_4-H was titrated against ANS, only a small increase in fluorescence was observed, which increased linearly across the concentration range that was used (Figure 16).

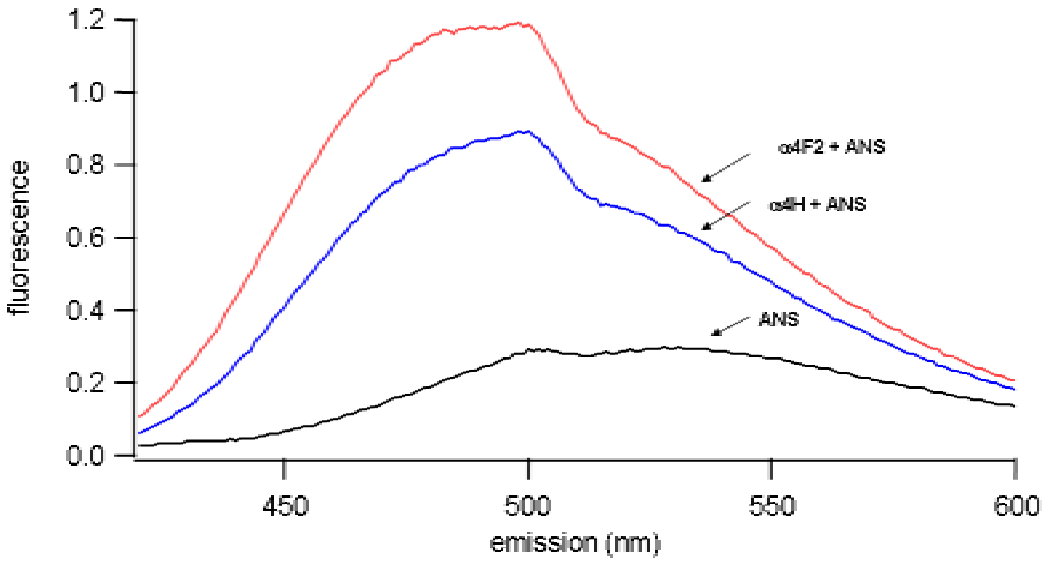


Fig. 16. Hydrophobic dye binding by peptides. (A) Fluorescence emission spectra of α_4-H and α_4-F_2 in the presence of 2 μM ANS, with the emission spectrum of ANS alone shown for comparison. The buffer was 100 mM potassium phosphate (pH 7.0), and the excitation wavelength was 370 nm.

This result indicates weak, nonspecific binding of ANS, rather than intercalation of the dye into the hydrophobic core of the peptide. This suggests that the core is well-packed, rather than molten globule-like. In this respect, α_4-H differs from Coil-LL, the peptide upon which it is based. Coil-LL has the same all-leucine core but was found to be molten globule-like. A similar result was obtained when α_4-F_2 was titrated against ANS (Figure 17). Although the fluorescence intensity is slightly higher with this peptide, there is no evidence for specific binding of ANS. As a positive control, the fluorescence spectrum of ANS (2 μM) was also recorded in the presence of 30 μM α-lactalbumin in 0.1 M KCl at pH 2.0, where it is known to adopt a molten globule state. Whereas ANS fluorescence is only increased 3-5-fold in the presence of the peptides, fluorescence was increased ~135-fold and the emission maximum was blue-shifted to 470 nm by α-lactalbumin. This further suggested the peptides were not molten globule-like.

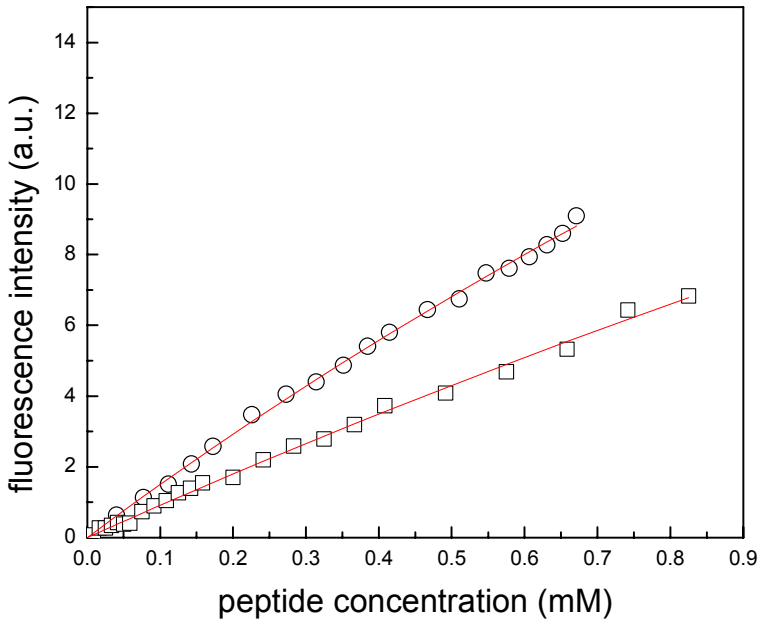


Fig. 17. ANS binding assay of α_4H (□) and α_4F_2 (○) into a 5 μM solution of ANS (Excitation: 370nm, Emission: 460nm)

Denaturation Experiments. The GuHCl-induced unfolding of the peptides was followed using CD spectroscopy by monitoring the ellipticity of the peptides at 222 nm as a function of increasing GuHCl concentration. These peptides exhibit well-defined unfolding transitions, as shown in Figure 18, that appear to be well-fitted by assuming a two-state equilibrium between unstructured monomeric peptides and folded tetrameric helical bundle.

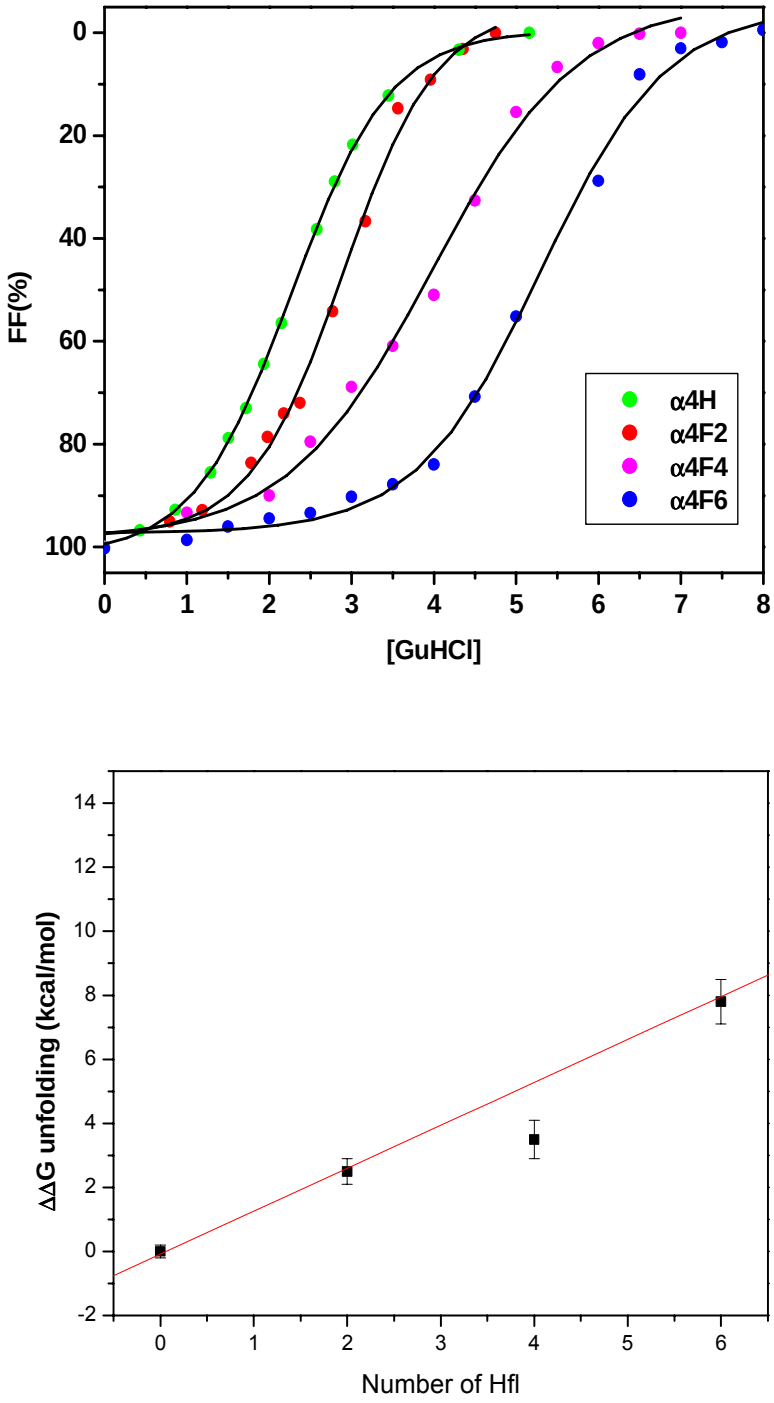


Figure 18 Unfolding of peptides in GuHCl. (left) Plots of ellipticity versus GuHCl concentration for α_4-H (◆), α_4-F_2 (▲), α_4-F_4 (●), and α_4-F_6 (■).

As shown in Figure 18 and Table 1, there is an initial increase in the stability of the 4-helix bundle of about 2.5 kcal/mol upon incorporating the first

two hFLeu residues corresponding to an additional stabilization approximately 0.3 kcal/mol per hFLeu residue upto 6 hFLeu.

Table 1. Summary of thermodynamic parameters determined from GuHCl-induced unfolding of peptides.

peptide	G_{unfold} (kcal/mol)	m ((kcal/mol)/M_{GuHCl})
α_4-H	20.3 ± 0.2	2.4 ± 0.1
α_4-F_2	22.8 ± 0.3	2.6 ± 0.2
α_4-F_4	23.8 ± 0.6	1.6 ± 0.1
α_4-F_6	28.1 ± 0.6	1.9 ± 0.1

NMR Spectroscopy. The CD and ultracentrifuge data described above established that the substitution of hFLeu for Leu in the α_4 peptides does not result in gross structural changes to the peptides in that they remain tetrameric and extensively helical. To obtain more detailed information on the effects of fluorination on the structure of the peptide, we turned to NMR spectroscopy. Initially, 1-D proton NMR spectra of α_4-H, α_4-F_2, α_4-F_4, and α_4-F_6 were recorded in 10% D_2O at pH 7.0. These spectra are shown in Figure 19. All the spectra are reasonably well dispersed and typical of folded small proteins.

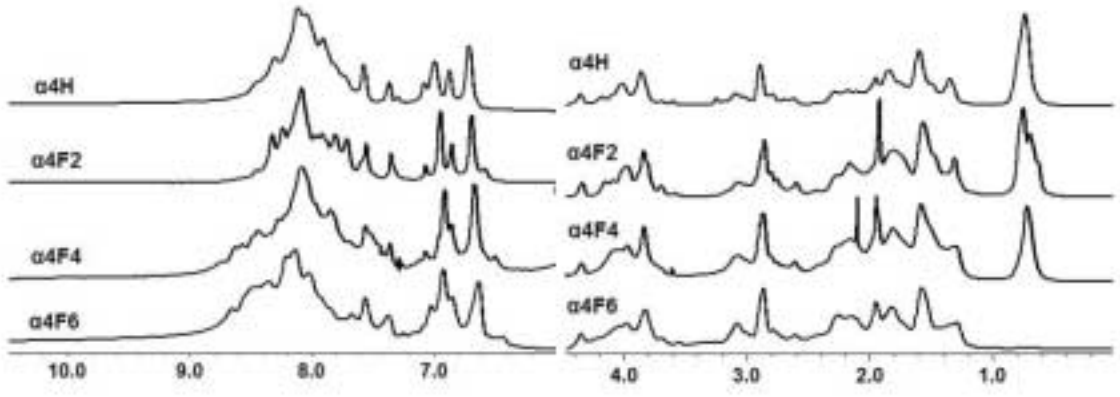


Figure 19 Proton NMR spectra of the amide and aromatic region (left panel) and aliphatic region (right panel) of the α_4 peptides. Spectra from top to bottom are of α_4-H, α_4-F_2, α_4-F_4, and α_4-F_6 and were recorded at 25°C, pH 7.0 in 10% D_2O.

To better compare the spectral changes in the amide region caused by repacking the hydrophobic core with hFLeu, we recorded natural abundance 2-D ^{1}H-^{15}N HSQC spectra of α_4-H and α_4-F_6. The spectra were recorded at 600 MHz at room temperature in 10% D_2O, at pH 7.0, with peptide concentration ~10 mM and are shown in Figure 20. Both spectra are characteristic of well-structured small proteins. Many of the peaks in the two spectra either overlap one another, as might be expected given the similarity between both the sequences and the overall structures of the two peptides. However there are a number of differences between the two spectra.

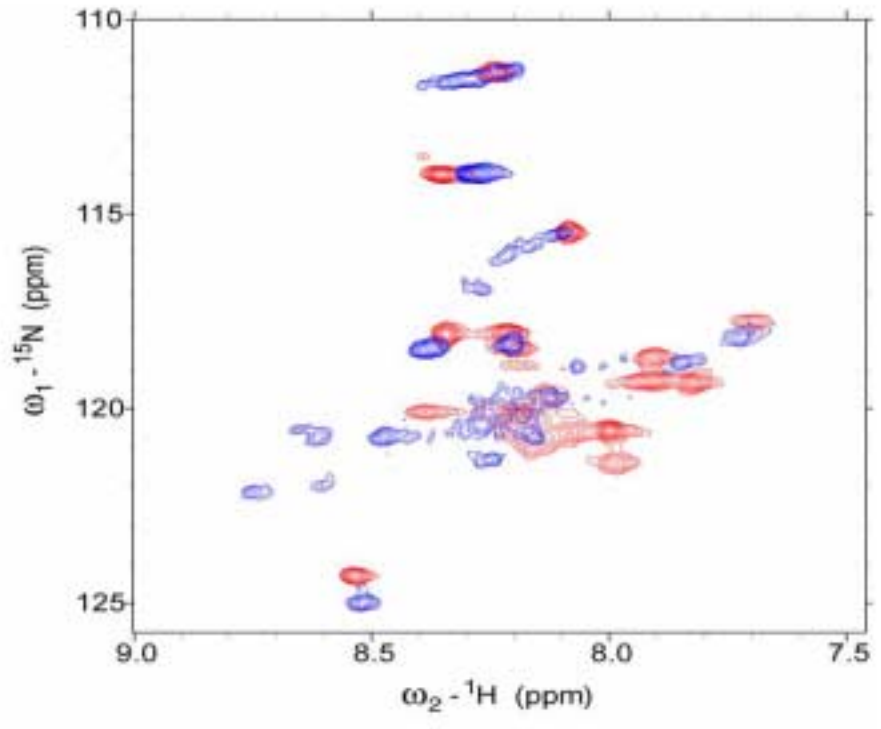


Fig. 20. Overlaid 2D ^{1}H-^{15}N HSQC NMR spectra comparing the amide regions of α_4-H (red) and α_4-F_6 (blue). The spectra were recorded using unlabeled (natural abundance ^{15}N) peptides at 25° C, pH 6.0 in 10% D_2O using a cryogenic probe (These spectra were taken by Dr. Qi Zhang).

Overall, 18 amide peaks are clearly resolved in the α_4-H spectrum, whereas for α_4-F_6 ~24 amide resonances are visible, which is in better agreement with expected 26 backbone amide resonances. The large number of resonances and greater chemical shift dispersion tends to suggest that the α_4-F_6 peptide is more structured than α_4-H. There are also 5-6 peaks that exhibit clear differences in chemical shifts, indicating that these sites experience different chemical environments. These are evident in the α_4-F_6 spectrum between 8.8 and 8.5 ppm, and in the α_4-H spectrum between 7.8 and 8.0 ppm in the ^{1}H dimension. These differences could either arise from differences in the conformation of the peptides or be a direct effect on the chemical shift of some

residues caused by the substitution of hydrogen by fluorine. Last, I have used ^{19}F NMR spectroscopy to examine the effect of fluorination on the hydrophobic core of the peptides. ^{19}F nucleus exhibits a wide range of chemical shifts that are sensitive to changes in environment, a property that has been exploited in various studies on protein structure. One-dimensional ^{19}F spectra were recorded for α_4-F_2, α_4-F_4, and α_4-F_6 at pH 7.0 at 25 °C and referenced to trifluoroacetate as an internal standard. To aid in interpreting these spectra, the ^{19}F spectrum of a simple dipeptide, hFLeu-Ser, was also recorded under the same conditions. The spectra are shown in Figure 21.

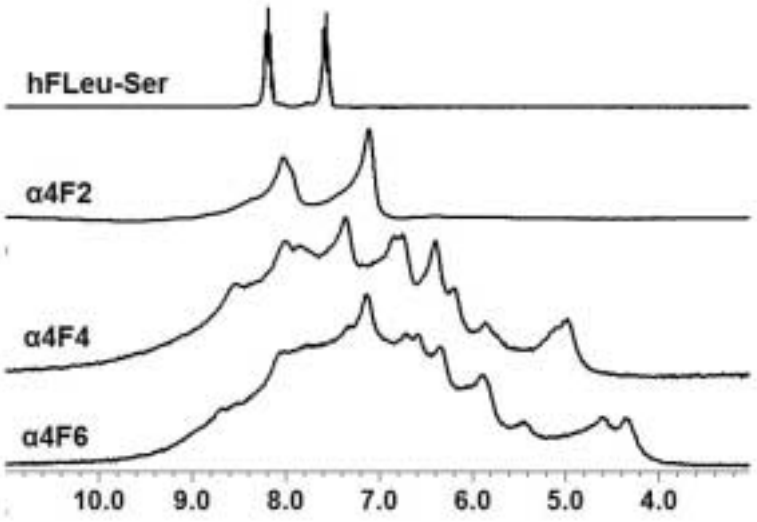


Fig. 21. ^{19}F spectra of (from top to bottom) hFLeu-Ser dipeptide, α_4-F_2, α_4-F_4, and α_4-F_6 recorded at 25 °C and neutral pH. All the spectra are referenced to TFA.

The spectrum of hFLeu-Ser shows two resonances for the diastereotopic trifluoromethyl groups at 7.8 and 8.2 ppm, which are split into multiplets due to coupling between the fluorines and the γ-hydrogen. The ^{19}F spectrum of α_4-F_2 is quite simple, showing only two peaks that presumably correspond to the two diastereotopic CF_3 groups. The peaks are broadened, as expected for fluorine incorporated in a macromolecule, and somewhat asymmetric-this may be due to

the two hFLeu residues being in slightly different environments with distinct rotational averaging properties. The centers of the peaks are shifted to slightly higher field (7.1 and 8.0 ppm relative to TFA) than in the dipeptide, consistent with the fluorines now being in a more hydrophobic environment. The incorporation of additional hFLeu residues dramatically changes the spectra of α_4-F_4 and α_4-F_6. The spectra become much more complex, even allowing for the fact that there are more fluorinated residues in the peptides. Most interestingly, the range of chemical shifts increases to cover the region from 4 ppm to ~9 ppm in the case of α_4-F_6. These changes suggested that the additional hFLeu residues lead to a situation in which side chains adopt multiple well-defined conformations that interchange slowly at the NMR time scale. To gain more insight into the unexpectedly complex ^{19}F spectra of α_4-F_4 and α_4-F_6 the spectra of all three peptides were recorded at 50 and 70 ° C (Figure 22).

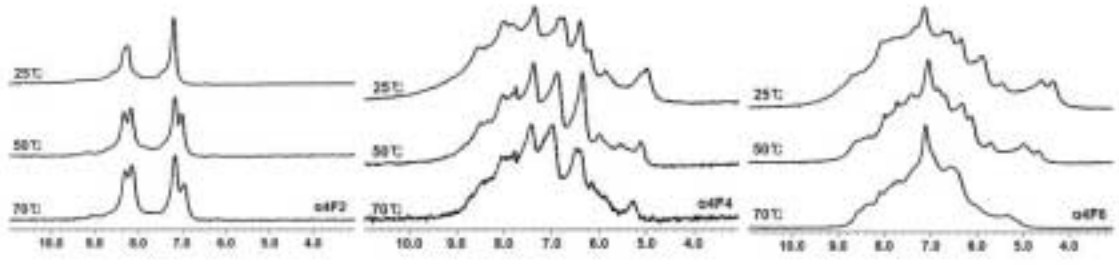


Fig. 22. ^{19}F spectra of α_4-F_2, α_4-F_4, and α_4-F_6 recorded at 25, 50, and 70 ° C. (Top) α_4-F_2; (middle) α_4-F_4; (bottom) α_4-F_6. All the spectra are referenced to TFA.

At higher temperatures the spectra of all the peptides become sharper, as would be expected due to decreased solvent viscosity, and the spectral envelope is narrowed. For α_4-F_2 the two sets of hFLeu resonances become much sharper and are clearly resolved at 50 ° C. The spectra of both α_4-F_4 and α_4-F_6 become somewhat sharper and some of the peaks begin to coalesce; this is more

noticeable for the spectrum of α_4-F_4. The dispersion becomes smaller, suggesting faster exchange (relative to NMR time scales) among the "frozen out" side chain conformations. None of the peptides undergo thermal unfolding, as judged by CD spectroscopy below 95°C. Therefore these spectral changes are most consistent with the hydrophobic core becoming more dynamic at higher temperatures.

Thermodynamic Study of α_4H and α_4F_6 for ΔCp, ΔH and ΔS.

Thermodynamic studies of a *de novo* designed peptide, α_4H, and its fluorous analog peptide, α_4F_6, can provide insight into peptide structure and stability as well as a comparison of physical properties of the peptides. Both peptides were highly stable. They showed no sign of unfolding even when the temperature was increased up to 95°C without adding denaturant. By introducing different concentrations of GuHCl, thermodynamic parameters, such as ΔH, ΔS and ΔCp, can be obtained from thermal denaturation studies.[50]

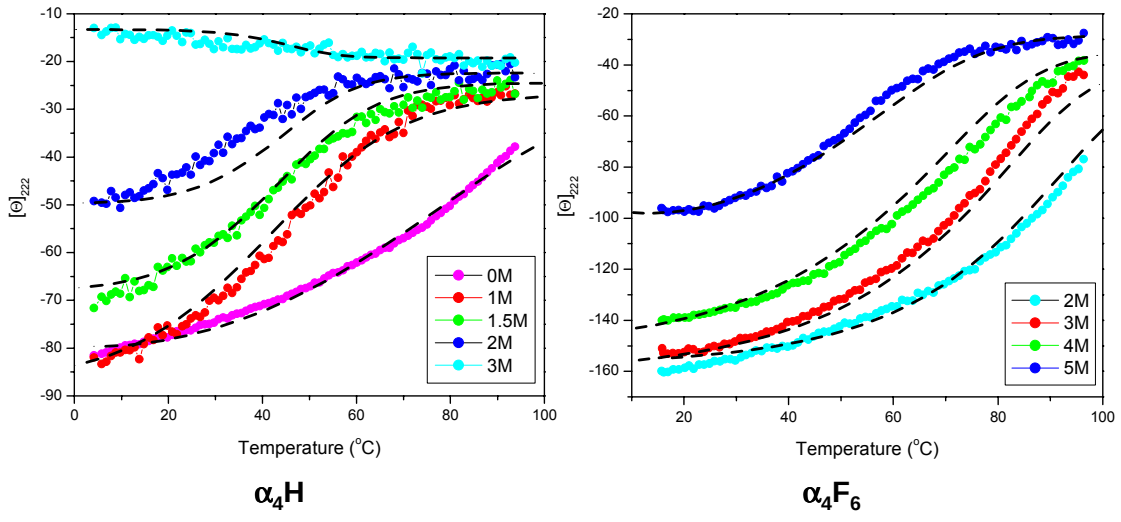


Fig. 23. Thermal unfolding curves and their fits of α_4H (200μM peptide, 0.1cm CD cell) and α_4F_6 (50μm peptide, 1cm CD cell) at various concentration of GuHCl. (--) stands for the fits at the concentrations of GuHCl.

Thermal unfolding curves and their fits were obtained at the various concentration of Gu.HCl by globally fitting the data to the Gibbs-Helmholz equation assuming temperature-independent ΔCp. Dr. Morris M. Slutsky performed all the work on global fitting with these data. Using the Gibbs-Helmholz equation,

$$\Delta G = (\Delta H - T * \Delta S) + \Delta Cp * ((T - T_0) - T * \ln (T/T_0)), \quad 1$$

this gives K = exp (-ΔG/RT) whereas T_0 is 298 K for simplicity. To compute the fraction unfolded for each point of the fit, a Newton's method solution of the equation 4U^4 + KU – KP = 0, where [U] is the concentration of unfolded monomer and [P] is total concentration was employed.

Table 2. Thermodynamic Parameters of the α_4 Peptides.

Peptides	ΔH (kcal mol^{-1})	ΔS (cal mol^{-1} K^{-1})	ΔC_p (cal mol^{-1} K^{-1})
α_4H	4.6 ± 0.4	- 37.3 ± 1.3	226 ± 19
α_4F_6	46.1 ± 16.0	58.7 ± 48.6	171 ± 251

Thermal unfolding curves were fitted globally using the Gibbs-Helmholtz equation to 0, 1, 1.5, 2, 3M GuHCl data for α_4H and to 2, 3, 4, 5M GuHCl date for α_4F_6 using 25°C as the reference temperature (Fig. 23). ΔC_p value of α_4H is 226 ± 19 cal $mol^{-1}K^{-1}$and 2.1 cal $mol^{-1}K^{-1}residue^{-1}$ (Table 2). Values were calculated from the data ranging from as high as 20.7 cal deg-1$mol^{-1}residue^{-1}$ for the VBP dimeric coiled coil[82] to as low as 3.8 cal $mol^{-1}K^{-1}residue^{-1}$ for the GCN4 dimeric coiled coil[83] has been reported. Fairman et al[84] reported a value from 8.8 to 14.8 cal $mol^{-1}K^{-1}residue^{-1}$ for tetrameric models. The value of 2.1 cal $mol^{-1}K^{-1}residue^{-1}$ for α_4H is lower than the reported value of 3.8 cal $mol^{-1}K^{-1}residue^{-1}$. The reliability of this value is however in question since it is rather difficult to accurately determine ΔC_p from the Gibbs-Helmholtz equation unless cold denaturation is observed which provides more constraint on fitting. Furthermore, the data for α_4F_6 was found to be not well titrated by the model as evidenced by

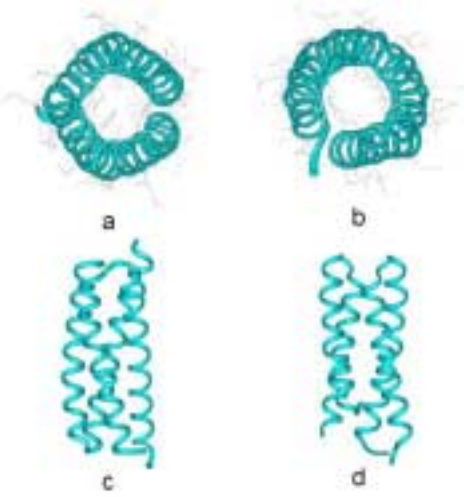
the large errors in the fits. No significance can be attached to the data. ΔG of α_4H calculated from thermal denaturation is 15.6 kcal/mol, which is smaller than 20.3 kcal/mol calculated from GuHCl denaturation experiments. Despite large errors on fitting for α_4F_6, ΔG of α_4F_6 obtained from thermal unfolding, the experimental value of 28.1 kcal/mol, is relatively consistent with the value, 28.1 kcal/mol, from GuHCl experiments. Further experiments will need to be performed to get more reliable data.

