Structural and Thermodynamic Investigation into the Effects of Protein Fluorination

by

Benjamin Curtis Buer

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in The University of Michigan
2012

Doctoral Committee:

Professor E. Neil G. Marsh, Chair
Professor Hashim M. Al-Hashimi
Professor Anna K. Mapp
Professor Janet L. Smith

Dedication

This dissertation is dedicated to my family and friends for their encouragement and support during my years at the University of Michigan, and to Lisa who has made this a fun journey.

Acknowledgements

I would like to thank my advisor, Prof. E. Neil G. Marsh, for his advice and guidance through my graduate school journey. I would also like to thank all members of the Marsh laboratory, past and present, for their helpful suggestions, thoughtful discussions and friendship. Lastly, this work would not be possible without the expertise and know-how of collaborators. I would like to thank Dr. Roberto de la Saud Bea for his mentorship and synthetic expertise, Dr. Jeetender Chugh and Prof. Hashim Al-Hashimi for expertise and guidance with NMR experiments, Dr. Jennifer Meagher and Prof. Jeanne Stuckey for X-ray crystallography expertise and guidance.

Table of Contents

List of Figures

Figures

List of Tables

List of Abbreviations

Abbreviations

AMP	Antimicrobial peptide
ANS	1-Anilino-8-napthalene sulfonate
Boc	*t*-Butyloxycarbonyl
CD	Circular dichroism
CPMG	Carr-Purcell-Meiboom-Gill
DCM	Dichloromethane
DMS	Dimethylsulfide
DIEA	*N*,*N*-Diisopropylethylamine
DHPC	1,2-Dihexanoyl-*sn*-glycero-3-phosphocholine
DMF	*N*,*N*-Dimethylformamide
DMPC	1,2-Dimyristoyl-*sn*-glycero-3-phosphocholine
DMPG	1,2-Dimyristoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol)
Fmoc	Fluorenylmethyloxycarbonyl
GuHCl	Guanidinium hydrochloride
HCTU	2-(6-chloro-1-H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HF	Hydrofluoric acid
hFLeu	Hexafluoroleucine
HPLC	High pressure liquid chromatography
ITC	Isothermal titration calorimetry
MALDI	Matrix assisted laser desorption ionization
MBHA	4-Methylbenzhydrylamine
MIC	Minimal inhibitory concentration
MS	Mass spectrometry
nFhSer	*O*-Perfluoro-*t*-butyl-L-homoserine
NMR	Nuclear magnetic resonance
PBS	Phosphate buffered saline
pFPhe	Pentafluorophenylalanine
POPC	1-Palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylcholine
POPG	1-Palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-(1'-*rac*-glycerol)
SDS	Sodium dodecyl sulfate
tBAla	β-*t*-Butyl-alanine
tFeG	Trifluoroethylglycine
tFIle	Trifluoroisoleucine
tFLeu	Trifluoroleucine
tFVal	Trifluorovaline

SUV	Small unilamellar vesicle
TFA	Trifluoroacetic acid

Chapter 1

Introduction

1.1 – Introduction

Protein engineering has relied heavily on mutagenesis, both site-directed and random, as a tool to modify the structure and function of enzymes and proteins. Until recently, this approach was limited to substitutions within the 20 natural (proteogenic) amino acids or post-translational chemical modification of protein structure. However, the development of various methods that allow a wide variety of non-natural, or non-proteogenic, amino acids to be incorporated into proteins has expanded the possibilities for modifying protein structure enormously. In particular, it is now possible to introduce a diverse range of chemical functionality into proteins that are not seen in Nature. Prominent among the non-natural amino acids that have been investigated are highly fluorinated analogs of hydrophobic amino acids (shown in Fig. 1.1). These have attracted interest because of the unusual physicochemical properties of perfluorocarbons and their potential to enhance the stability of natural proteins.

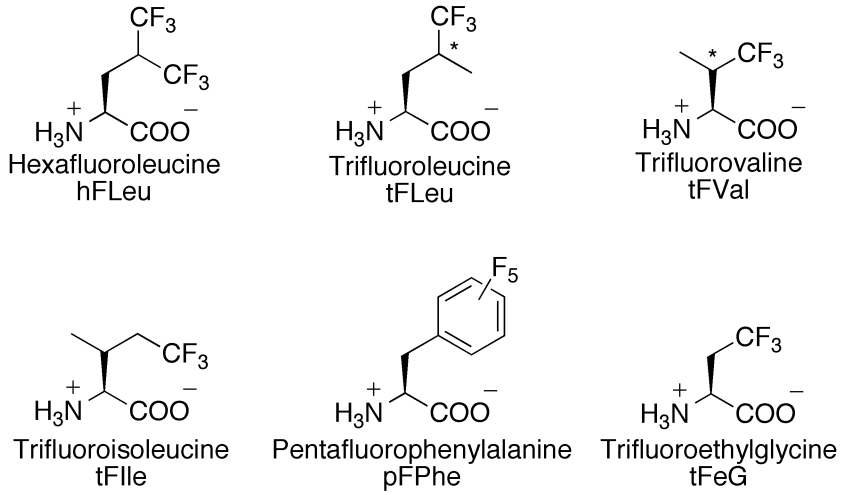

Figure 1.1. Fluorocarbon analogues of hydrophobic amino acids (asterisks denotes stereocenter, studies on both S and R isomers detailed in Chapter 1 section 4.1).

1.2 – Fluorocarbon Properties

The unique physical properties of fluorinated molecules derive, in part, from the extreme electronegativity of fluorine. A C-F bond is polarized in the opposite direction to a C-H bond, and is both more stable, by about 14 kcal/mol, and less polarizable than a C-H bond. Fluorine is often considered isosteric with hydrogen as the van der Waals radius of fluorine, 1.35 Å, is only slightly greater than that of hydrogen, 1.2 Å; however the C-F bond is significantly longer, ~1.4 Å, than a C-H bond, ~1.0 Å. Nevertheless, fluorine can frequently be substituted for hydrogen in small molecules, with minimal impact on their binding to proteins and enzymes. This property has been widely exploited to increase the hydrophobicity of pharmaceuticals and improve their bioavailability [1].

Perfluorocarbons are highly chemically inert and extremely hydrophobic; for example, solvent partitioning experiments have shown a trifluoromethyl group to be twice as hydrophobic as a methyl group [2]. They also exhibit unusual phase segregating properties; for example, water, hexane and perfluorohexane are each mutually

immiscible, and may therefore be described as both hydrophobic and lipophobic. This unusual property, which is known as the "fluorous effect", underlies the non-stick properties of materials such polytetrafluoroethylene (PTFE). It has also been exploited in organic synthesis to extract molecules equipped with fluorocarbon tags from multicomponent reaction mixtures into perfluorocarbon solvents (Fig. 1.2A) [3,4].

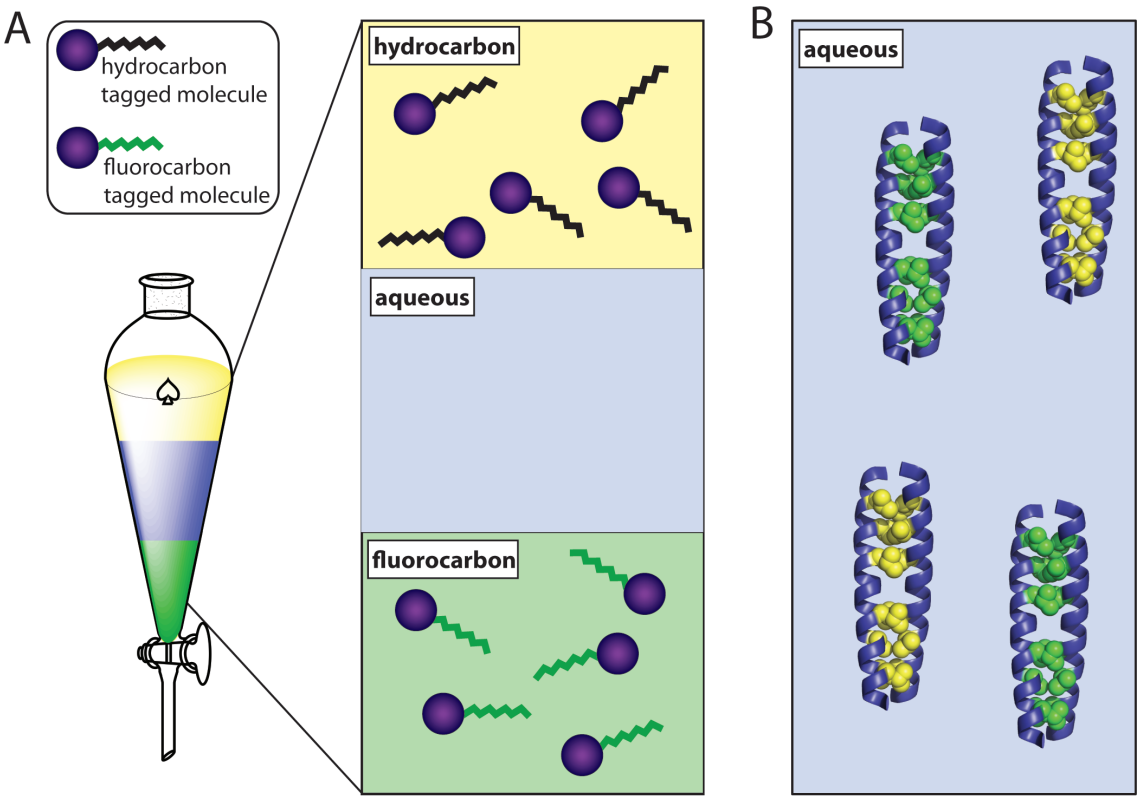

Figure 1.2. The fluorous effect in small molecules and proteins. (A) Triphasic mixtures are formed when fluorinated (green layer) solvents are mixed with aqueous (blue layer) and hydrocarbon (yellow layer) solvents. Solvent immiscibility can be used as a purification technique, when small molecules (purple spheres) are tagged with hydrocarbon (black) or fluorocarbon (green) tails. (B) Proposed self-segregation of proteins with fluorinated (green) and nonfluorinated (yellow) cores.

Fluorine is essentially absent from biology, making the introduction of man-made fluorinated amino acids a unique way to modify proteins. Fluorinated amino acids have

long been used as sensitive and non-perturbing NMR probes to examine changes in local protein environment and dynamics [5-12]. Incorporating nonperturbing fluorine probes into proteins has been readily accomplished through feeding auxotrophic bacterial strains monofluorinated analogues of tyrosine, tryptophan or phenylalanine. These ^{19}F nuclei are very sensitive to the local electronic environment of proteins, with chemical shift changes reporting on subtle changes in conformation or catalysis. The relaxation properties of the ^{19}F nucleus have also been exploited to inform on protein dynamics, which can be examined in a native-like state using solution NMR. A portion of my research has focused on using fluorinated amino acids to investigate the transient interactions between antimicrobial peptides (AMPs) and lipid membranes, this study is detailed in Chapter 6.

More recently, the use of fluorine to modulate the physicochemical properties of proteins by incorporating highly fluorinated analogues of hydrophobic amino acids, in particular leucine, isoleucine, valine and phenylalanine, into their structures has gained interest [13-15]. Such proteins exhibit increased stability towards unfolding by chemical denaturants, solvents and heat, and degradation by proteases. It has also been postulated that a protein-based fluorous effect could be created by incorporating highly fluorinated residues at protein interfaces, thereby introducing an interaction orthogonal to the hydrophobic effect with which to direct protein recognition and ligand binding (Fig. 1.2B).

1.3 – Synthesis of Highly Fluorinated Proteins

The incorporation of any non-proteogenic amino acid into a protein poses a synthetic challenge. It has been known for a long time that sparingly-fluorinated analogs of many hydrophobic amino acids can be incorporated biosynthetically with high efficiency using bacterial strains that are auxotrophic for the parent amino acid [16,17]. However, extensively fluorinated amino acids are not recognized by the endogenous protein synthesis machinery. Most studies on highly fluorinated proteins have focused on short proteins and peptides and utilized solid phase peptide synthesis to introduce fluorinated residues, which is straightforward and provides a great deal of flexibility.

Alternatively, Tirrell and coworkers have developed methods for residue specific incorporation of fluorinated amino acids such as trifluoroleucine (tFLeu), trifluoroisoleucine (tFIle) and trifluorovaline (tFVal) that can be activated by endogenous tRNA synthetases. The advantage of this method is that large proteins can be fluorinated, however, protein expression does not result in 100 % incorporation of fluorinated analogs due to the presence of natural amino acid substrate derived from cellular proteins; efficiencies of 70 – 90 % are typical. *In vivo* protein incorporation also results in global substitution of a particular amino acid, which limits some applications. Highly fluorinated analogs such as hexafluoroleucine (hFLeu) are not efficiently recognized by tRNA synthetases, and thus not incorporated *in vivo*. However, this limitation has been overcome by enhancing leucyl-tRNA synthetase activity through overexpression in *E. coli* of leucyl-tRNA synthetase. [17-20]

In principle, fluorous amino acids could be introduced biosynthetically in a site-specific manner using an evolved orthogonal aminoacyl-tRNA synthetase/tRNA pair, an approach which has been pioneered by Schultz and co-workers [21,22]. To my knowledge,

it has not been used so far to produce highly fluorinated proteins, presumably because of the technical barrier presented by the need to evolve the requisite tRNA synthetase. Similarly, expressed protein ligation techniques [23,24] offer a way to produce semi-synthetic proteins that contain segments of non-natural residues, but again this technique has not yet been used for the production of extensively fluorinated proteins.

1.4 – Stabilizing Proteins Through Fluorination

The hydrophobic effect is the major driving force in protein folding, so it is not surprising that fluorinated amino acids, being more hydrophobic than their hydrocarbon counterparts, are generally effective in stabilizing protein structure. The replacement of a side chain methyl group for a trifluoromethyl group increases volume by ~15 $Å^3$ but hydrophobic packing is only minimally perturbed because side chain shape is retained. For example, solvent partitioning studies predict that the increased hydrophobicity of hFLeu can stabilize a protein by ~0.4 kcal/mol/residue over Leu, although predicted stability increases for proteins as high as 1.1 kcal/mol/hFLeu residue have been reported [25,26]. Most studies have focused on the incorporation of fluorinated analogs of valine, leucine and isoleucine into the hydrophobic core of small α-helical proteins. In addition, the effect of fluorination on the stability of β-sheet proteins, transmembrane proteins, and other small globular proteins has also been studied, some of these studies also investigating fluorinated phenylalanine analogues.

1.4.1 – Studies on Parallel Coiled-Coil Proteins

The first reports of fluorous amino acids enhancing the stability of proteins came

from laboratories of Kumar, Tirrell and DeGrado, who studied the effects of fluorination on the coiled-coil domain of GCN4 and a de novo-designed coiled-coil dimer, A1. Substitution of the four Leu and three Val residues in GCN4 (Fig. 1.3A) by tFLeu and tFVal respectively (which were racemic at the side chain stereocenter as seen in Fig. 1.1) resulted in a relatively modest stabilization of ~1 kcal/mol over the non-fluorinated version [27]. Substitution of the six core **d**-position leucine residues of A1 by tFLeu resulted in a protein that was 0.4 kcal/mol/tFLeu residue more stable [17]. Increasing the fluorine content of A1 by substituting hFLeu for Leu resulted in a further stabilization of A1 by ~0.6 kcal/mol/hFLeu residue [28]. Fluorinated versions of the coiled-coil DNA binding protein GCN4-bZip and its dimerization subdomain GCN4-p1d have been produced synthetically [18]. In this case, substituting tFLeu for the four **d**-position Leu residues of GCN4-p1d substantially increased the thermal stability of the protein and provided a modest increase in the free energy of unfolding, $\Delta\Delta G_{unfold} \sim 0.6$ kcal/mol. Importantly, the fluorinated GCN4-bZip retained the ability of the wild-type protein to bind DNA, suggesting that fluorination may be a general strategy for increasing stability without compromising biological activity.

Further studies on the biosynthetic incorporation of fluorous amino acids examined the stereochemical preference of tRNA-synthetases for stereoisomers of tFVal, tFLeu and tFIle. Studies using purified valyl- and isoleucyl-tRNA synthetase demonstrated, somewhat surprisingly, that (2*S*,3*R*)-tFVal was recognized by both enzymes with similar efficiency, whereas the (2*S*,3*S*)-isomer was inactive [19,29]. *In vivo* incorporation of tFVal into murine dihydrofolate reductase gave similar results with (2*S*,3*R*)-tFVal being incorporated into both valine and isoleucine positions in the enzyme

[19]. Similarly, it was shown that the isoleucine analog 5-tFIle was efficiently recognized by isoleucyl-tRNA synthetase and incorporated into proteins, whereas the structural isomer 3-tFIle was not recognized [20]. The leucyl-tRNA synthetase appears to be less discriminating towards side chain fluorination as both (2*S*,4*S*)-tFLeu and (2*S*,4*R*)-tFLeu were biosynthetically incorporated into the coiled-coil protein A1 with similar efficiency [30].

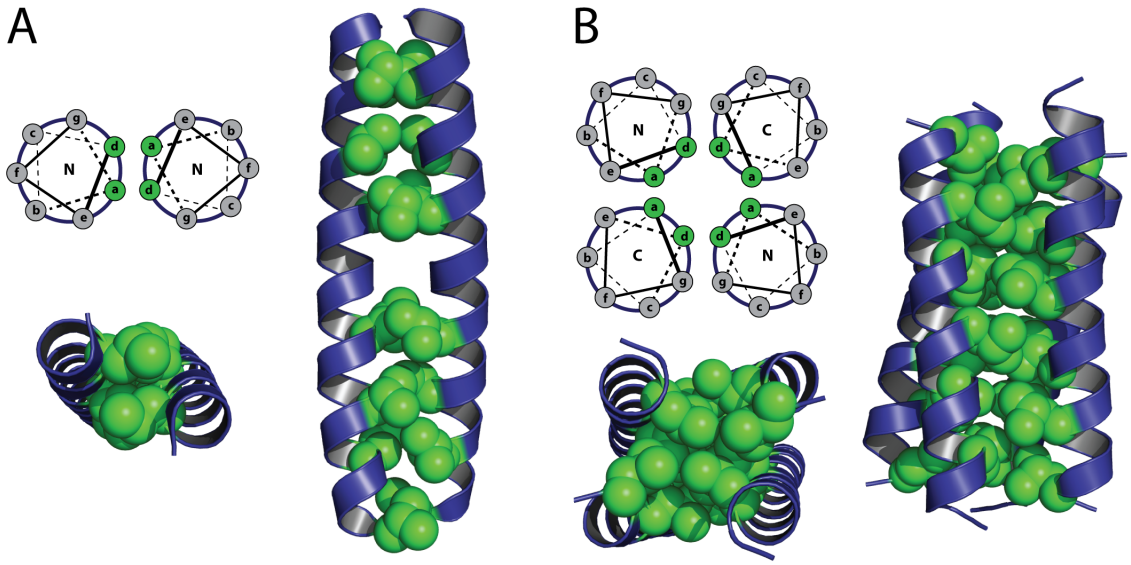

Figure 1.3. Coiled-coil proteins used as model systems to study fluorination. (A) Helix wheel diagram demonstrating the heptad repeat and hydrophobic packing of the parallel coiled-coil GCN4. Three dimensional representation of GCN4 indicating the seven **a** and **d** positions which have been modified with fluorinated residues. (B) Helix wheel diagram demonstrating the heptad repeat and hydrophobic packing of the antiparallel coiled-coil α_4. Three-dimensional representation of α_4 indicating the six **a** and **d** positions which have been modified with fluorinated residues.