Chapter 3 Designing Fluorinated ROP Proteins

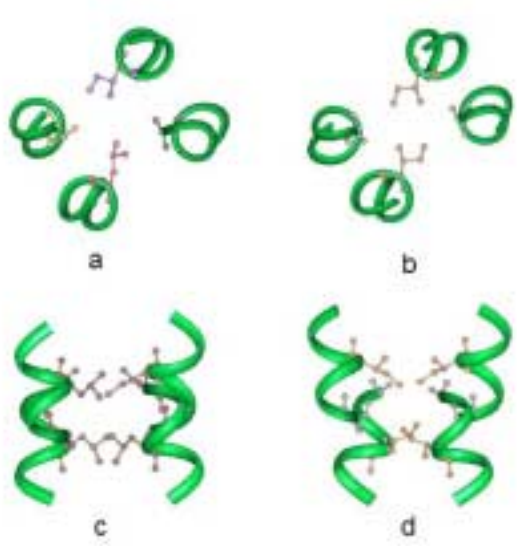
3.1. Introduction

ROP is an attractive model protein to test design concepts in four-stranded coiled-coils and has been used in precious protein design studies. It is a small and soluble protein with a well-defined, regular hydrophobic core.[34] ROP is an RNA binding protein that is involved in regulation of the copy number of ColE1 plasmid in *Escherichia coli*. The anti-parallel, homodimeric four-helix bundle topology of, ROP, is well characterized from the crystal structure,[34] NMR structure[34] and systematic mutations to the RNA-binding[9] and hydrophobic core.[10,85] Biological RNA binding activity of ROP variants can be easily investigated by using a simple electrophoretic mobility shift assay,[86,87] and a cell-based screen.[88]

ROP variants in which the hydrophobic core has been redesigned have been investigated and characterized. Residues at the 'a' and 'd' positions of the heptads were repacked with Ala, Leu, Met, Ile, or Val. Ala_2Leu_2-6 (ROP variant having 6 layers of the hydrophobic core packed with 2 Ala and 2 Leu) and Ala_2Leu_2-8 (ROP variant having 8 layers of the hydrophobic core packed with 2 Ala and 2 Leu) have enhanced thermal stability and native-like thermodynamic properties with similar wild-type RNA binding affinity and Ala_2Met_2-8 (ROP variant having 8 layers of the hydrophobic core packed with 2 Ala and 2 Met) maintains native-like properties, but has reduced thermal stability. Leu_4-8 variant (ROP variant having 8 hydrophobic cores packed with 4 Leu) has an over-packed core and show resistance to thermal and chemical denaturation and non-nativelike thermodynamic properties. Ala_4-8 (ROP variant having 8 layers of the hydrophobic core packed with 4 ala) and Ala_2Val_2-8 (ROP variant having 8 layers of the hydrophobic core packed with 2 ala and 2 val) have a small core volume so that they fail to fold 6-7.[10,85,89]



A) Wt-ROP (a: top view, c: side view; PDB:1ROP) and Ala_2Ile_2-6 ROP variant (b: top view, d: side view; PDB:1F4N)



B). One layer of wild-type ROP (a: top view, c: side view) and Ala_2Ile_2-6 ROP variant (b: top view, d: side view). Hydrophobic core of ROP (a and c) and Ala_2Ile_2-6 mutant (b and d) was dissected to show their packing structure.

Fig. 24. Crystal structure of wild-type ROP and Ala_2Ile_2-6 ROP variant.

Fig. 24 shows the packing structures of wild-type ROP and a repacked Ala_2Ile_2-6 (ROP variant having 6 layers of the hydrophobic core packed with 2 Ala and 2 Ile) with Ala at the 'a' positions and Ile at the 'd' positions of the heptads. Even though the packing topology for ROP Ala_2Ile_2-6 is only slightly

altered by substitution of Leu with Ile at the 'd' position of the hydrophobic core as compared to Ala_2Leu_2-6 or Ala_2Leu_2-8, which do bind RNA, this Ala_2Ile_2-6 variant does not have biological activity and its structure is quite different from wild-type ROP due to the structural changes. The four-helix bundle is still all anti-parallel but one of the protomers has flipped 180°. The structural variation arises due to the diagonal interaction between isoleucines in the core.[89]

3.2. Model Studies of Short-ROP Proteins

It would be interesting to investigate how fluorine changes the physical properties of real natural proteins such as ROP protein, as well as their biological activities. Short versions of repacked ROP protein, 35AL (ROP variant having 35 amino acids packed with Ala and Leu at the 'a' and 'd' positions of the core respectively), was first designed from the sequence of wild-type ROP proteins. This model study will hopefully provide fundamental information about the physical properties of short, fluorinated and dimeric folded peptides (Fig. 25).

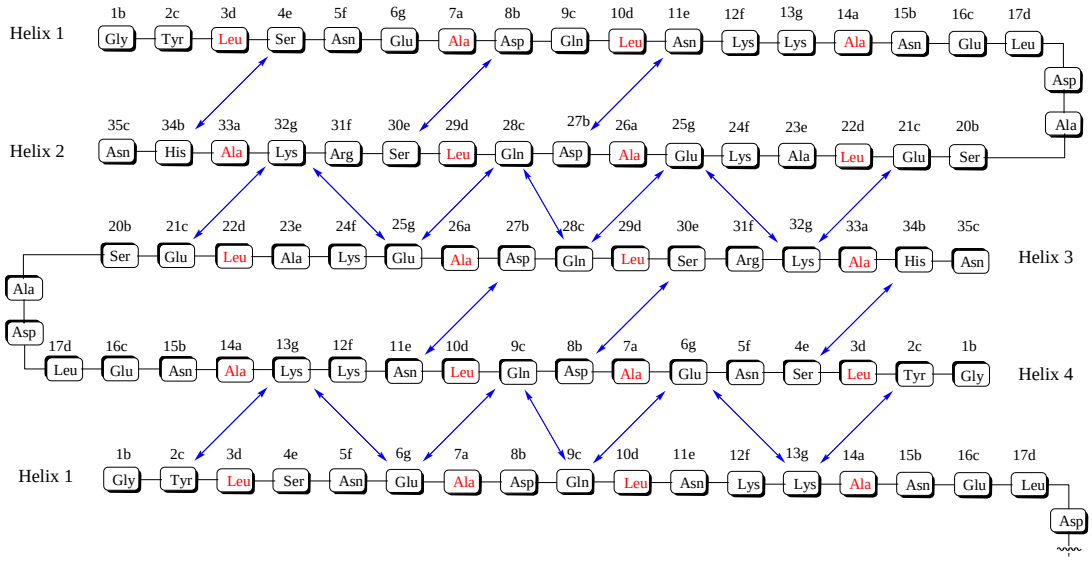


Fig. 25 Sequence of 35AL with showing possible electrostatic and hydrogen bonding interactions.

To design these peptides, residues at the ‘a’ and ‘d’ positions of the heptads were repacked with Ala and Leu residues respectively to maintain native-like properties and then, Leu residues at ‘d’ positions of heptads will be substituted with hFleu to introduce fluorous amino acids into the protein core as shown in Fig. 26.

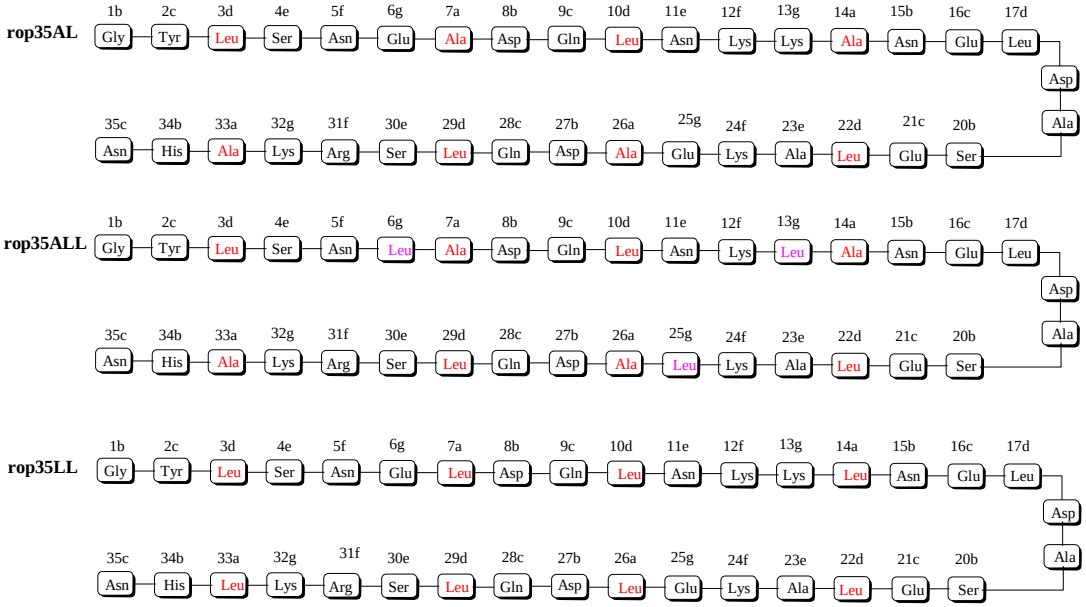


Fig. 26 Sequences of 35AL, 35ALL and 35LL.

A series of ROP-like peptides has been designed to be folded in anti-parallel dimer topologies. ROP variants, 35AL and 35ALL, have Ala and Leu at the hydrophobic ‘a’ and ‘d’ positions that are expecting to be packed into the core of the ROP-like fold. 35ALL was designed to replace Leu at three ‘g’ positions of heptad of 35AL to test the idea that hydrophobic residues at either ‘e’ or ‘g’ positions near the hydrophobic core stabilize the structure. In order to enhance the stability of the ROP peptides, 35LL, which have all Leu at the hydrophobic core, was designed similar to Leucine zipper. Three variants (35AL, 35ALL and 35LL) were synthesized and characterized their physical properties by CD and analytical ultracentrifugation.

The sequence is based on a coiled coil heptad repeat Ac-[Ya-Xb-Xc-Ld-Xe-Xf-Xg]$_2$-turn-[Ya-Xb-Xc-Ld-Xe-Xf]$_2$-$CONH_2$, where Y is either Ala or Leu and turn sequences are -Asp-Ala-. The overall topology is designed to be similar to that of *Escherichia coli* protein ROP. Although the sequences of the turn are not the primary topological determinant, near the turn, the sequence of the repacked ROP protein, Ala_2Leu_2-8, was conserved except that a Cys was changing to a Ser near the loop area which is -Asn-Glu-Leu-Asp-Ala-**Ser**-Glu-.

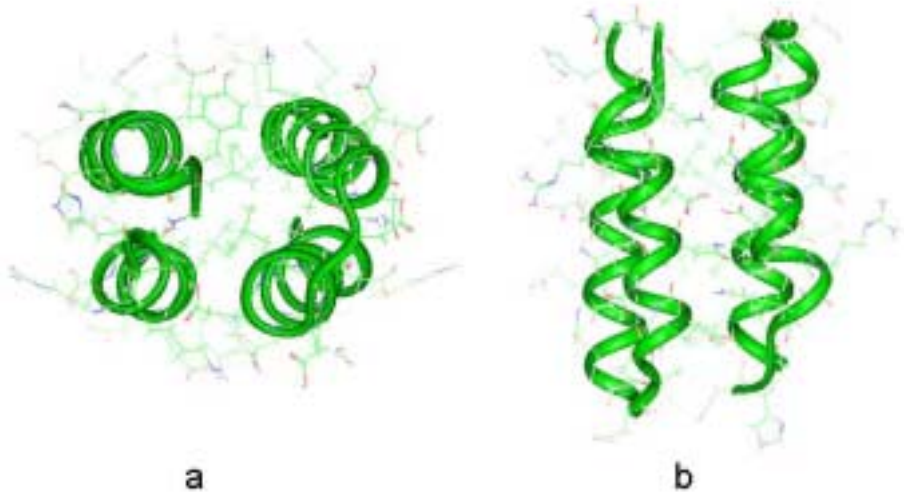
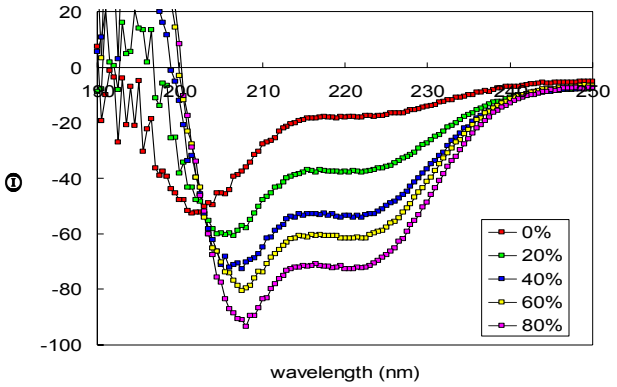


Fig. 27. Structure of 35AL derived from energy minimization using insight II program; A) top view B) side view.

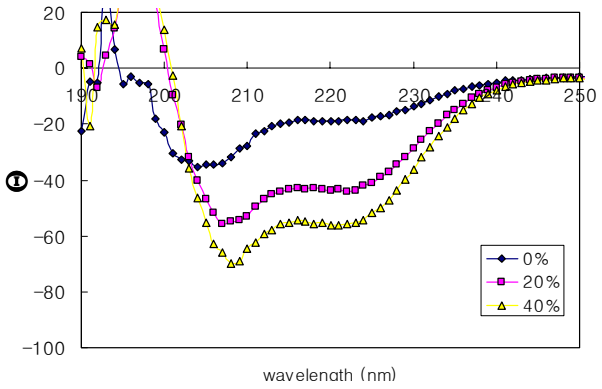
The sequence of 35AL was modeled on the structure of ROP starting with the coordinates of wild-type ROP. In this model, hydrophobic ‘a’ residues face ‘d’ residues in the core layer and there are 4 hydrophobic layers and a possible 3 intramolecular-electrostatic interactions between ‘b-e’ residues, which potentially help the protein fold. There are also 7 possible intermolecular electrostatic interactions between ‘g-c’, ‘g-g’ and ‘c-c’ residues to stabilize the bundles in anti-parallel pattern. My first designed model ROP, 35AL, derived from energy minimization using Insight II is shown in Fig. 27. In order to design a model short-ROP 35AL, I downloaded crystal structure of wild-type ROP (PDB:1ROP) to use it as a template, I replaced Ala and Leu hydrophobic residues at ‘a’ and ‘d’ positions of heptad. Afterward, hydrophilic residues were placed at the ‘e-g’ and

'b-c' positions of heptads to form an anti-parallel dimer. The sequence of 35AL was all re-designed except for the turn sequence of ROP, -Asn-Glu-Leu-Asp-Ala-Ser-Glu-.

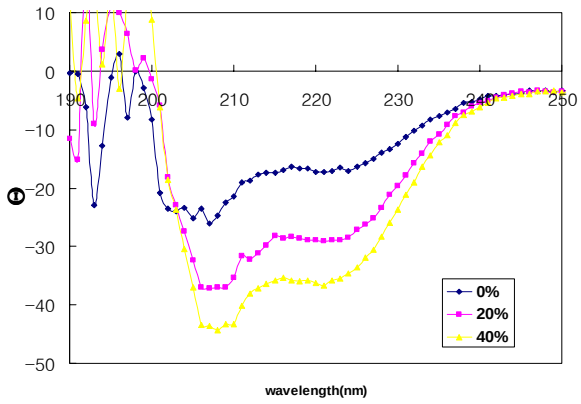
Comparison and Characterization of Peptides. The CD spectra obtained at neutral pH for all three peptides, 35AL, 35LL and 35ALL, are shown in Fig. 28. The three ROP35 peptides were synthesized by standard BOC chemistry and characterized by CD.



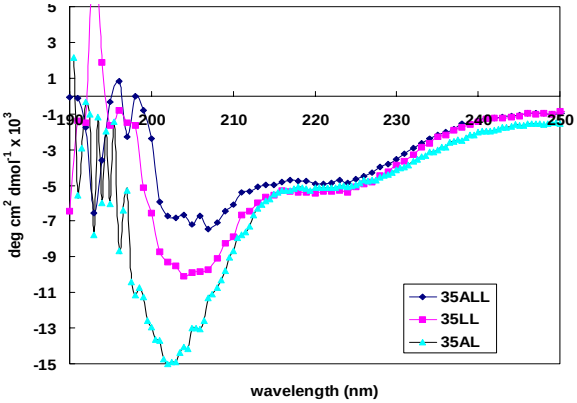
A (35AL)



B (35LL)



C (35ALL)



D (All three peptides)

Fig. 28. C.D. spectra of 35AL, 35LL and 35ALL in 10 mM sodium phosphate buffer, pH 7.0 and all three peptides concentrations were 100μM. A: C.D. spectra of 35AL in various concentration of TFE. B: C.D. spectra of 35LL in various concentration of TFE. C: C.D. spectra of 35ALL in the presence of TFE. D: C.D. spectra of all three peptides 35AL, 35LL and 35ALL without TFE.

All three peptides have random coil content and ~ 14% helical content calculated from the theoretical MRE value measured at 222nm. The percentage of helicity of the peptides may be approximated as % helicity = Θ/[40000/(1-2.5/n)], when n equals the number of residues.[90] α-Helical propensities of the peptides were significantly enhanced up to 72% in the presence of 80% TFE. The turn sequence has 7 residues out of 35 residues, which may not be helical and count about 20% of the whole chain.

Analytical Ultracentrifugation. Analytical sedimentation equilibrium centrifugation was used to investigate the oligomerization states of the peptides as shown in Fig 29. The initial peptide concentrations varied between 75 μM, 150 μM and 300 μM.

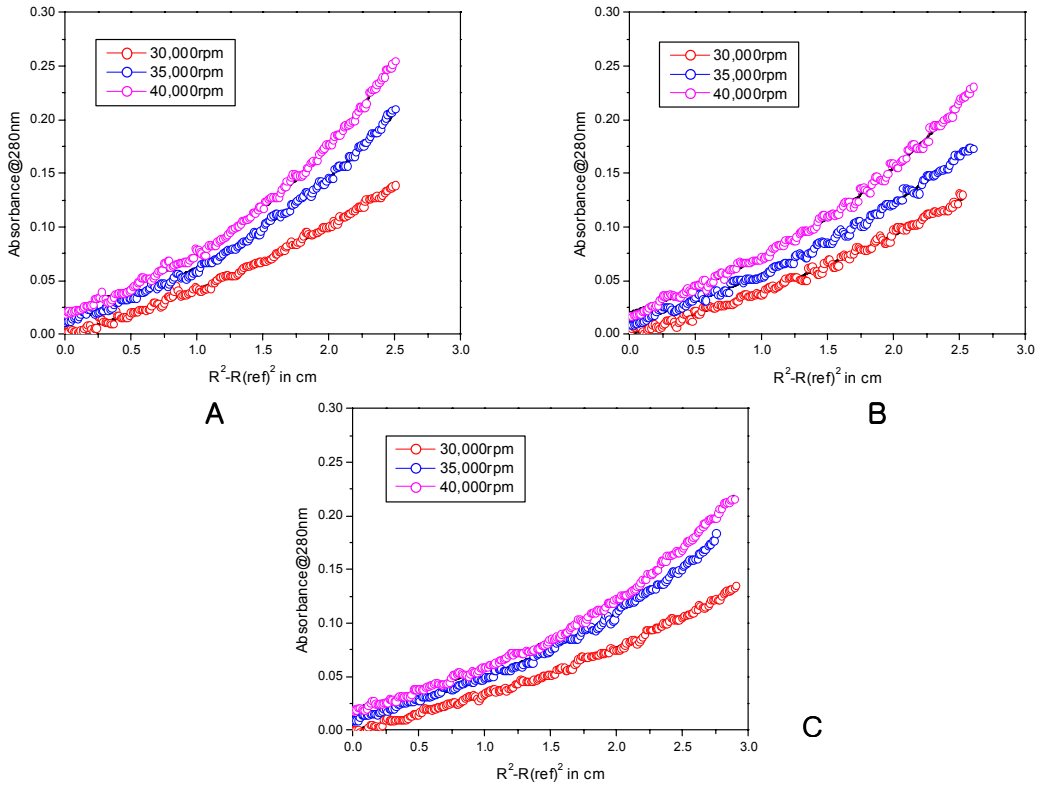


Fig. 29. Representative sedimentation equilibrium traces of 35AL (**A**), 35LL (**B**) and 35ALL (**C**) for monomer-dimer fitting with their fits(-line) at 150μM peptide concentration.

At lower concentrations (75μM) the sedimentation curves were excluded because they were unsuitable to fit. The averaged molecular weights of the three peptides were obtained by fitting the data as a single species to give 6087 ± 15 Da (calc. 3913.2 Da as a monomer) for 35AL, 4821 ± 13 Da (calc. 4081.5 Da as a monomer) for 35LL, 5886 ± 12 Da (calc. 3867.3 Da as a monomer) for 35ALL. The apparent molecular weights of 35AL, 35LL and 35ALL are all higher

than expected for a monomer, which imply that they are in monomer-dimer equilibrium; however their association interactions are apparently weak at the concentration used in the experiment. Fits to monomer-dimer equilibrium did not yield reliable numbers for the association constants.

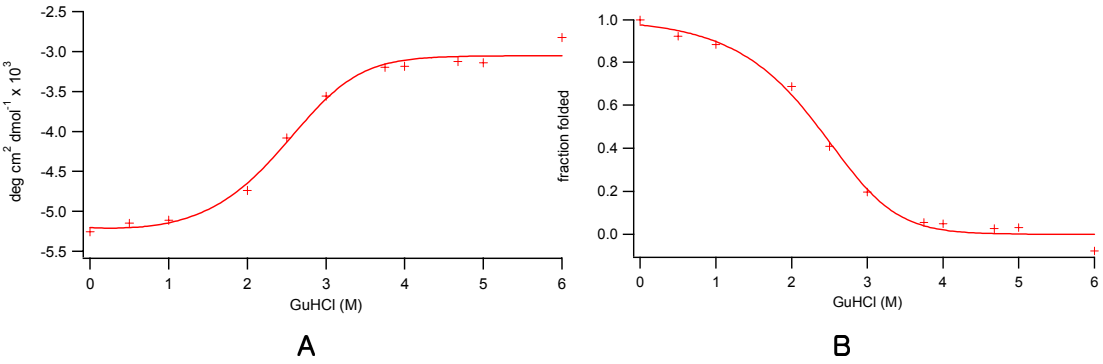


Fig. 30. Unfolding of 35ALL peptide in GuHCl. Unfolding was followed by C.D. spectroscopy by monitoring changes in ellipticity at 222 nm, the buffer was 100 mM sodium phosphate, pH 7.0 and the peptide concentration was 50 μM. A: Plot fit to MRE data. B: Plot fit to fraction folded. Temp = 298K, ΔG = 9.7 ± 1.9 kcal/mol, mDenat = 1.6 ± 0.6 kcal/mol.

The unfolding of the ROP35 series of peptides was investigated by titrating the peptides with GuHCl and following the decrease in ellipticity at 222 nm. Only the ROP 35ALL peptide showed a cooperative folding transition characteristic of a properly folded protein. The data are shown in Fig. 30.

This GuHCl denaturation data was taken by Dr. Roberto de la Salad-Bea and I have fitted the data to a monomer-dimer equilibrium model. Fitting the unfolding curve assuming equilibrium between dimeric folded protein and monomeric unfolded protein yielded a free energy of unfolding of 9.7 kcal/mol ± 1.9 kcal/mol, which is within the range seen for the $\Delta G_{unfolding}$ of many natural proteins.

3.3. Antiparallel Four-Helix Bundle Protein for Fluorous ROP.

Despite a large number of technical advances in the synthesis of peptides and proteins, it is still challenging to synthesize such big macromolecules by solid-phase peptide synthesis. One of the most important developments in the synthesis of the large peptides (over 50 amino acids) is the chemical ligation method. Native chemical ligation is an important step toward the general application of chemistry to proteins. Chemoselective reaction of two unprotected peptide segments gives an initial thioester-linked species. Spontaneous rearrangement of this transition intermediate yields a full-length product with a native peptide bond at the ligation site.

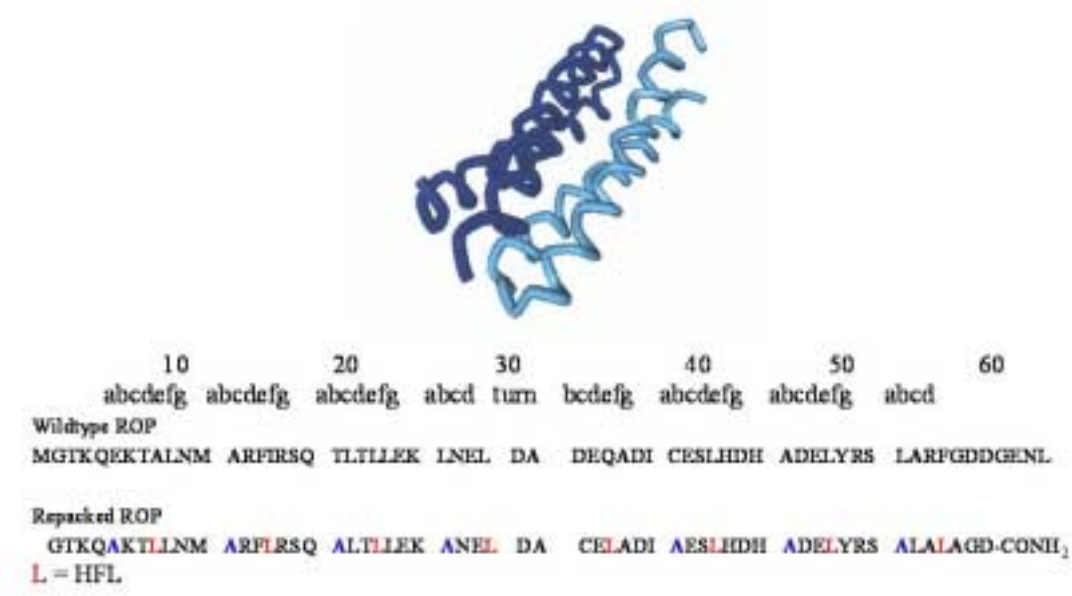


Fig. 31. The sequence of the wild type ROP and repacked ROP (E.coli repressor of primer) protein.

As shown in Fig. 31., wt-ROP is a relatively large protein, which has 63 amino acids so that it may be easier to synthesize the protein by using the native chemical ligantion (NCL) method. To test our long term goal of understanding how fluorine changes the physical properties of real natural proteins such as ROP protein, as well as their biological activities, first I repacked wt-ROP with Ala

and Leu because this repacked ROP variant's biological RNA binding activity has been known to be conserved. And then all of the Leu at the 'd' positions of the cores of the repacked ROP will be replaced with hFL for fluorinated ROP to compare their properties in future.

For the test reaction, short length of N-termial peptide (sequence 24-31) and C-terminus peptide (sequence 32-58) were synthesized chemically. The synthetic segment, peptide 1 (N-terminus peptide), which contains a thioester at the carboxyl group, undergoes nucleophilic attack by the side chain of the Cys residue at the amino terminal of peptide 2 (C-terminus peptide) to give a desired peptide, ROP5 (24-58).

A common required element of native chemical ligation is a cysteine residue at the ligation site. Aspartic acid (D_{32}) between the loop and the second sequence (32-58) was choosen for mutation with cysteine since it is known to be a conservative mutation.

The first sequence (24-31) was prepared by conventional Boc-based protocol since thioester linker is labile to Fmoc-based chemistry, which uses base to deprotect Fmoc groups. The repeated exposure of the resin-bound peptide to base results in cleavage of thioesters. However, the second sequence can be achieved by automated Fmoc-based protochol unless there are any hFL at d postions since it is not easy to prepare peptide with hFL by Fmoc-based chemistry

3.3.1. Materials and Methods

Materials. MBHA resin LL(100-200 mesh).HCl (for Boc SPPS, substitution: 0.67 mmole/g), Rink Amide resin (for Fmoc SPPS), Fmoc-protected and *t*-Boc-protected amino acids, and 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetra-methyluronium hexafluorophosphate (HBTU) were purchased from NovaBiochem. Peptide synthesis grade *N*-methylpyrrolidinone (NMP), *N*,*N*-dimethylformamide (DMF), and trifluoroacetic acid (TFA) and ACS grade *N*,*N*-diisopropylethylamine (DIEA) and piperidine were purchased from Fisher

Synthesis of Trtyl-3-Mercaptopropionic Acid

3-Mercaptopropionic acid (100μl, 1.15 mmol) and trityl chloride (481mg, 1.73 mmol) were stirred in 0.65ml DMF for 2 days at room temperature. A 10% sodium acetate aqueous solution (7ml) was then added and the precipitate was filtered and washed with distilled water. Afterward, the residue was stirred in acetone at 50^{o}C for 30 min and filtered after cooling. The residue was washed with little acetone and diethyl ether to get the final product of 370mg in 87% yield.

Preparation of Preloaded MBHA-Leu-3-Mercaptopropionic Acid Resin.

A pre-activated solution of 231mg of Boc-Leu (1mmol, 4eq), 500 μl of DIEA (~6eq) and 380mg of HBTU (1mmol, 4eq) in 2.5ml of DMF was added to 0.37g of MBHA resin (0.25 mmol, 1eq, substitution:0.67 mmol/g) and shaken for 30 min. After Kaiser test, resin was washed with DMF and treated with 3 ml of TFA for 1min twice. After washing with DMF, a pre-activated solution of 3-mercaptopropionic acid (141mg, 0.38 mmol, 1.5eq), DIEA (400μl) and 143mg of HBTU (0.38 mmol, 1.5 eq) in 2.5ml DMF was added to MBHA-Leu resin and mixed for 30min. And then after washing with about 30ml of DMF, 3ml of 2.5%

triisopropylsilane, 2.5% water and 95% TFA solution was added to the resin and shaked to remove Boc for 1min twice and then rinsed the resin with about 30ml of DMF and Diethyl Ether to get the final product, MBHA-Leu-3-Mercaptopropionic acid resin.

Synthesis of ROP5 (24-58) by Native Chemical Ligation (NCL).

A part of C-thioester peptide sequence (24-31) was synthesized on MBHA-Leu-3-Mercaptopropionic acid resin by Boc-based protocol due to the labile thioester linker in Fmoc-based chemistry and N-peptide (32-58) was synthesized by general Fmoc-based protocol. 1mg of ROP1 (N-terminal peptide, 0.34 umol) and 0.35mg of ROP4 (C-termial thioester peptide, 0.31 umol) were dissolved in 100mM sodium phosphate (pH = 8.5) to give a final concentration of 3 mM of each peptide. 100μl of 4% mercaptophenol in 100mM phosphate buffer solution was added to the ROP1 and ROP4 peptides and incubated at room temperature with stirring for 1day. Desired peptide, ROP5 (24-58), was successfully ligated in reasonably good yield (~60%).

3.3.2. Results.

Thioester peptide was treated with excess of thiol cofactor, thiophenol, to convert the weakly activated thioester to a more reactive species by thiol exchange. This effectively catalizes the ligation reaction to the thermodynamically favored amide bond at the ligation site as shown in Fig 32.

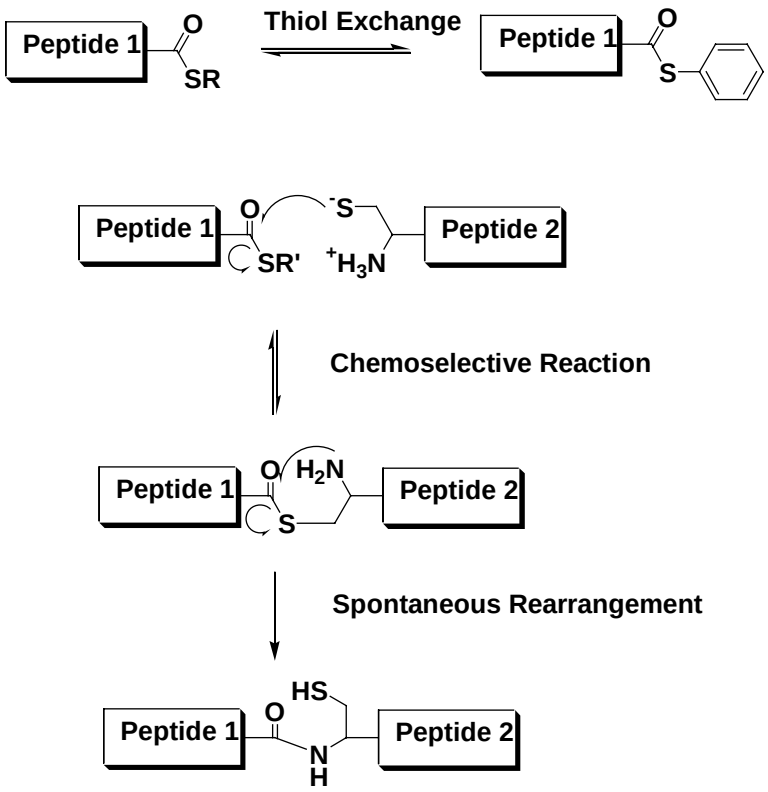


Fig. 32. Native Chemical Ligation

For the proof of concept studies, a part of thioester peptide (C-peptide) sequence (24-31) was synthesized by Boc-based protocol. The thioester peptide (24-31) and N-peptide (32-58) solution (pH = 8.5) was incubated with excess of thiophenol at room temperature and monitored by analytical HPLC after 1day. The desired peptide was successfully ligated in reasonably good yield (~60%) with remaining unreacted peptides (Fig. 33).

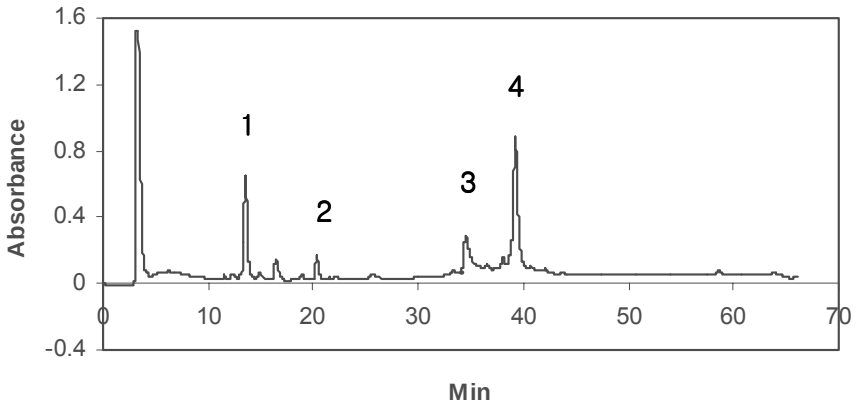


Fig. 33. HPLC Chromatograph of ROP5. (eluent a: 95% water, 5% MeCN, 0.1% TFA, eluent B: 90% MeCN, 10% water, 0.1% TFA, linear gradient 0-70% for 60min). 1: Benzyl thioester(~13min). 2: C-peptide (~20min). 3: N-peptide (~35 min). 4: Ligated peptide (~40 min).

To confirm the identity of the ligated peptide, the products of the reaction were purified by reverse phase HPLC. The major peak at ~40min was collected from HPLC and taken by MALDI-TOF to identify the product. The new peptide (calculated exact mass; 3811.1131 D) was detected at 3812.3 D by MALDI-TOF confirming that those peptides, ROP5, were ligated (Fig. 34).

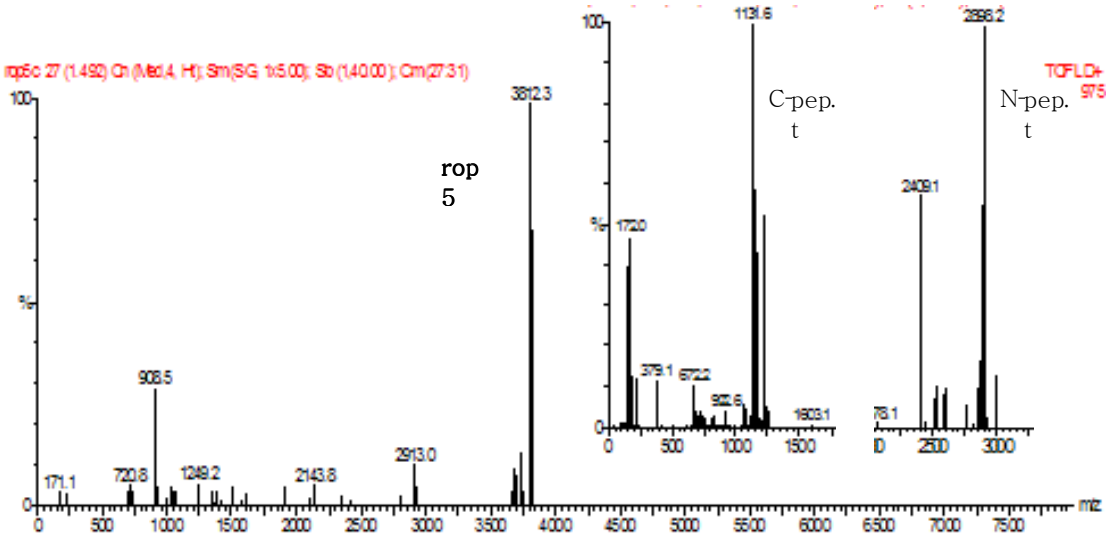


Fig. 34. MALDI-TOF spectrum of ROP5 (calc. exact mass; 3811.1131, det. exact mass; 3812.3).

Chapter 4 Discussion

4.1. *De novo* Designed Fluorinated α_4 Peptides.

Several studies have examined the effects of introducing extensively fluorinated amino acid residues into peptides and largely confirmed the prediction that fluorination would increase the stability of the peptides against unfolding, and that fluorous interaction could mediate specific protein-protein interactions. This study is the first in which the effects of systematically increasing the degree of fluorination on the stability and structure of a folded peptide have been investigated.

Firstly, I have demonstrated that it is possible to significantly increase the volume of the hydrophobic core of the four-helix bundle without grossly perturbing the structure. Thus, the introduction of 6 hFLeu residues in α_4-F_6 increases the volume by about 30 %, but the peptide retains its helical tetrameric structure. As I have discussed previously, the anti-parallel four-helix bundle is more tolerant of amino acid substitution in the hydrophobic core than parallel coiled coil peptides in which subtle changes in hydrophobic packing have been shown to result in changes to the oligomerization state.

Previous studies have produced estimates for the additional stability imparted by incorporating a trifluoromethyl group into a peptide that vary quite significantly. Much of the variation may be explained by differences in experimental design and probably depend on the context in which the fluorinated residue is introduced. By systematically increasing the number of hFLeu residues in the four helix bundle we have been able to obtain a better appreciation for how the degree of fluorination affects the stability of the bundle.

As expected, increasing the number of fluorous residues increases $\Delta G^{\circ}_{unfold}$ in an approximately linear fashion by introducing every two hFLeu residues ($\Delta\Delta G^{\circ}_{unfold}$ = 0.3 (kcal/mol)/residue) up to six hFLeu residues. We previously measured the free energy of partitioning of the HFLeu side chain from n-heptanol into water as $\Delta\Delta G$ = 0.4 kcal/mol. The increased stability imparted by

fluorination appears to be readily ascribed to the more hydrophobic nature of hFLeu, rather than specific “fluorous” interactions between the hFLeu side chains in the hydrophobic core. Indeed, one might expect that fluorous interactions might provide additional stability to the more extensively fluorinated peptides resulting in an *increase,* rather than a decrease, in the per-residue stability for α_4-F_4 and α_4-F_6.

Whereas the stabilizing effect of increasing the number of hFLeu residues was predicted, the structuring effect on the hydrophobic core was not anticipated. Two lines of evidence point to the hydrophobic core of α_4-F_6 being less conformationally dynamic than α_4-H. Firstly, HSQC spectra show that α_4-F_6 has more well-resolved amide resonances than α_4-H, Secondly, ^{19}F spectra of α_4-F_6 and α_4-F_4 exhibit a complex and spectrally disperse envelop of resonances that become dramatically sharper and simpler at higher temperatures. The changes in the spectra are consistent with the core melting at higher temperature to give a more dynamic molten globule-like interior. Interestingly, the incorporation of 2 hFLeu residues does not seem to be sufficient to introduce this structural change, as α_4-F_2 exhibits a fairly simple ^{19}F spectrum even at room temperature.

I previously examined the binding of ANS to α_4-H and α_4-F_2, which is a widely applied test for the molten globule state. I saw no evidence for intercalation of ANS into either peptide and concluded that neither peptide was molten globule-like. However, ANS-binding is not a definitive test for molten globules, and indeed proteins may exhibit a range of dynamic behavior from completely ordered ‘natural’ fold through to the highly mobile and poorly specified molten globule state. It seems reasonable that the α_4 series of peptides may exhibit a range of internal dynamic behavior while maintaining well formed secondary structure. And although the α_4-F_4 and α_4-F_6 peptides seem to exhibit better-packed, less dynamic interiors this does not necessarily imply that the hFLeu residues pack in a unique arrangement as would be expected in a natural protein.

The unexpected ordering imposed by the inclusion of 4 or 6 hFLeu

residues may explain why the α_4-F_4 and α_4-F_6 peptides show smaller per-residue increases in stability than α_4-F_2. The hFLeu side-chains in α_4-F_2 retain their mobility in the folded structure, thereby maximizing the free energy of transfer from water into the hydrophobic core. In contrast, the burying of hFLeu side-chains in the hydrophobic cores of α_4-F_4 and α_4-F_6 is accompanied by a loss of entropy associated with reduced side-chain mobility so that free energy of folding is relatively less favorable.

It is unclear why the incorporation of multiple hFLeu residues should induce ordering of the hydrophobic core. One explanation is that there is simply less room for the larger hFLeu side chains to rotate within the hydrophobic core. Thus, packing 2 layers of the core with hFLeu does not restrict rotation sufficiently, but packing further layers does impede the ability of the central layers to rotate, leading to a less dynamic structure.

4.2. Model Studies on Short ROP Peptides.

All three short ROP peptides (35AL, 35LL and 35ALL) have random coil content and ~ 14% helical content calculated from the theoretical MRE value measured at 222nm. Although hydrophobic core was packed in three different ways, their helical propensities observed at 222nm were not significantly affected, different from what I predicted that more hydrophobicity would enhance the helical propensity as well as stability. Of the three peptides synthesized only ROP 35ALL exhibited a cooperative unfolding transition when denatured in GuHCl. I reason that 35ALL is the most hydrophobic ROP variant among all three peptides due to the three additional Leu so that its folding may be more stabilized to be observed the transition state in the presence of increasing GuHCl. Fitting the unfolding curve of ROP 35ALL assuming an equilibrium between dimeric folded protein and monomeric unfolded protein yielded a free energy of unfolding of 9.7 kcal/mol ± 1.9 kcal/mol.

Extremely fluorinated analogs can be introduced into the proteins. These

thermally, chemically stabilized proteins may potentially have industrial applications. Enzymes are attractive “green” catalysts for industrial process, but often the substrates are poorly soluble in water at room temperature. This problem may be overcome by inclusion of organic co-solvents and raising the temperature, but this often results in denaturation of the enzyme. Fluorous proteins hold the potential to be stable at higher temperatures and in the presence of organic solvents.

By using native chemical ligation method, as I have demonstrated, it will be possible to synthesize the hyper-stabilized proteins such as fluorinated ROP variants in good yields. These ROP variants will be useful in such environments, using organic solvents or higher temperature for practical purpose.

4.3. Fluorinated Antimicrobial Peptide (AMP) for Practical Applications.

As a trial for practical applications, hFLs were incorporated in the hydrophobic cores of MSI-78, which is a magainin variant.[91,92] Magainin is an α-helical antimicrobial peptide (AMP) that is isolated from the African clawed frog (*Xenopus laevis*, also known as Platanna) as shown in Fig. 35.

Fig. 35. African clawed frog (*Xenopus laevis*)

Antimicrobial peptides (AMPs) have the potential to become important antibiotics that may combat bacterial resistance more efficiently. By substituting

hydrophobic residues in AMPs with extremely fluorinated amino acids such as hFl, the fluorinated magainin variant (Fluorogainin-1) was much less susceptible to proteolytic cleavage. When nonfluorinated peptide (MSI-78) and fluogainin-1 are structured in the presence of liposomes, fluorogainine-1 is much more resistant to proteolysis by two common proteases, trypsin and chymotrypsin. Proteolytic stability is attributed to increased structural stability, not incorporation of non-natural amino acids because in the presence of liposome where both MSI78 and fluorogainin-1 are unstructured both peptides are degraded by proteases. No hemolytic activity were found at 250 μg/mL for fluorinated or non-fluorinated peptides.

The minimum inhibitory concentrations (MIC) of MSI-78 and fluoroganin-1 were determined for a panel of 11 bacterial strains that included both Gram positive and Gram negative strains of common pathogenic bacteria. Overall, the fluorous AMP showed a similar spectrum of antibiotic activity to the non-fluorous AMP, however, against two important pathogenic bacteria the fluorous AMP was significantly ($p < 0.05$) more potent. Fluorogainin-1 had an MIC of 16 μM against *Klebsiella pneumoniae*, whereas MSI-78 showed no activity, and the MIC of fluorogainin-1 was ~ four times lower than MSI-78 against *Staphylococcus aureus*. This result suggests that, by introducing fluorinated amino acids, antimicrobial activity of the peptides can be modulated to increase the potency and selectivity towards bacteria.

Moreover, fluorine substituents have become a widespread and important drug components. Nowadays, up to ~ 20% of all pharmaceuticals are fluorinated drugs.[93] Introduction of fluorinated amino acids to known drugs may lead to the development of active compounds for a wide variety of biochemical applications such as inhibitors of several enzymes, anticancer and antiviral agents, antimicrobials, antidepressants and anorectic agents, hypolipidemic drugs and antidiabetics.

4.4. Future Directions

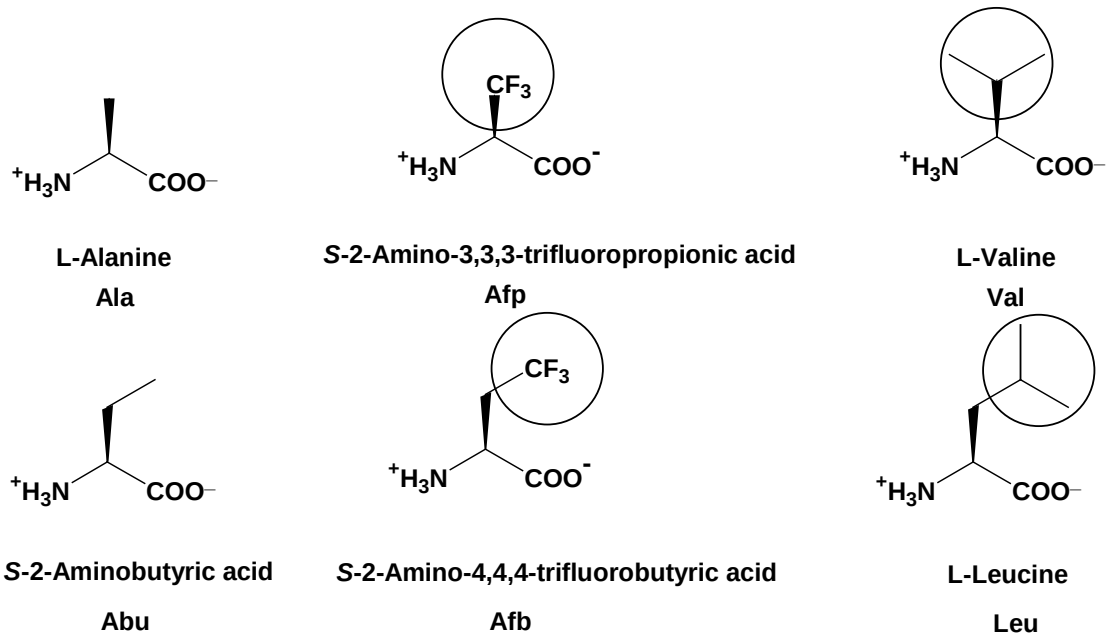


Fig. 36. Fluorinated amino acids and natural amino acids. Ala has a similar shape to Afp and Val has a similar volume to Afp. Abu has a similar shape to Afb and Leu may have a similar volume to Afb.

I have demonstrated that replacing Leu with highly fluorinated hFLeu enhanced the stabilities of model proteins towards thermal and chemical denaturation. In this work the methyl group of Leu was considered as isosteric with the trifluoromethyl group of hFLeu. However, the size of trifluoromethyl group has been a controversial issue, as I have shown in this dissertation. Literature values of steric demand of a CF_3 group vary from being equivalent to ethyl[94] or as large as t-butyl[95] depending on the experimental methods and interpretation of the data. Recent reports suggest that the volume of CF_3 may close to that of isopropyl group[96]. In any case, the volume of fluorous core volume should be larger than that of natural hydrophobic core (Fig. 36).

From the results in my dissertation, it is impled that the substitution of hFLeu with Leu may alter the folding structures due to 1) steric effect on bulky side chain of hFLeu, 2) the strong inductive effect on amide backbone hydrogen bonding structure 3) dipole-dipole interactions on side chains of hFLeu.

However, recent studies on steric effect on CF_3 group reveals that substitution of isosteric Atb with Leu in monomeric peptides still results in a lower helix propensity, implying that steric effects from the bulky side chain may not the reason.[64] It was also claimed that substituting Leu with Atb decreases the overall stability of the protein whereas substituting hFLeu with Leu increases stability despite its low helix propensity; however this could be misunderstanding. Jackel *et al*[97,98] reported that incorporation of a Atb with a Leu at the hydrophobic core of dimeric peptides decreases the peptide's melting point which suggest that the fluorine-hydrocarbon interaction may be weaker than the fluorine-fluorine and the fluorine-hydrocarbon interactions. Therefore if Abu residues are placed at the appropriate positions in the hydrophobic core, it could also stabilize the peptide toward thermal and chemical denaturation. Due to the isosteric side chain of Afb, when metal ions bind with metalloproteins, it may also not alter the structure or binding geometries possessing with higher stabilities.

From the assumption based on literature that trifluoromethyl group is isosteric of isopropyl, it may be possible to investigate the properties of highly fluorinated peptides having a similar core volume to the native hydrophobic core of natural proteins as shown in Fig. 37.

Ac——GNADEL^dYKEL^aEDL^dQERL^aRKL^dRKKL^aRSG——NH_2 **α_4-H**

Ac—GNADEAfp^dYKEAfp^aEDAfp^dQERAfp^aRKAfp^dRKKAfp^aRSG—NH_2 **α_4-Afp**

Ac—GNADEAtb^dYKEAtb^aEDAtb^dQERAtb^aRKAtb^dRKKAtb^aRSG—NH_2 **α_4-Atb**

Fig. 37. The sequence of α_4-H, α_4-Afp and α_4-Atb fluorinated peptides.

Based on the sequence of α_4-H peptide, Afp could be replaced with all 6 Leu of the core to have a Val-like hydrophobic core peptide and Atb will be incorporated to the same position so that the volume of the fluoro-core will be similar to α_4-H. Volumewise, Afp and Atb are isosteric of Val and Leu

respectively, however, shapewise, they look like Ala and 2-aminobutyric acid. It would be interesting to investigate the physical and structural properties of fluorinated analogues systematically packed to resemble α_4-H. Heterogeneity of two different fluorinated peptides, α_4-Afp and α_4-Atb, could be tested by NMR and chemical crosslinking method detected by mass spectroscopy.

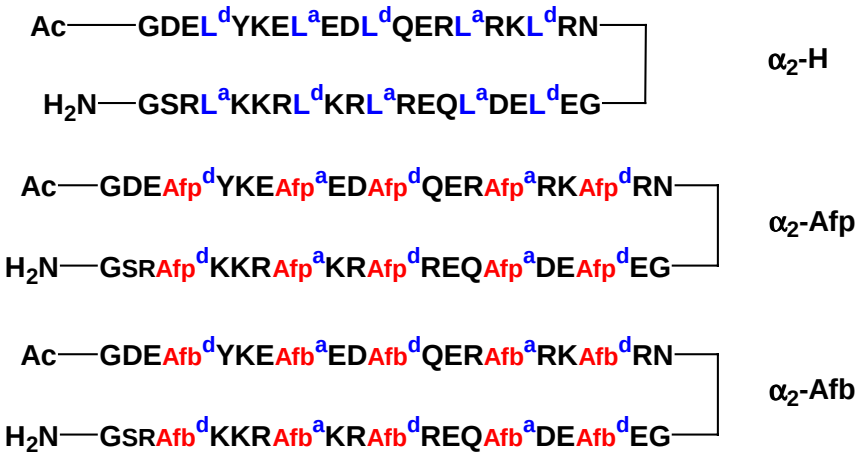


Fig. 38. The sequence of α_2-H, α_2-Afp and α_2-Atb fluorinated peptides.

An extention of this work would be to design antiparallel dimeric helix-turn-helix peptides packed with Leu at the cores, based on our α_4 series peptide. Some of residues at the N- and C- terminus sequence could be removed to design short peptide for easier synthesis. The well known Asn-Gly turn sequence could be introduced to fold the peptide in a dimeric hairpin structure. Incorporation of fluorine into peptides will provide an opportunity to study their conformational properties by ^{19}F-NMR as I have shown in this dissertation.

Apendix A) Covalent Metal-Peptide Framework Compounds that Extend in One and Two Dimensions.

Hyang-Yeo Lee, Jeffery W. Kampf, Kyo-Sung Park and E. Neil G. Marsh

Crystal Growth & Design, 2007, in press.

- This is the modified paper published from *Crystal Growth & Design*.