1.4.2 – Studies on Anti-Parallel Coiled-Coil Proteins

Studies in our laboratory to understand the effects of fluorination on the physical and biological properties of proteins have utilized a de novo-designed anti-parallel 4-α-helix bundle protein, α_4. The α-helix coiled-coil structural motif, embodied by α_4, was

first described by Crick in 1953. In this landmark study, the packing of dimeric and trimeric coiled-coils was described in which residues of adjacent helices pack in a knobs into holes fashion resulting in a helix offset of ~20° from parallel. In general, α-helices are defined by an arrangement of 3.6 residues per turn with a regular repeat of 7 residues, designated the heptad repeat. The heptad repeat comprises residues at **a** through **g** positions, with the packing of knobs into holes being primarily in **a** and **d** positions of the same helix face. Residues in the **a** and **d** positions are generally hydrophobic residues that in aqueous solution drive the association of α-helices into higher order tertiary and quaternary structures, as exemplified by the tetrameric α_4H (Fig. 1.3B). Residues in the **b**, **c**, **e**, **f** and **g** positions of the heptad repeat are polar with **b**, **c**, **e** and **g** residues stabilizing arrangement of helices through complimentary charge – charge interactions. For parallel coiled-coils, salt bridges form between **e – g** residues (Fig. 1.3A) and for antiparallel coiled-coils between **b** – **e** and **c** – **g** residues (Fig. 1.3B). The **f** position faces away from the coiled-coil center and is generally a polar residue to aid in solubility.

The sequence of α_4 is based on the closely related Coil-LL, which was designed and studied by Betz and DeGrado [31]. α_4H, the parent protein, contains Leu at the three **a** and three **d** positions of the heptad repeat, to form the hydrophobic core of the folded tetramer (Fig. 1.3B) [25,32]. We chose the α_4 protein to study fluorination due to the robustness of the anti-parallel arrangement, where two distinct interfaces (**b** – **e** and **c** – **g**) reinforce the anti-parallel structure through charge-charge complementarity, preventing altered topology due to changes in hydrophobic volume. Greater oligomeric variation occurs in the related parallel coiled-coils, where there is only one unique interface composed of **e** – **g** interactions; slight changes in hydrophobic volume can lead to entirely

different oligomeric states [33,34]. In previous studies, a number of fluorinated versions of α_4, designated α_4F_n have been synthesized that incorporate hFLeu at different positions within the core and examined their physical and biological properties. In all cases these proteins retain well-folded, native-like properties despite the fact that the hFLeu side chain is ~ 30 % larger than Leu.

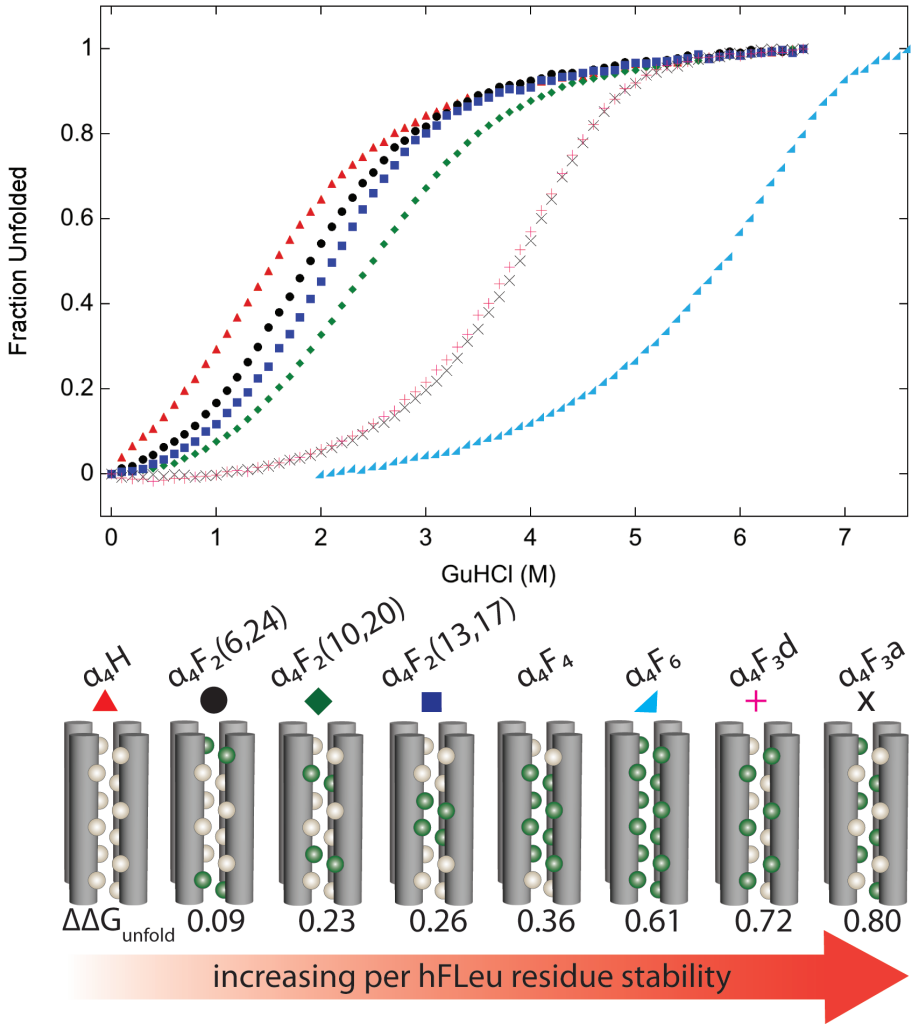

Figure 1.4. Thermodynamic stability of hFLeu substituted α_4 proteins from Chapter 2. (Top) GuHCl induced unfolding of α_4 proteins followed by circular dichroism at 222 nm, protein identities are listed in the center (α_4F_4 not shown). (Bottom) Cartoons illustrating the packing of α_4 proteins with Leu as tan spheres and hFLeu as green spheres. Fluorination greatly increases protein stability, $\Delta\Delta G_{unfold}$ (kcal/mol/hFLeu residue) shown

as increasing from left to right.

The stability of the α_4F proteins progressively increases as the number of hFLeu residues increases, so that α_4F_6, in which all the Leu residues are replaced by hFLeu, is 14.8 kcal/mol more stable than α_4H; a per residue increase of 0.6 kcal/mol/hFLeu. As described in Chapter 2, the position and pattern of the hFLeu substitutions was also found to effect the stability of the protein (Fig. 1.4). For the series of α_4F_2 proteins, each of which contain two hFLeu residues per strand, the stability increases from 0.09 to 0.26 kcal/mol/hFLeu as the hFLeu residues are progressively moved from the more solvent-exposed ends of the bundle to the solvent-excluded center of the bundle [35]. The most stabilizing arrangement of hFLeu and Leu residues appears to be an alternating pattern in which hFLeu is incorporated at **a** positions and Leu at **d** positions, or *vice versa*. Thus α_4F_3a, which contains hFLeu residues in all **a** positions, is more stable than α_4H by 0.8 kcal/mol/hFLeu residue. The importance of considering packing effects, and not just the degree of fluorination, when designing extremely stable fluorinated proteins is further discussed in Chapter 2.

Focusing on the α_4F_6 protein, previous members of the Marsh lab have also examined how fluorination alters its resistance to solvent denaturation and degradation by proteases, properties which may have practical applications. In water-alcohol mixtures containing methanol, ethanol or 2-propanol α_4F_6 retains its helical structure whereas α_4H, which is itself quite a stable protein, becomes increasingly more unfolded as the hydrophobic nature of the solvent increases [36]. Contrary to the behavior predicted by the fluorous effect, fluorinated solvents, such as trifluoroethanol or hexafluoro-2-propanol do not preferentially unfold α_4F_6 but cause it to dissociate into highly helical monomers;

these fluorinated solvents have a similar effect on α_4H, consistent with their well-documented ability to increase the helicity of a large number of peptides.

Fluorination has also been found to protect α_4F_6 against proteolysis. Whereas α_4H was nearly completely degraded in ~ 2 hours by either trypsin or chymotrypsin, far less proteolysis of α_4F_6 was observed under the same conditions. This likely reflects a much slower rate of unfolding by the more thermodynamically stable fluorinated protein rather than the inability of proteases to act on fluorinated substrates.

1.4.3 – Context Effects

Whereas studies on coiled-coil proteins have found that fluorinated leucine analogs invariably confer greater stability, an interesting study by Cheng and coworkers has concluded that, in the context of a single helix, hFLeu is actually destabilizing relative to Leu [37]. Using a monomeric, alanine-based model helix, various fluorinated and hydrocarbon side chains were introduced into a central "guest" position. The helicity of these peptides was then compared relative to alanine at the guest position. Comparing the helicity of ethylglycine with trifluoroethylglycine (tFeG), Leu with hFLeu, and Phe with pFPhe, the fluorocarbon amino acids were found to be significantly less helical than their hydrocarbon counterparts. In the case of hFLeu, the helix propensity is decreased 8-fold compared to Leu, corresponding to an energetic penalty of 1.15 kcal/mol/hFLeu. The decrease in helical content is surprising given that fluorinated amino acids are stabilizing in helical coiled-coils.

Equally surprising is that in the context of a β-sheet, fluorinated residues in solvent-exposed positions appear to stabilize the folded state [38]. Evidence for this comes

from experiments in which hydrocarbon and fluorocarbon residues were introduced at a solvent-exposed position on an internal strand of a β-sheet in the small protein GB1. The fluorinated residues tFeG, hFLeu and pFPhe each increased the protein's stability by ~ 0.3 kcal/mol over their hydrocarbon counterparts. The reason that fluorination seems to exert opposite effects on protein stability in the context of an α-helix versus in a β-sheet is unclear. Moreover, the stabilizing potential of fluorinated amino acids in β-sheets has largely been looked-over, making this an interesting avenue for future research.

Koksch and coworkers have studied how the spatial demands and polarity of fluorinated residues influence the properties of proteins in a model antiparallel coiled-coil protein [39-41]. They examine the effects of ethylglycine and its fluorinated analogues: difluoroethylglycine (dFeG), trifluoroethylglycine (tFeG) and difluoropentylglycine (dFpG) on protein stabilities. This set of small fluorinated amino acids allowed the hydrophobic volume and side chain polarity to be varied. Analogs of the native antiparallel dimer showed decreased stability when any of the fluorinated amino acids were substituted for Leu9 in the hydrophobic core or solvent exposed Lys8. These results demonstrate a decrease in stability due to both decreased hydrophobic volume and changes in polarity of the hydrophobic core.

1.4.4 – Studies on More Complex Protein Structures

Although most studies have focused on α-helices, the effects of fluorination on other protein folds have also been investigated. In one case, the effect of substituting valines by tFVal on the folding kinetics and stability of the globular α-β protein NTL9 (Fig. 1.5) was investigated. The small isosteric change to the core of the globular protein

NTL9 did not disrupt the native fold but significantly changed the stability and folding kinetics. At one position, introduction of tFVal increased the stability of the protein by 1.4 kcal/mol/tFVal residue [42]. Fluorination resulted in a marked decrease in the unfolding rate and a slight increase in the folding rate. The change in folding kinetics was attributed to the increased hydrophobicity of the trifluoromethyl group stabilizing the transition state for folding.

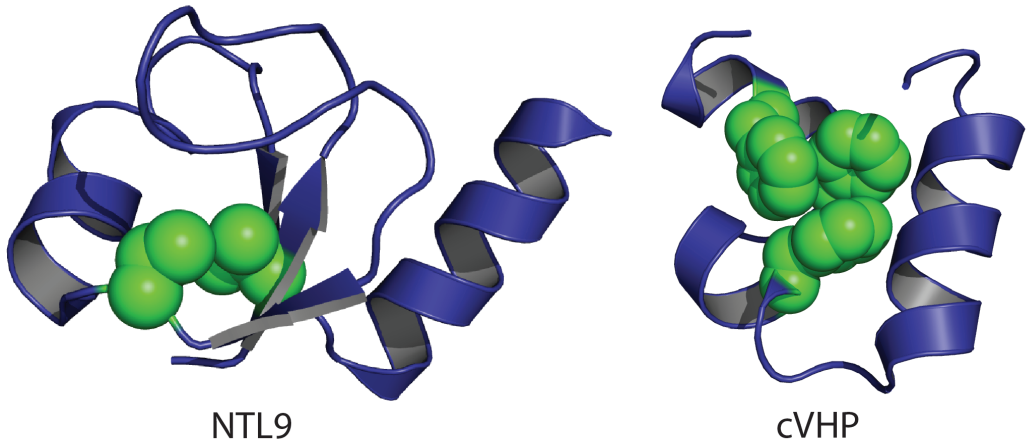

Figure 1.5. Fluorinated proteins with more complex folds. (Left) Model of NTL9 illustrating positions Val3 and Val21 in green, which were substituted for tFVal. (Right) Model of cVHP demonstrating packing interactions of core Phe residues in green.

Other studies have investigated the 35 residue, independently folded ‘‘headpiece’’ subdomain of chicken villin protein (cVHP), which has three phenylalanine residues in the hydrophobic core (Fig. 1.5) [43,44]. When these were substituted by pentafluorophenylalanine (pFPhe) [44], only at one position did the substitution stabilize the protein, whereas at the other two positions pFPhe was actually destabilizing. This could be due to steric effects caused by the larger volume of pFPhe or due to changes in the quadrupole moment of the aromatic ring induced by fluorination. Whereas a phenyl ring

has an electron-rich center and correspondingly electron-poor periphery, the high electronegativity of fluorine results in the center of the aromatic ring being electron-poor and periphery being electron-rich. Quadrupolar interactions between stacked pairs of perfluorinated and perhydrogenated aromatic rings is highly favorable, exemplified by a calculated estimate of facial interaction strength between benzene and hexafluorobenzene being 3.7 kcal/mol [33]. Engineering protein stability using pFPhe is dependent on side chain orientation, with greatest stability achieved when pFPhe is stacked face to face with Phe. Large increases in protein stability can also be achieved with tetrafluorophenylalanine where the single hydrogen substituent on the ring periphery is partially positively charged. This was observed when residues in the aromatic core of cVHP were replaced by tetrafluorophenylalanine thereby introducing favorable edge–face interactions between the tetrafluorophenylalanine aromatic hydrogen and the π-electrons of Phe [43].

Quadupole-quadrupole interactions can make important contributions to protein structure. The energetic contribution of the quadrupole interaction between a Phe-pFPhe pair in a de novo designed, dimeric, 4-helix bundle protein designated α_2D was investigated by Gao and coworkers [45]. The protein was designed to assemble from two peptides; one containing two Phe residues at core positions, the other containing two pFPhe residues. Mixing the peptides resulted in a single, stable species with assembly directed by the introduced quadrupole interactions. By analyzing the folding energies through the use of a double-mutant cycle, the stabilization due to the quadrupole interaction, ΔG_{quad} was estimated to be ~1.0 kcal/mol/Phe-pFPhe pair. Further stabilization studies of α_2D by systematic substitution of the stacked, core phenylalanine

residues with mono-, di-, tri-, tetra- and pentafluorophenylalanine demonstrated that a combination of dipole-dipole interactions and hydrophobic interactions contribute to stability [46]. The substitutions led to the discovery of the phenylalanine/*ortho*-tetrafluorophenylalanine being the most stable aromatic pair with $\Delta\Delta G_{fold}$ of 6.7 kcal/mol. These studies demonstrate how modification of aromatic residues with fluorine can be used to investigate how van der Waals, hydrophobic and electrostatic forces to contribute to protein stability.

1.4.5 - The Potential for Fluorous Effects in Fluorinated Proteins

The unusual tendency of fluorocarbons to self-associate, leading to phase separation of fluorocarbon and hydrocarbon solvents, gave rise to the intriguing possibility that highly fluorinated proteins might possess similar properties. In proteins, the fluorous effect might result in specific protein-protein interactions through fluorous contacts between side chains that would be orthogonal to normal protein-protein interactions. Evidence for self-segregating properties of fluorinated proteins is mixed and may be protein fold dependent. There is much debate in the realm of fluorous proteins as to whether the fluorous effect is a driving force for stability or if the increased hydrophobic volume of fluorinated residues is the main contributor to enhanced stability.

Kumar and coworkers have demonstrated the preferential interaction of a fluorinated parallel coiled-coil dimer in both aqueous and membrane environments [47-50]. These experiments used peptides that contained either Leu or hFLeu at the hydrophobic **a** and **d** positions that comprise the core of the coiled-coil and Cys residues at their N-termini. Disulfide bond formation allowed the two partner peptides in the coiled-coil to

be covalently cross-linked and analyzed. It was found that the peptides preferentially self-segregated into fluorinated (**FF**) and non-fluorinated (**HH**) coiled-coils with less than 3% of peptides forming heterodimers (Fig. 1.6) [49]. However, the interpretation of this experiment is complicated by the fact that the fluorinated peptide formed a tetramer rather than the intended dimer. It may be simply that the bulkier hFLeu sidechain was not compatible with the hydrophobic core of a dimeric coiled-coil.

The self-association of fluorinated peptides designed to form transmembrane helices has also been demonstrated. Again, the peptides were designed to form parallel coiled-coils containing six Leu or hFLeu at **a** and **d** positions, but in this case the peptides were labeled with a fluorophore and FRET used to determine whether the fluorinated and non-fluorinated peptides interacted. Similar to the soluble coiled-coils, the fluorinated and non-fluorinated transmembrane peptides appeared to preferentially self-associate; however, again, whereas the non-fluorinated peptide was dimeric, the fluorinated peptide adopted a tetrameric structure. The results were interpreted as the fluorocarbon side chains forming an interface orthogonal to that of hydrocarbon lipid chains and protein side chains

Our laboratory has investigated the segregation of fluorinated and hydrocarbon versions of the *de novo* designed α_4 proteins, which form anti-parallel 4-helix bundles (Fig. 1.6). This motif has proved highly robust, and α_4 tolerates fluorination without changing its quaternary structure. In this system we found little or no evidence that these peptides undergo self-segregation, contrary to the predictions of the fluorous effect. Two experiments in particular point to the absence of preferential fluorous interactions.