Introduction

Metal-Organic Frameworks (MOFs)

The design of new molecular architectures based on the coordination of metal ions with a variety of bridging organic ligands, to produce either discrete molecular complexes or extended molecular solids (metal-organic frameworks, MOFs), has been an area of explosive growth in recent years (Fig.1)[1a-5]. As well as the intrinsic intellectual interest associated with the design and characterization of new compounds, such materials may have important applications [6]. In particular, much interest has focused on porous materials for applications such as gas storage, separations and catalysis [7-10].

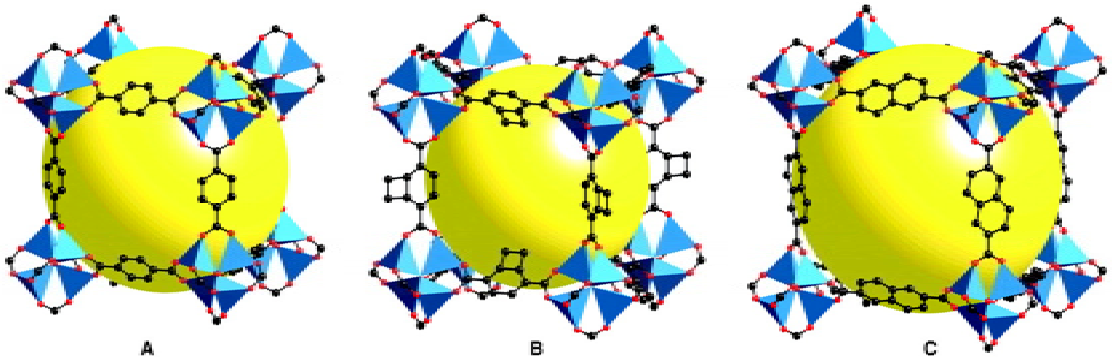


Fig.1. Single-crystal x-ray structures of MOFs shown for 3 dimensional structures with large porosity.

There has also been considerable interest in expanding the functional repertoire of MOFs by incorporating more complex organic ligands into MOF assemblies.

Functionalized Porous Frameworks

In the recent years, zeolitic porous materials have been studied for many catalytic applications as crystalline micro-porous zeolites in petroleum industries

due to their functionalized pores with various transition metals. The zeolitic imidazolate frameworks (ZIF) crystal structures are base on the aluminosilicate zeolites networks and they have exceptional chemical and thermal stability in organic or basic aqueous conditions.

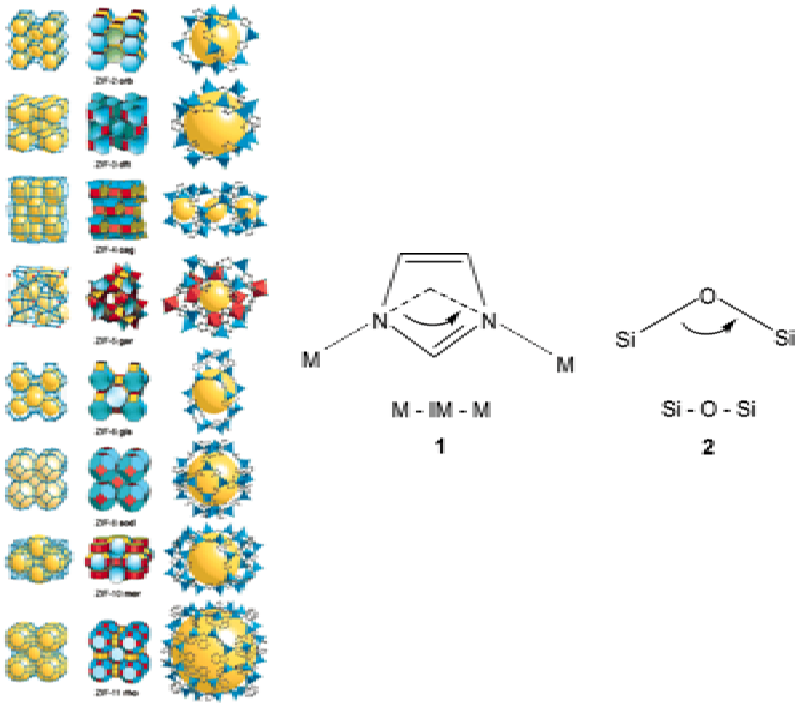


Fig. 2. The crystal structures of Zifs (left) and the bridging angles in metal IMs (1) and zeolites (2). This figure is taken from Park et al.[1b]

Using multi-dentate metal-coordinating ligands, instead of simple organic ligands, it has been possible to introduce new magnetic, optical and electronic properties, into MOFs [11-16]. Furthermore, by using ligands that retain an open coordination site at the metal centre, it has been possible to design catalytic properties into these materials. However non-porous materials may exhibit interesting magnetic, optical, and electronic properties and may have valuable pharmacological properties.

Metal-Peptide Frameworks (MPFs)

One of the simplest examples of peptide frameworks was reported by et al using achiral gly-gly dipetide and metal ions such as Zn(II) or Cd(II). Gly-Gly dipeptides used as organic linkers in MOFs were ligated by Cd metal ions and extended in 2 dimensional networks as shown in Fig. 3. It shows simple tridentate binding pattern, two from C-terminal carboxylates and one from N-terminal nitrogens in the crystal structure. The hydrogen bonding from amides in between the peptide layers further extend the networks to 3 dimensions.

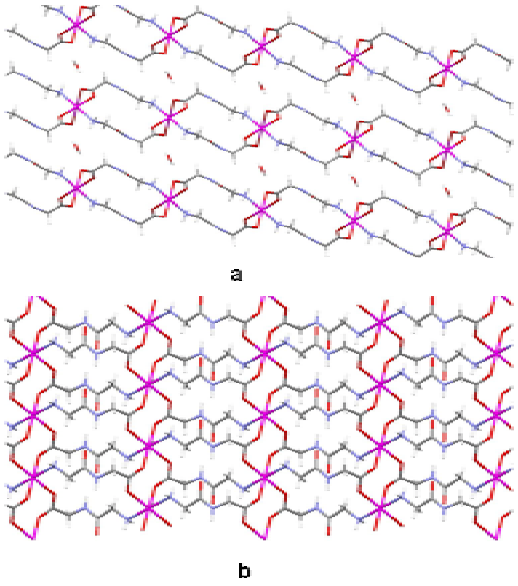


Figure 3 Network formed by the prototypical metal-peptide framework complex $Cd(Gly_2)_2.H_2O$[19] **A** viewed along the crystallographic *b* axis, illustrating the extended network formed by the bridging di-glycine linkers. **B** viewed along crystallographic *c* axis illustrating the 8-membered rings formed by Cd-bridging carboxylate groups.

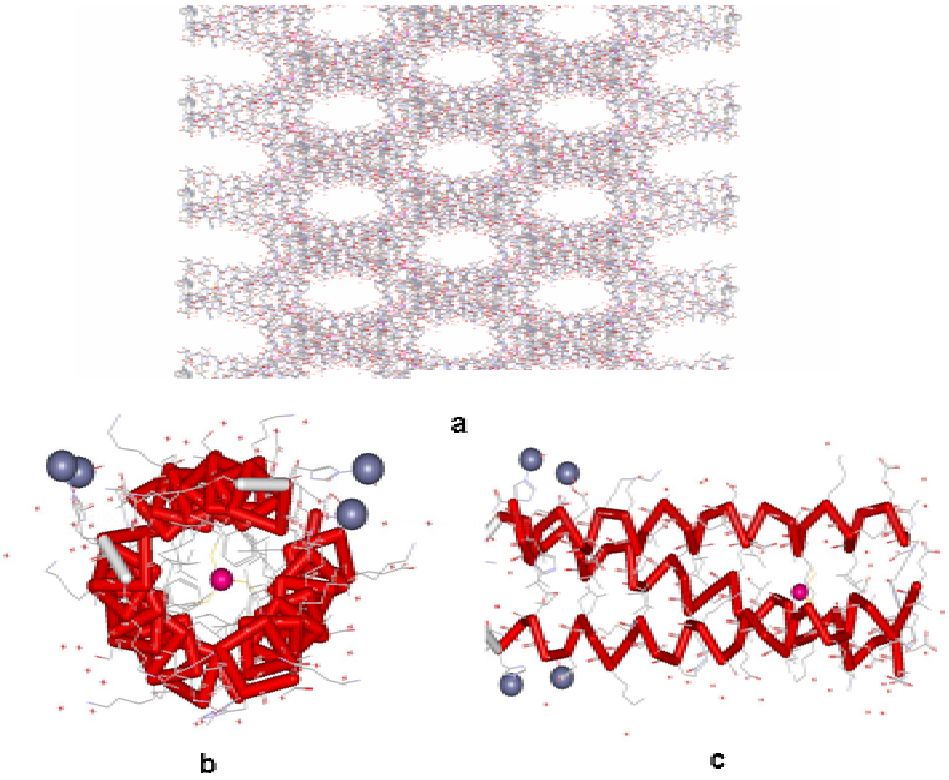


Fig. 4. Crystal structure of metal binding peptide, Coil-Ser, in 3-demensional networks (a). Parallel three-stranded Coil-Ser peptide binding with As(III) in the core and Zn(II) at the near c-terminus as a top view (b) and a side view (c). PDB:2JGO. [2b]

The crystal structure of coiled-coil peptide, Coil-Ser, binding with As(III) and Zn(II) has been reported. The *de novo* designed metal-binding peptide, Coil-Ser, is structurally more complicated example of Metal-Peptide Frameworks. Parallel three-stranded Coil-Ser is designed to bind with heavy metals such as Cd(II), Hg(II) and AS(III) interior of the coiled coils. This metal binding peptide with histidine residue at near the c-terminus to facilitate the nucleation used as building blocks is crystallized in 3 dimensional networks with large porosity as shown in Fig. 4.

I describe the structures of one- and two-dimensional Metal-Peptide Frameworks (MPFs) in which the ligands are short peptides. Peptides are attractive ligands for designing MPFs as they are easy to synthesize, and

extremely versatile. The peptide sequence can be varied in a predictable and modular fashion allowing a high degree of structural and functional diversity to be incorporated into the material. Peptides readily bind metals through their amino- and carboxyl-termini and through various side-chains. Lastly, hydrogen-bonding between backbone amide groups provides the potential for "secondary" structure-forming interactions between peptides.

Metal-peptide frameworks have the potential to be a large and novel class of compounds with potential pharmacological applications [17]. However, only a few scattered examples are reported that are limited to di-peptide linkers [17-19].

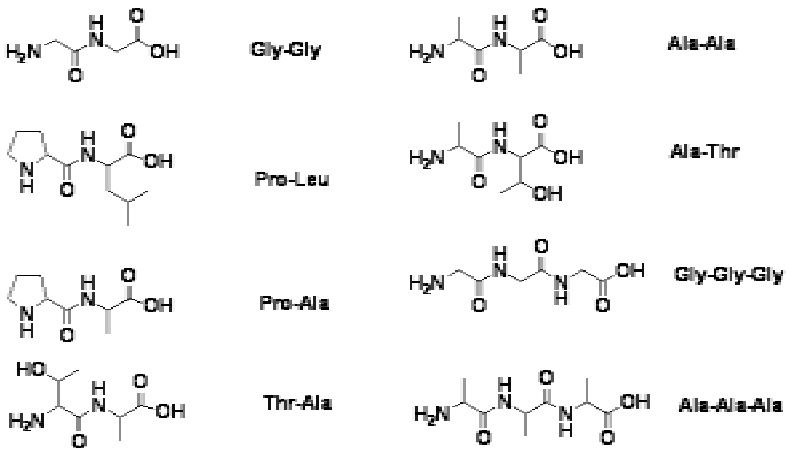


Fig. 5. Di- and tri-peptide ligands that are used for crystallization.

Here I describe a series of extended peptide-metal frameworks structures based on di-peptides and tri-peptides. Peptide ligands that are used for the structural study shown in Fig. 5. Their structures may be rationalized in much the same way as protein structures: by invoking the need to satisfy hydrogen bonding interactions between backbone amide groups and the formation of hydrophobic cores by packing together hydrophobic amino acid side-chains. These structures may suggest some basic design "rules" for constructing 1-D, 2-D and 3-D frameworks.

Experimental Section

Materials Peptides either were purchased from Sigma Chemical Co, or synthesized by standard liquid phase peptide coupling and deprotection methods using boc-protected monomers. All other reagents were purchased from commercial suppliers and used without further purification.

Preparations. All complexes were crystallized by the base diffusion method [20]

Synthesis of $Cd(Gly_3)_2.H_2O$ (1). A solution of $CdAc_2$ (14 mg, 0.05 mmol) and Gly_3 tri-peptide (10 mg, 0.05 mmol) in 2.0 ml of DMF, water and ethanol (1:1:1.5) was prepared, and 1.0 ml of a solution of 2 % triethylamine in ethanol added. The solution was sonicated, and placed in a 5 ml glass vial, which was in turn was placed in 20 ml screw-cap vial. 1.0 ml of 2 % triethylamine in ethanol solution was added to 20 ml vial and sealed with a cap. After about 3 months colorless needles had grown; these were separated, washed with ethanol, and air-dried (59 % yield, based on cadmium). FTIR data (cm^{-1}): 3453, 3293, 3100, 2938, 2816, 1675, 1649, 1577, 1423, 1381, 1290, 1258, 1061, 1027, 1015, 715, 596, 477.

Synthesis of $Cd(_L\text{-}Ala_3)_2$ (2). A solution of $CdAc_2$ (40 mg, 0.15 mmol) and $_L$-Ala_3 tripeptide (50 mg, 0.15 mmol) in 10ml DMF, water and ethanol solution (2:2:1) was prepared, and 10 ml of a solution of 2 % triethylamine in ethanol added. The solution was sonicated and placed in a 20 ml vial, which was in turn placed in 100 ml screw-capped jar. 10 ml of 2 % triethylamine solution was added to 20 ml vial and sealed with a cap. After about 3 weeks transparent needles had grown; these were washed with ethanol, and air-dried (43 % yield, based on cadmium). FTIR data (cm^{-1}): 3376, 3258, 3223, 3067, 2986, 2937, 1679, 1639, 1592, 1516, 1444, 1413, 1142, 1110, 1029, 694.

Synthesis of Cd($_L$-Ala$_2$)$_2$ (3). A solution of $CdAc_2$ (16mg, 0.06mmol) and $_L$-Ala$_2$ di-peptide (10mg, 0.12 mmol) in 2 ml DMF, water and ethanol solution (1:1:1.5) was prepared, and 2.0 ml of a solution of 2% triethylamine in ethanol added. The solution was sonicated, and placed in a 5 ml glass vial, which was in turn was placed in 20 ml screw-cap vial. 2 ml of 2 % triethylamine solution was added to 20 ml vial and sealed with a cap. After about 4 weeks transparent needles had grown; these were separated, washed with ethanol, and air-dried (62% yield, based on cadmium). FTIR data (cm^{-1}): 3346, 3322, 3287, 3242, 3131, 3078, 2969, 2933, 1644, 1563, 1513, 1442, 1401, 1147, 1101, 1001, 857, 686.

Crystallization of Cd($_L$-Ala-$_L$-Thr)$_2$.4H$_2$0 (4). A solution of $CdAc_2$ (16mg, 0.06mmol) and $_L$-Ala-$_L$-Thr di-peptide (10mg, 0.12mmol) in 2 ml of water was prepared, and a solution of 2 % triethylamine in ethanol added. The solution was sonicated, and placed in a 5 ml glass vial, which was in turn was placed in 20 ml screw-cap vial. 2 ml of 2 % triethylamine solution was added to 20 ml vial and sealed with a cap. After about 3 weeks transparent needles had grown; these were washed with ethanol, and air-dried (60% yield, based on cadmium). FTIR data (cm^{-1}): 3447, 3372, 3316, 3271, 3170, 2968, 2921, 1668, 1645, 1604, 1584, 1533, 1418, 1395, 1242, 1123, 1083, 1060, 1040, 801.

Synthesis of Cd($_L$-Pro-$_L$-Ala)$_2$ (5). Compound 5 was synthesized by the base diffusion method from a mixture of $CdAc_2$ (13mg, 0.05mmol) and $_L$-Pro-$_L$-Ala peptide (9mg, 0.05mmol) in 1.5ml DMF and water solution (1:1) and a mixture of 2% TEA in ethanol (1ml). After being sonicated cadmium and dipeptide solution in 5ml vial, it was placed in 20ml vial. 1ml of 2% TEA solution was added to 20ml vial and sealed with a cap. About 1 month later, the transparent crystals were separated, washed with ethanol, and air-dried (ca. 25% yield, based on cadmium).

Synthesis of Cd($_L$-Pro-$_L$-Leu)$_2$ (6). Compound 6 was synthesized by the base diffusion method from a mixture of $CdAc_2$ (16mg, 0.06mmol) and $_L$-Pro-$_L$-Leu

peptide (10mg, 0.12mmol) in 2ml DMF, water and ethanol solution (1:1:1.5) and a mixture of 2% TEA in ethanol (2ml). After being sonicated cadmium and dipeptide solution in 5ml vial, it was placed in 20ml vial. 2ml of 2% TEA solution was added to 20ml vial and sealed with a cap. About 2 months later, the transparent needle was separated, washed with ethanol, and air-dried (ca. 10% yield, based on cadmium).

Crystal Structure Determination Crystals of complexes **1**-**6** were mounted on a standard Bruker SMART 1K CCD-based X-ray diffractometer equipped with a LT-2 low-temperature device and normal focus Mo-target X-ray tube (λ = 0.71073 Å) operated at 2000 W power (50 kV, 40 mA). The final cell constants (Table 1) were based on the *xyz* centroids of reflections above 10σ(*I*). Analysis of the data showed negligible decay during data collection; the data were processed with SADABS and corrected for absorption. The structures were solved and refined with the Bruker SHELXTL (version 6.12) software package. All of the non-hydrogen atoms were refined anisotropically with the hydrogen atoms placed in idealized positions. Figures of the molecular structures were produce using the Mercury program [21]. Further crystal structure and refinement data for 1-4 of the complexes are summarized in Table 1.

Results

My starting point in this project was prototypical metal-peptide framework of di-glycine coordinated by cadmium, $Cd(Gly_2)_2.H_2O$, which was previously shown by Takayama et al [19] to form a 2-D network comprising carboxylate-bridged cadmium ions in one dimension and N-to-C-terminal peptide-linked cadmium ions in the second dimension (Figure 1). I therefore decided to examine the how the peptide may influence the type of network formed by either extending the length of the peptide linker, or changing the side-chains of the peptide. Initially, we took a conservative approach to modifying the peptide linker, focusing on tri-glycine, di-alanine and tri-alanine as linkers. These peptides were screened for their ability to form crystals in complex with various divalent metal ions

including Cd, Zn, Cu, Mn, Cr, Co, Fe, Mg, Pb and Ag. However, apart from the Gly_3 peptide, which also gave crystals with zinc, only cadmium produced the high quality diffracting crystals in our hands. The crystallographic data for each of the complexes described below is summarized in Table 1.

Structure of Cd $(Gly_3)_2.H_2O$ (1) To determine effects of extending the peptide linker on structure and dimensionality, I crystallized the cadmium tri-glycine complex, Cd $(Gly_3)_2.H_2O$. The structure, illustrated in Figure 6, is almost analogous to the di-peptide complex except for distortion of the peptide ligand to maximize the hydrogen bonding interaction. Each Cd ion is octahedrally ligated by 6 peptides: the N-termini of two peptides supply two apical nitrogen ligands; one oxygen from each of four C-terminal carboxylates supply four equatorial ligands. Each carboxylate group bridges two Cd ions to form chains of eight membered rings that extend the molecular framework in one dimension. The second dimension of the framework is formed by the peptide backbone: the tri-glycine peptides adopt an extended conformation, with two peptides running anti-parallel to each other to form a ring that is closed by ligation to two cadmium ions.

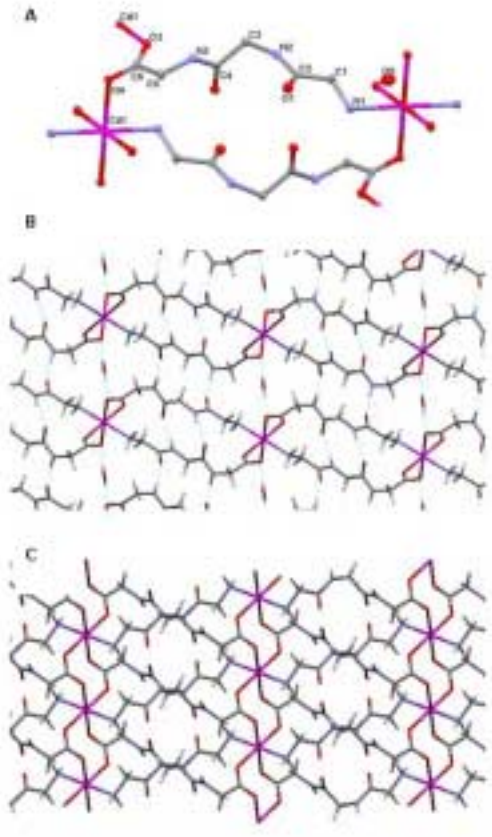


Figure 6. **A** Structure of Cd $(Gly_3)_2.H_2O$ (**1**). **B** viewed along the crystallographic *b* axis (*inter*- and *intra*-molecular hydrogen-bonds in blue) illustrating the similarities of the 2-D network to that of Cd $(Gly_2)_2.H_2O$ shown in figure 1. **C** viewed along crystallographic *c* axis showing the 8-membered rings formed by Cd-bridging carboxylate groups, similar to figure 1B.

Hydrogen bonding plays an important role in the structure, with all four potential peptide backbone hydrogen bonds being satisfied. Each peptide forms two hydrogen bonds with its partner across the covalently closed ring. Two more hydrogen bonds are formed with peptides in adjacent layers that serve to extend the crystal in the third dimension.

Structure of Cd $(_L\text{-Ala}_3)_2$ (2) I next sought to examine how addition of a simple side-chain (i.e. the methyl group of alanine), and hence chirality, would modulate the structure of the metal-peptide framework. We therefore crystallized the complex of tri-alanine with cadmium and determined its structure. Cd $(_L\text{-Ala}_3)_2$ (**2**) forms a one-dimensional extended molecular chain comprised of an infinite

series of rings linked together through the Cd ions (Figure 7). In this case each Cd is octahedrally coordinated by four tri-alanine peptides, as steric hindrance from the methyl side-chains prevent the carboxylate groups from bridging between Cd ions (c.f. the tri-glycine complex). Carboxylate groups from two peptides coordinate each Cd in a mono-dentate fashion; the N-termini of the other two peptides supply two nitrogen ligands, and N-terminal amide carbonyl oxygens provide the final two ligands to the metal.

Similarly to complex **1**, two tri-alanine peptide linkers bridge between two Cd ions in an anti-parallel orientation to form a ring. Most interestingly, the tri-peptide is seen to fold around the methyl group of the N-terminal alanine residue so that it points into the center of the ring. Together with the side-chain from the opposing tri-peptide, the two methyl groups form a hydrophobic core that fills the central cavity of the ring (Figure 7). This folding disposes the methyl groups of the second and third residues to face outwards where they make hydrophobic contacts with methyl groups of adjacent metal-peptide chains. These inter-chain hydrophobic contacts propagate the structure in the second dimension (Figure 7D).

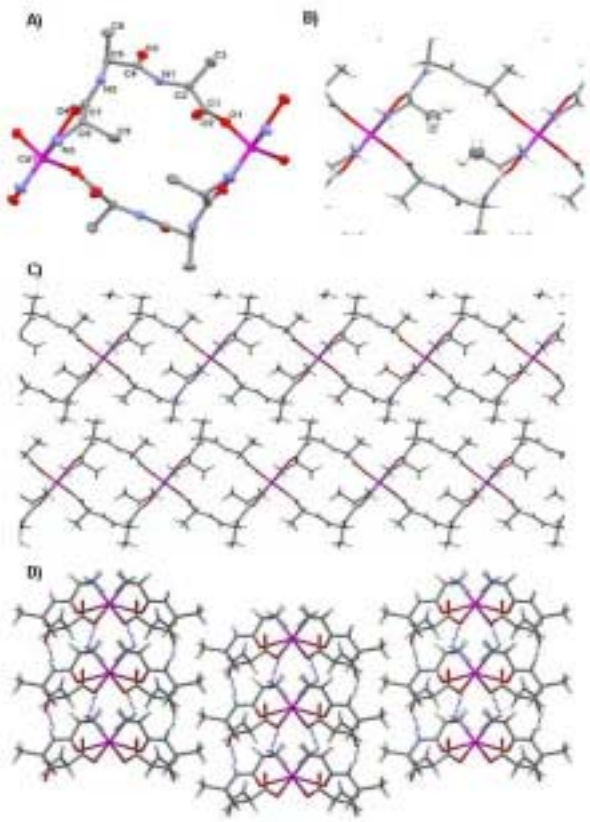


Figure 7 **A** Structure of Cd($_L$-Ala$_3$)$_2$. (**2**) **B** Packing of two methyl groups within the centre of each ring to form a hydrophobic core. **C** view along the crystallographic *b* axis illustrating the 1-D chain structure. **D** view along crystallographic *c* axis illustrating the packing of chains within the crystal mediated by interdigitation of the alanine methyl groups and hydrogen bonding between backbone amide groups (hydrogen bonds shown in blue).

Hydrogen bonding interactions between amide N-H and carbonyl oxygens propagate the structure in the third dimension. Two hydrogen bonds, between the first and second N-H and carbonyl groups, connect each peptide to its counterpart in the next layer of the structure. Steric hindrance from the side-chains means that potential hydrogen bonding interactions with the third N-H group and the unligated carboxylate oxygen are not satisfied.

Structure of Cd ($_L$-Ala$_2$)$_2$ (3) Given the key role that the side-chains play in the structure of Cd ($_L$-Ala$_3$)$_2$, I next investigated the effect of reducing the linker length

by crystallizing the Cd di-alanine complex, Cd $(_L\text{-Ala}_2)_2$. The result is a completely different 2-D peptide-metal framework in which four di-peptide linkers to form a highly distorted square lattice, as shown in Figure 8. Surprisingly, the peptide adopts three different bridging modes between Cd ions in which results in two chemically distinct modes of cadmium ligation (Figure 8D, 8E). In one bridging mode ("long" mode) the N-terminus amino group acts as a mono-dentate ligand and the C-terminal carboxylate is also a mono dentate ligand. In another bridging mode ("short" mode) the N-terminus acts as a bi-dentate ligand with both the NH_2- and amide carbonyl groups coordinating the metal, whereas the C-terminal carboxylate group is now a bi-dentate ligand. The third bridging mode ("medium" mode) involves the N-terminus acting as a bi-dentate ligand through the NH_2- and amide carbonyl groups, but the C-terminal carboxylate group is mono-dentate. In the structure one set of Cd ions is ligated by the N-termini of two "short" linkers and the C-termini of two "long" linkers, whereas the other set is ligated by the N-terminus of a "long" linker, the C-terminus of a "short" linker, the N-terminus of a "medium" linker and the C-terminus of a second "medium" linker. Both sets of Cd ions are pseudo-octahedrally coordinated with the two nitrogen ligands occupying adjacent positions.