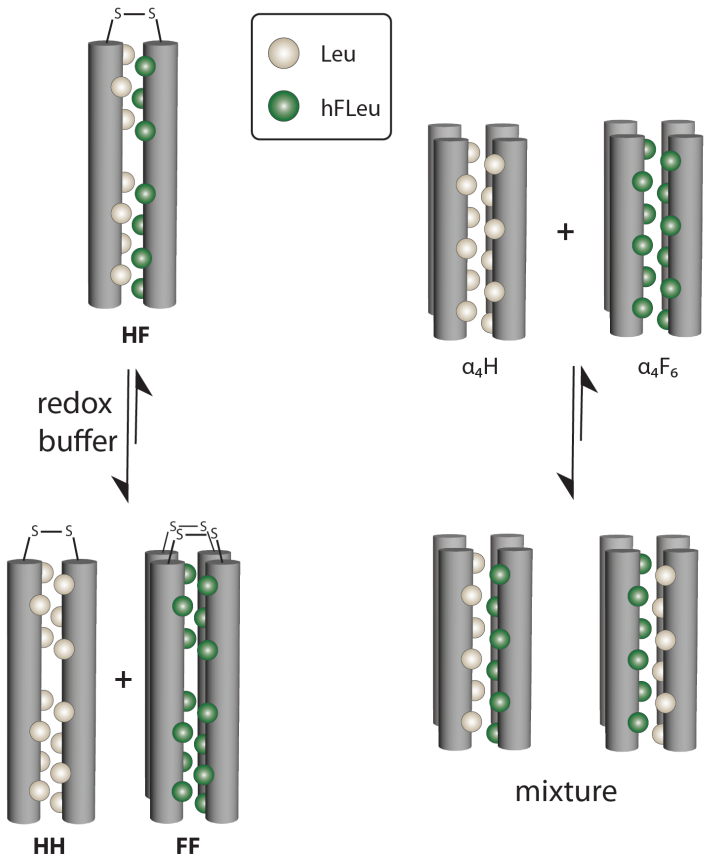

Figure 1.6. Self-assembly of two different protein systems. (Left) Hybrid protein mixtures of **HF** self-segregate into equilibrium populations of the Leu core containing **HH** and the hFLeu core containing **FF** in the presence of a redox buffer. (Right) Upon combining Leu core containing α_4H and hFLeu core containing α_4F_6, mixtures of protein tetramers are observed.

In one experiment ^{19}F NMR was used to examine the interactions between α_4H, which contains Leu at all the **a** and **d** positions, and α_4F_6, which contains hFLeu at all the **a** and **d** positions [36]. α_4F_6 has a complex ^{19}F NMR spectrum that reflects the high sensitivity of the ^{19}F nucleus to slight differences in the chemical environments of the hFLeu $-CF_3$ groups. Titrating α_4F_6 with increasing concentrations of α_4H resulted in progressive changes to the ^{19}F NMR spectrum, with the signals becoming less disperse and shifting downfield as α_4H ratio was increased. Sedimentation equilibrium analytical

ultracentrifugation measurements indicated that the peptide mixtures remained tetrameric. Clearly, no change in the ^{19}F spectrum would be expected unless the α_4F_6 and α_4H peptides were interacting, so these results indicate that the peptides form heterotetramers in which the mononers exchange on the NMR timescale.

Further evidence against the idea that fluorine-fluorine contacts *per se* can be used to engineer orthogonal interactions between proteins comes from studies on proteins with mixed hydrocarbon-fluorocarbon cores, which are discussed in more detail in Chapter 2 [35]. α_4F_3a and α_4F_3d have hFLeu in either all the **a** or all the **d** positions respectively. This gives rise to a hydrophobic core in which fluorocarbon residues are interposed with hydrocarbon residues. These proteins are very stable and, on a per-hFLeu-residue basis, exceed the stability of the "all fluorine" protein α_4F_6. These results suggest that optimizing core packing to reduce steric hindrance and accommodate changes in side chain volumes is more important for stability than potential self-segregating effects of fluorinated residues. The extent of any "fluorous effect" in fluorinated proteins is complicated by the fact that proteins rely on numerous weak interactions to specify their folded structures.

1.5 – Applications: Modulating the Properties of Bioactive Peptides

Fluorination has also been used as a tool to modify the properties of biologically active peptides and investigate their mechanism of action. In particular, some classes of peptides, notably antimicrobial peptides (AMPs) and venom peptides, exert their biological effect through direct disruption of cell membranes, rather than specific peptide-protein or peptide-nucleic acid interactions. This disruptive effect depends on the

overall balance of positively-charged and hydrophobic residues, rather than sequence-specific interactions, making fluorination an ideal method to alter the hydrophobicity of these peptides in a non-disruptive manner (Fig. 1.7). The incorporation of fluorinated residues also allows peptide-membrane interactions to be followed by ^{19}F NMR [51,52].

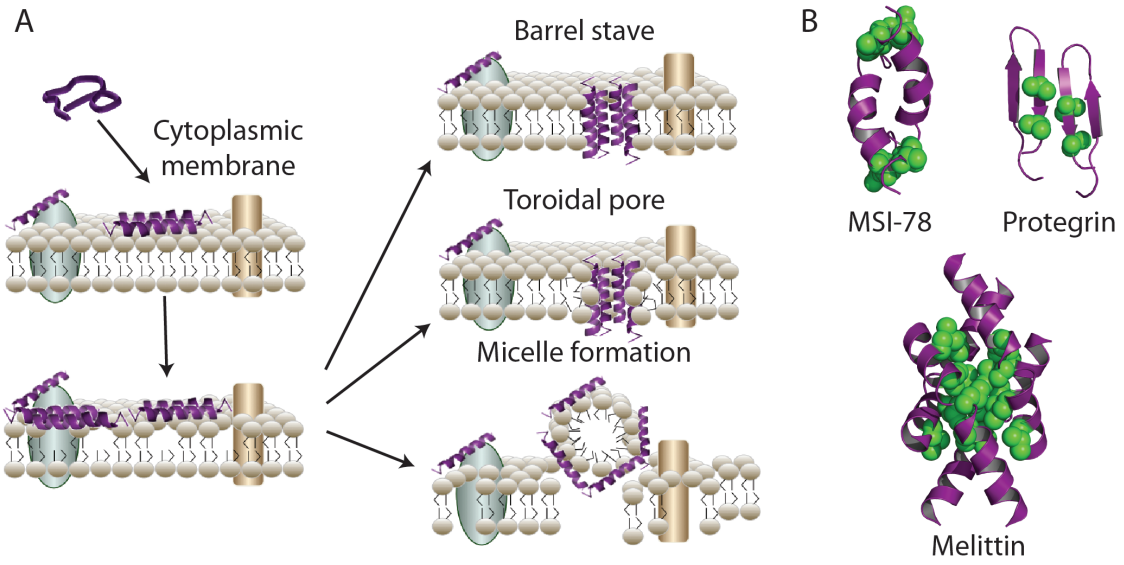

Figure 1.7. Antimicrobial peptides (AMPs) are potent therapeutic agents that act by disrupting bacterial membranes. (A) Membrane disruption by AMPs is initiated by attraction of the positively charged peptide with the negatively charged bacterial membrane lipid headgroups. Loss of membrane integrity may result from three distinct pore forming mechanisms. (B) Models of the AMPs, MSI-78 and protegrin-1, and venom peptide, melittin, with positions of fluorinated amino acid substitution shown in green.

Studies on an analog of the bee venom peptide, melittin, were among the first to show that incorporating fluorinated amino acids could modulate the biological activity of membrane-active peptides. Replacing four Leu residues with tFLeu in melittin resulted in increased partitioning of the fluorinated peptide into liposomes [53]. In our laboratory we have used fluorinated amino acids to modulate the biological activity of the potent synthetic AMP MSI-78 and probe its interactions with membranes.

In one study by Gottler *et al*., two Leu and two Ile residues in MSI-78 were substituted for hFLeu [54]. Overall, the resulting highly fluorinated AMP, dubbed fluorogainin-1, exhibited very similar antimicrobial activity to MSI-78 against a broad range of bacteria. Interestingly, fluorogainin-1 displayed a significantly lower MIC against *Klebsiella pneumoniae* and *Staphylococcus aureus* than MSI-78. Whereas antimicrobial activity was retained, fluorination appeared to alter the mechanism of membrane disruption. Differential scanning calorimetry measurements indicated that the parent peptide, MSI-78, induces positive membrane curvature consistent with a toroidal pore mechanism; in contrast, fluorogainin-1 induced negative membrane curvature indicative of the barrel-stave mechanism for membrane disruption.

In another study, fluorination was used to assess the effects of increasing hydrophobicity in protegrin-1 (PG-1), which is a potent member of the β-hairpin-forming class of antimicrobial peptides. By substituting two valine residues on the hydrophobic face of protegrin-1 for either two Leu or hFLeu residues [55] it was possible to progressively increase hydrophobicity whilst minimally perturbing secondary structure. The Leu containing-peptide was significantly more active than wild-type protegrin against several common pathogenic bacterial strains, whereas the hFLeu-substituted peptide, in contrast, showed significantly diminished activity against several bacterial strains. Isothermal titration calorimetry measurements revealed significant changes in the interaction of the peptides binding to liposomes that mimic the lipid composition of the bacterial membrane. Notably both these substitutions appear to alter the stoichiometry of the lipid-peptide interaction, suggesting that these substitutions may stabilize

oligomerized forms of protegrin that are postulated to be intermediates in the assembly of the β-barrel membrane pore structure.

One significant obstacle to the therapeutic use of AMPs is the inherent susceptibility of peptides towards proteolysis. Strategies to increase proteolytic stability of peptide-based therapeutics include use of D-peptides, β-peptides and arylamide polymers [56-58]. It appears, also, that incorporation of fluorinated amino acids provides a further means of stabilizing bio-active peptides that could increase their therapeutic index. Thus, when bound to liposomes, fluorogainin-1 proved far more resistant to proteolysis than MSI-78. Whereas MSI-78 was degraded by either either trypsin or chymotrypsin in about 30 min, under the same conditions fluorogainin-1 exhibited no degradation after 10 hours. Similar results have been obtained with other membrane-active peptides, such as buforin and magainin [59], suggesting that fluorination may be a general strategy for prolonging the lifetime of peptides *in vivo*.

Kumar and coworkers have used hFLeu to stabilize glucagon-like peptide-1 (GLP-1), which is a promising therapeutic to regulate blood glucose homeostasis to treat type 2 diabetes. Clinical applications of GLP-1 are severely limited due to degradation by the regulatory serine protease, dipeptidyl peptidase IV. Substituting any of the hydrophobic positions 8, 9 and 10 with hFLeu conferred resistance to proteolysis [60]. However, fluorination seemed to reduce the affinity of the peptide for its receptor, possibly due to the increased volume of the hFLeu residues.

1.6 – Goals

The extensive research from our lab and others has shown that the unique

physicochemical properties of fluorocarbons can be introduced into proteins, imparting them with novel and useful characteristics. These previous studies have established the utility of incorporating fluorinated amino acids towards stabilizing proteins. The goals of my research include understanding in detail how fluorination alters the structure and thermodynamic properties of proteins, using the de novo-designed protein, α_4, as a model system. Using a combination of techniques including X-ray crystallography, circular dichroism and NMR, my research has provided insights into why fluorinated proteins are more stable than their natural hydrocarbon counterparts. These structural studies also shed light on whether the self-segregating properties of fluorocarbons can be introduced into proteins through incorporation of hFLeu. Lastly, the biocompatibility and NMR sensitive properties of fluorine have been exploited by using fluorinated amino acids to serve as probes for elucidating peptide–membrane interactions that are vital to the mechanism of AMPs. This research aims to contribute to the basic knowledge of protein design using fluorine, with a long-term objective of being able to design more stable therapeutics and biomaterials.

1.7 – References

(1) Müller, K.; Faeh, C.; Diederich, F. *Science* **2007**, *317*, 1881.
(2) Resnati, G. *Tetrahedron* **1993**, *49*, 9385.
(3) Luo, Z.; Zhang, Q.; Oderaotoshi, Y.; Curran, D. P. *Science* **2001**, *291*, 1766.
(4) Horvath, I. T.; Rabai, J. *Science* **1994**, *266*, 72.
(5) Gakh, Y. G.; Gakh, A. A.; Gronenborn, A. M. *Magn Reson Chem* **2000**, *38*, 551.
(6) Gerig, J. T. *Prog Nucl Mag Res Spec* **1994**, *26*, 293.
(7) Khan, F.; Kuprov, I.; Craggs, T. D.; Hore, P. J.; Jackson, S. E. *J Am Chem Soc* **2006**, *128*, 10729.
(8) Danielson, M. A.; Falke, J. J. *Annu. Rev. Biophys. Biomolec. Struct.* **1996**, *25*, 163.
(9) Ropson, I. J.; Frieden, C. *Proc Nat Acad Sci USA* **1992**, *89*, 7222.
(10) Luchette, P. A.; Prosser, R. S.; Sanders, C. R. *J Am Chem Soc* **2002**, *124*, 1778.

(11) Long, G. J.; Rosen, J. F.; Schanne, F. A. *J Biol Chem* **1994**, *269*, 834.
(12) Duewel, H.; Daub, E.; Robinson, V.; Honek, J. F. *Biochemistry* **1997**, *36*, 3404.
(13) Marsh, E. N. G. *Chem Biol* **2000**, *7*, R153.
(14) Yoder, N. C.; Yüksel, D.; Dafik, L.; Kumar, K. *Curr Opin Chem Biol* **2006**, *10*, 576.
(15) Jäckel, C.; Koksch, B. *Eur J Org Chem* **2005**, *2005*, 4483.
(16) Rennert, O. M.; Anker, H. S. *Biochemistry* **1963**, *2*, 471.
(17) Tang, Y.; Ghirlanda, G.; Petka, W. A.; Nakajima, T.; DeGrado, W. F.; Tirrell, D. A. *Angew Chem Int Ed* **2001**, *40*, 1494.
(18) Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, W. A.; DeGrado, W. F.; Tirrell, D. A. *Biochemistry* **2001**, *40*, 2790.
(19) Wang, P.; Fichera, A.; Kumar, K.; Tirrell, D. A. *Angew Chem Int Ed* **2004**, *43*, 3664.
(20) Wang, P.; Tang, Y.; Tirrell, D. A. *J Am Chem Soc* **2003**, *125*, 6900.
(21) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498.
(22) Wang, L.; Xie, J.; Schultz, P. G. *Annu. Rev. Biophys. Biomolec. Struct.* **2006**, *35*, 225.
(23) Muir, T. W. *Annu. Rev. Biochem.* **2003**, *72*, 249.
(24) Muir, T. W.; Sondhi, D.; Cole, P. A. *Proc Nat Acad Sci USA* **1998**, *95*, 6705.
(25) Lee, K.-H.; Lee, H.-Y.; Slutsky, M. M.; Anderson, J. T.; Marsh, E. N. G. *Biochemistry* **2004**, *43*, 16277.
(26) Pendley, S. S.; Yu, Y. B.; Cheatham, T. E. *Proteins* **2009**, *74*, 612.
(27) Bilgiçer, B.; Fichera, A.; Kumar, K. *J Am Chem Soc* **2001**, *123*, 4393.
(28) Tang, Y.; Tirrell, D. A. *J Am Chem Soc* **2001**, *123*, 11089.
(29) Son, S.; Tanrikulu, I. C.; Tirrell, D. A. *ChemBioChem* **2006**, *7*, 1251.
(30) Montclare, J. K.; Son, S.; Clark, G. A.; Kumar, K.; Tirrell, D. A. *ChemBioChem* **2009**, *10*, 84.
(31) Cui, T.; Bondarenko, V.; Ma, D.; Canlas, C.; Brandon, N. R.; Johansson, J. S.; Xu, Y.; Tang, P. *Biophysical journal* **2008**, *94*, 4464.
(32) Lee, H.-Y.; Lee, K.-H.; Al-Hashimi, H. M.; Marsh, E. N. G. *J Am Chem Soc* **2006**, *128*, 337.
(33) West Jr, A. P.; Mecozzi, S.; Dougherty, D. A. *Journal of physical organic chemistry* **1997**, *10*, 347.
(34) Van Deventer, J. A.; Fisk, J. D.; Tirrell, D. A. *ChemBioChem* **2011**, *12*, 700.
(35) Buer, B. C.; de la Salud-Bea, R.; Al Hashimi, H. M.; Marsh, E. N. G. *Biochemistry* **2009**, *48*, 10810.
(36) Gottler, L. M.; de la Salud-Bea, R.; Marsh, E. N. G. *Biochemistry* **2008**, *47*, 4484.
(37) Chiu, H.-P.; Suzuki, Y.; Gullickson, D.; Ahmad, R.; Kokona, B.; Fairman, R.; Cheng, R. P. *J Am Chem Soc* **2006**, *128*, 15556.
(38) Chiu, H.-P.; Kokona, B.; Fairman, R.; Cheng, R. P. *J Am Chem Soc* **2009**, *131*, 13192.
(39) Jäckel, C.; Salwiczek, M.; Koksch, B. *Angew Chem Int Ed* **2006**, *45*, 4198.
(40) Jäckel, C.; Seufert, W.; Thust, S.; Koksch, B. *ChemBioChem* **2004**, *5*, 717.
(41) Salwiczek, M.; Koksch, B. *ChemBioChem* **2009**, *10*, 2867.
(42) Horng, J.-C.; Raleigh, D. P. *J Am Chem Soc* **2003**, *125*, 9286.
(43) Zheng, H.; Comeforo, K.; Gao, J. *J Am Chem Soc* **2008**, *131*, 18.

(44) Woll, M. G.; Hadley, E. B.; Mecozzi, S.; Gellman, S. H. *J Am Chem Soc* **2006**, *128*, 15932.
(45) Zheng, H.; Gao, J. *Angew Chem Int Ed* **2010**, *49*, 8635.
(46) Pace, C. J.; Zheng, H.; Mylvaganam, R.; Kim, D.; Gao, J. *Angew Chem Int Ed* **2011**, In Press.
(47) Bilgiçer, B.; Kumar, K. *Tetrahedron* **2002**, *58*, 4105.
(48) Bilgiçer, B.; Kumar, K. *Proc Nat Acad Sci USA* **2004**, *101*, 15324.
(49) Bilgiçer, B.; Xing, X.; Kumar, K. *J Am Chem Soc* **2001**, *123*, 11815.
(50) Naarmann, N.; Bilgiçer, B.; Meng, H.; Kumar, K.; Steinem, C. *Angew Chem Int Ed* **2006**, *45*, 2588.
(51) Buer, B. C.; Chugh, J.; Al-Hashimi, H. M.; Marsh, E. N. G. *Biochemistry* **2010**, *49*, 5760.
(52) Suzuki, Y.; Buer, B. C.; Al-Hashimi, H. M.; Marsh, E. N. G. *Biochemistry* **2011**, *50*, 5979.
(53) Niemz, A.; Tirrell, D. A. *J Am Chem Soc* **2001**, *123*, 7407.
(54) Gottler, L. M.; Lee, H.-Y.; Shelburne, C. E.; Ramamoorthy, A.; Marsh, E. N. G. *ChemBioChem* **2008**, *9*, 370.
(55) Gottler, L. M.; de la Salud Bea, R.; Shelburne, C. E.; Ramamoorthy, A.; Marsh, E. N. G. *Biochemistry* **2008**, *47*, 9243.
(56) Bessalle, R.; Kapitkovsky, A.; Gorea, A.; Shalit, I.; Fridkin, M. *FEBS Lett* **1990**, *274*, 151.
(57) Choi, S.; Isaacs, A.; Clements, D.; Liu, D.; Kim, H.; Scott, R. W.; Winkler, J. D.; DeGrado, W. F. *Proc Nat Acad Sci USA* **2009**, *106*, 6968.
(58) Liu, D.; DeGrado, W. F. *J Am Chem Soc* **2001**, *123*, 7553.
(59) Meng, H.; Kumar, K. *J Am Chem Soc* **2007**, *129*, 15615.
(60) Meng, H.; Krishnaji, S. T.; Beinborn, M.; Kumar, K. *J Med Chem* **2008**, *51*, 7303.

Chapter 2

Optimizing Hydrophobic Packing of Fluorinated Proteins to Enhance Stability

2.1 – Introduction

The work described in this chapter has been published as: Engineering Protein Stability and Specificity Using Fluorous Amino Acids: The Importance of Packing Effects. Buer BC, de la Salud-Bea R, Al Hashimi HM, & Marsh ENG (2009) *Biochemistry* 48(45):10810-10817. Coauthors were very helpful in conducting this research and analyzing the results, Dr. Roberto de la Salud Bea assisted me with the synthesis of hFLeu and also synthesized and purified the α_4F_2 series of peptides. Helpful suggestions and guidance on NMR experiments came from Prof. Hashim Al-Hashimi.

The extreme hydrophobicity and low polarizability of perfluorocarbon molecules render them poorly soluble in both polar solvents such as water and apolar solvents such as hexane, a phenomenon known as the fluorous effect [1-4]. This property has been exploited to great effect in the development of methods to purify organic compounds by tagging them with perfluorocarbon 'tails' that allow molecules to be selectively extracted from reaction mixtures into perfluorocarbon solvents [1,3].

Much interest has focused on whether unusual properties of simple perfluorocarbons can be engineered into proteins by incorporating extensively fluorinated analogs of hydrophobic amino acids into their structures [4-7]. In particular, fluorination is

predicted to **a**: confer a general increase in the thermodynamic stability of folded proteins, because fluorocarbon side-chains are more hydrophobic than their hydrocarbon counterparts; **b**: confer self-segregating properties on proteins in which the hydrophobic core is packed with fluorous side-chains, by analogy with the self-segregating properties of hydrocarbon and fluorocarbon solvents [2,5].

The first prediction has proved to be true in many cases, based on studies of a number of peptides designed to adopt α-helical coiled-coil structures that incorporate fluorinated analogs of leucine and valine at **a** and **d** positions of the canonical coiled-coil heptad repeat [8-16]. However, it should be noted that there are exceptions to this prediction, as discussed in Chapter 1 sections 4.3 and 4.4. One study found that the intrinsic helix-forming propensities of several fluorinated amino acids were actually lower than that of their non-fluorinated counterparts [17]. Experiments on the independently-folded headpiece sub-domain of chicken villin protein found that substituting pentafluorophenylalanine for phenylalanine was either stabilizing or destabilizing, depending upon the position [18]. Other experiments have examined the spatial demands of various fluorinated amino acids [19]. All of these studies serve to emphasize the importance of steric contributions, which are changed to a certain degree by fluorination, to the stability of proteins.

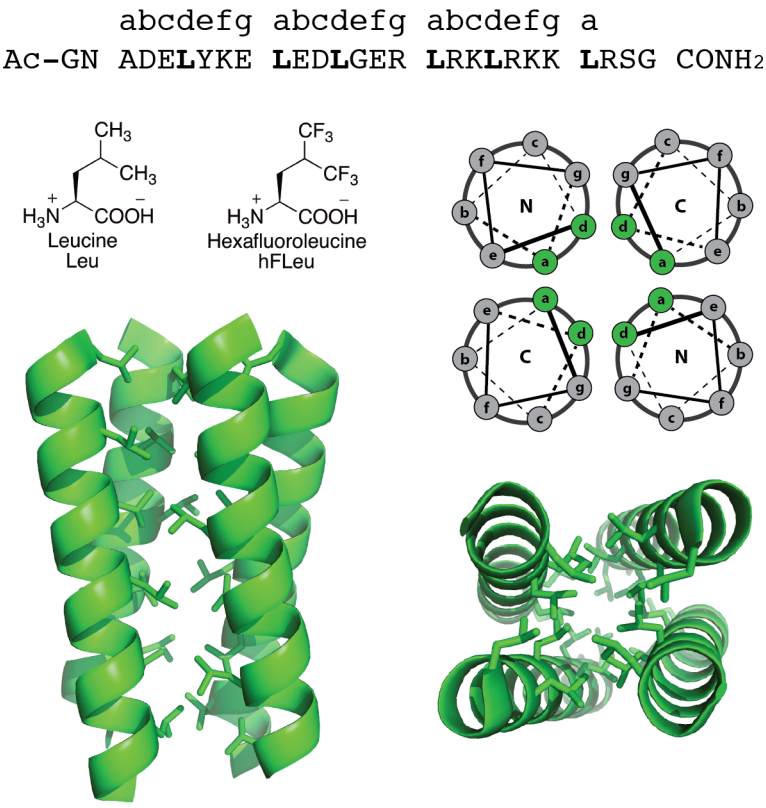

Figure 2.1. *Top*: Sequence of α_4; Leu at **a** and **d** positions can be substituted for hFLeu. Helical wheel diagram illustrating hydrophobic packing of an antiparallel 4-helix bundle with **a** and **d** residues in the core. *Bottom*: Side and top views of α_4 cartoon illustrating the three heptad repeats which form six discreet layers of hydrophobic residues.

Regarding the second prediction, there is some evidence that fluorinated side chains, incorporated at the hydrophobic interface between helices, can mediate the specific self-association of α-helical peptides in solution [8] and in the context of membrane-spanning α-helical peptides [20,21]. However a recent study in our laboratory that used ^{19}F NMR to investigate the interaction between two 4-α-helix bundle proteins, one packed with Leu and the other with hFLeu at the **a** and **d** positions, concluded that the fluorinated and non-fluorinated peptides did not form exclusive interactions [22].

Our lab described previously the design of a model anti-parallel 4-α-helix bundle protein designated α_4, in which all six layers of the hydrophobic core can be packed with combinations of Leu and hFleu at the buried **a** and **d** positions of the canonical heptad repeat as demonstrated in Figure 2.1 [15,16]. In the fluorinated series of proteins we have investigated the effects of packing 2, 4 or 6 layers of the hydrophobic core with hFLeu. All the fluorine-substituted proteins were well-folded and retained the intended 4-helix bundle structure. Increasing the number of hFLeu residues increased the free energy of unfolding for the peptides in an almost linear fashion, with a per-residue stabilization of $\Delta\Delta G \sim -0.3$ kcal/mol/residue, suggesting that fluorination could be used to fine-tune the stability of proteins.

In this chapter I describe the experiments to investigate the effect on protein stability of introducing hFLeu residues at different positions within the hydrophobic core, whilst maintaining the same total number of residues within the 4-helix bundle. The interaction of various combinations of peptides with each other has also been examined. The main conclusion is that the fluorous effect, as originally envisioned, does not appear to be primarily responsible for the self-segregating properties observed with some of the peptides that have been investigated. Rather, a favorable alternating packing arrangement between Leu residues at **a** positions and hFLeu residues at **d** positions (or *vice versa*) within the hydrophobic core determines the specificity of helix:helix interactions.

2.2 – Experimental Procedures

2.2.1 - Materials and Synthesis

Boc-protected and Fmoc-protected amino acids, MBHA resin, MBHA Rink Amide resin and 2-(6-chloro-1-H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) were purchased from Novabiochem. Peptide synthesis grade *N*,*N*-dimethylformamide (DMF), ACS grade *N*,*N*-diisopropylethylamine (DIEA) and molecular biology grade guanidinium hydrochloride (GuHCl) were purchased from Fisher. ACS grade trifluoroacetic acid (TFA) was purchased from Acros Organics. Piperidine was purchased from Sigma-Aldrich.

L-5,5,5,5',5',5'-hexafluoroleucine was synthesized according to the procedure of Anderson *et al*. [23] and converted to the Boc- protected derivative using 1.5 equivalents of di-*t*-butyl dicarbonate and 3.0 equivalents of sodium bicarbonate in 50:50 water/THF. After extraction and purification via silica gel chromatography the resultant light pink-colored waxy solid was used for peptide synthesis.

The sequences of the peptides used in this study are shown in Figure 2.1. These were synthesized by manual Fmoc procedures (α_4H) or manual Boc procedures (hFLeu containing peptides) as described previously with some minor modifications [15,16]. The hFLeu containing peptides were synthesized with Boc-protected amino acids on MBHA resin using the in situ neutralization protocol described by Schnolzer *et al*. These were synthesized on a 0.125 or 0.25 mmol scale using DMF as the main solvent with HCTU as the coupling reagent and solvent washes with dichloromethane (DCM) before and after TFA deprotection (2 x 1 minute) of residues Asn2 and Gln14 to prevent aspartamide failure sequences. Peptides were side chain deprotected and cleaved from the resin by

CS Bio Company using hydrofluoric acid (HF) with 6% vol/vol anisole and 2% vol/vol dimethylsulfide (DMS).

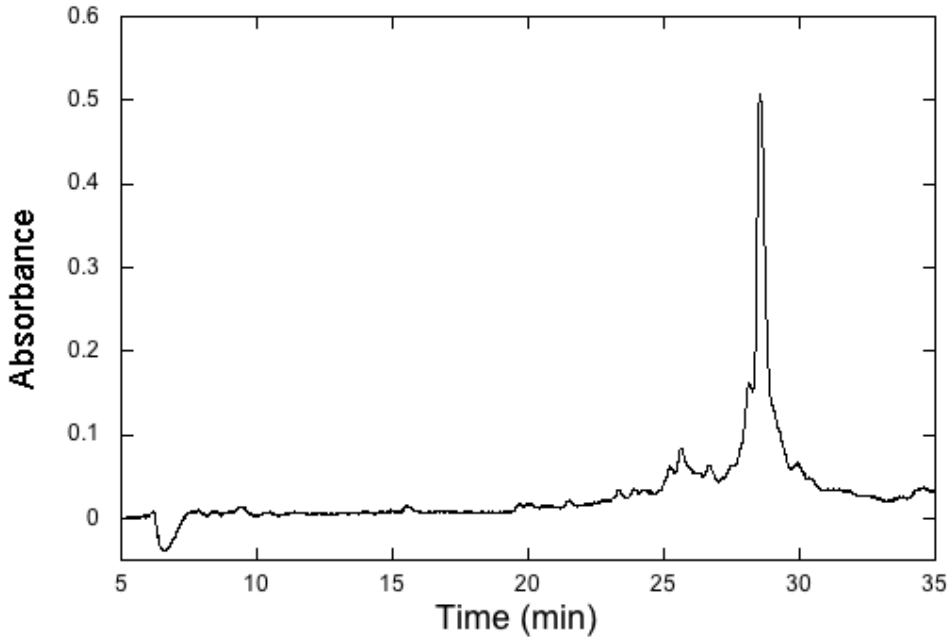

Figure 2.2. UV trace at 222 nm of α_4F_3a HPLC purification from crude. The major peak at 28.5 minutes corresponds to α_4F_3a. Solvent gradient is 0% to 100% B in 50 minutes.

α_4H was synthesized in a 0.25 mmol scale with Fmoc-protected amino acids on MBHA Rink Amide resin using DMF as the main solvent with HCTU as the coupling reagent. Fmoc deprotection at each step was accomplished with 1 x 8 minutes and 1 x 15 minutes 20% vol/vol piperidine in DMF. Resin cleavage and side chain deprotection was performed at room temperature for 2 hours using 15 mL of a mixture of 95% TFA, 2.5% water and 2.5% triisopropylsilane. TFA was evaporated and the crude peptide was precipitated with cold ether.

Similar purification results were achieved when the peptide was dissolved in water then filtered through a 0.22 μm syringe before HPLC injection as to when the peptide was first lyophilized, then filtered and purified. All peptides were purified on a

Waters preparatory HPLC using a linear gradient of 95% water, 4.9% acetonitrile and 0.1% TFA for solvent A and 9.9% water, 90% acetonitrile and 0.1% TFA for solvent B, with a flow rate of 10 mL/min on either a Waters or Vydac C18 preparatory column. A representative HPLC trace for purification of crude peptide is shown in Figure 2.2. Peptide identity was confirmed using MALDI-MS with a matrix of α-cyano-4-hydroxycinnamic acid.

2.2.2 - Circular Dichroism

CD spectra of peptides were recorded with an Aviv 62DS spectropolarimeter at 25 °C. Wavelength scan measurements were taken between 250 and 190 nm using a 0.1 cm cuvette. Mean residue ellipticities, $[\theta]$, were calculated using Equation 1,

$$[\theta] = \theta_{obsd} / 10lcn \qquad (1)$$

where θ_{obsd} is the ellipticity measured in millidegrees, c is the molar concentration, l is the cell path length in cm and n is the number of residues in the protein. To examine the unfolding of the peptide by GuHCl, stock solutions were prepared containing 40 μM peptide (concentration of monomer), as determined by tyrosine absorbance at 280 nm, in 10 mM potassium phosphate buffer, pH 7.0, both with and without 8.0 M GuHCl. 2.0 mL of sample without denaturant was placed in a 1 cm cuvette, and 9.0 mL of 8.0 M GuHCl containing sample was prepared to titrate into the native protein. An auto-titrator was used to mix the two solutions to incrementally increase the concentration of GuHCl while keeping the volume constant at 2.0 mL. After equilibration was achieved in the sample cuvette via stirring for one minute, the ellipticity at 222 nm was measured. 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 [15]. Igor Pro software (Wavemetrics, Inc.) was used to fit the denaturation curves.

2.2.3 - Analytical Ultracentrifugation

Sedimentation equilibrium experiments were performed using a Beckman XLA analytical ultracentrifugation equipped with scanning u.v.-visible optics [24]. Initial peptide concentration was 200 μM in 10 mM potassium phosphate buffer, pH 7.0. The temperature was 293 K. The samples were centrifuged at 35,000, 37,500, 40,000, 42,500 and 45,000 rpm and were judged to have obtained equilibrium when successive radial scans were indistinguishable. The data were fitted 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 [25]: the partial specific volumes of the peptides were calculated as: α_4H = 0.74 cm^3g^{-1}; α_4F_6 = 0.66 cm^3g^{-1}; α_4F_2 = 0.71 cm^3g^{-1}; α_4F_3 = 0.70 cm^3g^{-1} (the positional variants of α_4F_2 and α_4F_3 were assumed to each have the same partial specific volumes). The calculated tetramer molecular weight for all α_4F_2 peptides is 14060 Da and 14490 Da for both α_4F_3 peptides. The experimentally determined molecular weights were α_4F_2(6,24) = 14300 ± 1000, α_4F_2(10,20) = 13400 ± 1000, α_4F_2(13,17) = 13400 ± 1000, α_4F_3a = 14500 ± 1000, α_4F_3d = 16500 ± 1000. The α_4H and α_4F_6 peptides have previously been shown to be tetramers (16).

2.2.4 - ^{19}F NMR

^{19}F NMR spectra were recorded using a Varian 400 MHz NMR spectrometer equipped with a ^{19}F probe. Peptide samples (0.5 – 4.5 mM) were prepared in 10% D_2O in a final volume of 0.5 mL and buffered to pH 7.0 with potassium phosphate buffer. Spectra were recorded at 25 °C and were referenced to trifluoroacetate ion = 0 ppm.

2.3 – Results

The parent peptide for these experiments was α_4H [15], a 27 residue peptide containing 3 canonical heptad repeats that is designed to fold into a tetrameric anti-parallel 4-helix bundle structure as shown in Figure 2.1. Previous investigations in the Marsh laboratory have studied the effects of incorporating increasing numbers of hFLeu residues into α_4H by synthesizing a series of peptides designed to pack progressively more layers of the hydrophobic core with hFLeu [15,16]. These peptides were α_4F_2, which incorporates hFLeu at positions 13 and 17 to pack the central two layers of the hydrophobic core with hFLeu; α_4F_4, which incorporates hFLeu at positions 10, 13, 17 and 20 to pack the central 4 layers of the core; and α_4F_6, which incorporates hFLeu at positions 6, 10, 13, 17, 20 and 24 so that the entire hydrophobic core comprises fluorinated residues.

For this study, 4 additional peptides were synthesized to compare the effects of introducing hFLeu at different positions within the hydrophobic core (Fig. 2.3). To be consistent with previous nomenclature these peptides were designated α_4F_n(position) where n refers to the number of hFLeu residues and their position within the sequence is indicated in parenthesis. α_4F_2(6,24) contains hFLeu at the N-terminal **d** position-6 and the C-terminal **a** position-24 so that first and last layers of the hydrophobic core are

packed by hFLeu; $\alpha_4F_2(10,20)$ contains hFLeu at the second **a** position-10 and the C-terminal **d** position-20 so that the second and fifth layers of the hydrophobic core are packed with hFLeu. These two peptides were designed to complement $\alpha_4F_2(13,17)$, which has been previously characterized [15], and in which the third and fourth layers of the hydrophobic core are packed with hFLeu. Two peptides were also synthesized to investigate the effects of introducing hFLeu at only the **a** positions or at only the **d** positions: α_4F_3a contains hFLeu at the **a** positions 10, 17 and 24, whereas α_4F_3d contains hFLeu at the **d** positions 6, 13 and 20. Figure 2.3 illustrates the arrangement of Leu and hFLeu residues in these various peptides and their packing within the hydrophobic core of the 4-helix bundle.

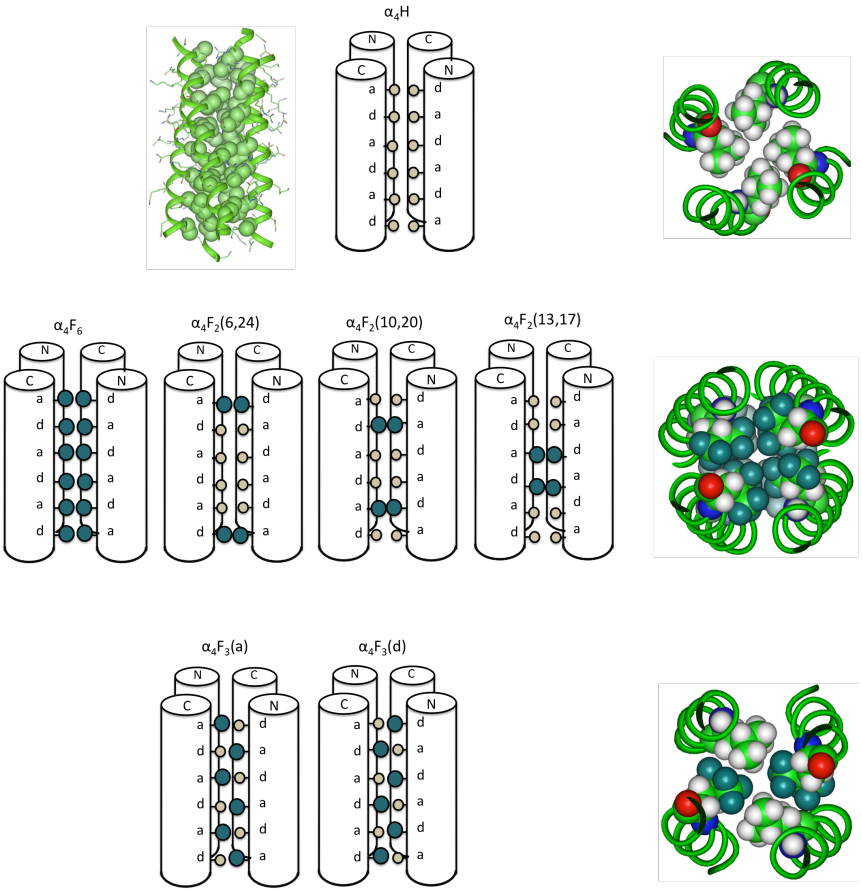

Figure 2.3. *Top*: Models of α_4H illustrating the packing of Leu in the hydrophobic core of the antiparallel four-helix bundle. *Middle*: Cartoons illustrating the pattern of fluorinated residues in α_4F_6, $\alpha_4F_2(6,24)$, $\alpha_4F_2(10,20)$, and $\alpha_4F_2(13,17)$ and model illustrating the packing arrangement of 4 hFLeu side-chains within one layer of the hydrophobic core. *Bottom*: Cartoons illustrating the pattern of fluorinated residues in α_4F_3a and α_4F_3d and model illustrating the packing of alternating Leu and hFLeu residues within one layer of the hydrophobic core. (hFLeu represented by large green spheres, Leu represented by small grey spheres)

2.3.1 - Initial Characterization

Unless noted otherwise, all the experiments were performed in 10 mM potassium phosphate buffer, pH 7.0. The peptides all exhibited CD spectra with strong minima at 208 and 222 nm; the mean residue ellipticities of the peptides were in the range 15100 – 20400 deg cm^2 $dmol^{-1}$ $residue^{-1}$, which are characteristic of extensively helical proteins.