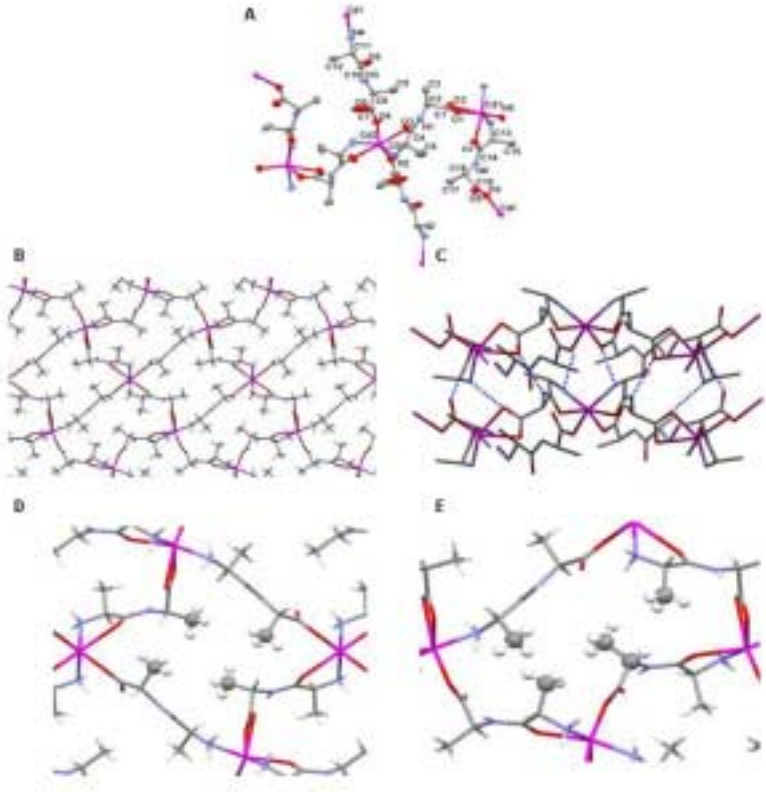


Figure 8 **A** Structure of Cd($_L$-Ala$_2$)$_2$. (**3**) **B** view along the crystallographic *b* axis showing the 2-D lattice framework. **C** *Inter*-peptide hydrogen bonds (shown in blue) between two adjacent molecules of the asymmetric unit, viewed along crystallographic *c* axis. **D**, **E** stabilization of the lattice structure by hydrophobic packing of methyl groups within the centre of each ring. The different linking modes adopted by the di-alanine units within the crystal are indicated by the labels *l* (long), *s* (short) and *m* (medium) as discussed in the text.

The resulting combinations of “long”, “‘medium” and “short” di-alanine linkers give rise to two types of lattice spaces in the crystal. The interior space of each lattice element is filled by four methyl side-chains, one from each of the di-peptides that define the lattice. The alanine methyl groups pack together to produce a hydrophobic core that most likely stabilizes the structure and explains the preference for a 2-D “square” network over the 1-D linear chain formed by the Cd ($_L$-Ala$_3$)$_2$ complex.

As with the other peptide-metal frameworks, hydrogen bonding between the

layers of the 2-D lattice allows the structure to propagate in the 3rd dimension. Each peptide donates and accepts two hydrogen bonds to other peptides in adjacent layers. However, the hydrogen-bonding pattern is complicated because the “long’”, “short’ and “medium” peptide linkers each make slightly different hydrogen bonding interactions.

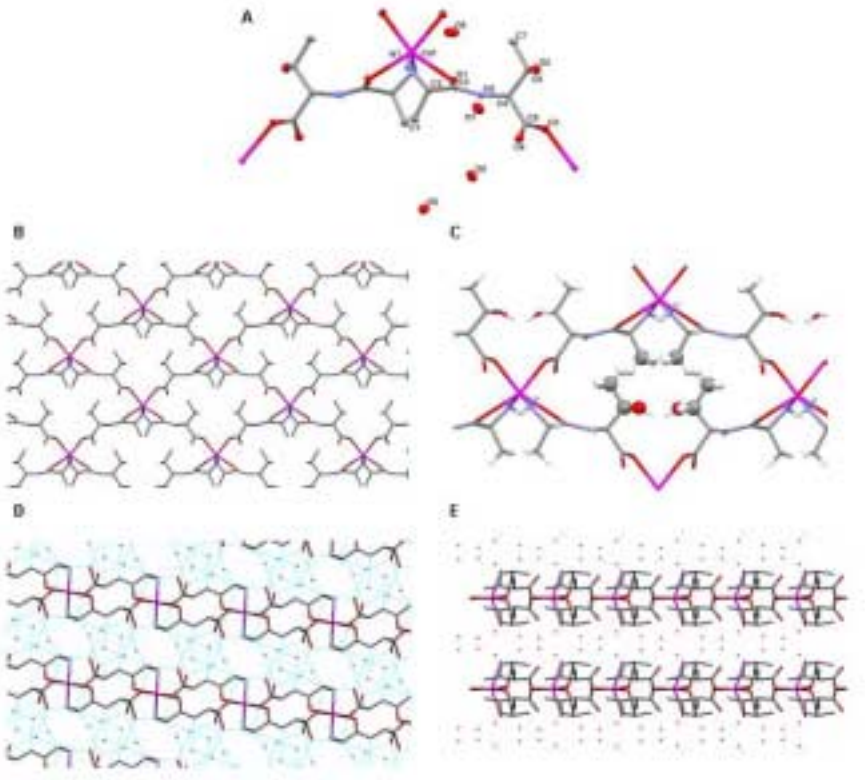


Figure 9 **A** Structure of Cd($_L$-Ala-$_L$-Thr)$_2$.4H_2O. (**4**) **B** view along the crystallographic *c* axis showing the 2-D lattice framework (water molecules and hydrogen atoms have been omitted for clarity). **C** packing of alanine and threonine side-chains with the lattice. **D**, **E** views along crystallographic *b* and *a* axes respectively, illustrating the layer of water molecules (red dots) between each layer of the lattice. The network of hydrogen bonds is shown in blue.

Structure of Cd ($_L$-Ala-$_L$-Thr)$_2$ (4) We were able to modify the 2-D di-alanine framework, by substituting one alanine residue by threonine. The –CH(OH)CH_3 side-chain of threonine is amphiphillic, possessing both methyl and hydroxyl groups. The structure of Cd ($_L$-Ala-$_L$-Thr)$_2$ exhibits the same 2-dimensional topology as Cd ($_L$-Ala$_2$)$_2$, again forming a “square” lattice in which four peptides

are linked together by four cadmium ions (Figure 9). In this case each di-peptide acts as a “medium” length linker, coordinating Cd ions in a bi-dentate manner through the N-terminus (NH_2- and amide carbonyl ligation) and at the C-terminus through mono-dentate carboxylate ligation. This results in a more regular structure. Again, each lattice space is filled by four side-chains (two Ala, two Thr), one from each peptide linker.

Most strikingly, the hydroxyl group of threonine points out of the lattice and recruits a layer of water molecules between the adjacent sheets of the lattice, as shown in Figures 9D and 9E. Two water molecules per peptide (four per asymmetric unit) are incorporated into the structure where they form an extensive hydrogen-bonding network with the backbone oxygen and NH groups, and the threonine hydroxyl. The result is that, unlike all the other structures, there are no direct hydrogen-bonding interactions between adjacent layers of the 2-D network.

Structure of Cd ($_L$-Pro-$_L$-Ala)$_2$ (5) To further study on the effects of various side chains of the peptide linker on structure and dimensionality, I manipulated the structure by substituting one alanine residue by proline from the Cd ($_L$-Ala-$_L$-Ala)$_2$.H_2O complex. As illustrated in Figure 10, the structure of Cd ($_L$-Pro-$_L$-Ala)$_2$ also exhibits 2-dimensional topology, forming hydrophobic core between side chains of praline and alanine. Peptides are linked together by two cadmium ions, coordinating Cd ions in a bi-dentate manner through the N-terminus and at the C-terminus through mono-dentate carboxylate ligation. Hydrogen bonding plays also important role to further extend the crystals in third dimension.

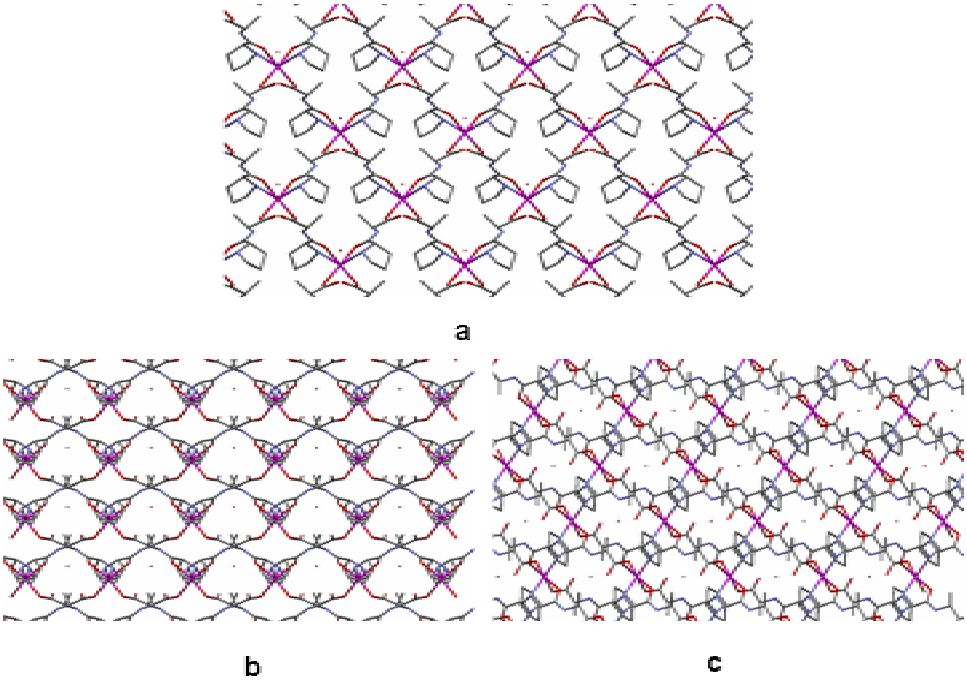


Fig. 10. Crystal structure of Cd ($_L$-Pro-$_L$-Ala)$_2$.H_2O (5) **B**, **C** views along crystallographic *b* and *a* axes respectively, illustrating the layer of water molecules (red dots).

Structure of Cd (Pro-Leu)$_2$ (6) By introducing more bulky side chain on substituting leucine with alanine, I next investigated the steric effect of the linker by crystallizing the Cd ($_L$-Pro-$_L$-Leu)$_2$. The result is a completely different peptide-metal framework from the crystal complex, Cd ($_L$-Pro-$_L$-Ala)$_2$. Di-peptide linkers adopt in 1-demensional structure; however by linking with bridged acetate ligands, the structure further extends to 2-demensions. Side chains of leucine contact with cyclopentane rings of proline to maximize the hydrophobic interaction.

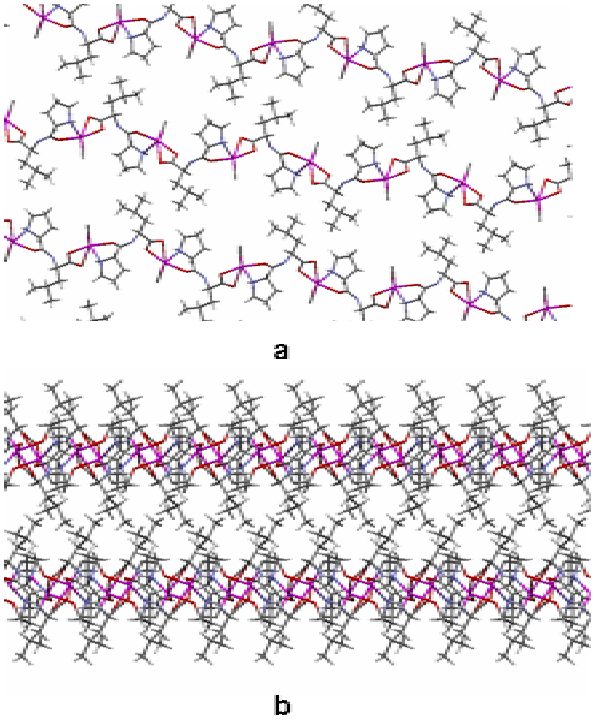


Fig. 11. Crystal structure of Cd ($_L$-Pro-$_L$-Leu)$_2$ (6). **B** view along the crystallographic *c* axis showing the hydrophobic packing of leucine side-chains with the lattice.

Due to the bulky side chain of leucine, the structure may only allow to adopt in 1-demension network, which is completely different from that of Cd ($_L$-Pro-$_L$-Ala)$_2$, which forms square lattice in 2-demensions, so that bridging acetate ligands are needed to further stabilize the structure in two dimensions compensating the lack of hydrogen bonding at the peptide backbone (Fig.11).

Discussion

To date, the vast majority of MOF-like structures have been synthesized using very rigid organic ligands that incorporate hydrophobic aromatic rings or C-C double or triple bonds as spacers. These typically result in open 3-D networks characterized by a degree of porosity. The synthesis of metal-peptide framework compounds introduces ligands with very different properties from those of conventional MOFs. The peptide backbone is intrinsically flexible, and

each amide linkage incorporates both a hydrogen bond donor (N-H) and acceptor (C=O) into the ligand that will need to be satisfied to minimize the energy of any given structure. The nature of the side-chain, its size and whether hydrophobic or hydrophilic, will also strongly influence the properties of the peptide. In the structures described above, all of these properties of peptides are seen to play important roles in defining the networks that are formed. In general, these properties of peptides will mediate against the formation of classical open, 3-D MOF-like structures in favor of tightly packed structures in which amide hydrogen bonds, and side-chain hydrophobic and hydrophilic interactions can be maximized, as are found in proteins.

The structures we have characterized allow some general conclusions to be drawn regarding the assembly of metal-peptide frameworks. The cadmium serves to assemble the peptide linkers together and the requirement for electroneutrality, combined with the preference for a six-coordinate [2N + 4O] ligand set results in 4 peptides coordinating the metal, two by their N-termini and two by their C-termini (the bridging carboxylate mode seen in the tri-glycine complex, **1**, would appear to be atypical, probably because glycine lacks a side-chain that would otherwise prevent the coordination of the additional peptide chains). Cadmium coordination appears to be a robust means of assembling peptides, in that the [2N + 4O] ligand set can be achieved in a variety of ways and distortions from octahedral geometry are tolerated, as illustrated by the $Cd(_{L}\text{-}Ala_2)_2$ complex, **3**. Other than with glycine-based peptides, where zinc and magnesium complexes have been reported [17,19], attempts to obtain crystals with a variety of divalent metal ions were unsuccessful. This suggests that a relatively large metal is needed to accommodate the distorted octahedral geometry imposed by the coordinating peptide ligands.

Although arguing from a limited number of examples, it appears that di-peptide linkers are more likely to form 2-D covalent frameworks because steric crowding imposed by the amino acid side-chains prevents two peptides bridging between the same two metals. However, longer peptides can fold around the side-chain, as illustrated by the structure of Cd $(_{L}\text{-}Ala_3)_2$, **2**, to form 1-D chains

that allow the peptide to pack efficiently together.

In all cases inter-peptide hydrogen bonds play an important role in the structure. The 1-D and 2-D frameworks exhibited by these compounds maximize the number of hydrogen bonding interactions, resulting in highly anisotropic layered materials. The need to satisfy hydrogen bonds may well explain why peptide-metal frameworks appear restricted to 1-D and 2-D networks. It is difficult to imagine how a 3-D network could form that would allow all the potential peptide hydrogen bonding interactions to be satisfied.

Our studies suggest that these structures may serve as templates that can be modified in a somewhat predictable manner, as illustrated by the incorporation of water molecules into the structure when the hydrophobic alanine side-chain is substituted for the amphiphilic threonine side-chain. The Cd $(_{L}\text{-Ala}_3)_2$ structure, **2**, is particularly interesting in this regard, as the first alanine residue appears to nucleate the folding of the tri-peptide leaving residues 2 and 3 oriented to make inter-chain contacts. This would suggest that there is considerable scope for modulating the properties of this material by changing the side-chains at these positions, while leaving the basic chain connectivity intact.

Acknowledgement

We thank Profs. Omar Yaghi and Mark Banaszak Holl for helpful discussions

Supporting information available Crystallographic data for complexes **1** – **4** in CIF format

References

(1) Ockwig, N. W.; Delgado-Friedrichs, O.; O'Keeffe, M.; Yaghi, O. M. *Accounts of Chemical Research* **2005**, *38*, 176-182; 1b) Park, K. S., Ni, Z., Cote, A.; Choi, J. Y.; Huang, R.; Uribe-Romo F.J.; Chae H.K.; O'Keeffe M., Yaghi O.M. PNAS, **2006**, 103, 10186

(2) Yaghi, O. M.; O'Keeffe, M.; Ockwig, N. W.; Chae, H. K.; Eddaoudi, M.; Kim, J. *Nature* **2003**, *423*, 705-714.; 2b) Touw D.S. Nordman, C.E.; Stuckey J.A.; Pecoraro, V.L.,PNAS, **2007**, 104, 11969

(3) Seidel, S. R.; Stang, P. J. *Accounts of Chemical Research* **2002**, *35*, 972-983.

(4) Holliday, B. J.; Mirkin, C. A. *Angewandte Chemie-International Edition* **2001**, *40*, 2022-2043.

(5) Leininger, S.; Olenyuk, B.; Stang, P. J. *Chemical Reviews* **2000**, *100*, 853-907.

(6) Hupp, J. T.; Poeppelmeier, K. R. *Science* **2005**, *309*, 2008-2009.

(7) Dybtsev, D. N.; Nuzhdin, A. L.; Chun, H.; Bryliakov, K. P.; Talsi, E. P.; Fedin, V. P.; Kim, K. *Angewandte Chemie-International Edition* **2006**, *45*, 916-920.

(8) Chen, B. L.; Liang, C. D.; Yang, J.; Contreras, D. S.; Clancy, Y. L.; Lobkovsky, E. B.; Yaghi, O. M.; Dai, S. *Angewandte Chemie-International Edition* **2006**, *45*, 1390-1393.

(9) Rowsell, J. L. C.; Yaghi, O. M. *Angewandte Chemie-International Edition* **2005**, *44*, 4670-4679.

(10) Chae, H. K.; Siberio-Perez, D. Y.; Kim, J.; Go, Y.; Eddaoudi, M.; Matzger, A. J.; O'Keeffe, M.; Yaghi, O. M. *Nature* **2004**, *427*, 523-527.

(11) Sioncke, S.; Verbiest, T.; Persoons, A. *Materials Science & Engineering R-Reports* **2003**, *42*, 115-155.

(12) Verbiest, T.; Van Elshocht, S.; Kauranen, M.; Hellemans, L.; Snauwaert, J.; Nuckolls, C.; Katz, T. J.; Persoons, A. *Science* **1998**, *282*, 913-915.

(13) Marder, S. R.; Kippelen, B.; Jen, A. K. Y.; Peyghambarian, N. *Nature* **1997**, *388*, 845-851.

(14) Decurtins, S.; Schmalle, H. W.; Pellaux, R.; Schneuwly, P.; Hauser, A. *Inorganic Chemistry* **1996**, *35*, 1451-1460.

(15) Decurtins, S.; Schmalle, H. W.; Schneuwly, P.; Ensling, J.; Gutlich, P. *Journal of the American Chemical Society* **1994**, *116*, 9521-9528.

(16) Meijer, E. W.; Havinga, E. E.; Rikken, G. *Physical Review Letters* **1990**, *65*, 37-39.

(17) Ueda, E.; Yoshikawa, Y.; Kishimoto, N.; Tadokoro, M.; Sakurai, H.; Kajiwara, N.; Kojima, Y. *Bulletin of the Chemical Society of Japan* **2004**, *77*, 981-986.

(18) Ferrari, R.; Bernes, S.; de Barbarin, C. R.; Mendoza-Diaz, G.; Gasque, L. *Inorganica Chimica Acta* **2002**, *339*, 193-201.

(19) Takayama, T.; Ohuchida, S.; Koike, Y.; Watanabe, M.; Hashizume, D.; Ohashi, Y. *Bulletin of the Chemical Society of Japan* **1996**, *69*, 1579-1586.

(20) Yaghi, O. M.; Davis, C. E.; Li, G.; Li, H. *J. Am. Chem. Soc.* **1997**, *119*, 2861-2868.

(21) Macrae, C. F.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Shields, G. P.; Taylor, R.; Towler, M.; van de Streek, J. *J. Appl. Cryst.* **2006**, *39*, 453-457.

Table 1 Crystallographic Data and Details of Refinement for Complexes **1** – **4**

	1	**2**	**3**	**4**
empirical formula	$C_{12}H_{22}CdN_6O_9$	$C_{18}H_{32}CdN_6O_8$	$C_{36}H_{66}Cd_3N_{12}O_{18}$	$C_{14}H_{42}CdN_4O_{16}$
Formula weight	506.76	572.90	1292.21	634.92
T (K)	123(2)	123(2)	258(2)	123(2)
λ(Å)	0.71073	0.71073	0.71073	0.71073
Crystal system	monoclinic	monoclinic	orthorhombic	monoclinic
Space group	C 2/c	C 2	P 2_1 2_1 2	C 2
A (Å)	23.241(8)	26.188(3)	13.8710(14)	15.257
B (Å)	4.7300(16)	5.4915(6)	33.898(4)	10.489(3)
C (Å)	15.601(5)	9.0351(10)	5.4168(6)	8.092(2)
α (°)	90.00	90.00	90.00	90.00
β (°)	96.234(5)	107.3240(10)	90.00	96.809(4)
γ (°)	90.00	90.00	90.00	90.00
V (Å^3)	1704.88	1240.41	2546.97	1285.8(6)
Z	4	2	2	2
D_c(g/cm3)	1.974	1.534	1.685	1.640
μ(mm^{-1})	1.347	0.933	1.320	0.928
R_{int}	0.0202	0.0639	0.0411	0.0264
GOF on F^2	1.082	1.059	1.076	1.073
R1[a] [I > 2r(I)]	0.0263	0.0319	0.0354	0.0118
$_W$R2[b] [I > 2r(I)]	0.0629	0.0678	0.0757	0.0310
R1 (all data)	0.0359	0.0356	0.0433	0.0118
$_W$R2 (all data)	0.0671	0.0697	0.0847	0.0310
Largest diff. peak and hole (e/ Å^3)	1.012 and -0.646	0.445 and -0.778	1.028 and -0.498	0.307 and -0.184

[a] $R1 = \Sigma||F_0| - |F_C||/\Sigma|F_0|$. [b] $wR2 = [w(F_0^2 - F_C^2)^2]/[w(F_0^2)^2]^{1/2}$.

Table 2. Selected Metal-Ligand Bond Lengths (Å) and Angles (deg) for the Complexes **1** – **4** [a]

		$Cd(Gly_3)_2.H_2O$ **1**			
Cd(1)-O(3)#1	2.3045(18)	O(3)#2-Cd(1)-N(1)#3	86.24(7)	O(3)#2-Cd(1)-O(4)#5	93.21(7)
Cd(1)-N(1)	2.333(2)	O(3)#2-Cd(1)-O(4)#4	86.79(7)	N(1)#3-Cd(1)-O(4)#5	84.88(7)
Cd(1)-O(4)#4	2.3374(19)	O(3)#1-Cd(1)-O(4)#4	93.21(7)	C(6)-O(3)-Cd(1)#6	129.49(16)
O(3)#1-Cd(1)-N(1)	86.23(7)	N(1)-Cd(1)-O(4)#4	84.88(7)	C(6)-O(4)-Cd(1)#7	119.17(16)
O(3)#2-Cd(1)-N(1)	93.77(7)	N(1)#3-Cd(1)-O(4)#4	95.12(7)	C(1)-N(1)-Cd(1)	108.31(16)
O(3)#1-Cd(1)-N(1)#3	93.76(7)				
		$Cd(Ala_3)_2$ **2**			
Cd(1)-O(1)	2.224(3)	O(1)#1-Cd(1)-N(3)#2	121.02(12)	N(3)#2-Cd(1)-N(3)#3	79.50(18)
Cd(1)-N(3)#2	2.300(3)	O(1)-Cd(1)-N(3)#2	124.84(11)	C(8)-N(3)-Cd(1)#4	121.2(3)
O(1)-Cd(1)-O(1)#1	90.08(15)	O(1)#1-Cd(1)-N(3)#3	124.84(11)	C(1)-O(1)-Cd(1)	106.4(3)
		$Cd(Ala_2)_2$ **3**			
Cd(1)-O(9)#1	2.260(3)	O(9)#1-Cd(1)-N(4)#2	103.37(12)	O(7)-Cd(1)-O(2)	89.62(10)
Cd(1)-N(5)	2.265(3)	N(5)-Cd(1)-N(4)#2	99.41(11)	O(4)-Cd(2)-O(4)#3	96.47(15)
Cd(1)-O(1)	2.312(2)	O(1)-Cd(1)-N(4)#2	93.99(10)	O(4)-Cd(2)-N(2)#3	140.69(12)
Cd(1)-N(4)#2	2.317(3)	O(9)#1-Cd(1)-O(7)	92.87(10)	O(4)#3-Cd(2)-N(2)#3	97.24(13)
Cd(1)-O(7)	2.432(3)	N(5)-Cd(1)-O(7)	71.78(9)	N(2)#3-Cd(2)-N(2)	95.03(19)
Cd(1)-O(2)	2.537(3)	O(1)-Cd(1)-O(7)	85.50(9)	O(4)-Cd(2)-O(3)	96.14(11)
Cd(2)-O(4)	2.223(3)	N(4)#2-Cd(1)-O(7)	163.71(11)	O(4)#3-Cd(2)-O(3)	75.49(11)
Cd(2)-N(2)	2.280(4)	O(9)#1-Cd(1)-O(2)	148.86(10)	N(2)#3-Cd(2)-O(3)	122.95(12)
Cd(2)-O(3)	2.547(3)	N(5)-Cd(1)-O(2)	94.16(11)	N(2)-Cd(2)-O(3)	66.53(11)
O(9)#1-Cd(1)-N(5)	116.09(11)	O(1)-Cd(1)-O(2)	53.58(9)	N(2)-Cd(2)-O(3)#3	122.95(12)
O(9)#1-Cd(1)-O(1)	95.66(9)	N(4)#2-Cd(1)-O(2)	77.16(11)	O(3)-Cd(2)-O(3)#3	167.62(13)
N(5)-Cd(1)-O(1)	141.11(11)				
		$Cd(Ala\text{-}Thr)_2.4H_2O$ **4**			
Cd(1)-O(3)#1	2.2833(10)	O(3)#2-Cd(1)-N(1)	102.02(4)	O(3)#2-Cd(1)-O(1)	85.82(5)
Cd(1)-N(1)	2.2985(12)	N(1)#3-Cd(1)-N(1)	141.70(5)	N(1)#3-Cd(1)-O(1)	86.07(4)
Cd(1)-O(1)	2.3834(10)	O(3)#1-Cd(1)-O(1)#3	85.82(5)	N(1)-Cd(1)-O(1)	72.93(3)
O(3)#1-Cd(1)-O(3)#2	75.42(6)	O(3)#2-Cd(1)-O(1)#3	161.06(3)	O(1)#3-Cd(1)-O(1)	113.03(6)
O(3)#1-Cd(1)-N(1)#3	102.02(4)	N(1)#3-Cd(1)-O(1)#3	72.93(3)	C(3)-O(1)-Cd(1)	113.04(7)
O(3)#2-Cd(1)-N(1)#3	108.11(3)	N(1)-Cd(1)-O(1)#3	86.07(4)	C(8)-O(3)-Cd(1)#4	119.66(8)
O(3)#1-Cd(1)-N(1)	108.11(3)	O(3)#1-Cd(1)-O(1)	161.06(3)		

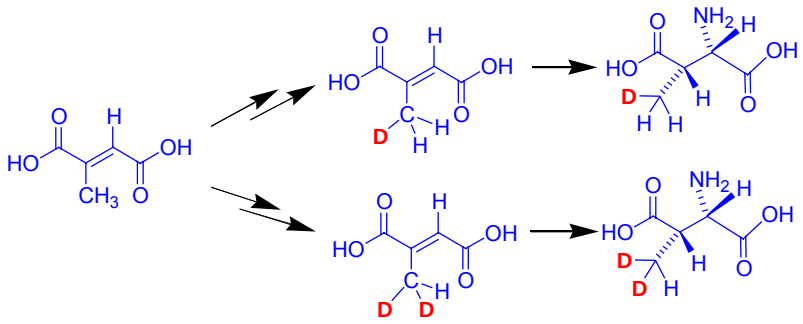
[a]Symmetry transformations used to generate equivalent atoms. For **1**: #1 -x+1/2,-y+1/2,-z+1; #2 x-1/2,y-1/2,z; #3 -x,-y,-z+1; #4 x-1/2,y+1/2,z; #5 -x+1/2,-y-1/2,-z+1; #6 x+1/2,y+1/2,z; #7 x+1/2,y-1/2,z. For **2**: #1 -x,y,-z; #2 -x,y,-z+1; #3 x,y,z-1; #4 x,y,z+1. For **3**: #1 x+1/2,-y+1/2,-z+1; #2 x+1,y,z; #3 -x+1,-y+1,z; #4 x-1/2,-y+1/2,-z+1; #5 x-1,y,z. For **4**: #1 x-1/2,y-1/2,z; #2 -x+1/2,y-1/2,-z+2; #3 -x,y,-z+2; #4 x+1/2,y+1/2,z.

Apendix B) Synthesis of mesaconic acid (methylfumaric acid) and (2*S*, 3*S*)-3-methylaspartic acid specifically mono- or di-deuterated at the methyl group

Hyang-Yeol Lee, Miri Yoon and E. Neil G. Marsh*

Tetrahedron 63 (2007) 4663-4668.

Graphical Abstract



Abstract

Kinetic isotope effects provide a powerful method to investigate the mechanisms of enzyme-catalyzed reactions, but often other slow steps in the reaction such as substrate binding or product release suppress the isotopically sensitive step. For reactions at methyl groups, this limitation may be overcome by measuring the isotope effect by an *intra*-molecular competition experiment. This requires the synthesis of substrates containing regio-specifically mono- or di-deuterated methyl groups. To facilitate mechanistic investigations of the adenosylcobalamin-dependent enzyme, glutamate mutase we have developed a synthesis of mono- and di-deuterated (2S, 3S)-3-methylaspartic acids. Key intermediates are the correspondingly labeled dimethyl mesaconic acids that potentially provide starting materials for a variety of isotopically labeled molecules.

1. Introduction

Isotope effects provide an extremely powerful tool to probe the mechanisms of chemical reactions and have proved particularly useful for investigating of enzyme mechanisms [1-4]. Most isotope effect measurements rely on *inter*-molecular competition between labeled and unlabeled molecules. However, for an isotope effect to be measured in this way, it must be associated with the rate-determining step in reaction, or in the case of isotope effects on V_{max}/K_m occur either at or before the rate-determining step. In many enzyme reactions slow steps that are not isotopically sensitive, such as substrate binding, product release, or protein conformational changes, completely mask the intrinsic isotope effects, limiting our ability to learn about the chemical steps.

However, for chemical reactions that occur at methyl groups it is possible to measure *intrinsic* deuterium isotope effects in enzymes by specifically labeling the methyl carbon with one or two deuterium atoms. The isotope effect can be measured, even when the isotopically sensitive step is *not* rate determining, because it is manifested though *intra-molecular* competition between protium and deuterium atoms, which remain chemically equivalent even in the enzyme active site due to the rapid rotation of the methyl group. The principle of this experiment is illustrated in Figure 1.

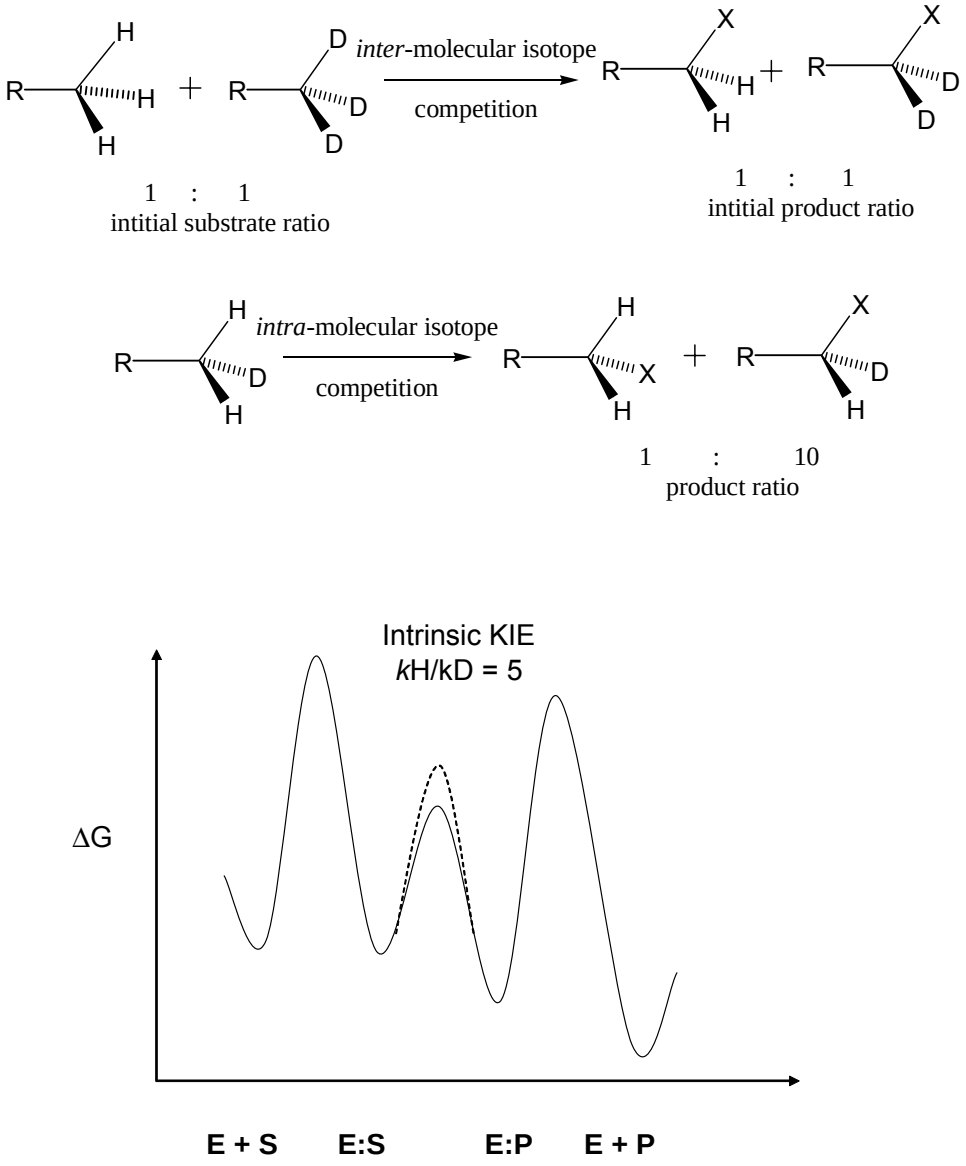
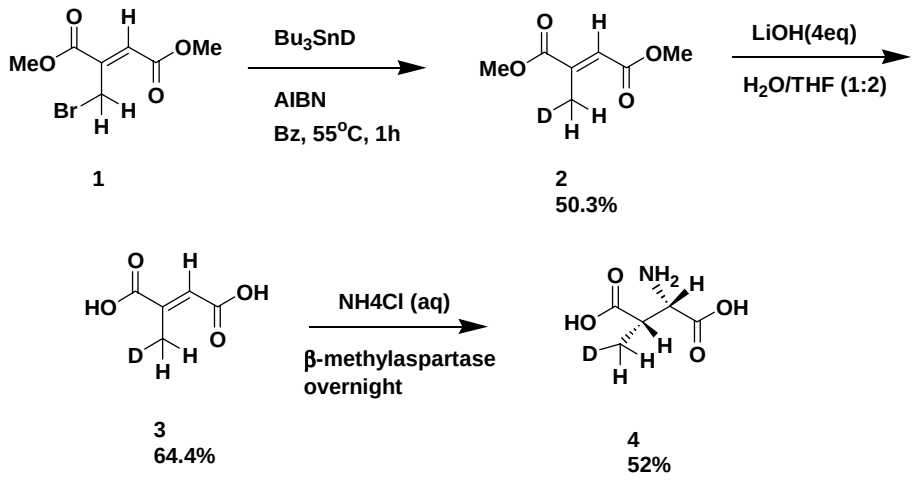


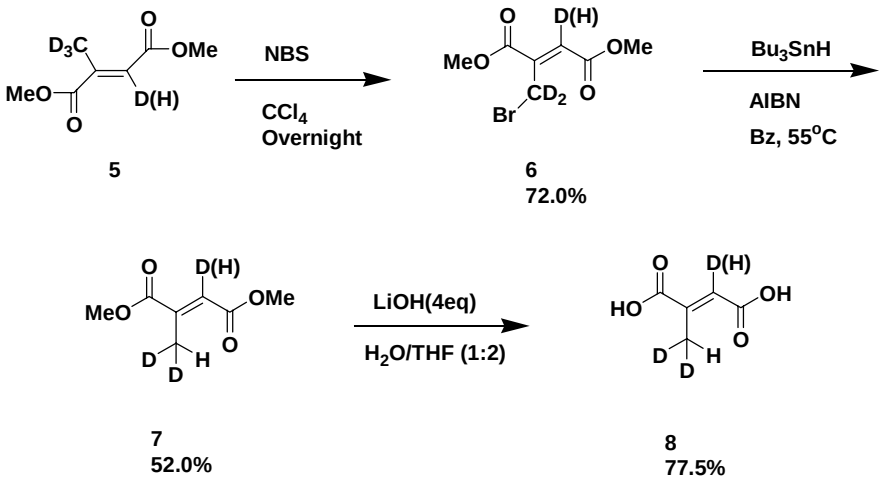
Figure 1 Illustration of the different outcomes obtained when measuring isotope effects in an enzyme catalyzed reaction using *inter*-molecular versus *intra*-molecular competition experiments. Top: Hypothetical free energy profile for an enzyme catalyzing the substitution of a methyl group hydrogen atom with functional group, X, in which substrate binding and product release mask the intrinsic deuterium kinetic isotope effect of 5 on the chemical step. Bottom: a) the relative distribution of isotopically labeled products formed in an *inter*-molecular competition experiment: the ratio of deuterated to undeuterated products is the same as the starting materials, i.e. no isotope effect is expressed. b) the relative distribution of isotopically labeled products formed in an *intra*-molecular competition experiment: here the full isotope effect is expressed (the ratio of deuterium-containing products is 10 : 1, and not 5 : 1, because there are twice as many protons as deuterons in the methyl group).

This method of measuring isotope effects has proved especially useful for investigating enzyme reactions at unactivated methyl groups, for example oxygenation reactions catalyzed by cytochrome p450 enzymes [5]. As part of our efforts to understand the mechanism of hydrogen atom transfer in the coenzyme B_{12}-dependent enzyme glutamate mutase [6-10], we sought to synthesize the substrate (*2S, 3S*)-3-methylaspartate that was regio-specifically mono- or di-deuterated in the methyl group. These substrates allow us to measure the intrinsic kinetic isotope effects on hydrogen transfer between the substrate methyl group and the 5'-carbon of coenzyme B_{12} under single turn-over conditions by setting up an *intra*-molecular competition between protium and deuterium atoms in the methyl group.

Our synthesis is based on the regio-specific deuteration of mesaconic acid (methylfumaric acid), which is an intermediate in the fermentation of glutamate by many anaerobic bacteria. Mesaconate is a versatile intermediate that can be readily converted to 3-methylaspartate through the action of β-methylaspartase (methylaspartate ammonia-lyase), an enzyme that has been used to synthesize a variety of aspartatic acid analogs [11,12]. The enzyme-catalyzed reaction stereospecifically introduces an asymmetric centre adjacent to the labeled methyl group. Mesaconate is also amenable to numerous chemical transformations, potentially allowing a wide range of labeled, branched-chain compounds to be synthesized.



Scheme 2 Strategy for the stereospecific synthesis of regiospecifically mono-deuterated (*2S, 3S*)-3-methylaspartic acid.



Scheme 3 Strategy for the stereospecific synthesis of regiospecifically di-deuterated (*2S, 3S*)-3-methylaspartic acid.

2. Results and Discussion

The strategy for the synthesis of the regio-specifically mono- and di-deuterated methylaspartic acids is shown in Figure 2. Mesaconic acid **1** was first protected as its dimethyl ester **2** in 80 % yield by refluxing overnight in methanol in the presence of 1.5 % (vol/vol) sulfuric acid. As described

previously[13], these conditions yielded predominantly the *cis*-dimethylester **2a** with a small amount of the *trans*-isomer **2b**, the relative proportions varying somewhat from reaction to reaction. Both stereoisomers of **2** were converted, without separation, to 3-(bromomethyl)-fumarate dimethylester, **3**, by reaction with 1.5 equivalents of N-bromosuccinimide and a catalytic amount (10 %) of AIBN as a radical initiator. The reaction proceeded smoothly overnight to produce **3** in good yields (72 %), with only the mono-brominated product being detected. During the reaction, the *cis*- stereoisomer is converted to the *trans*-form. This points to the formation of an allylic radical intermediate during the reaction, which would permit rotation around the double bond.

Compound **3** was carefully purified from unreacted **2** and other by-products by chromatography on silica gel. It is, of course, most important to remove any traces of **2** at this point, because otherwise the isotopic purity of the final product will be diluted with unlabeled material. Introduction of deuterium was accomplished by reductive debromination using tributyl-tin deuteride in dry benzene at 55 ^{o}C with 10 % AIBN as a radical initiator. This gave the mono-deuterated dimethyl methylfumarate **4** in 50 % yield. Lastly, the ester was hydrolyzed using lithium hydroxide in aqueous tetrahydrofuran, room temperature, 30 h, to yield after acidification mono-deuterated mesaconic acid in 64 % yield.

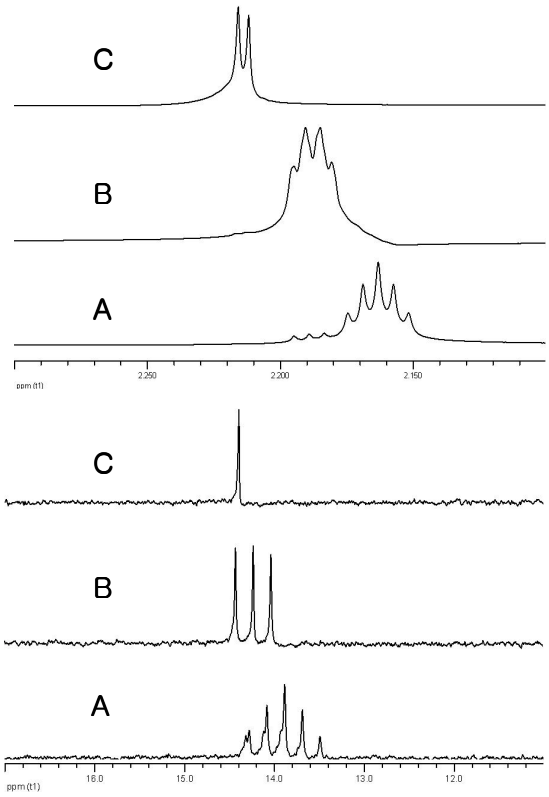


Figure 4 Analysis of the deuterium content of mesaconic acids by NMR. Top: proton NMR (400 MHz) showing the deuterium-induced changes in chemical shift and coupling pattern for the methyl resonance of the mono- and di-deuterated compounds. Bottom: ^{13}C NMR (100 MHz) showing the deuterium-induced changes in chemical shift and coupling pattern of the same compounds. Spectrum A: unlabelled material; spectrum B: mono-deuterated at the methyl group; spectrum C: di-deuterated at the methyl group.

The isotopic composition of the mesaconic acid was determined from the proton NMR spectrum, by taking advantage of the fact that incorporation of deuterium introduces an up-field shift in the protons of the methyl group by about 0.03 ppm (Figure 4). The methyl protons of the non-deuterated compound appear as a doublet, ^4J = 2 Hz, due to long-range coupling with the vinylic proton. The methyl group of the mono-deuterated material appears as an overlapping doublet of triplets that arises from the coupling of the protons to the spin 1 deuterium nucleus, $^4J_{HH} = {}^2J_{HD}$ = 2 Hz. Integration of the peak areas for these

two signals in a sample of mono-deuterated mesaconic acid indicated that the deuterium content was at least 98 mol %. The proton decoupled ^{13}C NMR spectrum showed the expected triplet signal, $^{1}J_{CD}$ = 20 Hz, for the ^{13}C – deuterium coupling (Figure 4).

Mesaconic acid incorporating 2 deuterium atoms in the methyl group was synthesized by an analogous strategy starting with d_4-mesaconic acid. d_4-Mesaconic acid was synthesized by dissolving itaconic acid (methylenesuccinic acid) in 40 % NaOD / D_2O and heating at 120 ^{o}C for 90 min[14]. This resulted in the isomerisation of itaconate to mesaconate and the complete exchange of all 4 protons. After neutralization, the resulting d_4-mesaconic acid was converted to a mixture of *cis*- and *trans*- dimethyl esters (**7a** and **7b**), and then to deuterated 3-(bromomethyl)-fumarate dimethylester (**8**) as described above. The bromination reaction proceeds noticeably slower with the deuterated material, pointing to a significant deuterium kinetic isotope effect for this step. Reduction with tributyl-tin hydride yielded the dimethyl ester of mesaconate, **9**, containing two deuterium atoms in the methyl group, which was then hydrolysed to give (2-d_1, methyl-d_2)-mesaconic acid, **10**.

The proton NMR of **10** showed a multiplet at 2.16 ppm exhibiting the expected quintet coupling pattern for two deuterium atoms coupling to the single proton $^{2}J_{HD}$ = 2 Hz (Figure 3). A weak signal is also present arising from the mono-deuterated methyl group. After accounting for the fact that the relative intensity of this signal is twice that of the dideuterated methyl group, the mono-deuterated material represents less than 2 mol % of the total. A signal from the mono-deuterated material is also evident in the ^{13}C NMR spectrum (Figure 3), where its intensity is enhanced by relaxation through the proton spin. It was also evident from the proton spectrum of **9** that a small amount of protium, ~ 5%, was present at C-2 (data not shown), which appears to have exchanged during bromination. For our purposes, however, this does not present a problem since the α-hydrogen of amino acids can readily be exchanged with water enzymatically, as discussed below.

The isotopically labeled mesaconic acids were readily converted to the corresponding (*2S, 3S*)-3-methylaspartic acids through the action of the enzyme methylaspartase in the presence of 0.5 M ammonium chloride. The deuterated amino acids were produced in 52 % yield and were shown to be active as substrates for glutamate mutase [15]. At this point deuterium present at C-2 of the methylaspartate can easily be removed by exchange with H_2O through the action of glutamate:aspartate amino transferase [10].

3. Conclusion

We have developed a synthesis of mesaconic acids that are specifically mono- or di-deuterated in the methyl group. These compounds could be then enzymatically converted to the correspondingly deuterated (2*S*, 3*S*)-3-methylaspartic acids, in this case to facilitate mechanistic experiments on glutamate mutase. More generally, mesaconic acid, or its dimethyl ester, provides a versatile 5-carbon fragment that may readily be elaborated to more complex molecules through a variety of synthetic transformations. Compounds containing methyl groups that are specifically mono- or di-deuterated should prove useful mechanistic probes of enzymes that catalyze reactions at methyl groups.

4. Experimental

4.1. General

$(Bu)_3SnH$ and $(Bu)_3SnD$ were purchased from Acros Co., 40% sodium deuteroxide in D_2O was purchased from Cambridge Isotope Laboratories, Inc., mesaconic acid was purchased from Sigma, itaconic acid and N-bromosuccinimide were purchase from Aldrich. Recombinant β-methylaspartase was a gift from Prof. David Gani (St. Andrews University) and was purified from *E. coli* as described by Goda et al [16]

4.2. Dimethyl methylmaleate (2a). A solution of mesaconic acid (3.00 g, 23.1 mmol) in methanol (40 mL) and H_2SO_4 (0.5 mL) was refluxed for 24 h under nitrogen. The reaction mixture was concentrated by rotary evaporation. The product was extracted with ethyl acetate and washed with brine three times. The crude material was purified by chromatography on silica gel using 20% ethyl acetate in hexanes to give a colorless liquid comprising mainly **2a** with a small amount of the *trans* isomer (**2b**). Yield of **2a** + **2b** was 80%. **2a**: ^{1}H NMR ($CDCl_3$, 400 MHz) δ 2.01 (s, 3H), 3.67 (s, 3H), 3.78 (s, 3H), 5.81 (s, 1H); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 169.3, 165.3, 145.7, 120.5, 52.3, 51.8, 20.4.

4.3. Dimethyl bromomethylfumarate (3). The mixture of, **2a** and **2b**, (3.16g, 20mmol), N-bromosuccinimide (5.34g, 30mmol), and a catalytic amount of AIBN (0.33g, 2mmol) were mixed in carbon tetrachloride (45ml) and gently refluxed for 24 h in a 100mL round-bottom flask. The reaction was cooled to room temperature and filtered to remove unreacted N-bromosuccinimide. The crude product was purified by silica gel column chromatography using 20% ethyl acetate in hexanes to give pure **3**: Yield 78 %, ^{1}H NMR ($CDCl_3$, 400 MHz) δ 3.79 (s, 3H), 3.84 (s, 3H), 4.68 (s, 2H), 6.79 (s, 1H); ^{13}C NMR ($CDCl_3$, 100 MHz)δ 165.0, 142.7, 128.3, 53.0, 52.2, 22.4.

4.4. Dimethyl d_1-methylfumarate (4). The reduction of dimethyl bromomethylfumarate **3** (0.59g, 2.5 mmol) was carried out with tributyltin deutride (810 μL, 3mmol) and 10 mol % AIBN (45mg, 0.25 mmol) in 10mL of dried benzene. The mixture was stirred at 55^{o}C for 1 hour and cooled to room temperature and KF/celite was added. After stirring the mixture overnight, it was concentrated and purified by silica gel column chromatography. The column was first washed with 2 column volumes of 100% hexanes and the product then eluted with 10% ethyl acetate in hexanes to give pure **4**: Yield 50% , ^{1}H NMR ($CDCl_3$, 300 MHz) δ 2.17 (quartet, *J* = 2 Hz, 2H), 3.67 (s, 3H), 3.71 (s, 3H), 6.68 (s, 1H); ^{13}C NMR ($CDCl_3$, 75 MHz)δ 167.1, 165.8, 143.3, 126.1, 52.1, 51.2, 13.9,

13.6, 13.3; MS (ES, positive) m/z found MH^+, 160.1. $C_7H_9DO_4$ requires M, 159.1580.

4.5. d_1-Mesaconic acid (5). A solution of dimethyl d_1-methylfumarate **4** (200 mg, 1.26 mmol) in 3 mL of THF was added to a solution of LiOH hydrate (196 mg, 4.66 mmol) in 1.5 mL of water and stirred for 30 hours. The reaction mixture was concentrated to remove THF and the pH was adjusted to ~ pH 1.0 by adding 3 M aqueous HCl. The solution was extracted with ethyl acetate three times, dried with sodium phosphate and the solvent removed by rotary evaporation. The solid residue was triturated with warm hexanes three times to give **5**: Yield 64% ([D]:[H] = 98:2). ^{1}H NMR ($CDCl_3$, 400 MHz) δ 6.75 (bs, 1H), 2.18 (quartet, *J* = 2 Hz, 2H); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 170.3, 169.2, 144.8, 128.2, 14.4, 14.2, 14.0.; MS (EI, positive) m/z found 132.0, 113.0, 85.0, 68.0, 40.2.

4.6. d_1-methyl-(2S, 3S)-3-methylaspartic acid (6). Compound **5** was converted to (2S, 3S)-methylaspartic acid, **6**, through the action of β-methylaspartase. 17.1 mg of deuterated mesaconic acid (130 μmol) was dissolved in 300 μL of 250 mM potassium phosphate buffer, pH 8.0, containing 20 mM potassium chloride, 2 mM magnesium chloride and approximately 1.1 M ammonium hydroxide and converted to the diammonium salt. 60 units of β-methylaspartase (360μL of enzyme solution stored in 25 mM potassium phosphate buffer, pH 7.0, 50 % glycerol) were added and the reaction mixture was incubated at 37°C for 6 hours. An additional 27 units (160 μL) of β-methylaspartase and 80 μL of 2 M ammonium chloride were added and the incubation was continued at 37°C for a further 12 hours. The decrease in 240 nm absorbance of the reaction mixture indicated that 105 μmol of mesaconate had been consumed. 50 μL of 12N HCl was added to quench the reaction and the mixture and heated at 94°C for 5 min to precipitate the β-methylaspartase. The precipitated protein was removed by centrifugation of the suspension (12,000 rpm x 3 min). The supernatant solution was extracted five times with a mixture of 4 mL water and 24 mL ethyl ether to remove unreacted mesaconate.

The solution was then adjusted to neutral pH and the concentration of **6** was determined to be 14 mM by assay with β-methylaspartase [17]. The final volume of the solution was 4.8 mL corresponding to a yield of 52 %.