The oligomerization state of the peptides was examined by equilibrium analytical ultracentrifugation. The molecular weights determined from the equilibrium sedimentation traces were in each case consistent with the peptides adopting a tetrameric structure, as designed.

2.3.2 - Positional Effects on Peptide Stability

The effect on the free energy of folding, ΔG^{o}_{fold}, of incorporating hFLeu at different positions within the hydrophobic core of α_4 was investigated by titrating the protein with guanidinium hydrochloride (GuHCl) and using CD to monitor ellipticity at 222 nm to determine the extent of unfolding (Fig. 2.4). ΔG^{o}_{fold} was determined by fitting the denaturation curves assuming a two-state transition between unfolded monomer and folded tetramer, as described previously [15]. Previous measurements relied on manual mixing to titrate the peptides with GuHCl, however in these experiments an auto-titrator was used, allowing significantly more data to be collected per unfolding experiment with less experimental error.

The denaturation curves and fits for α_4H, $\alpha_4F_2(6,24)$, $\alpha_4F_2(10,20)$, $\alpha_4F_2(13,17)$, α_4F_3a, α_4F_3d and α_4F_6 are shown in Figure 2.4 and the values for ΔG^{o}_{fold} and *m* reported in Table 2.1.

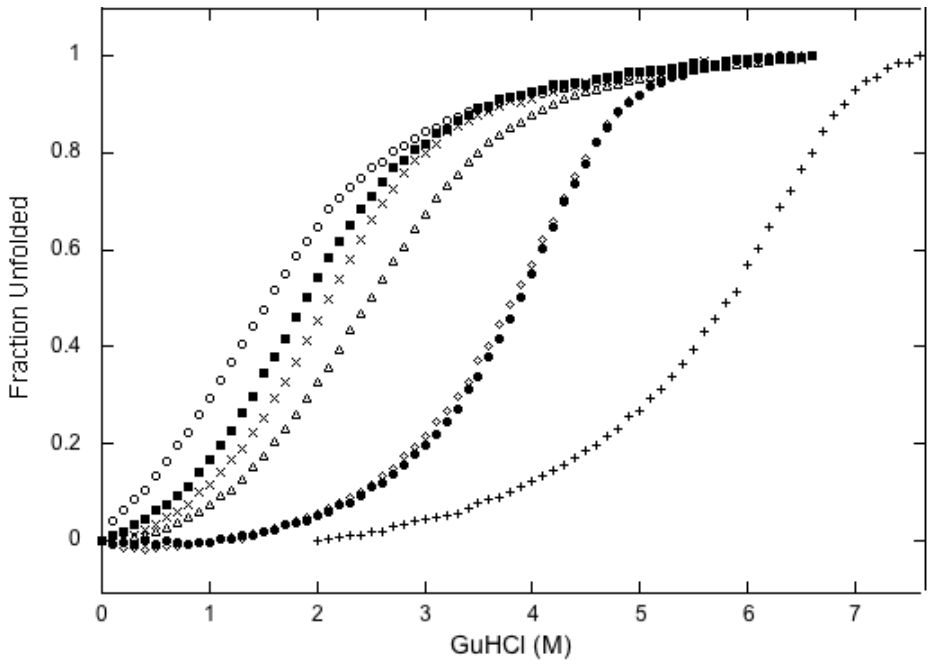

Figure 2.4. Unfolding of peptides in GuHCl. Plots of fraction unfolded versus GuHCl concentration for α_4H (○), $\alpha_4F_2(6,24)$ (■), $\alpha_4F_2(10,20)$ (△), $\alpha_4F_2(13,17)$ (×), α_4F_3a (●), α_4F_3d (◇) and α_4F_6 (+). Unfolding was followed by measuring changes in ellipticity at 222 nm. The peptide concentration was 40 μM in 10 mM potassium phosphate buffer, pH 7.0, 25 °C.

All the fluorinated peptides were more stable than the parent α_4H peptide, as would be expected simply from the more hydrophobic nature of the hFLeu side-chain. Of note is ΔG^o_{fold} for α_4F_6 being significantly larger than the previously published estimate for this peptide [16]; this is due to previous measurements being made at higher peptide concentrations where the upper baseline for the unfolding transition could not be established. Typically this leads to an under-estimate of ΔG^o_{fold}.

Peptide	$\Delta G°_{fold}$ (kcal/mol)	m (kcal/mol/M_{GuHCl})	$\Delta\Delta G°_{fold}$ [a] (kcal/mol)	$\Delta\Delta G°_{fold}$ (kcal/mol/hFLeu residue)
α_4H	-18.0 ± 0.2	-1.04 ± 0.04	NA	NA
$\alpha_4F_2(6,24)$	-18.7 ± 0.1	-1.18 ± 0.03	-0.7 ± 0.2	-0.09 ± 0.03
$\alpha_4F_2(10,20)$	-19.8 ± 0.1	-1.22 ± 0.02	-1.8 ± 0.2	-0.23 ± 0.03
$\alpha_4F_2(13,17)$	-20.1 ± 0.1	-1.53 ± 0.04	-2.1 ± 0.2	-0.26 ± 0.03
α_4F_3a	-27.6 ± 0.1	-2.47 ± 0.03	-9.6 ± 0.2	-0.80 ± 0.02
α_4F_3d	-26.6 ± 0.1	-2.24 ± 0.03	-8.6 ± 0.2	-0.72 ± 0.02
α_4F_6	-32.6 ± 0.3	-2.38 ± 0.04	-14.6 ± 0.4	-0.61 ± 0.02
Peptide Mixtures				
$\alpha_4F_2(6,24):(10,20)$	-20.0 ± 0.1	-1.37 ± 0.03	-2.0 ± 0.02	-0.25 ± 0.03
$\alpha_4F_2(6,24):(13,17)$	-19.5 ± 0.1	-1.47 ± 0.04	-1.5 ± 0.02	-0.19 ± 0.03
$\alpha_4F_2(10,20):(13,17)$	-21.0 ± 0.1	-1.61 ± 0.04	-3.0 ± 0.02	-0.38 ± 0.03

[a] $\Delta\Delta G^{o}_{fold}$ values calculated relative to α_4H

Table 2.1. Summary of the thermodynamic parameters determined from GuHCl-induced unfolding of peptides.

For the α_4F_2 series of peptides it is evident that the position within the hydrophobic core at which hFLeu is introduced does influence the stability of the protein. Thus, the $\alpha_4F_2(6,24)$ variant is only slightly more stable than the parent peptide containing no fluorinated residues. This observation is consistent with the residues at positions 6 and 24 occupying the least well buried positions in the hydrophobic core at the termini of the helical bundle where, even when folded, they are still significantly exposed to solvent. The $\alpha_4F_2(10,20)$ and $\alpha_4F_2(13,17)$ peptides both exhibit similar increases in stability, as would be expected because these positions are fully buried in the hydrophobic core. However, one might have expected the $\alpha_4F_2(13,17)$ variant to have been considerably *more* stable than $\alpha_4F_2(10,20)$, had favorable "fluorous interactions" between fluorocarbon side-chains been operating, because this arrangement allows the

two fluorous layers of the core to pack adjacent to each other and maximize fluorine-fluorine contacts, whereas for $\alpha_4F_2(10,20)$ the fluorinated layers are separated by hydrocarbon layers.

2.3.3 - ^{19}F NMR Studies

To obtain further insights into side-chain interactions within the α_4F_2 peptides the ^{19}F NMR spectra were recorded. The fluorine nucleus is very sensitive and exhibits large changes in chemical shift in response to changes in environment making it an excellent probe for examining interactions between fluorinated residues and adjacent side chains, and dynamic behavior of the peptides [26,27].

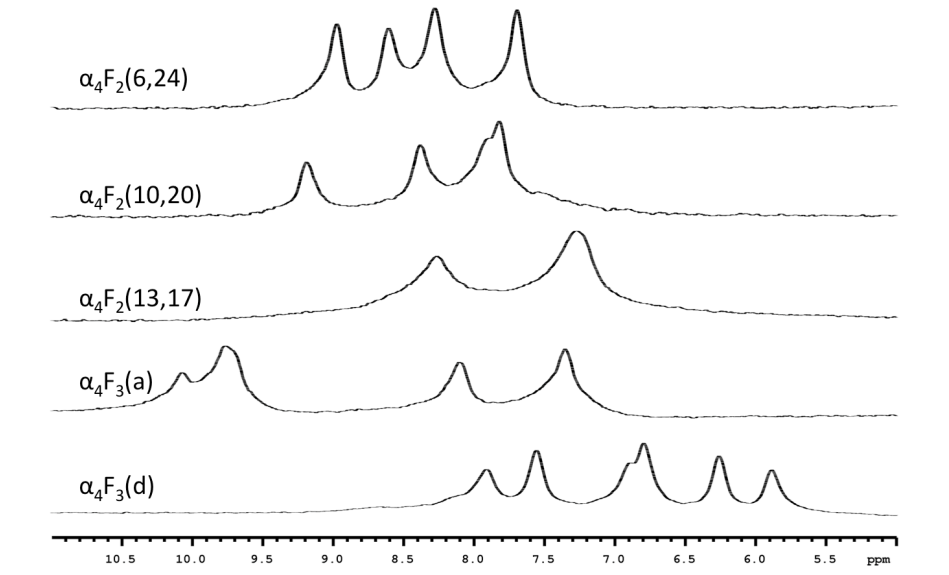

Figure 2.5. ^{19}F NMR spectra of (from top to bottom) $\alpha_4F_2(6,24)$, $\alpha_4F_2(10,20)$, $\alpha_4F_2(13,17)$, α_4F_3a, α_4F_3d recorded at 25 °C, pH 7.0 in 10% D_2O. All spectra are referenced to TFA.

The spectra of the α_4F_2 series of peptides are shown in Figure 2.5; all the spectra were recorded under identical conditions and chemical shifts referenced to TFA. The

spectrum of $\alpha_4F_2(6,24)$ exhibited the sharpest peaks, with all 4 resonances for the diastereotopic trifluoromethyl groups of the two hFLeu residues clearly distinguished. This is consistent with the hFLeu residues at the ends of the bundle occupying less buried positions and being more mobile. The spectrum of $\alpha_4F_2(10,20)$ exhibits distinctly broader peaks, which is consistent with these residues being packed more tightly within the hydrophobic core and being less mobile.

$\alpha_4F_2(13,17)$, however, exhibits much broader peaks so that the resonances due to the hFLeu residues at the 13 and 17 positions are no longer distinguishable and are shifted up-field of the other two peptides. In part, this may be attributable to further reduced mobility of the hFLeu residues because two adjacent layers of the hydrophobic core are packed with the more bulky hFLeu side-chain. However, it is suspected that part of the broadening may be attributable to the phenomenon of chemical exchange broadening. This would occur if the hFLeu side-chains inter-converted, on the time scale of μs to ms, between rotational conformations so that the trifluoromethyl groups sample different chemical environments with different chemical shifts.

2.3.4 - Effects of Fluorination at **a** and **d** Positions

Surprisingly, the introduction of hFLeu only at **a** positions, α_4F_3a, or only at **d** positions, α_4F_3d, stabilized the proteins by a much greater extent than expected based on previous studies of α_4F_2, α_4F_4 and α_4F_6. Furthermore, as evident from Figure 2.4, these proteins exhibited a more cooperative unfolding transition (*m* values significantly increased; Table 2.1) indicative of a better-packed hydrophobic core. Indeed the stability of α_4F_3a and α_4F_3d approaches that for α_4F_6, in which the entire hydrophobic core is

fluorinated. The per-residue stabilization afforded by this arrangement of hFLeu side chains is ~3 times that observed in the α_4F_2 series of peptides and ~1.5 times that observed in α_4F_6.

The ^{19}F NMR spectra of α_4F_3a and α_4F_3d are shown in Figure 2.5 and are quite distinct; this emphasizes the sensitivity of the ^{19}F nucleus to what are quite subtle differences between the **a** and **d** positions in the hydrophobic core. The spectrum of α_4F_3d is somewhat sharper than that of α_4F_3a and all six trifluoromethyl groups can be distinguished, whereas for α_4F_3a the signals for two of the residues are closely overlapping. Neither spectrum exhibits the pronounced peak broadening seen in the spectrum of $\alpha_4F_2(13,17)$, implying that the hFLeu residues in these proteins are either less restricted in their motion, or populate fewer restricted conformations.

It is evident that steric or packing effects, (as opposed to simple differences in hydrophobicity) are playing an important role in the stability of α_4F_3a and α_4F_3d. As illustrated in Figure 2.3, placing hFLeu at **a** positions and Leu at **d** positions (or *vice versa*) within an anti-parallel 4-helix bundle results in an arrangement in which hFLeu and Leu residues alternate both within layers and between layers. Notably, this arrangement provides for the *maximum* separation between hFLeu side-chains thus the dramatic increase in stability observed for this arrangement is *opposite* from what would be expected if fluorous interactions between hFLeu residues were important. Suggesting that the alternating arrangement of Leu and hFLeu is energetically favorable because it relieves steric crowding, allowing the larger hFLeu residues to be better accommodated within the core of the 4-helix bundle.

2.3.5 - Stability of Peptide Mixtures

The stability of α_4F_3a and α_4F_3d provided by the alternate packing of 2 hFLeu and 2 Leu within the core suggested that the formation of hetero-tetrameric bundles might be favorable, if the hetero-tetramer could replicate this packing arrangement. This can potentially be achieved through the formation of hetero-tetramers of any two of the α_4F_2 peptides, illustrated in Figure 2.6, which allows for alternate Leu-hFLeu packing of four of the six layers of the core. Each combination of the α_4F_2 peptides were mixed together in 1:1 ratios and the ΔG^{o}_{fold} of the mixtures was determined. It is expected that if the peptides formed favorable hetero-tetramers then this should be reflected in the mixtures either unfolding more cooperatively and/or unfolding at higher concentrations of GuHCl than either of the pure peptides. Of course, a statistical mixture of hetero-tetrameric bundles could arise even if there is no additional stability associated with them but this would not change the apparent free energy of folding.

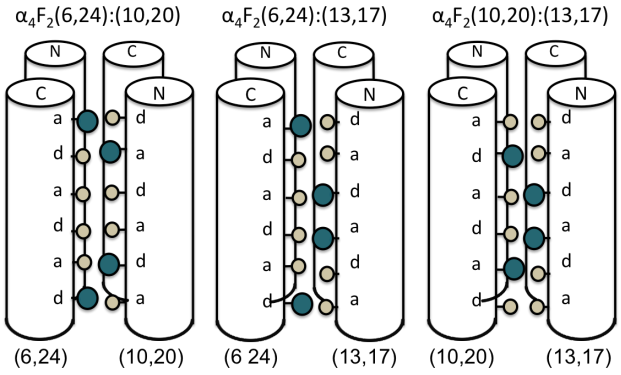

Figure 2.6. Cartoons illustrating the potential packing arrangement of alternating Leu and hFLeu side-chains in hetero-tetramers of $\alpha_4F_2(6,24):(10,20)$, $\alpha_4F_2(6,24):(13,17)$ and $\alpha_4F_2(10,20):(13,17)$. (hFLeu represented by large green spheres, Leu represented by small grey spheres)

The GuHCl denaturation curves for 1:1 mixtures of $\alpha_4F_2(6,24)$ and $\alpha_4F_2(10,20)$, $\alpha_4F_2(6,24)$ and $\alpha_4F_2(13,17)$, and $\alpha_4F_2(10,20)$ and $\alpha_4F_2(13,17)$ were obtained and compared with the denaturation curves that would be expected if the peptides were independently unfolding (i.e. the curve calculated by averaging the two unfolding curves obtained for the pure peptides from data in Figure 2.4). The apparent ΔG^o_{fold} for these mixtures are given in Table 2.1. Only the $\alpha_4F_2(10,20)$:(13,17) mixture exhibited an unfolding transition that occurred at significantly higher GuHCl concentrations from that predicted by the calculated curve (Figure 2.7). Correspondingly, ΔG^o_{fold} calculated for the mixture was significantly larger than that for either of the two pure components, pointing to the formation of more stable hetero-tetrameric species.

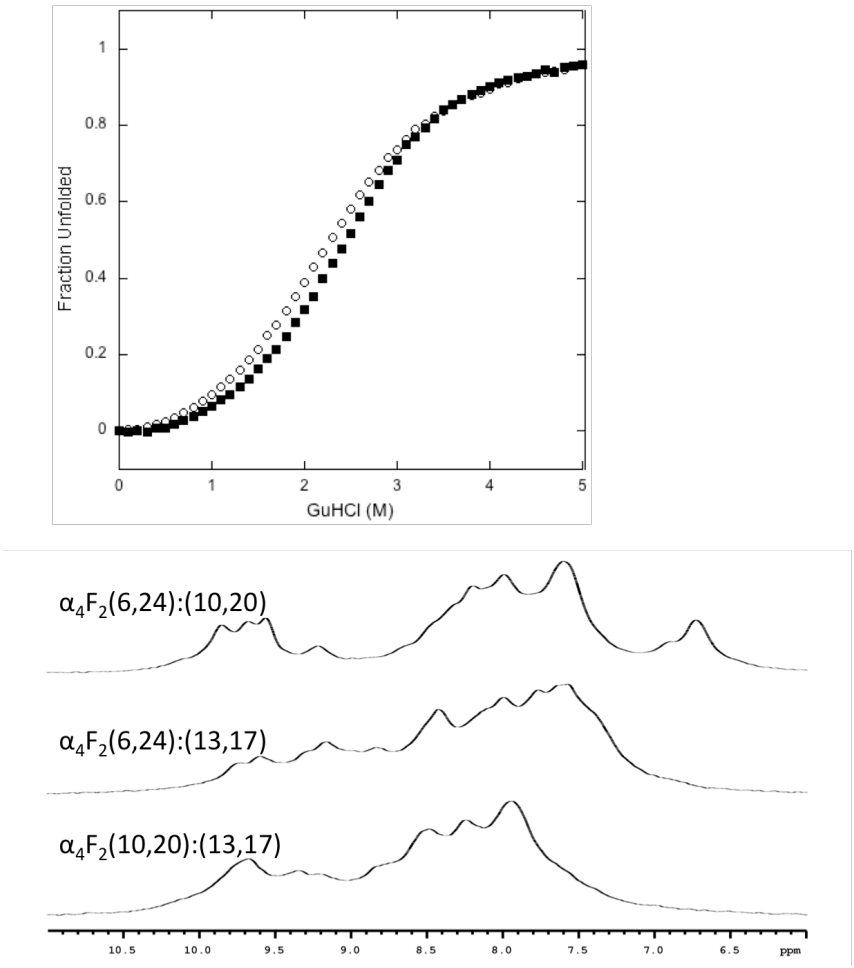

Figure 2.7. *Top*: GuHCl induced unfolding of a 1:1 mixture of α_4F_2(10,20):(13,17) (■) compared to a theoretical unfolding of a 1:1 mixture of α_4F_2 (10,20) and α_4F_2(13,17) (○). Conditions are as noted in Figure 2.4. *Bottom*: ^{19}F NMR spectra of (from top to bottom) α_4F_2(6,24):(10,20), α_4F_2(6, 24):(13,17) and α_4F_2(10,20):(13,17) recorded at 25 °C, pH 7.0.