4.7. Dimethyl d_4-methylmaleate (7a). A solution of itaconic acid (5.00 g, 38.46 mmol) in 30 mL of $NaOD/D_2O$ was placed in the Parr reactor and heated to 120 ^oC for 1 h. After cooling to room temperature, the reaction mixture was poured into a 500 mL Erlenmeyer flask in an ice bath. 6 M aqueous HCl was slowly added to the reaction mixture to bring the pH to ~1. Water was removed by freeze-drying in a lyophilizer. To the solid residue 30 mL of dilute HCl was added and the mixture extracted with ethyl acetate (3 x 30 ml), the combined extracts were dried with sodium phosphate and the solvent was then removed by rotary evaporation. The resulting white solid was dissolved in 40 mL of methanol followed by addition of 0.5 mL of concentrated H_2SO_4 and refluxed for 12 hr. The methanol was removed under vacuum and the crude material purified by silica gel column chromatography with 20 % ethyl acetate in hexanes to give a colorless liquid mixture **7a** (as major) and the *trans* isomer **7b** in total 87 % yield. **7a**: ^1H NMR ($CDCl_3$, 300 MHz) δ 3.76 (s, 3H), 3.72 (s, 3H); ^{13}C NMR ($CDCl_3$, 75 MHz) δ 167.2, 165.9, 143.4, 126.2, 126.0, 125.8, 52.3, 51.4, 13.6, 13.5, 13.45, 13.36, 13.3, 13.2.; MS (ES, positive) m/z found MH^+, 163.1. $C_7H_6D_4O_4$ requires M, 162.0826). **7b**: ^1H NMR ($CDCl_3$, 300 MHz) δ 3.83(s, 3H), 3.72 (s, 3H).

4.8. Dimethyl bromo-d_2-methyl-d_1-fumarate (8). The mixture of **7a** and **7b** (1.5 g, 9.3 mmol), N-bromosuccinimide (2.48 g, 13.95 mmol), and a catalytic amount of AIBN (150 mg, 0.93 mmol) in carbon tetrachloride (45 mL) was gently refluxed for 24 h in a 100 mL round-bottom flask. AIBN (75 mg, 0.46 mmol) in carbon tetrachloride (1 mL) was added and refluxed for an additional 12 hours. The reaction was cooled to room temperature and filtered to remove unreacted N-bromosuccinimide. The crude product was purified by silica gel column chromatography using 10% ethyl acetate in hexanes to give pure **8**: Yield 72%,

^{1}H NMR ($CDCl_3$, 400 MHz) δ 3.82 (s, 3H), 3.77 (s, 3H); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 165.1, 164.9, 128.2, 127.9, 127.7, 52.9, 52.2, 22.4, 22.3, 22.1, 21.9.

4.9. Dimethyl d_2-methyl-d_1-fumarate (9). The reduction of dimethyl bromo-d_2-methyl-d_1-fumarate **7** (1.00 g, 4.18 mmol) was carried out with tributyltin deutride (1.35 mL, 5.02 mmol) and 10 mol % AIBN (69 mg, 0.42 mmol) in 20 mL of dried benzene at 55 ^{o}C for 1 hour. The mixture was cooled down to room temperature and KF/celite was added. After stirring the mixture overnight, it was concentrated and purified by silica gel column chromatography. The column was first washed with 2 column volumes of 100% hexanes and the product then eluted with 10% ethyl acetate in hexanes to give pure **9**: Yield 52% , ^{1}H NMR ($CDCl_3$, 400 MHz) δ 3.65 (s, 3H), 3.61 (s, 3H), 2.08 (bs, 1H); ^{13}C NMR ($CDCl_3$, 100 MHz) δ167.2, 165.9, 143.4, 125.9(triplet, J = 25 Hz), 52.3, 51.3, 28.0, 26.5, 17.0, 13.8, 13.6, 13.4, 13.2, 13.0.

4.10. (2-d_1, d_2-methyl)-mesaconic acid (10). A solution of dimethyl d_2-methyl-d_1-fumarate **8** (0.33 g, 2.05 mmol) in 6 mL of THF was added to a solution of LiOH hydrate (0.32 g, 7.59 mmol) in 3 mL of water and stirred for 24 hours. The reaction mixture was concentrated and the pH was adjusted to ~ 1 by adding 3 M HCl aqueous solution. The solution was extracted with ethyl acetate three times and combined extracts dried with sodium phosphate and the solvent removed under vacuum. The solid residue was triturated with warm hexanes three times to give pure **9**: 78% yield ([D]:[H] = 92:8). ^{1}H NMR ($CDCl_3$, 400 MHz) δ 2.15 (quintet, *J* = 2 Hz, 1H); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 170.4, 169.3, 144.6, 128.1, 127.8, 127.6, 14.3, 14.1, 13.9, 13.7, 13.5.; MS (EI, positive) m/z found 134.1, 115.0, 87.0, 70.0, 41.5.

References

1. Cleland, W. W. *CRC Crit. Rev. Biochem.* **1982,** *13*, 385-428.
2. Cleland, W. W. Secondary isotope effects on enzymatic reactions. In *Isotopes in Organic Chemistry;* Buncel, E.; Lee, C. C. Eds.; Elsevier: Amsterdam, 1987; Vol. 7.
3. Schramm, V. L. *Curr. Opin. Chem. Biol.* **2001,** *5*, 556-563.
4. Nagel, Z. D.; Klinman, J. P. *Chem. Rev.* **2006,** *106*, 3095-3118.
5. Iyer, K. R.; Jones, J. P.; Darbyshire, J. F.; Trager, W. F. *Biochemistry* **1997,** *36*, 7136-7143.
6. Cheng, M.-C.; Marsh, E. N. G. *Biochemistry* **2005,** *44*, 2686-2691.
7. Cheng, M.-C.; Marsh, E. N. G. *Biochemistry* **2004,** *43*, 2155-2158.
8. Chih, H.-W.; Marsh, E. N. G. *Biochemistry* **2001,** *40*, 13060-13067.
9. Marsh, E. N. G.; Drennan, C. L. *Curr. Opin. Chem. Biol.* **2001,** *5*, 499-505.
10. Marsh, E. N. G.; Ballou, D. P. *Biochemistry* **1998,** *37*, 11864-72.
11. Barker, H. A.; Smyth, R. D.; Wilson, R. M. *Fed. Proc.* **1958,** *17*, 185-185.
12. Gulzar, M. S.; Akhtar, M.; Gani, D. *J. Chem. Soc. Perkin Trans. 1* **1997**, 649-655.
13. Kar, A.; Argade, N. P. *J. Org. Chem.* **2002,** *67*, 7131-7114.
14. Eagar, R. G., Jr.; Herbst, M. M.; Barker, H. A.; Richards, J. H. *Biochemistry* **1972,** *11*, 253-64.
15. Chen, H. P.; Marsh, E. N. G. *Biochemistry* **1997,** *36*, 14939-45.
16. Goda, S. K.; Minton, N. P.; Botting, N. P.; Gani, D. *Biochemistry* **1992,** *31*, 10747-10756.
17. Barker, H. A.; Smyth, R. D.; Wilson, R. M.; Weissbach, H. *J. Biol. Chem.* **1959,** *234*, 320-328.

Apendix C) Enhancing the biological stability of an antimicrobial peptide using fluorous amino acids

Lindsy M. Gottler, Hyang-Yeol Lee, Charles E. Shelburne, A. Ramamoorthy and E. Neil G. Marsh

ChemBioChem, 2007, in press

* Lindsey has done most of this work and I was involved in the synthesis of HFI.

Introduction

The emergence of bacterial strains resistant to most of the clinically useful antibiotics has provided the impetus to develop new classes of antibiotics that may combat bacterial resistance more effectively. Antimicrobial peptides (AMPs) are small peptides (typically 15 – 30 residues) that show promise as therapeutic agents against bacteria, fungi and viruses. Widely distributed in plants and animals, they form part of the innate immune system's defense against microbes. Although highly diverse in sequence and structure, almost all AMPs share the property of being highly amphiphathic, with one face of the peptide being hydrophobic and the other face presenting a cluster of positively charged residues. AMPs function by disrupting bacterial membranes, which contain predominantly negatively charged phospholipids. Eukaryotic membranes, which contain predominantly neutral phospholipids, are not targeted. Although promising as broad spectrum antibiotics, AMPs are susceptible to proteolysis in vivo by endogenous or bacterial proteases, which may considerably diminish their effectiveness. Attempts to overcome this by increasing the concentration of AMP often leads to toxic side effects, most notably lysis of red blood cells, which has been attributed to non-specific hydrophobic interactions between the peptide and the eukaryotic cell membrane. Here we describe an approach to overcome these limitations, by replacing hydrophobic residues in AMPs with extensively fluorinated (fluorous) analogs. Previously it has been shown that extensively fluorinated analogs of leucine and

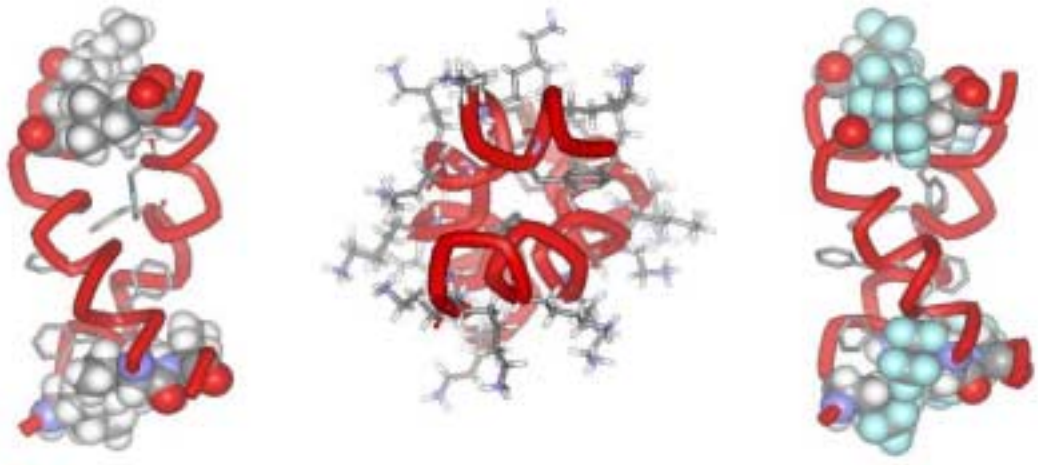
valine can significantly stabilize small proteins against thermal and chemical denaturation, an effect that may be attributed to the extremely hydrophobic nature of fluorocarbons. Furthermore, fluorocarbon solvents exhibit the unusual self-segregating properties that have been exploited in organic synthesis and in the design of self-segregating coiled-coil peptides. We reasoned that if these properties of fluorocarbons could be designed into AMPs, then the expected increased structural stability might provide resistance to proteolysis, whereas the "Teflon" nature of fluorocarbons might reduce non-specific hydrophobic interactions with eukaryotic cells.

Materials and Methods

Reagents. Rink Amide resin, t-Boc-protected amino acids, and 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophophate (HBTU) were purchased from NovaBiochem. L-5,5,5,5',5',5'-Hexafluoroleucine (HFL) was synthesized as described previously1 and converted to the t-Boc protected derivative by standard procedures. 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphotidylcholine (POPC) was purchased from Avanti Polar Lipids. Chymotrypsin and trypsin were purchased from Boeringer Manheim GmbH. All other chemicals were purchased from Fisher Chemical Co. and used without any further purification.

Peptide Synthesis and Purification. Peptide fluorogainin-1 (sequence below)

was synthesized using t-Boc-protected amino acids for Merrifield manual solid-phase synthesis on MBHA resin; couplings were performed using in situ neutralization/HBTU protocol described by Schnolzer et al.2 on a 0.25 mM scale. The peptide was cleaved from the resin using “high”-HF conditions. MSI-78 was synthesized as described previously3.



Fluorogainin-1 GXGKFXKKAKKFGKAFVKXXKK (X = hFLeu)

MSI-78 GIGKFLKKAKKFGKAFVKILKK

Figure 1 Left: structure of MSI-78 dimer showing Leu and Ile residues in CPK rendering. Center: view along the helical axis of MSI-78 dimer showing disposition of Lys residues. Right: model of fluorogainin-1 based on structure of MSI-78 showing hFleu residues in CPK rendering.

The peptide was redissolved at ~10mg/mL in Milli-Q water and purified by reverse-phase HPLC on a Waters semipreparative C18 column equilibrated in 0.1% TFA and eluted with a linear gradient from 5 to 90% acetonitrile containing 0.1% TFA. The peptide was determined to be pure by analytical HPLC and

TOF MS ES+ mass spectrometry: expected mass for MSI-78 4HFL of 2950 amu, detected mass of 493(+6), 591(+5), 739(+4) and 985(+3). The concentration of the peptides was determined by the absorbance at 257 nm due to three phenylalanine residues, using and extinction coefficient of 585 cm-1 M-1.

Circular dichroism spectroscopy. CD spectra of peptides were recorded with an Aviv 62DS spectropolarimeter at 25°C in triplicate and averaged. Mean residue ellipticities, ΘM, were calculated using equation 1

$$\Theta M = \Theta obs/10lcn \qquad (1)$$

where Θobs is the ellipticity measured in millidegrees, c is the molar concentration, l is the cell path length in centimeters, and n is the number of residues in the protein.

To examine the secondary structure of MSI-78 and fluorogainin-1 samples containing 200 µL of each peptide (68 µM) were made up in Buffer A (100mM Tris-HCl, 10mM CaCl2, pH 7.8). POPC liposomes were prepared freshly in Buffer A to make a 136 mM solution of uni-lamellar liposomes by sonication (microtip, 10sec intervals, level 3) with Fisher Scientific 550 Sonic Dismembrator until clarification of the solution. 20 µL of the liposome solution was added to the peptide to make a solution containing peptide to lipid in a 1:200 molar ratio. The CD spectrum for each peptide/liposome solution was recorded in triplicate between 200 and 250 nm.

Proteolytic digestion. Stock solutions of peptides at a concentration of 680 μM in Milli-Q water were prepared and stored at -20°C. 2x Buffer A. A 0.5mg/mL solution of tryptophan (to act as an internal standard) were prepared and stored at 4°C. Protease solutions were prepared at a concentration of 1mg/mL in 1mM HCl immediately prior to use. Liposome solutions were prepared as described above immediately prior to use.

Proteolytic digestions were set up by mixing 10 μL of peptide stock, 20 μL 2x Buffer A (or 20 μL 2x Buffer A containing liposomes), and 0.5 μL tryptophan solution. Reactions were initiated by the addition of 0.5μL of stock protease solution, or 0.5μL of a 10x dilution of stock protease solution, to give a 1:80 or 1:800 ratio of peptide to enzyme respectively. Aliquots (5-10μL) were removed at different times and quenched using an equal volume of 1 M HCl. Samples were stored at -20°C until analysis by HPLC. Reactions were performed at 0°C or 25°C. All reactions were repeated 3 times. Controls containing all reagents except protease were performed and no degradation of the parent peptide was observed for MSI-78 or fluorogainin-1 over a 10 hour period (data not shown).

Samples were diluted to ~40μL with 5% acetonitrile, 0.15% acetic acid and analyzed by reverse phase HPLC using analytical C18 column. Samples were eluted with a linear gradient from 5 to 90% acetonitrile, 0.15% acetic acid. Parent peak heights (peak heights corresponding to the original peptide) were normalized based on the tryptophan peak as an internal standard. Plots of the parent peak height vs. time were generated to obtain a linear fit of the data. Time zero was extrapolated from the linear fits and % peptide remaining was

calculated from eq 2

(peak height at tn/peak height at t0) x 100 (2)

where tn is any time point and t0 is time zero.

Bacterial strains and Growth. E. coli D5α was obtained from Invitrogen (Invitrogen, Carlsbad, CA). Enterococcus feacalis cultures were a gift from Dr. Donald B. Clewell, University of Michigan. Bacillus subtilus (ATCC 663), Kocheria rhizophila (ATCC 9341), Earobacter aerogenese (ATCC 13408), Klebsiella pneumoniae ATCC 4352, Proteus mirabilis (ATCC 25933), Salmonella enteritis Typhimerium ATCC 14028, Streptococcus pyogenes (ATCC 19615), Staphylococcus aureus (ATCC 6538) and Shigella sonnei (ATCC 25931) were obtained from MicroBiologics, St. Cloud, Minnesota. All species were maintained by weekly transfer on Trypicase Soy Agar and broth cultures grown directly from individual colonies in Trypticase Soy Broth.

MIC determinations. A sterile, 96- well micro-titer plate (NUNC) was used as the platform for the assay. An overnight culture of each bacterium was diluted to 106/mL in sterile phosphate buffered saline and seeded into each well (100 µl) of the plate. Doubling dilutions (500 µg/mL-7.8 µg/mL) of MSI-78 or F-MSI-78 (100µL) in replicates of 8 were then added and the cultures covered with a sterile adhesive plastic film. After centrifugation for 1 minute at 800 x g to collect the entire inoculum to the bottom of the plate it was incubated at 37 ^{o}C

overnight in air. Growth was determined by measuring the OD595 of each well using a microwell plate reader (GENious, Techan, Manendorff, Switzerland). The MIC for each organism was determined as the lowest dilution of each peptide without significant growth above the inoculum ($p < 0.01$, T-test). The differences between peptides for each bacterial strain were evaluated using a t-test.

Hemolytic Assay. A sterile, 96- well round bottom micro-titer plate (Costar) was used as the platform for the assay. 2.5% bovine red blood cells in 100 uL of PBS were added to the wells of the plate. Doubling dilutions (500 µg/mL-7.8 µg/mL) of MSI-78 or fluorogainin-1 (100µL) in replicates of 8 were then added and the cultures covered with a sterile adhesive plastic film. The plate was incubated at 37 ^{o}C for 2 hours at room temperature and examined by eye. The highest dilution of peptides without an intact cell pellett was noted. No hemolysis was noted with either MSI-78 or fluorogainin-1.

To test this concept we have synthesized a fluorous analog of the potent AMP, MSI-78 (also called pexiganan). MSI-78 is a synthetic analogue of magainin-2, an α-helical AMP originally isolated from *Xenopus laevis*, that has potent antibacterial activity [7,18,19]. The peptide is unstructured in free solution but forms a dimeric antiparallel α-helical coiled coil on association with lipid bilayers. The peptide is unstructured in free solution but forms a dimeric antiparallel α-helical coiled coil on association with lipid bilayers (Figure 1) and is believed to exert its antibacterial effect by forming toroidal pores in the bacterial

membrane. We have replaced the two leucine and isoleucine residues in MSI-78 with the fluorous amino acid L-5,5,5,5',5',5'-hexafluoroleucine (hFLeu) to produce a molecule we call fluorogainin-1 (Figure 1). We have compared the antimicrobial activity, hemolytic activity and resistance to proteolysis of fluorogainin-1 and MSI-78.

Fluorogainin-1 was synthesized manually using boc-protected amino acids by standard protocols; boc-protected hFLeu was synthesized as described previously. As a control, MSI-78 was synthesized by standard automated methods as described previously. Both peptides were purified by reverse phase HPLC and their identities confirmed by MALDI-MS.

The secondary structure of the peptides (70 μM, in 100 mM Tris-Cl buffer, 10mM $CaCl_2$, pH 7.8, at 25 ^{o}C) was examined by C.D. spectroscopy. In the absence of liposomes both peptides were unstructured (data not shown), consistent with their being highly positively charged. Incubation of the peptides with uni-lamella liposomes freshly prepared from 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphotidyl-choline (POPC, 14 mM final conc.) resulted in both peptides adopting an α-helical structure (Figure 2). Interestingly, fluorogainin-1 appeared significantly less helical than MSI-78 as judged by C.D. with a mean residue ellipticity at 222 nm of only about 2/3 that of the non-fluorinated peptide. This is in accord with recent studies that show hFLeu to have a poor helix propensity, despite the fact that it has been found to stabilize α-helical proteins against unfolding.

We next examined the stability of fluorogainin-1 and MSI-78 towards

proteolysis by two common proteases, trypsin and chymotrypsin. These proteases provide a stringent test for structural stability as both peptides have multiple potential cleavage sites for both enzymes. Peptides (350 μM in 100 mM Tris-Cl buffer, 10mM $CaCl_2$, pH 7.8) were incuubated in the presence of POPC liposomes and 1.25 % (w/w) protease at 25 ^{o}C for various times and the extent of proteolysis determined by reverse phase HPLC. Under these conditions MSI-78 was almost completely degraded by either protease within 30 min, however fluorogainin-1 showed no signs of degradation even after 10 hours (Figure 2). In the absence of liposomes, where the peptides are unstructured, both peptides were equally rapidly degraded by either trypsin or chymotrypsin. Therefore, the resistance to proteolysis of fluorogainin-1 appears to be due either to it forming a more stable α-helix, or a stronger interaction with the liposome, and not because the incorporation of hFLeu per se prevents the peptide from being digested.

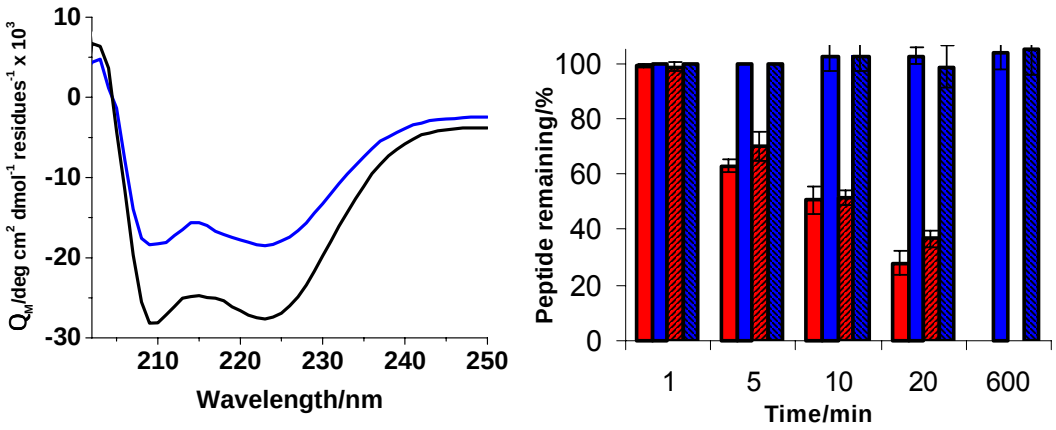


Figure 2 Left C.D. spectra for MSI-78 (solid circles) and fluorogainin-1 (open circles) in the presence of POPC. Right: Stability of MSI-87 (red) and fluorogainin-1 (blue) towards digestion with trypsin (solid bars) or chymotrypsin (hashed bars).

Antimicrobial activity. The minimum inhibitory concentrations (MIC) of MSI-78 and fluorogainin-1 were determined for a panel of 11 bacterial strains that included both Gram positive and Gram negative strains of common pathogenic bacteria. The MIC values were determined by two-fold serial dilutions of each AMP into cultures of each bacterial strain and noting the concentration at which bacterial growth ceased. The results are summarized in Table 1. Overall, the fluorous AMP showed a similar spectrum of antibiotic activity to the non-fluorous AMP, however, against two important pathogenic bacteria the fluorous AMP was significantly ($p < 0.05$) more potent. Fluorogainin-1 had an MIC of 16 μM against Klebsiella pneumoniae, whereas MSI-78 showed no activity, and the MIC of fluorogainin-1 was ~ four times lower than MSI-78 against Staphylococcus aureus. The only bacterium tested for which fluorogainin-1 was significantly ($p < 0.05$) less effective than MSI-78 was Streptococcus pyogenes.

The hemolytic activity of both AMPs was tested against sheep erythrocytes. Neither peptide exhibited any hemolytic activity at concentrations up to 250 μg/mL, the highest concentration tested.

The thermograms for each peptide binding to liposomes are shown in Figure 3. The enthalpies (calculated from the peak areas of each injection of peptide) for each peptide binding to the lipids are similar: ΔH for MSI-78 is -14.4 ± 0.2 kcal/mol, whereas ΔH for fluorogainin-1 is slightly less exothermic at -12.5 ± 0.3 kcal/mol. These enthalpies are similar to those measured previously for magainin-2 peptide binding to SUVs[27]. The titration of fluorogainin-1 appears to show a slight increase in heat released with increasing injection number

indicating some dependence on peptide concentration, although the significance of this unclear. The enthalpic contribution to liposome binding has been determined to arise primarily from electrostatic interactions between the positively charged peptide and negatively charged lipid head groups[27]. Since MSI-78 and fluorogainin-1 contain identical cationic residues, the electrostatic interactions are expected to be very similar. This result strongly suggests that the increase in hydrophobicity imparted by the fluorous residues is primarily responsible for modifying the biological properties of fluorogainin-1.

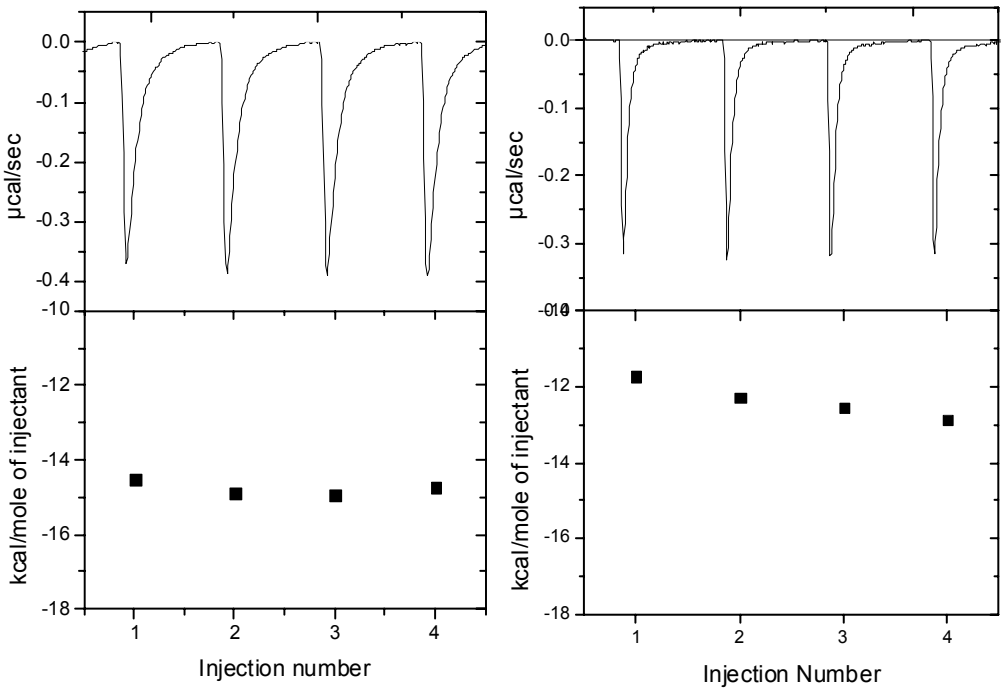


Figure 3 *Left*: Isothermal titration calorimetry of MSI-78 into SUVs. *Right:* Isothemal titration calorimetry of fluorogainin-1 into SUVs. In each case the upper panel shows heat of reaction for 4 consecutive 4 µL injections of 200 µM peptide solution into a large molar excess (1.2 mL, 15 mM) of liposomes, and the lower panel the integration of these peaks.