The ^{19}F NMR spectra of the three α_4F_2 peptide combinations exhibited spectra that were each more complex than those of the individual peptides. Notably, additional resonances at higher and lower fields than those due to the pure peptides, were observed that can only be attributed to the formation of hetero-tetrameric species. As noted above, this is to be expected because there are, in principle, six different anti-parallel 4-helix bundles that can be formed from two peptides, and some hetero-tetramer combinations

may actually be slightly more stable than the corresponding homo-tetramers. However, the spectra give no indication that the peptide mixtures are assembling into a single, specific hetero-terameric structure; and indeed this would not be expected in view of the very modest increase in ΔG^{o}_{fold} measured for the peptide mixtures.

The unfolding of 1:1 mixtures of α_4H and α_4F_3a, α_4H and α_4F_3d and of α_4H and α_4F_6 were also examined. Here the difference in the thermodynamic stabilities of α_4H and α_4F_3a, α_4F_3d or α_4F_6 is very large, so if they formed stable hetero-tetramers the unfolding curves of the mixtures should be distinctly different. Whereas stable hetero-tetramers of α_4H and α_4F_3a peptides were not expected to form, it was reasoned that α_4H and α_4F_6 could form a hetero-tetramer in which each layer of the hydrophobic core was packed with two Leu and two hFLeu residues. This would mimic the alternating packing arrangement that appears to stabilize α_4F_3a and α_4F_3d, although in a hypothetical α_4H:α_4F_6 hetero-tetramer the hFLeu residues occupy alternating **a** and **d** positions throughout the core, rather than only **a** or only **d** positions. On the other hand, if the introduction of extensively fluorinated residues into the hydrophobic core resulted in self-segregating behavior, this should also be reflected in the unfolding curves of these peptide mixtures.

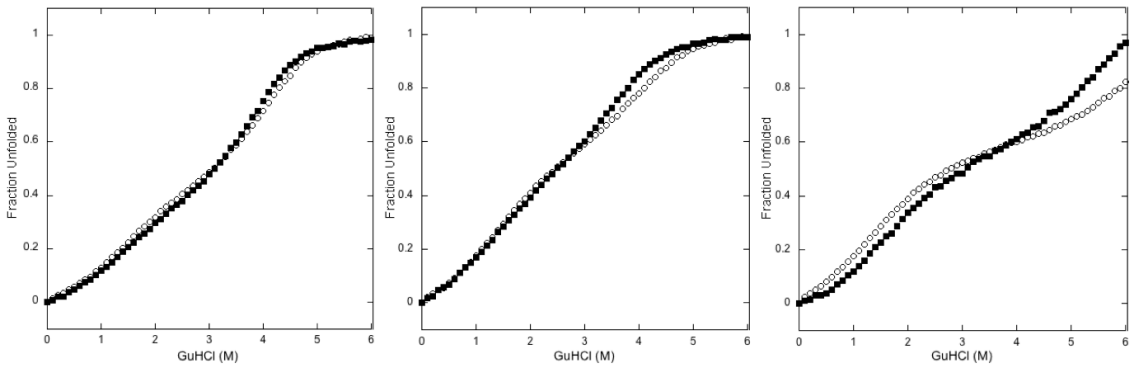

Figure 2.8. GuHCl-induced unfolding of mixtures of fluorinated and non-fluroinated peptides. *Left*: 20 μM α_4F_3a with 20 μM α_4H, *Middle*: 20 μM α_4Fd with 20 μM α_4H, *Right*: 20 μM α_4F_6 with 20 μM α_4H. For comparison, the theoretical unfolding curve for a 1:1 mixture of peptides that would result if there were no interaction between peptides is also shown in each case (○). Conditions are as noted in Figure 2.4.

Figure 2.8 shows the unfolding transitions of the peptide mixtures, and for comparison the calculated unfolding curves for a 1:1 mixture of the pure peptides is superimposed. The α_4H:α_4F_3a mixture most closely follows the unfolding curve for the ideal, non-interacting peptide mixture. The fit is less good for the α_4H:α_4F_3d mixture, and the α_4H:α_4F_6 mixture deviates significantly from the unfolding curve expected if the two peptides were completely self-segregating. However, in none of the cases were the curves characteristic of a 2-state unfolding transition of a single species that would be expected if a stable, uniquely folded hetero-tetramer was being formed. This indicates that these peptides exhibit some preference to self-segregate, but at least in the case of α_4F_3d and α_4F_6 there is a significant population of hetero-tetrameric peptides.

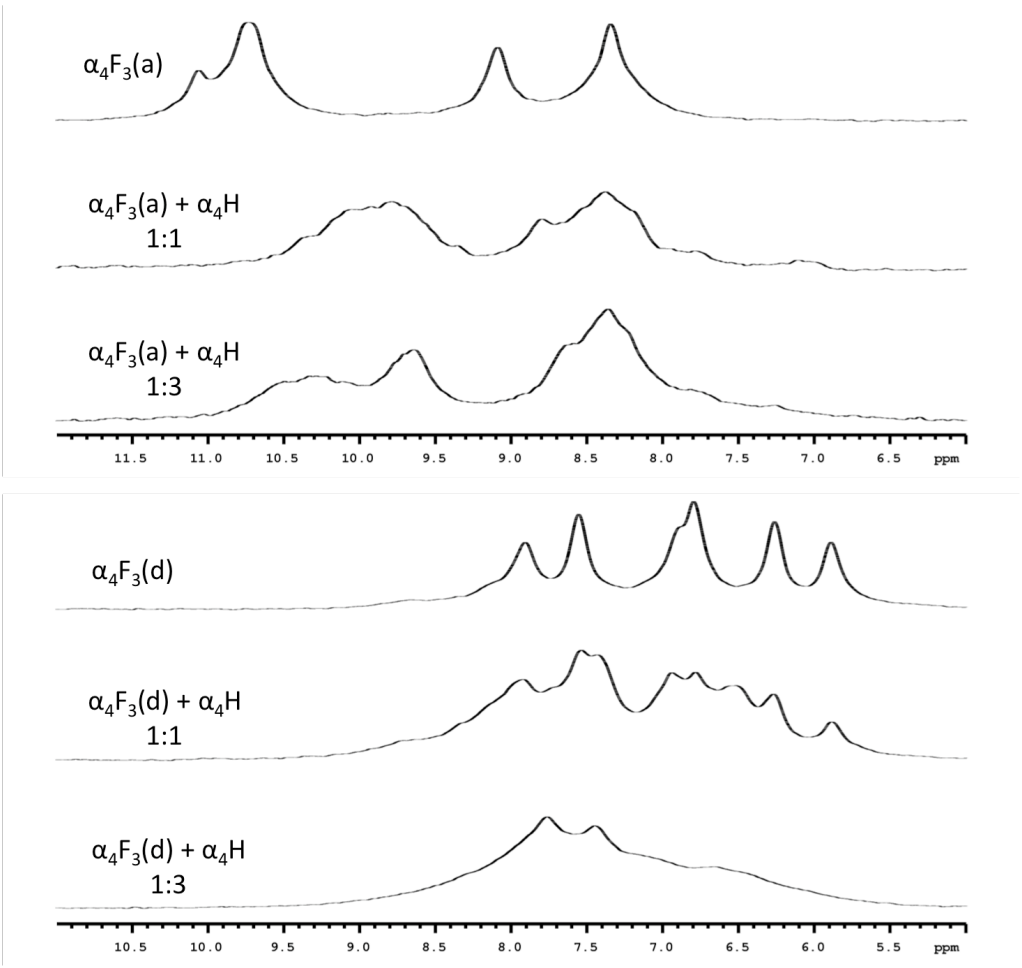

Figure 2.9. ^{19}F spectra of α_4F_3a (top panel) and α_4F_3d (bottom panel) in the presence of increasing concentrations of α_4H. Spectra, from top to bottom, are 1.5 mM α_4F_3 peptide, 1.5 mM α_4F_3 in the presence of 1.5 mM α_4H, and 1.5 mM α_4F_3 in the presence of 4.5 mM α_4H.

To provide further insight into the interactions between α_4H and α_4F_3a and α_4F_3d, the ^{19}F NMR spectra of these peptides in the presence of α_4H were recorded (Figure 2.9). As the concentration of α_4H increased, some new peaks emerged in the spectra of α_4F_3a and α_4F_3d and some changes in the chemical shifts of some peaks were apparent. However, the most apparent change was increased broadening of the peaks at higher concentrations of α_4H. This is characteristic of different peptide complexes in chemical exchange with each other on the NMR time-scale. Our laboratory has previously

performed experiments in which the interaction of α_4H with α_4F_6 was investigated by ^{19}F NMR and obtained qualitatively similar results to the present case [22].

We note that NMR spectra are very sensitive to the dynamics of peptide folding, and even small changes in the chemical environment and mobility of the ^{19}F nucleus will significantly affect chemical shift and peak shape due to chemical exchange [28]. In particular, if the peptides undergo exchange between hetero-tetrameric and homo-tetrameric species on the NMR (μs - ms) timescale the spectra will exhibit exchange broadening due to the differences in chemical shifts and relaxation times. If the chemical shift differences are large, as is the case for the fluorine nucleus, an equilibrium population of hetero-tetrameric species comprising only a few percent would be sufficient to cause significant broadening of the spectra of the homo-tetramer. Thus, whereas the unfolding experiments point to the presence of at least two populations of discretely folded peptides with very different unfolding characteristics, the NMR experiments reveal the existence of sub-populations of hetero-tetrameric peptides and highlight the dynamic nature of the 4-helix bundle structure.

2.4 – Discussion

These results show that the incorporation of fluorine into proteins, in this case using hFLeu, can be used as a design strategy to produce very stable proteins. Whereas there is some tendency for these peptides to self-segregate, as predicted by the partitioning of small molecule fluorocarbons into fluorous solvents [29] and demonstrated in some other extensively fluorinated de novo-designed peptides [8,20,21], the selectivity is not great. Furthermore, it appears that packing effects within the hydrophobic core are

more important than the segregation of fluorinated residues in contributing to the stability of these proteins. Thus, α_4F_3a and α_4F_3d, in which the fluorinated side-chains are interspersed with non-fluorinated side-chains exhibit greater per-residue stability than α_4F_6 in which the core is fully fluorinated. The large differences in stability between α_4H and α_4F_3a, α_4F_3d or α_4F_6 appear to provide an adequate explanation for why these peptides prefer to form homo-tetramers rather than hetero-tetramers, without the need to invoke fluorous self-segregation *per se*. Indeed, if anything, α_4F_6 (which is the most stably folded peptide) shows a greater tendency to interact with α_4H than either α_4F_3a or α_4F_3d.

The ^{19}F NMR spectra provide insight into the dynamic behavior of the peptides. Thus, although the fluorinated peptides are thermodynamically very stable, they are also dynamic and the NMR spectra provide evidence of exchange of individual peptides between helical bundles on the μs – ms time-scale. The spectra also suggest that packing adjacent layers of the hydrophobic core with hFLeu may result in over-packing of the core of $\alpha_4F_2(13,17)$ resulting in slow inter-conversion of different side-chain conformers that give rise to the broadening of the spectrum. It is worth noting that similarly broad, and more complex ^{19}F NMR spectra have previously been observed for α_4F_4 and α_4F_6 peptides [16] where the overcrowding of the protein core with hFLeu would be even greater.

These experiments further exemplify the usefulness of fluorine as a label to probe protein-protein interactions, protein structure and dynamics using NMR; further investigations into using fluorine NMR as a biological probe are discussed in Chapter 6. These experiments also establish the packing of hFLeu as a large stability determinant;

further investigations into the packing and structural accommodation of fluorinated amino acids are presented in Chapter 4. Lastly, the incorporation of hFLeu greatly increases the free energy of unfolding for α_4 proteins, contributions of thermodynamic parameters to the stability enhancement are presented in Chapter 5.

2.5 – References

(1) Horvath, I. T.; Rabai, J. *Science* **1994**, *266*, 72.
(2) Marsh, E. N. G. *Chem. Biol.* **2000**, *7*, R153.
(3) Luo, Z. Y.; Zhang, Q. S.; Oderaotoshi, Y.; Curran, D. P. *Science* **2001**, *291*, 1766.
(4) Yoder, N. C.; Yuksel, D.; Dafik, L.; Kumar, K. *Current Opinion in Chemical Biology* **2006**, *10*, 576.
(5) Yoder, N. C.; Kumar, K. *Chemical Society Reviews* **2002**, *31*, 335.
(6) Jackel, C.; Koksch, B. *European Journal of Organic Chemistry* **2005**, 4483.
(7) Smits, R.; Koksch, B. *Curr. Top. Med. Chem.* **2006**, *6*, 1483.
(8) Bilgicer, B.; Xing, X. C.; Kumar, K. *J. Am. Chem. Soc.* **2001**, *123*, 11815.
(9) Bilgicer, B.; Fichera, A.; Kumar, K. *J. Am. Chem. Soc.* **2001**, *123*, 4393.
(10) Tang, Y.; Tirrell, D. A. *J. Am. Chem. Soc.* **2001**, *123*, 11089.
(11) Niemz, A.; Tirrell, D. A. *J. Am. Chem. Soc.* **2001**, *123*, 7407.
(12) Tang, Y.; Ghirlanda, G.; Petka, W. A.; Nakajima, T.; DeGrado, W. F.; Tirrell, D. A. *Angew. Chem.-Int. Edit.* **2001**, *40*, 1494.
(13) Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, W. A.; DeGrado, W. F.; Tirrell, D. A. *Biochemistry* **2001**, *40*, 2790.
(14) Bilgicer, B.; Kumar, K. *Tetrahedron* **2002**, *58*, 4105.
(15) Lee, K.-H.; Lee, H.-Y.; Slutsky, M. S.; Anderson, J. T.; Marsh, E. N. G. *Biochemistry* **2004**, *43*, 16277.
(16) Lee, H. Y.; Lee, K. H.; Al-Hashimi, H. M.; Marsh, E. N. G. *J. Am. Chem. Soc.* **2006**, *128*, 337.
(17) Chiu, H. P.; Suzuki, Y.; Gullickson, D.; Ahmad, R.; Kokona, B.; Fairman, R.; Cheng, R. P. *J. Am. Chem. Soc.* **2006**, *128*, 15556.
(18) Woll, M. G.; Hadley, E. B.; Mecozzi, S.; Gellman, S. H. *J. Am. Chem. Soc.* **2006**, *128*, 15932.
(19) Jackel, C.; Salwiczek, M.; Koksch, B. *Angew. Chem.-Int. Edit.* **2006**, *45*, 4198.
(20) Bilgicer, B.; Kumar, K. *Proc. Natl. Acad. Sci. (USA)* **2004**, *101*, 15324.
(21) Naarmann, N.; Bilgicer, B.; Meng, H.; Kumar, K.; Steinem, C. *Angew. Chem.-Int. Edit.* **2006**, *45*, 2588.
(22) Gottler, L. M.; de la Salud-Bea, R.; Marsh, E. N. G. *Biochemistry* **2008**, *47*, 4484.
(23) Anderson, J. T.; Toogood, P. L.; Marsh, E. N. G. *Org. Lett.* **2002**, *4*, 4281.
(24) Harding, S. E.; Rowe, A. J.; Horton, H. C. *Analytical Ultracentrifugation in Biochemistry and Polymer Science*; The Royal Society of Chemistry: Cambridge, UK, 1992.

(25) Cohn, E. J.; Edsall, J. T. *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*; Reinhold: New York, 1943.
(26) Yu, J. X.; Kodibagkar, V. D.; Cui, W. N.; Mason, R. P. *Current Medicinal Chemistry* **2005**, *12*, 819.
(27) Gerig, J. T. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 293.
(28) Mizuno, T.; Hasegawa, C.; Tanabe, Y.; Hamajima, K.; Muto, T.; Nishi, Y.; Oda, M.; Kobayashi, Y.; Tanaka, T. *Chemistry – A European Journal* **2009**, *15*, 1491.
(29) Studer, A.; Hadida, S.; Ferritto, R.; Kim, S. Y.; Jeger, P.; Wipf, P.; Curran, D. P. *Science* **1997**, *275*, 823.

Chapter 3

Crystallization of Hydrocarbon and Fluorocarbon α_4 Proteins

3.1– Introduction

Designing proteins with enhanced physical properties, such as chemical and metabolic stability, is of great interest. These desirable properties have been introduced into proteins through the incorporation of unnatural amino acid side chains or noncanonical backbones [1-4]. Highly fluorinated analogues of the amino acids valine, leucine and phenylalanine have shown promise in enhancing conformational stability of several protein folds. Although fluorine is larger than hydrogen, van der Waals radius of 1.35 Å versus 1.2 Å and C-F bond length of ~1.4 Å versus C-H bond length of ~1.0 Å, fluorine substitution for hydrogen is largely considered nonperturbing. The bulkier fluorous amino acids retain the shape of their hydrocarbon counterparts and are generally considered isosteres of natural residues as exemplified by tFLeu, tFVal and tFIle being activated by endogenous (and highly specific) tRNA synthetases [4-8]. Fluorine is also frequently substituted for hydrogen in small molecules, with minimal detriment to binding proteins and enzymes [9]. The high fluorine content of hFVal, hFLeu and pFPhe makes these amino acids of exceptional interest towards modifying the properties of proteins.

The structures of folded proteins can be very sensitive to alterations brought upon by mutating residues; disrupting a tightly packed core or a salt bridge can greatly destabilize the folded state. Engineering proteins with natural amino acids requires an understanding of structural consequences to a given substitution; this is made easier by the thousands of high resolution protein structures deposited in the Protein Data Bank that offer insight into the folds and amino acid interactions necessary to stabilize natural proteins. Straying beyond the natural complement of 20 amino acids for mutational studies becomes increasingly challenging due to a lack of structural information. Fluorinated protein structures determined by X-ray crystallography would provide an excellent method to dissect molecular interactions and help predict how these substitutions may have structural consequences thereby altering stability and function.

Currently, structural data for fluorous proteins is lacking. There are a few structures demonstrating the incorporation of monofluorinated side chains into a protein environment [10-18]. Of these studies, global replacement of tyrosine for 3-fluorotyrosine in glutathione transferase M1-1 represents the highest degree of fluorine substitution in a protein structure so far. The incorporation of 14 3-fluorotyrosine residues into the interior, exterior and active site disrupts local packing interactions; however, overall structure and function is retained. These studies confirm the nonperturbing nature and biocompatibility of incorporating small numbers of fluorine atoms.

Of greater interest for protein design are amino acids with perfluorinated sidechains, which have demonstrated the largest potential benefit for stabilizing natural proteins. Recently, studies from the Gellman group have determined both NMR and X-ray structures of chicken villin headpiece subdomain (cVHP) with pFPhe replacing one

of three Phe residues in the core [19,20]. cVHP is a small, 35-residue protein which forms a discrete tertiary structure with residues Phe6, Phe10 and Phe17 forming the hydrophobic core. Crystal structures for pFPhe mutants at positions 10 and 17 were obtained and displayed very similar aryl side chain rmsd values compared to native structures. Previous investigations determined that not all substitutions to the Phe containing core are stabilizing. Out of the seven possible pFPhe mutants with one, two or all three Phe positions substituted, only one was stabilizing (pFPhe10). The compact nature of cVHP appears unable to accommodate the increased volume of the pFPhe side chain in the other two positions [21].

Although most fluoroaryl substitutions in cVHP are destabilizing, fluoroalkyl substitutions have seen great success in stabilizing a variety of protein folds. Analogous crystal structures of both native and fluorinated states of cVHP, as well as other proteins, are key to understanding the structural consequences of incorporating fluorinated amino acids. These studies highlight the ability of fluorous proteins to be crystallized in conditions similar to those of natural proteins. Of notable interest to our laboratory is hFLeu, which no protein crystal structure has yet been reported, although a low resolution structure of a β-peptide bundle containing hexafluoro-β^3-leucine in the core has been described by Kumar and Schepartz [22]. Due to resolution limitations (2.75 Å) arising from twinning in the crystal, distance measurements and side chain orientations of fluorinated residues were not reported and would be nearly impossible to ascertain.

The high rate of occurrence of Leu in natural proteins (9.1% compared to 3.9% for Phe) makes it an ideal residue for fluorous substitution. Our laboratory and several others have investigated the physical and chemical properties of hFLeu incorporation into

a variety of model proteins [23-36]. In most cases, these modified proteins display increased stability towards unfolding by chemical denaturants, solvents [30] and increased resistance towards proteolytic degredation, which is of particular importance to AMPs [30,31,34,35]. The overall protein structure does not appear to be altered by these substitutions; however, the local interactions and structural consequences of accommodating two or more trifluoromethyl groups in lieu of methyl groups are unknown.

		abcdefg	abcdefg	abcdefg	a
α_4H	Ac–GN	ADE**L**YKE	**L**ED**L**QER	**L**RK**L**RKK	**L**RSG–NH_2
α_4Ht	Ac–GN	ADE**X**YKE	**X**ED**X**QER	**X**RK**X**RKK	**X**RSG–NH_2
α_4F_2(6,24)	Ac–GN	ADE**X**YKE	LEDLQER	LRKLRKK	**X**RSG–NH_2
α_4F_2(10,20)	Ac–GN	ADELYKE	**X**EDLQER	LRK**X**RKK	LRSG–NH_2
α_4F_2(13,17)	Ac–GN	ADELYKE	LED**X**QER	**X**RKLRKK	LRSG–NH_2
α_4F_3(6-13)	Ac–GN	ADE**X**YKE	**X**ED**X**QER	LRKLRKK	LRSG–NH_2
α_4F_3(17-24)	Ac–GN	ADELYKE	LEDLQER	**X**RK**X**RKK	**X**RSG–NH_2
α_4F_3a	Ac–GN	ADELYKE	**X**EDLQER	**X**RKLRKK	**X**RSG–NH_2
α_4F_3d	Ac–GN	ADE**X**YKE	LED**X**QER	LRK**X**RKK	LRSG–NH_2
$\alpha_4F_3$at	Ac–GN	ADE**X**YKE	**X**ED**X**QER	**X**RK**X**RKK	**X**RSG–NH_2
$\alpha_4F_3af_3d$	Ac–GN	ADE**X**YKE	**X**ED**X**QER	**X**RK**X**RKK	**X**RSG–NH_2
α_4F_6	Ac–GN	ADE**X**YKE	**X**ED**X**QER	**X**RK**X**RKK	**X**RSG–NH_2

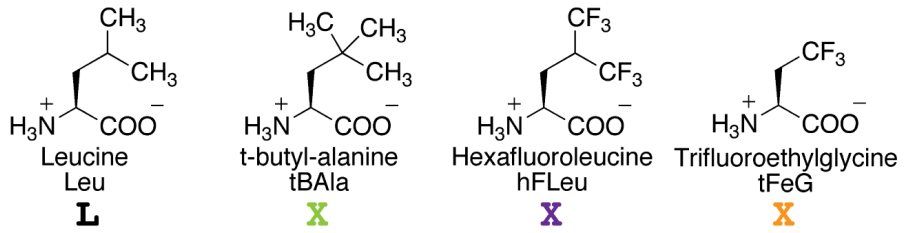

Figure 3.1. Sequences of α_4 proteins with hydrocarbon and fluorocarbon analogues of Leu that are substituted at **a** and **d** positions.

To determine in detail the structural changes induced by fluorination with the intention of gaining understanding of the mechanism by which fluorine stabilizes proteins, I undertook protein crystallization experiments of α_4H and its various fluorinated analogues listed in Figure 3.1 and described previously in Chapter 2. α_4H, is a

de novo designed anti-parallel 4-α-helix bundle, containing Leu at the three **a** and three **d** positions of the heptad repeat, which forms the hydrophobic core of the folded tetramer. In various studies, our laboratory has synthesized a number of fluorinated versions of α_4, designated α_4F_n, which incorporate hFLeu at different positions within the core and examined their physical and biological properties. Proteins containing the nonnatural Leu analogues tBAla and tFeG have also been synthesized. The core Leu residues of α_4H are very tolerant to hFLeu substitution, with all fluorous analogues being well-folded tetramers as assessed by circular dichroism and analytical ultracentrifugation. The high thermodynamic stability of all analogues and retention of native structure (even though hFLeu has a side chain volume ~30 % greater than that of Leu) suggested that crystallization conditions and subsequent structure determination would be similar for all proteins. To this end crystallization trials were undertaken on 12 α_4 analogues, of which, 11 crystallized and 6 yielded high-resolution structures.

3.2 – Experimental Procedures

3.2.1 - Materials and Peptide Synthesis

L-5,5,5,5′,5′,5′-hexafluoroleucine was synthesized and converted to Boc-protected derivative by standard procedures as described in Chapter 2 section 2.1. 4,4,4-trifluoroethylglycine was purchased from SynQuest Laboratory and enzymatically resolved as described in Chapter 6 section 2.1 [38]. Boc- and Fmoc-protected β-t-butyl-L-alanine were purchased from AnaSpec Inc. Peptides were synthesized by manual Fmoc procedures (α_4H and α_4Ht) or manual Boc procedures ($\alpha_4F_2(6,24)$, $\alpha_4F_2(10,20)$, $\alpha_4F_2(13,17)$, α_4F_3a, α_4F_3d, $\alpha_4F_3(6\text{-}13)$, $\alpha_4F_3(17\text{-}24)$, $\alpha_4F_3af_3d$, α_4F_3at and α_4F_6) as described

in Chapter 2 section 2.1 [26,33,39]. All peptides were purified on a Waters preparatory HPLC using a linear gradient of 95% water, 4.9% acetonitrile and 0.1% TFA for solvent A and 9.9% water, 90% acetonitrile and 0.1% TFA for solvent B, with a flow rate of 10 mL/min on either a Waters or Vydac C18 preparatory column. Peptide identity was confirmed using MALDI-MS with a matrix of α-cyano-4-hydroxycinnamic acid.

3.2.2 - Crystallization Conditions

Peptides were dissolved in ddH_2O to a concentration of 6 mM (monomer) as determined by absorbance at 280 nm of the solvent exposed tyrosine residue at position 7. Initial protein crystal hits came from screening BCB-SP conditions (Table 3.1) in a Greiner 96-well sitting drop plate with well volume of 75 µL and drop volume of 1 µL peptide and 1 µL precipitant.

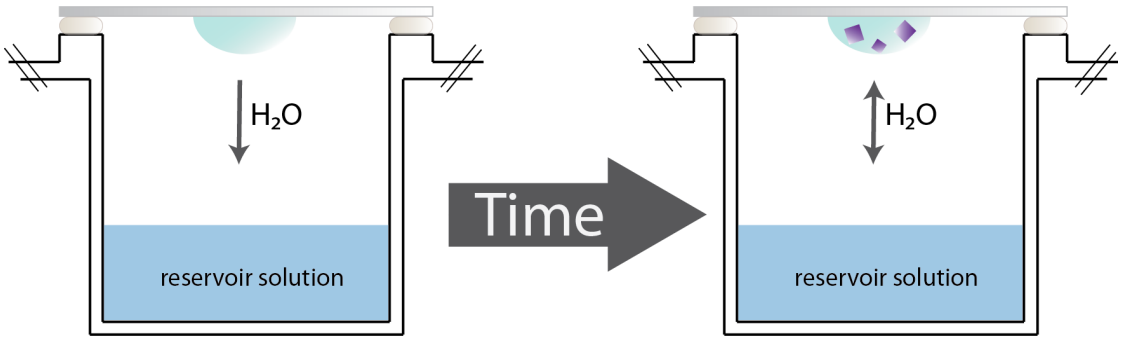

Figure 3.2. Hanging drop vapor diffusion. Initial precipitant concentration in the hanging drop is half that of the reservoir. Over time water diffuses out of the drop thereby slowly doubling the precipitant concentration and protein concentration.

Crystals for data collection were grown by vapor diffusion at 20 °C in a 24-well grid screen plate, with 1 mL precipitant well volume, and a hanging drop containing 2 µl peptide and 2 µl precipitant on a siliconized glass cover slide (Fig 3.2). All peptide

crystals formed under high concentrations of either PEG 400 or PEG 600, providing cryoprotection and allowing them to be flashed-cooled using liquid N_2 in their mother liquor for data collection.

3.3 – Results

3.3.1 - Initial Attempts – α_4F_6 Crystals

Initial attempts to crystallize fluorinated α_4 proteins began with trials of α_4F_6. The high solubility of α_4F_6 (>10 mM of monomer) complicated determination of appropriate conditions because crystallization relies on protein supersaturation to initiate crystal growth. Supersaturation can manifest as precipitation or under ideal conditions as protein crystal formation. Little protein precipitation was seen for initial trials, even at high precipitant concentrations, making determination of conditions for supersaturation difficult. Commercially available screens in 96-well format were used for initial trials. These included Wizard I and II screens from Emerald BioSystems, Crystal Screen HT and 2 HT from Hampton Research, Cryo Screen I and II from Emerald BioSystems and PEG I and II from Qiagen. Use of these diverse screens allowed a broad sampling of salts, polymers, organic compounds, pH (4.2-10.5) and cryoprotectants to obtain the greatest chance of finding an initial crystallization condition.

The only condition to afford a non-salt crystal of α_4F_6 was 100 mM phosphate-citrate buffer (pH 4.2) and 40 % PEG 600 from Cryo Screen I. Further 24-well grid screens around conditions of the initial hit gave the largest and most consistent crystals with 6 mM α_4F_6, 100 mM phosphate-citrate buffer (pH 4.2) and 53 % PEG 600 (Fig. 3.3). Since the resulting crystals appeared to have rough edges and were $\leq$ 50 μm in any one

dimension, additive screening was used to improve crystal quality. Hampton Research Additive Screen allowed 96 combinations of various metals and small molecules to be screened while keeping 100 mM phosphate-citrate buffer (pH 4.2) and 53 % PEG 600 consistent. The screen identified 0.4 µL 30 % MeOH and 0.4 µL 0.1 M NaBr added to the 4 µL drop of 6 mM α_4F_6 in 100 mM phosphate-citrate buffer (pH 4.2) and 53 % PEG 600 as producing the best crystals (Fig. 3.3). These hexagonal crystals measured 75 x 30 µm, however, they only diffracted to ~3.6 Å and the data collected did not yield structural information.

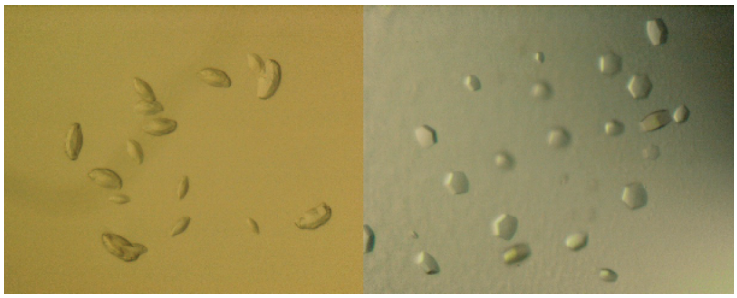

Figure 3.3. Crystals of fully fluorinated α_4F_6 proteins. *Left*: α_4F_6: 100 mM phos-cit buffer (pH 4.2) 53% PEG 600 *Right*: α_4F_6: 100 mM phos-cit buffer (pH 4.2) 53% PEG 600, 0.4µL 30% MeOH, 0.4µL 0.1M NaBr

After minimal success crystallizing α_4F_6 from exhaustive screening using a variety of commercial screens and additives, I assumed the difficulties were due to the inherent structure of α_4F_6 and proceeded to focus my effort on other α_4 proteins. The ~30% larger side-chain volume of hFLeu compared with Leu results in a much larger hydrophobic core volume for α_4F_6 versus α_4H. This increase in volume, although seen to be greatly stabilizing, may result in overpacking of the core giving some deviation from the idealized coiled-coil helix design of α_4H.

Uniformity within a crystal lattice is paramount for well-diffracting crystals; disruption of the three dimensional repeating protein arrangement can result in poorly formed crystals, giving low-resolution diffraction. Screening α_4H and α_4F_3a, which have a lower content of hFLeu resulted in well-formed crystals giving high-resolution diffraction patterns.

3.3.2 – Matrix Screen for α_4 Proteins

α_4F_3a crystallized in a variety of conditions containing PEG 200, PEG 400 and PEG 600 from the Emerald BioSystems Cryo Screen I and II. This led to the development of a 96-well, α_4-protein-specific screen. Named BCB-Small PEG (BCB-SP), this screen is shown in Table 3.1 and uses high concentrations of various small molecular weight PEGs (200-600) across five common buffers, spanning pH values from 4.5 to 9.0.

A1	0.1 M	Sodium acetate pH 4.5	50% PEG 200	E1	0.1 M	Tris pH 8.0	55% PEG 400
A2	0.1 M	MES pH 5.5	50% PEG 200	E2	0.1 M	CHES pH 9.0	55% PEG 400
A3	0.1 M	HEPES pH 7.0	50% PEG 200	E3	0.1 M	Sodium acetate pH 4.5	60% PEG 400
A4	0.1 M	Tris pH 8.0	50% PEG 200	E4	0.1 M	MES pH 5.5	60% PEG 400
A5	0.1 M	CHES pH 9.0	50% PEG 200	E5	0.1 M	HEPES pH 7.0	60% PEG 400
A6	0.1 M	Sodium acetate pH 4.5	55% PEG 200	E6	0.1 M	Tris pH 8.0	60% PEG 400
A7	0.1 M	MES pH 5.5	55% PEG 200	E7	0.1 M	CHES pH 9.0	60% PEG 400
A8	0.1 M	HEPES pH 7.0	55% PEG 200	E8	0.1 M	Sodium acetate pH 4.5	40% PEG 550 MME
A9	0.1 M	Tris pH 8.0	55% PEG 200	E9	0.1 M	MES pH 5.5	40% PEG 550 MME
A10	0.1 M	CHES pH 9.0	55% PEG 200	E10	0.1 M	HEPES pH 7.0	40% PEG 550 MME
A11	0.1 M	Sodium acetate pH 4.5	60% PEG 200	E11	0.1 M	Tris pH 8.0	40% PEG 550 MME
A12	0.1 M	MES pH 5.5	60% PEG 200	E12	0.1 M	CHES pH 9.0	40% PEG 550 MME
B1	0.1 M	HEPES pH 7.0	60% PEG 200	F1	0.1 M	Sodium acetate pH 4.5	45% PEG 550 MME
B2	0.1 M	Tris pH 8.0	60% PEG 200	F2	0.1 M	MES pH 5.5	45% PEG 550 MME
B3	0.1 M	CHES pH 9.0	60% PEG 200	F3	0.1 M	HEPES pH 7.0	45% PEG 550 MME
B4	0.1 M	Sodium acetate pH 4.5	45% PEG 300	F4	0.1 M	Tris pH 8.0	45% PEG 550 MME
B5	0.1 M	MES pH 5.5	45% PEG 300	F5	0.1 M	CHES pH 9.0	45% PEG 550 MME
B6	0.1 M	HEPES pH 7.0	45% PEG 300	F6	0.1 M	Sodium acetate pH 4.5	50% PEG 550 MME
B7	0.1 M	Tris pH 8.0	45% PEG 300	F7	0.1 M	MES pH 5.5	50% PEG 550 MME
B8	0.1 M	CHES pH 9.0	45% PEG 300	F8	0.1 M	HEPES pH 7.0	50% PEG 550 MME
B9	0.1 M	Sodium acetate pH 4.5	50% PEG 300	F9	0.1 M	Tris pH 8.0	50% PEG 550 MME
B10	0.1 M	MES pH 5.5	50% PEG 300	F10	0.1 M	CHES pH 9.0	50% PEG 550 MME
B11	0.1 M	HEPES pH 7.0	50% PEG 300	F11	0.1 M	Sodium acetate pH 4.5	55% PEG 550 MME
B12	0.1 M	Tris pH 8.0	50% PEG 300	F12	0.1 M	MES pH 5.5	55% PEG 550 MME
C1	0.1 M	CHES pH 9.0	50% PEG 300	G1	0.1 M	HEPES pH 7.0	55% PEG 550 MME
C2	0.1 M	Sodium acetate pH 4.5	55% PEG 300	G2	0.1 M	Tris pH 8.0	55% PEG 550 MME
C3	0.1 M	MES pH 5.5	55% PEG 300	G3	0.1 M	CHES pH 9.0	55% PEG 550 MME
C4	0.1 M	HEPES pH 7.0	55% PEG 300	G4	0.1 M	Sodium acetate pH 4.5	40% PEG 600
C5	0.1 M	Tris pH 8.0	55% PEG 300	G5	0.1 M	MES pH 5.5	40% PEG 600
C6	0.1 M	CHES pH 9.0	55% PEG 300	G6	0.1 M	HEPES pH 7.0	40% PEG 600
C7	0.1 M	Sodium acetate pH 4.5	60% PEG 300	G7	0.1 M	Tris pH 8.0	40% PEG 600
C8	0.1 M	MES pH 5.5	60% PEG 300	G8	0.1 M	CHES pH 9.0	40% PEG 600
C9	0.1 M	HEPES pH 7.0	60% PEG 300	G9	0.1 M	Sodium acetate pH 4.5	45% PEG 600
C10	0.1 M	Tris pH 8.0	60% PEG 300	G10	0.1 M	MES pH 5.5	45% PEG 600
C11	0.1 M	CHES pH 9.0	60% PEG 300	G11	0.1 M	HEPES pH 7.0	45% PEG 600
C12	0.1 M	Sodium acetate pH 4.5	45% PEG 400	G12	0.1 M	Tris pH 8.0	45% PEG 600
D1	0.1 M	MES pH 5.5	45% PEG 400	H1	0.1 M	CHES pH 9.0	45% PEG 600
D2	0.1 M	HEPES pH 7.0	45% PEG 400	H2	0.1 M	Sodium acetate pH 4.5	50% PEG 600
D3	0.1 M	Tris pH 8.0	45% PEG 400	H3	0.1 M	MES pH 5.5	50% PEG 600
D4	0.1 M	CHES pH 9.0	45% PEG 400	H4	0.1 M	HEPES pH 7.0	50% PEG 600
D5	0.1 M	Sodium acetate pH 4.5	50% PEG 400	H5	0.1 M	Tris pH 8.0	50% PEG 600
D6	0.1 M	MES pH 5.5	50% PEG 400	H6	0.1 M	CHES pH 9.0	50% PEG 600
D7	0.1 M	HEPES pH 7.0	50% PEG 400	H7	0.1 M	Sodium acetate pH 4.5	55% PEG 600
D8	0.1 M	Tris pH 8.0	50% PEG 400	H8	0.1 M	MES pH 5.5	55% PEG 600
D9	0.1 M	CHES pH 9.0	50% PEG 400	H9	0.1 M	HEPES pH 7.0	55% PEG 600
D10	0.1 M	Sodium acetate pH 4.5	55% PEG 400	H10	0.1 M	Tris pH 8.0	55% PEG 600
D11	0.1 M	MES pH 5.5	55% PEG 400	H11	0.1 M	CHES pH 9.0	55% PEG 600
D12	0.1 M	HEPES pH 7.0	55% PEG 400	H12	0.1 M	HEPES pH 7.0	45% PEG 200

Table 3.1. BCB-SP screen for crystallizing α_4 proteins.