In conclusion, by introducing fluorous amino acid residues into an AMP we have conferred almost complete resistance to proteolysis of the fluorous AMP under conditions where the non-fluorinated AMP is rapidly degraded, while retaining broad-spectrum antimicrobial activity. Protection against proteolysis is only observed in the presence of lipids, suggesting that lipid-peptide interactions are important. Since MSI-78 has been shown to dimerise to form a coiled coil in a membrane environment, it is plausible that incorporation of hFLeu would strengthen the interactions between peptides because it is significantly more hydrophobic than Leu. This would, in turn, promote the formation of structured dimers that are resistant to proteolysis. It is also noteworthy that fluorogainin-1 shows improved potency against against *K. pneumoniae* and *S. aureus*, with MICs of 16 μg/mL against both bacteria. Although the selectivity of AMPs for some bacteria and not others is poorly understood, it is known that the resistance of *S. aureus* to AMPs is due, at least in part, to secretion of proteases, it is known that the resistance of *S. aureus* to AMPs is due, at least in part, to secretion of proteases. The resistance of fluorogainin-1 to proteolysis may explain its improved potency against this important bacterial pathogen. More generally, our results suggest the strategy of incorporating fluorous residues may enhance the efficacy of other biologically active membrane-associated peptides; e. g. those involved in membrane fusion or ion-channel formation or those that have anti-cancer or anti-viral activity.

Bacterial Strain	MIC (μg/mL)	
	MSI-78	fluorogainin-1
Bacillus subtilis	< 4	8
Kocheria rhizophila	< 4	8
Enterobacter aerogenes	>250.00	>250.00
Klebsiella pneumoniae	>250.00	16[a]
Proteus mirabilis	>250.00	>250.00
Salmonella enteritis	16	32
Streptococcus pyogenes	8	62[a]
Escherichia coli (DH5a)	< 4	8
Staphylococcus. aureus (UH-11)	62	16[a]
Shigella sonnei	16	32
Enterococcus fecaelis (OG1 X)	>250	>250

Table 1 MICs of MSI-78 and fluorogainin 1 against various bacteria
[a] MIC of fluorogainin-1 significantly lower than MSI-78 ($p<0.05$); [b]MIC of fluorogainin-1 significantly higher than MSI-78 ($p<0.05$)

References

[1] Y. Shai, *Curr. Pharmaceutical Design* **2002,***8,* 715-725.

[2] M. Zasloff, *Nature* **2002,***415,* 389-395.

[3] H. G. Boman, *Ann. Rev. Immunol.* **1995,***13,* 61-92.

[4] Y. Shai, *Biochim. Biophys. Acta-Biomembranes* **1999,***1462,* 55-70.

[5] M. H. Wu, E. Maier, R. Benz, R. E. W. Hancock, *Biochemistry* **1999,***38,* 7235-7242.

[6] I. Zelezetsky, A. Tossi, *Biochim. Biophys. Acta-Biomembranes* **2006,***1758,* 1436-1449.

[7] W. L. Maloy, U. P. Kari, *Biopolymers* **1995,***37,* 105-122.

[8] A. Studer, S. Hadida, R. Ferritto, S. Y. Kim, P. Jeger, P. Wipf, D. P. Curran, *Science* **1997,***275,* 823-826.

[9] N. C. Yoder, K. Kumar, *Chem. Soc. Rev.* **2002,***31,* 335-341.

[10] E. N. G. Marsh, *Chem. Biol.* **2000,***7,* R153-R157.

[11] B. Bilgicer, K. Kumar, *Proc. Natl. Acad. Sci. (USA)* **2004,***101,* 15324-15329.

[12] K.-H. Lee, H.-Y. Lee, M. S. Slutsky, J. T. Anderson, E. N. G. Marsh, *Biochemistry* **2004,***43,* 16277-16284.

[13] H.-Y. Lee, K.-H. Lee, H. M. Al-Hashimi, E. N. G. Marsh, *J. Am. Chem. Soc.* **2006,***128,* 337-343.

[14] Y. Tang, G. Ghirlanda, N. Vaidehi, J. Kua, D. T. Mainz, W. A. Goddard, W. F. DeGrado, D. A. Tirrell, *Biochemistry* **2001,***40,* 2790-2796.

[15]a) A. Niemz, D. A. Tirrell, *J. Am. Chem. Soc.* **2001,***123,* 7407-7413; b) B. Bilgicer, A. Fichera, K. Kumar, *J. Am. Chem. Soc.* **2001,***123,* 4393-4399.

[16] D. Gimenez, C. Andreu, M. del Olmo, T. Varea, D. Diaz, G, Asensio, *Bioorg. Med. Chem*. **2006,** *14*, 6971-6978

[17] K. Matsuzaki, *Biochim. Biophys. Acta-Rev. Biomembranes* **1998,***1376,* 391-400.

[18] M. Zasloff, B. Martin, H. C. Chen, *Proc. Natl. Acad. Sci. (USA)* **1988,***85,* 910-913.

[19] F. Porcelli, B. A. Buck-Koehntop, S. Thennarasu, A. Ramamoorthy, G. Veglia, *Biochemistry* **2006,***45,* 5793-5799.

[20] K. J. Hallock, D. K. Lee, A. Ramamoorthy, *Biophys. J.* **2003,***84,* 3052-3060.

[21] M. Schnolzer, P. Alewood, A. Jones, D. Alewood, S. B. H. Kent, *Intl. J. Peptide and Protein Res.* **1992,***40,* 180-193.

[22] J. T. Anderson, P. L. Toogood, E. N. G. Marsh, *Org. Lett.* **2002,***4,* 4281-4283.

[23] A. Ramamoorthy, S. Thennarasu, D. K. Lee, A. M. Tan, L. Maloy, *Biophys. J.* **2006,***91,* 206-216.

[24] C. E. Shelburne, F. Y. An, V. Dholpe, A. Ramamoorthy, D. E. Lopatin, M. S. Lantz, *J.f Antimicrob. Chemotherapy* **2007,***59,* 297-300.

References

(1) Donald Voet, J. V.; John Wiley & Sons, Inc.; Vol. 1.

(2) Dobrodumov, A.; Gronenborn, A. M. Filtering and selection of structural models: Combining docking and nmr *Proteins* **2003**, *53*, 18-32.

(3) Halperin, I.; Ma, B. Y.; Wolfson, H.; Nussinov, R. Principles of docking: An overview of search algorithms and a guide to scoring functions *Proteins* **2002**, *47*, 409-443.

(4) Bilgicer, B.; Fichera, A.; Kumar, K. A coiled coil with a fluorous core *J Am Chem Soc* **2001**, *123*, 4393-4399.

(5) Harbury, P. B.; Kim, P. S.; Alber, T. Crystal-structure of an isoleucine-zipper trimer *Nature* **1994**, *371*, 80-83.

(6) Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T. A switch between 2-stranded, 3-stranded and 4-stranded coiled coils in gcn4 leucine-zipper mutants *Science* **1993**, *262*, 1401-1407.

(7) Munson, M.; Obrien, R.; Sturtevant, J. M.; Regan, L. Redesigning the hydrophobic core of a 4-helix-bundle protein *Protein Sci* **1994**, *3*, 2015-2022.

(8) Predki, P. F.; Nayak, L. M.; Gottlieb, M. B. C.; Regan, L. Dissecting rna-protein interactions - rna-rna recognition by rop *Cell* **1995**, *80*, 41-50.

(9) Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, W. A.; DeGrado, W. F.; Tirrell, D. A. Stabilization of coiled-coil peptide domains by introduction of trifluoroleucine *Biochemistry-Us* **2001**, *40*, 2790-2796.

(10) Tang, Y.; Tirrell, D. A. Biosynthesis of a highly stable coiled-coil protein containing hexafluoroleucine in an engineered bacterial host *J Am Chem Soc* **2001**, *123*, 11089-11090.

(11) DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A. De novo design and structural characterization of proteins and metalloproteins *Annu Rev Biochem* **1999**, *68*, 779-819.

(12) Lombardi, A.; Summa, C. M.; Geremia, S.; Randaccio, L.; Pavone, V.; DeGrado, W. F. Retrostructural analysis of metalloproteins: Application to the design of a minimal model for diiron proteins *P Natl Acad Sci USA* **2000**, *97*, 6298-6305.

(13) Severin, K.; Lee, D. H.; Kennan, A. J.; Ghadiri, M. R. A synthetic peptide

ligase *Nature* **1997**, *389*, 706-709.

(14) Walsh, S. T. R.; Cheng, H.; Bryson, J. W.; Roder, H.; DeGrado, W. F. Solution structure and dynamics of a de novo designed three-helix bundle protein *P Natl Acad Sci USA* **1999**, *96*, 5486-5491.

(15) West, M. W.; Wang, W. X.; Patterson, J.; Mancias, J. D.; Beasley, J. R.; Hecht, M. H. De novo amyloid proteins from designed combinatorial libraries *P Natl Acad Sci USA* **1999**, *96*, 11211-11216.

(16) Kamtekar, S.; Schiffer, J. M.; Xiong, H. Y.; Babik, J. M.; Hecht, M. H. Protein design by binary patterning of polar and nonpolar amino-acids *Science* **1993**, *262*, 1680-1685.

(17) Moffet, D. A.; Hecht, M. H. De novo proteins from combinatorial libraries *Chem Rev* **2001**, *101*, 3191-3203.

(18) Wei, Y. N.; Liu, T.; Sazinsky, S. L.; Moffet, D. A.; Pelczer, I.; Hecht, M. H. Stably folded de novo proteins from a designed combinatorial library *Protein Sci* **2003**, *12*, 92-102.

(19) Ho, S. P.; Degrado, W. F. Design of a 4-helix bundle protein - synthesis of peptides which self-associate into a helical protein *J Am Chem Soc* **1987**, *109*, 6751-6758.

(20) Dieckmann, G. R.; McRorie, D. K.; Lear, J. D.; Sharp, K. A.; DeGrado, W. F.; Pecoraro, V. L. The role of protonation and metal chelation preferences in defining the properties of mercury-binding coiled coils *J Mol Biol* **1998**, *280*, 897-912.

(21) Dieckmann, G. R.; McRorie, D. K.; Tierney, D. L.; Utschig, L. M.; Singer, C. P.; OHalloran, T. V.; PennerHahn, J. E.; DeGrado, W. F.; Pecoraro, V. L. De novo design of mercury-binding two- and three-helical bundles *J Am Chem Soc* **1997**, *119*, 6195-6196.

(22) Gibney, B. R.; Mulholland, S. E.; Rabanal, F.; Dutton, P. L. Ferredoxin and ferredoxin-heme maquettes *P Natl Acad Sci USA* **1996**, *93*, 15041-15046.

(23) Lee, K. H.; Lee, H. Y.; Slutsky, M. M.; Anderson, J. T.; Marsh, E. N. G. Fluorous effect in proteins: De novo design and characterization of a four-alpha-helix bundle protein containing hexafluoroleucine *Biochemistry-Us* **2004**, *43*, 16277-16284.

(24) Song, L. Y.; Caguiat, J.; Li, Z. R.; Shokes, J.; Scott, R. A.; Olliff, L.; Summers, A. O. Engineered single-chain, antiparallel, coiled coil mimics the merr metal binding site *J Bacteriol* **2004**, *186*, 1861-1868.

(25) Oshea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T. X-ray structure of the

gcn4 leucine zipper, a 2-stranded, parallel coiled coil *Science* **1991**, *254*, 539-544.

(26) Lo, T. P.; Komarpanicucci, S.; Sherman, F.; Mclendon, G.; Brayer, G. D. Structural and functional-effects of multiple mutations at distal sites in cytochrome-c *Biochemistry-Us* **1995**, *34*, 5259-5268.

(27) Murphy, G. A.; Spedale, E. J.; Powell, S. T.; Pillus, L.; Schultz, S. C.; Chen, L. The sinc-terminal coiled coil is required for telomeric and mating type silencing in saccharomyces cerevisiae *J Mol Biol* **2003**, *334*, 769-780.

(28) Thepaut, M.; Maiorano, D.; Guichou, J. F.; Auge, M. T.; Dumas, C.; Mechali, M.; Padilla, A. Crystal structure of the coiled-coil dimerization motif of geminin: Structural and functional insights on DNA replication regulation *J Mol Biol* **2004**, *342*, 275-287.

(29) Weis, W. I.; Drickamer, K. Trimeric structure of a c-type mannose-binding protein *Structure* **1994**, *2*, 1227-1240.

(30) Kowalski, K.; Merkel, A. L.; Booker, G. W. H-1, c-13 and n-15 resonance assignments of the third spectrin repeat of alpha-actinin-4 *J Biomol Nmr* **2004**, *29*, 533-534.

(31) Kammerer, R. A.; Kostrewa, D.; Progias, P.; Honnappa, S.; Avila, D.; Lustig, A.; Winkler, F. K.; Pieters, J.; Steinmetz, M. O. A conserved trimerization motif controls the topology of short coiled coils *P Natl Acad Sci USA* **2005**, *102*, 13891-13896.

(32) Bullough, P. A.; Hughson, F. M.; Skehel, J. J.; Wiley, D. C. Structure of influenza hemagglutinin at the ph of membrane-fusion *Nature* **1994**, *371*, 37-43.

(33) Stetefeld, J.; Jenny, M.; Schulthess, T.; Landwehr, R.; Engel, J.; Kammerer, R. A. Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer *Nat Struct Biol* **2000**, *7*, 772-776.

(34) Banner, D. W.; Kokkinidis, M.; Tsernoglou, D. Structure of the co1e1 rop protein at 1.7 a resolution *J Mol Biol* **1987**, *196*, 657-675.

(35) Lederer, F.; Glatigny, A.; Bethge, P. H.; Bellamy, H. D.; Mathews, F. S. Improvement of the 2.5 a resolution model of cytochrome-b562 by redetermining the primary structure and using molecular graphics *J Mol Biol* **1981**, *148*, 427-448.

(36) Chantalat, L.; Jones, N. D.; Korber, F.; Navaza, J.; Pavlovsky, A. G. The crystal-structure of wild-type growth-hormone at 2.5 angstrom resolution *Protein Peptide Lett* **1995**, *2*, 333-340.

(37) Keeler, C.; Dannies, P. S.; Hodsdon, M. E. The tertiary structure and backbone dynamics of human prolactin *J Mol Biol* **2003**, *328*, 1105-1121.

(38) Gibney, B. R.; Rabanal, F.; Reddy, K. S.; Dutton, P. L. Effect of four helix bundle topology on heme binding and redox properties *Biochemistry-Us* **1998**, *37*, 4635-4643.

(39) Gibney, B. R.; Rabanal, F.; Skalicky, J. J.; Wand, A. J.; Dutton, P. L. Design of a unique protein scaffold for maquettes *J Am Chem Soc* **1997**, *119*, 2323-2324.

(40) Salgado, E. N.; Faranone-Mennella, J.; Tezcan, A. F. Controlling protein-protein interactions through metal coordination: Assembly of a 16-helix bundle protein *J Am Chem Soc* **2007**, *129*, 13374.

(41) Pauling, L.; Corey, R. B. Compound helical configurations of polypeptide chains - structure of proteins of the alpha-keratin type *Nature* **1953**, *171*, 59-61.

(42) Crick, F. H. C. The fourier transform of a coiled-coil *Acta Crystallogr* **1953**, *6*, 685-689.

(43) Lupas, A. Coiled coils: New structures and new functions *Trends Biochem Sci* **1996**, *21*, 375-382.

(44) Lumb, K. J.; Kim, P. S. Measurement of interhelical electrostatic interactions in the gcn4 leucine-zipper *Science* **1995**, *268*, 436-439.

(45) Yu, Y.; Monera, O. D.; Hodges, R. S.; Privalov, P. L. Ion pairs significantly stabilize coiled-coils in the absence of electrolyte *J Mol Biol* **1996**, *255*, 367-372.

(46) Farrer, B. T.; Harris, N. P.; Balchus, K. E.; Pecoraro, V. L. Thermodynamic model for the stabilization of trigonal thiolato mercury(ii) in designed three-stranded coiled coils *Biochemistry-Us* **2001**, *40*, 14696-14705.

(47) Matzapetakis, M.; Farrer, B. T.; Weng, T. C.; Hemmingsen, L.; Penner-Hahn, J. E.; Pecoraro, V. L. Comparison of the binding of cadmium(ii), mercury(ii), and arsenic(iii) to the de novo designed peptides tri l12c and tri l16c *J Am Chem Soc* **2002**, *124*, 8042-8054.

(48) Bilgicer, B.; Kumar, K. Synthesis and thermodynamic characterization of self-sorting coiled coils *Tetrahedron* **2002**, *58*, 4105-4112.

(49) Bilgicer, B.; Xing, X.; Kumar, K. Programmed self-sorting of coiled coils with leucine and hexafluoroleucine cores *J Am Chem Soc* **2001**, *123*, 11815-11816.

(50) Boice, J. A.; Dieckmann, G. R.; DeGrado, W. F.; Fairman, R. Thermodynamic analysis of a designed three-stranded coiled coil *Biochemistry-Us* **1996**, *35*, 14480-14485.

(51) Ogihara, N. L.; Weiss, M. S.; Degrado, W. F.; Eisenberg, D. The crystal structure of the designed trimeric coiled coil coil-v(a)l(d): Implications for engineering crystals and supramolecular assemblies *Protein Sci* **1997**, *6*, 80-88.

(52) Kalsbeck, W. A.; Robertson, D. E.; Pandey, R. K.; Smith, K. M.; Dutton, P. L.; Bocian, D. F. Structural and electronic properties of the heme cofactors in a multi-heme synthetic cytochrome *Biochemistry-Us* **1996**, *35*, 3429-3438.

(53) Barkay, T.; Miller, S. M.; Summers, A. O. Bacterial mercury resistance from atoms to ecosystems *Fems Microbiol Rev* **2003**, *27*, 355-384.

(54) Wright, J. G.; Tsang, H. T.; Pennerhahn, J. E.; Ohalloran, T. V. Coordination chemistry of the hg-merr metalloregulatory protein - evidence for a novel tridentate hg-cysteine receptor-site *J Am Chem Soc* **1990**, *112*, 2434-2435.

(55) Zeng, Q. D.; Stalhandske, C.; Anderson, M. C.; Scott, R. A.; Summers, A. O. The core metal-recognition domain of merr *Biochemistry-Us* **1998**, *37*, 15885-15895.

(56) Caguiat, J. J.; Watson, A. L.; Summers, A. O. Cd(ii)-responsive and constitutive mutants implicate a novel domain in merr *J Bacteriol* **1999**, *181*, 3462-3471.

(57) Wendt, H.; Berger, C.; Baici, A.; Thomas, R. M.; Bosshard, H. R. Kinetics of folding of leucine-zipper domains *Biochemistry-Us* **1995**, *34*, 4097-4107.

(58) Bryson, J. W.; Betz, S. F.; Lu, H. S.; Suich, D. J.; Zhou, H. X. X.; Oneil, K. T.; Degrado, W. F. Protein design - a hierarchical approach *Science* **1995**, *270*, 935-941.

(59) Hill, R. B.; DeGrado, W. F. Solutions structure of alpha d-2, a nativelike de novo designed protein *J Am Chem Soc* **1998**, *120*, 1138-1145.

(60) Betz, S. F.; DeGrado, W. F. Controlling topology and native-like behavior of de novo-designed peptides: Design and characterization of antiparallel four-stranded coiled coils *Biochemistry-Us* **1996**, *35*, 6955-6962.

(61) Stephen, J.; Woodfin, R.; Corlett, J. E.; Paul, N. D.; Jones, H. G.; Ayres, P. G. Response of barley and pea crops to supplementary uv-b radiation *J Agr Sci* **1999**, *132*, 253-261.

(62) Lee, H. Y.; Lee, K. H.; Al-Hashimi, H. M.; Marsh, E. N. G. Modulating protein structure with fluorous amino acids: Increased stability and native-like structure conferred on a 4-helix bundle protein by hexafluoroleucine *J Am Chem Soc* **2006**, *128*, 337-343.

(63) Anderson, J. T.; Toogood, P. L.; Marsh, E. N. G. A short and efficient

synthesis of l-5,5,5,5',5',5'-hexafluoroleucine from n-cbz-l-serine *Org Lett* **2002**, *4*, 4281-4283.

(64) Chiu, H. P.; Suzuki, Y.; Gullickson, D.; Ahmad, R.; Kokona, B.; Fairman, R.; Cheng, R. P. Helix propensity of highly fluorinated amino acids *J Am Chem Soc* **2006**, *128*, 15556-15557.

(65) Horvath, I. T. Fluorous biphase chemistry *Abstr Pap Am Chem S* **1998**, *216*, U825-U825.

(66) Horvath, I. T.; Rabai, J. Facile catalyst separation without water - fluorous biphase hydroformylation of olefins *Science* **1994**, *266*, 72-75.

(67) Luo, Z. Y.; Zhang, Q. S.; Oderaotoshi, Y.; Curran, D. P. Fluorous mixture synthesis: A fluorous-tagging strategy for the synthesis and separation of mixtures of organic compounds *Science* **2001**, *291*, 1766-1769.

(68) Studer, A.; Hadida, S.; Ferritto, R.; Kim, S. Y.; Jeger, P.; Wipf, P.; Curran, D. P. Fluorous synthesis: A fluorous-phase strategy for improving separation efficiency in organic synthesis *Science* **1997**, *275*, 823-826.

(69) Deng, H.; O'Hagan, D.; Schaffrath, C. Fluorometabolite biosynthesis and the fluorinase from streptomyces cattleya *Nat Prod Rep* **2004**, *21*, 773-784.

(70) Dong, C. J.; Huang, F. L.; Deng, H.; Schaffrath, C.; Spencer, J. B.; O'Hagan, D.; Naismith, J. H. Crystal structure and mechanism of a bacterial fluorinating enzyme *Nature* **2004**, *427*, 561-565.

(71) Marsh, E. N. G. Towards the nonstick egg: Designing fluorous proteins *Chem Biol* **2000**, *7*, R153-R157.

(72) Yoder, N. C.; Kumar, K. Fluorinated amino acids in protein design and engineering *Chem Soc Rev* **2002**, *31*, 335-341.

(73) Tang, Y.; Ghirlanda, G.; Petka, W. A.; Nakajima, T.; DeGrado, W. F.; Tirrell, D. A. Fluorinated coiled-coil proteins prepared in vivo display enhanced thermal and chemical stability *Angew Chem Int Edit* **2001**, *40*, 1494-+.

(74) Bilgicer, B.; Kumar, K. De novo design of defined helical bundles in membrane environments *P Natl Acad Sci USA* **2004**, *101*, 15324-15329.

(75) Boulanger, M. J.; Garcia, K. C. Shared cytokine signaling receptors: Structural insights from the gp130 system *Adv Protein Chem* **2004**, *68*, 107-+.

(76) Hecht, M. H.; Das, A.; Go, A.; Bradley, L. H.; Wei, Y. N. De novo proteins

from designed combinatorial libraries *Protein Sci* **2004**, *13*, 1711-1723.

(77) Summa, C. M.; Rosenblatt, M. M.; Hong, J. K.; Lear, J. D.; DeGrado, W. F. Computational de novo design, and characterization of an a(2)b(2) diiron protein *J Mol Biol* **2002**, *321*, 923-938.

(78) Schnolzer, M.; Alewood, P.; Jones, A.; Alewood, D.; Kent, S. B. H. Insitu neutralization in boc-chemistry solid-phase peptide-synthesis - rapid, high-yield assembly of difficult sequences *Int J Pept Prot Res* **1992**, *40*, 180-193.

(79) Cohn, E. J.; Edsall, J. T., Proteins, amino acids and peptides as ions and dipolar ions.

(80) Harding, S. E.; Rowe, A. J.; Horton, H. C. *Analytical ultracentrifugation in biochemisty and polymer science*; The Royal Society of Chemistry, 1992.

(81) Oakley, M. G.; Hollenbeck, J. J. The design of antiparallel coiled coils *Curr Opin Struc Biol* **2001**, *11*, 450-457.

(82) Krylov, D.; Mikhailenko, I.; Vinson, C. A thermodynamic scale for leucine-zipper stability and dimerization specificity - e-interhelical and g-interhelical interactions *Embo J* **1994**, *13*, 2849-2861.

(83) Thompson, K. S.; Vinson, C. R.; Freire, E. Thermodynamic characterization of the structural stability of the coiled-coil region of the bzip transcription factor gcn4 *Biochemistry-Us* **1993**, *32*, 5491-5496.

(84) Fairman, R.; Chao, H. G.; Mueller, L.; Lavoie, T. B.; Shen, L. Y.; Novotny, J.; Matsueda, G. R. Characterization of a new 4-chain coiled-coil - influence of chain-length on stability *Protein Sci* **1995**, *4*, 1457-1469.

(85) Munson, M.; Balasubramanian, S.; Fleming, K. G.; Nagi, A. D.; OBrien, R.; Sturtevant, J. M.; Regan, L. What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties *Protein Sci* **1996**, *5*, 1584-1593.

(86) Eguchi, Y.; Tomizawa, J. Complexes formed by complementary rna stem-loops - their formations, structures and interaction with cole1 rom protein *J Mol Biol* **1991**, *220*, 831-842.

(87) Spyridaki, A.; Glykos, N. M.; Kotsifaki, D.; Fadouloglou, V. E.; Kokkinidis, M. Crystallization and diffraction to ultrahigh resolution (0.8 angstrom) of a designed variant of the rop protein *Acta Crystallogr D* **2000**, *56*, 1015-1016.

(88) Magliery, T. J.; Regan, L. A cell-based screen for function of the four-helix bundle protein rop: A new tool for combinatorial experiments in biophysics *Protein Eng Des Sel* **2004**, *17*, 77-83.

(89) Willis, M. A.; Bishop, B.; Regan, L.; Brunger, A. T. Dramatic structural and thermodynamic consequences of repacking a protein's hydrophobic core *Structure* **2000**, *8*, 1319-1328.

(90) Chakrabartty, 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-Us* **1993**, *32*, 5560-5565.

(91) Maloy, W. L.; Kari, U. P. Structure-activity studies on magainins and other host-defense peptides *Biopolymers* **1995**, *37*, 105-122.

(92) Zasloff, M.; Martin, B.; Chen, H. C. Antimicrobial activity of synthetic magainin peptides and several analogs *P Natl Acad Sci USA* **1988**, *85*, 910-913.

(93) Muller, K.; Faeh, C.; Diederich, F. Fluorine in pharmaceuticals: Looking beyond intuition *Science* **2007**, *317*, 1881-1886.

(94) Leroux, F. Atropisomerism, biphenyls, and fluorine: A comparison of rotational barriers and twist angles *Chembiochem* **2004**, *5*, 644-649.

(95) Mikami, K.; Itoh, Y.; Yamanaka, M. Fluorinated carbonyl and olefinic compounds: Basic character and asymmetric catalytic reactions *Chem Rev* **2004**, *104*, 1-16.

(96) Nagai, T.; Nishioka, G.; Koyama, M.; Ando, A.; Miki, T.; Kumadaki, I. Reactions of trifluoromethyl ketones .9. Investigation of the steric effect of a trifluoromethyl group based on the stereochemistry on the dehydration of trifluoromethyl homoallyl alcohols *J Fluorine Chem* **1992**, *57*, 229-237.

(97) Jackel, C.; Salwiczek, M.; Koksch, B. Fluorine in a native protein environment - how the spatial demand and polarity of fluoroalkyl groups affect protein folding *Angew Chem Int Edit* **2006**, *45*, 4198-4203.

(98) Jackel, C.; Seufert, W.; Thust, S.; Koksch, B. Evaluation of the molecular interactions of fluorinated amino acids with native polypepticles *Chembiochem* **2004**, *5*, 717-720.