The BCB-SP screen was used to identify initial crystallization conditions and subsequently obtained single crystals for all α_4 proteins were as a result of hits from the BCB-SP screen. Large (>50 micron), single crystals with sharp edges were obtained for

all α_4 proteins screened with the exception of $\alpha_4F_3$at. Optimal buffer and precipitant conditions for highest resolution data (or best appearing crystals if no data was collected) are listed in Table 3.2 and representative crystal images are shown in Figure 3.4.

protein	buffer (100mM)	pH	precipitant	diffraction limit (Å)
α_4H	CHES	9.0	48% PEG 400	1.36
α_4Ht	Tris	8.3	54% PEG 400	1.54
α_4F_2(6,24)	Tris	8.0	55% PEG 600	~2.5[b]
α_4F_2(10,20)	Na-acetate	4.5	62% PEG 600	~2.5[b]
α_4F_2(13,17)	Tris	8.0	48% PEG 600	~2.5[b]
α_4F_3a	Tris	7.8	55% PEG 400	1.54
α_4F_3d	Tris	8.3	51% PEG 600	1.19
α_4F_3(6-13)	HEPES	7.3	54% PEG 600	1.48
α_4F_3(17-24)	Tris	8.5	40% PEG 600	~2.5[b]
$\alpha_4F_3$at	-	-	-	-
$\alpha_4F_3af_3$d	Tris	8.5	48% PEG 600	1.72
α_4F_6	phos-cit	4.2	53% PEG 600[a]	~3.6[b]

Table 3.2. Crystallization conditions for α_4 proteins and crystal diffractions limits. [a] The best crystals of α_4F_6 also contained additives: 0.4µL 30% MeOH and 0.4µL 0.1M NaBr added to the 2µL hanging drop. [b] Estimates of the diffraction limit based upon observed spots from the diffraction pattern.

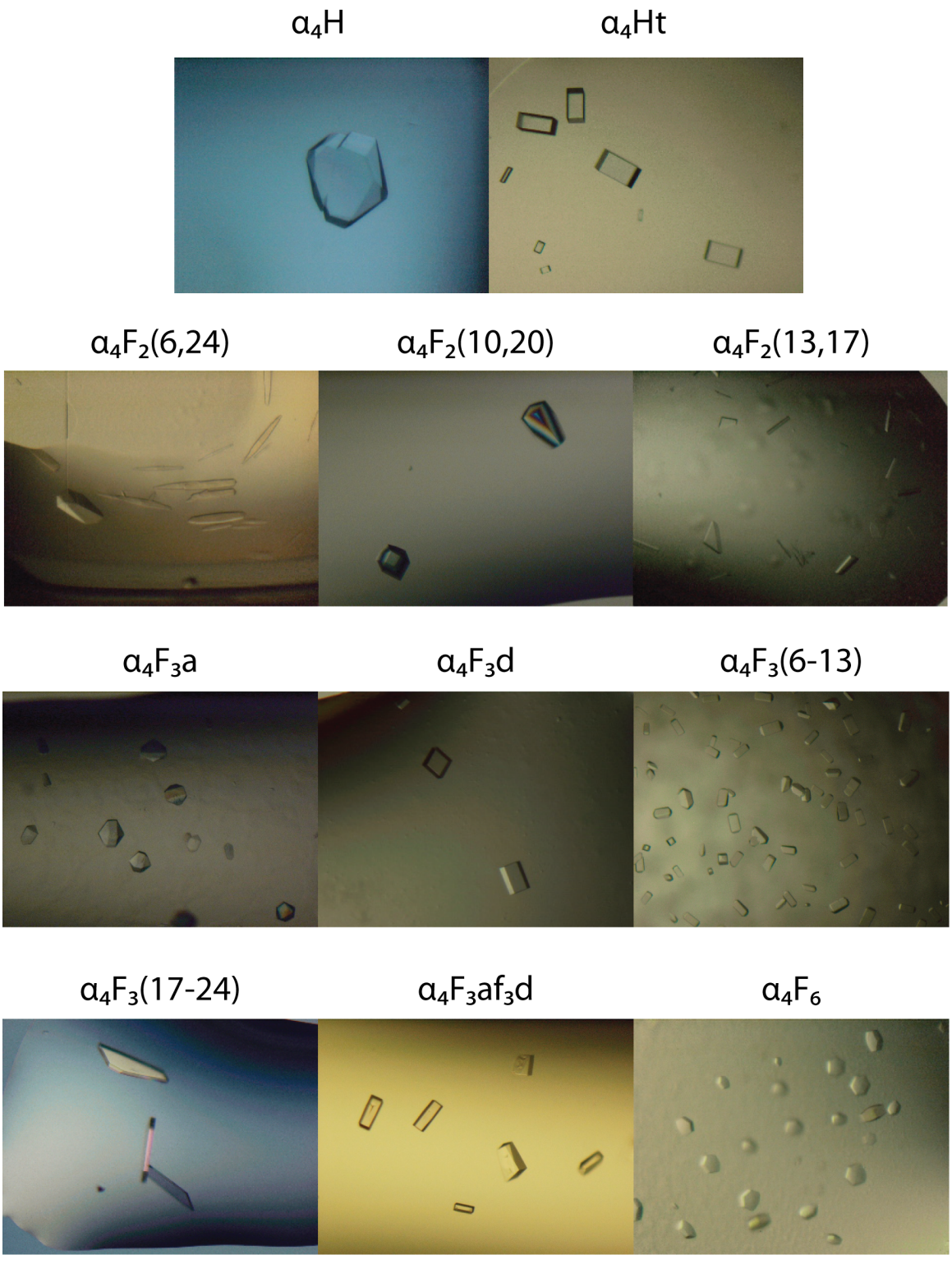

Figure 3.4. Representative crystals of α_4 proteins.

3.4 – Discussion

α_4H has been shown to be a very robust model system for analyzing the effects of hydrophobic residue substitution on stability and structure [26,32,33]. The core residues at **a** and **d** positions account for 6 out of 27 total residues, or about 22 %. Substituting any or all of the Leu residues, as in the cases of α_4Ht, $\alpha_4F_3af_3d$, α_4F_3at and α_4F_6, results in a well-folded, anti-parallel 4-helix bundle. This insensitivity of structural topology to repacking the hydrophobic core is likely due to the robustness of the anti-parallel arrangement, where two distinct interfaces (**b**–**e** and **c**–**g**) reinforce the anti-parallel structure through charge-charge complementarity. Greater oligomeric variation occurs in the related parallel coiled-coil, where there is only one unique interface composed of **e**–**g** interactions; slight changes in hydrophobic volume can lead to entirely different oligomeric states [40,41].

Analogues of α_4H display very similar structural properties: highly α-helical, a 2-state unfolding transition and folding into an anti-parallel tetramer [26,32,33]. Decreasing or increasing hydrophobic content or altering the chemical composition of the core appears to have little effect on the overall structure even though stability can be drastically altered. Studies from the Cheng laboratory have shown hFLeu and other fluorinated amino acid analogues to be destabilizing relative to hydrocarbon analogues – in the context of a single helix [28]. Our studies on 4-helix bundles show these hydrophobic substitutions to be stabilizing and having little impact on overall helical structure. The ability of nearly all α_4 proteins to crystallize attests to the well-folded nature of fluorinated analogues.

Although 11 out of 12 analogues crystallized, I was able to solve the structures for only 6 proteins. The other proteins gave poorly diffracting crystals and, interestingly, 4 of the 5 shared similar hFLeu packing arrangements. These are the α_4F_2 series and α_4F_6, which all have complete layers of the core packed with hFLeu. It is unwise to interpret these negative results with much weight, however, one could speculate that the increased volume of an entirely fluorinated layer may distort the helical structure thereby disrupting crystal packing. The crystal packing for the 6 solved structures is very similar, all have a dimer in the asymmetric unit, with three proteins in the $P2_12_12$ space group and three in the $I4_1$ space group.

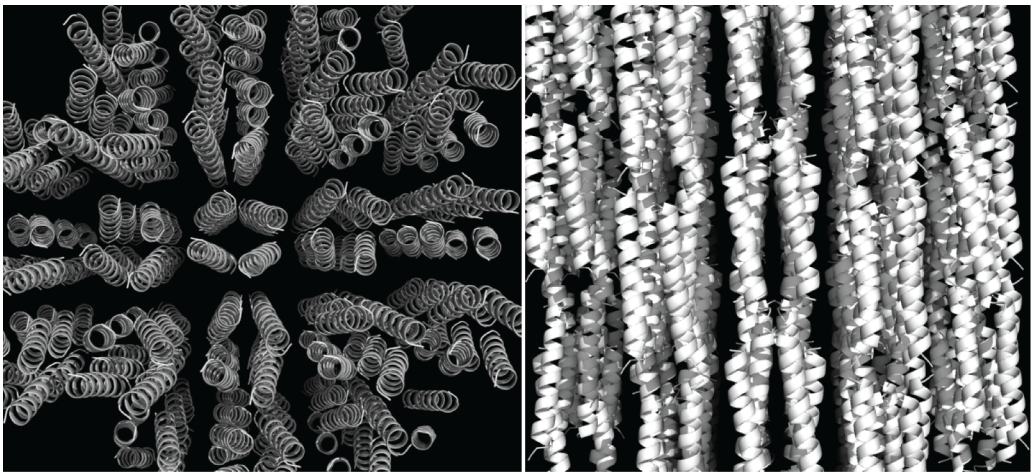

Figure 3.5. Crystal packing of α_4F_3(6-13) in $I4_1$ space group.

α_4 proteins pack in the crystal end to end with narrow solvent channels between tetramers and numerous contacts between charged residues of neighboring proteins (Fig. 3.5). Crystal packing of protein tetramers results in a solvent content volume of ~30% for all 6 α_4 proteins with determined structures. It is worth noting that the average fraction of crystal volume occupied by solvent for structures in the Protein Data Bank is ~50% with very few examples of protein crystals containing 30% or less. Due to the compact nature

of the 4-helix bundle and general rectangular prism shape of a tetrameric coiled-coil, these narrow solvent channels and extensive contacts between proteins are unavoidable. Accordingly, disruptions in crystal packing could result from small perturbations in protein structure that is propagated across the lattice.

α_4H provides a protein that is highly amenable to increases in hydrophobic amino acid volume of both hydrocarbon and fluorocarbon nature. The retention of structure and high stability of these mutants appears to be at least partially responsible for the ease in crystallization. The protein structures obtained from the crystallization trials described here are analyzed in depth in Chapter 4. These are the first structures of highly fluorinated proteins and provide important insights into enhanced stability of fluorinated proteins and the use of hydrophobic amino acids in protein design.

3.5 – References

(1) Baek, S.; Kutchukian, P. S.; Verdine, G. L.; Huber, R.; Holak, T. A.; Lee, K. W.; Popowicz, G. M. *Journal of the American Chemical Society* **2012**.

(2) Boersma, M. D.; Haase, H. S.; Peterson-Kaufman, K. J.; Lee, E. F.; Clarke, O. B.; Colman, P. M.; Smith, B. J.; Horne, W. S.; Fairlie, W. D.; Gellman, S. H. *Journal of the American Chemical Society* **2011**.

(3) Sawada, T.; Gellman, S. H. *Journal of the American Chemical Society* **2011**.

(4) Son, S.; Tanrikulu, I. C.; Tirrell, D. A. *ChemBioChem* **2006**, *7*, 1251.

(5) Montclare, J. K.; Son, S.; Clark, G. A.; Kumar, K.; Tirrell, D. A. *ChemBioChem* **2009**, *10*, 84.

(6) Tang, Y.; Ghirlanda, G.; Petka, W. A.; Nakajima, T.; DeGrado, W. F.; Tirrell, D. A. *Angew Chem Int Ed* **2001**, *40*, 1494.

(7) Wang, P.; Fichera, A.; Kumar, K.; Tirrell, D. A. *Angew Chem Int Ed* **2004**, *43*, 3664.

(8) Wang, P.; Tang, Y.; Tirrell, D. A. *J Am Chem Soc* **2003**, *125*, 6900.

(9) Müller, K.; Faeh, C.; Diederich, F. *Science* **2007**, *317*, 1881.

(10) Alexeev, D.; Barlow, P. N.; Bury, S. M.; Charrier, J.-D.; Cooper, A.; Hadfield, D.; Jamieson, C.; Kelly, S. M.; Layfield, R.; Mayer, R. J.; McSparron, H.; Price, N. C.; Ramage, R.; Sawyer, L.; Starkmann, B. A.; Uhrin, D.; Wilken, J.; Young, D. W. *ChemBioChem* **2003**, *4*, 894.

(11) Bae, J. H.; Paramita Pal, P.; Moroder, L.; Huber, R.; Budisa, N. *ChemBioChem* **2004**, *5*, 720.
(12) Campos-Olivas, R.; Aziz, R.; Helms, G. L.; Evans, J. N. S.; Gronenborn, A. M. *FEBS Letters* **2002**, *517*, 55.
(13) Duewel, H.; Daub, E.; Robinson, V.; Honek, J. F. *Biochemistry* **1997**, *36*, 3404.
(14) Gakh, Y. G.; Gakh, A. A.; Gronenborn, A. M. *Magn Reson Chem* **2000**, *38*, 551.
(15) Long, G. J.; Rosen, J. F.; Schanne, F. A. *J Biol Chem* **1994**, *269*, 834.
(16) Minks, C.; Huber, R.; Moroder, L.; Budisa, N. *Biochemistry* **1999**, *38*, 10649.
(17) Rozovsky, S.; Jogl, G.; Tong, L.; McDermott, A. E. *Journal of Molecular Biology* **2001**, *310*, 271.
(18) Xiao, G.; Parsons, J. F.; Tesh, K.; Armstrong, R. N.; Gilliland, G. L. *Journal of Molecular Biology* **1998**, *281*, 323.
(19) Cornilescu, G.; Hadley, E. B.; Woll, M. G.; Markley, J. L.; Gellman, S. H.; Cornilescu, C. C. *Protein Science* **2007**, *16*, 14.
(20) Mortenson, D. E.; Satyshur, K. A.; Guzei, I. A.; Forest, K. T.; Gellman, S. H. *Journal of the American Chemical Society* **2012**, *134*, 2473.
(21) Woll, M. G.; Hadley, E. B.; Mecozzi, S.; Gellman, S. H. *J Am Chem Soc* **2006**, *128*, 15932.
(22) Molski, M. A.; Goodman, J. L.; Craig, C. J.; Meng, H.; Kumar, K.; Schepartz, A. *Journal of the American Chemical Society* **2010**, *132*, 3658.
(23) Bilgiçer, B.; Kumar, K. *Tetrahedron* **2002**, *58*, 4105.
(24) Bilgiçer, B.; Kumar, K. *Proc Nat Acad Sci USA* **2004**, *101*, 15324.
(25) Bilgiçer, B.; Xing, X.; Kumar, K. *J Am Chem Soc* **2001**, *123*, 11815.
(26) Buer, B. C.; de la Salud-Bea, R.; Al Hashimi, H. M.; Marsh, E. N. G. *Biochemistry* **2009**, *48*, 10810.
(27) Chiu, H.-P.; Kokona, B.; Fairman, R.; Cheng, R. P. *J Am Chem Soc* **2009**, *131*, 13192.
(28) Chiu, H.-P.; Suzuki, Y.; Gullickson, D.; Ahmad, R.; Kokona, B.; Fairman, R.; Cheng, R. P. *J Am Chem Soc* **2006**, *128*, 15556.
(29) Gottler, L. M.; de la Salud Bea, R.; Shelburne, C. E.; Ramamoorthy, A.; Marsh, E. N. G. *Biochemistry* **2008**, *47*, 9243.
(30) Gottler, L. M.; de la Salud-Bea, R.; Marsh, E. N. G. *Biochemistry* **2008**, *47*, 4484.
(31) Gottler, L. M.; Lee, H.-Y.; Shelburne, C. E.; Ramamoorthy, A.; Marsh, E. N. G. *ChemBioChem* **2008**, *9*, 370.
(32) Lee, H.-Y.; Lee, K.-H.; Al-Hashimi, H. M.; Marsh, E. N. G. *J Am Chem Soc* **2006**, *128*, 337.
(33) Lee, K.-H.; Lee, H.-Y.; Slutsky, M. M.; Anderson, J. T.; Marsh, E. N. G. *Biochemistry* **2004**, *43*, 16277.
(34) Meng, H.; Krishnaji, S. T.; Beinborn, M.; Kumar, K. *J Med Chem* **2008**, *51*, 7303.
(35) Meng, H.; Kumar, K. *J Am Chem Soc* **2007**, *129*, 15615.
(36) Tang, Y.; Tirrell, D. A. *J Am Chem Soc* **2001**, *123*, 11089.
(37) Anderson, J. T.; Toogood, P. L.; Marsh, E. N. G. *Org. Lett.* **2002**, *4*, 4281.
(38) Buer, B. C.; Chugh, J.; Al-Hashimi, H. M.; Marsh, E. N. G. *Biochemistry* **2010**, *49*, 5760.
(39) Lee, H.-Y.; Lee, K.-H.; Al-Hashimi, H. M.; Marsh, E. N. G. *J. Am. Chem. Soc.* **2006**, *128*, 337.

(40) West Jr, A. P.; Mecozzi, S.; Dougherty, D. A. *Journal of physical organic chemistry* **1997**, *10*, 347.
(41) Van Deventer, J. A.; Fisk, J. D.; Tirrell, D. A. *ChemBioChem* **2011**, *12*, 700.