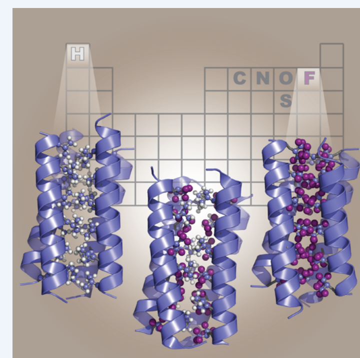
ACCOUNTS of chemical research
Article
pubs.acs.org/accounts

Fluorinated Proteins: From Design and Synthesis to Structure and Stability

E. Neil G. Marsh*

Departments of Chemistry and Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109, United States

CONSPECTUS: Fluorine is all but absent from biology; however, it has proved to be a remarkably useful element with which to modulate the activity of biological molecules and to study their mechanism of action. Our laboratory's interest in incorporating fluorine into proteins was stimulated by the unusual physicochemical properties exhibited by perfluorinated small molecules. These include extreme chemical inertness and thermal stability, properties that have made them valuable as nonstick coatings and fire retardants. Fluorocarbons also exhibit an unusual propensity to phase segregation. This phenomenon, which has been termed the "fluorous effect", has been effectively exploited in organic synthesis to purify compounds from reaction mixtures by extracting fluorocarbon-tagged molecules into fluorocarbon solvents. As biochemists, we were curious to explore whether the unusual physicochemical properties of perfluorocarbons could be engineered into proteins.



To do this, we developed a synthesis of a highly fluorinated amino acid, hexafluoroleucine, and designed a model 4-helix bundle protein, α_4H, in which the hydrophobic core was packed exclusively with leucine. We then investigated the effects of repacking the hydrophobic core of α_4H with various combinations of leucine and hexafluoroleucine. These initial studies demonstrated that fluorination is a general and effective strategy for enhancing the stability of proteins against chemical and thermal denaturation and proteolytic degradation.

We had originally envisaged that the "fluorous interactions", postulated from the self-segregating properties of fluorous solvents, might be used to mediate specific protein−protein interactions orthogonal to those of natural proteins. However, various lines of evidence indicate that no special, favorable fluorine−fluorine interactions occur in the core of the fluorinated α_4 protein. This makes it unlikely that fluorinated amino acids can be used to direct protein−protein interactions.

More recent detailed thermodynamic and structural studies in our laboratory have uncovered the basis for the remarkably general ability of fluorinated side chains to stabilize protein structure. Crystal structures of α_4H and its fluorinated analogues show that the fluorinated residues fit into the hydrophobic core with remarkably little perturbation to the structure. This is explained by the fact that fluorinated side chains, although larger, very closely preserve the shape of the hydrophobic amino acids they replace. Thus, an increase in buried hydrophobic surface area in the folded state is responsible for the additional thermodynamic stability of the fluorinated protein. Measurements of $\Delta G°$, $\Delta H°$, $\Delta S°$, and $\Delta C_p°$ for unfolding demonstrate that the "fluorous" stabilization of these protein arises from the hydrophobic effect in the same way that hydrophobic partitioning stabilizes natural proteins.

■ INTRODUCTION

Our research on fluorinated proteins and peptides was initially motivated by purely academic curiosity. We were intrigued by the idea of marrying the properties of synthetic molecules (fluorocarbons) with natural proteins.[1] Noting that perfluorination confers novel and useful properties upon simple hydrocarbon molecules, we wondered what would happen if fluorine, an abiological element, was introduced in large amounts into the structure of a complex macromolecule such as a protein?

A C−F bond is both stronger and longer than a C−H bond. Also, the high electronegativity of fluorine results in a bond that is both less polarizable and possesses a dipole moment that is the opposite of a C−H bond.[2] These properties combine to make perfluorinated molecules chemically inert and to endow them with unusual phase-segregating properties; thus, hexane and perfluorohexane, although both very hydrophobic solvents, are mutually immiscible. This self-segregating property has been termed the "fluorous effect".[3] These properties are most famously on display in nonstick polymer coatings such as polytetrafluoroethylene that repel both sticky, hydrophilic and greasy, hydrophobic molecules. Perfluorinated molecules have also found important uses as fire retardants, anesthetics, and, because they dissolve oxygen extremely well, blood substitutes.

The self-segregating properties of perfluorocarbons particularly intrigued us, as many protein−protein interfaces are formed by hydrophobic interactions. The substitution of F for H is sterically quite conservative, suggesting that extensively fluorinated analogues of hydrophobic amino acids might be accommodated within the hydrophobic core of a protein without disrupting the intricate packing of its side chains. Could one, therefore, exploit this "fluorous effect" to design a protein interface using fluorinated amino acids that would be

Received: March 21, 2014
Published: June 2, 2014

© 2014 American Chemical Society

dx.doi.org/10.1021/ar500125m | Acc. Chem. Res. 2014, 47, 2878−2886

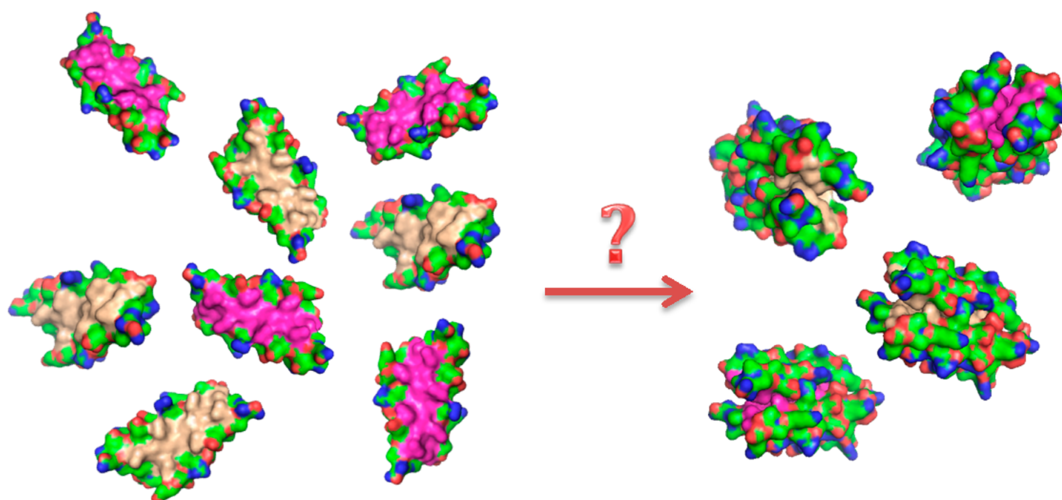


Figure 1. Can bio-orthogonal interactions potentially be designed using fluorous amino acids? In this hypothetical example, otherwise identical proteins containing complementary hydrocarbon (brown) or fluorocarbon (purple) interfaces form mutually exclusive dimers.

orthogonal to that of the natural protein on which it was based (as shown in Figure 1)?

In this Account, I focus on how our laboratory set about answering this question and our efforts to understand the effects of fluorination on protein structure and stability. We have not been the only laboratory to recognize the potential of fluorine in protein design; various other laboratories have published insightful contributions to the field of fluorinated proteins, including those of Tirrell and DeGrado,[4,5] Kumar,[6] Koksch,[7] Gellman,[8] Raines,[9] and Cheng.[10] One productive offshoot of our interest in fluorinated proteins has involved exploiting fluorinated amino acids as nonperturbing ^{19}F NMR reporters to investigate the interactions of membrane-active peptides[11−14] and, more recently, collaboration with A. Ramamoorthy to follow the aggregation pathways of amyloid-forming peptides.[15,16] These studies have been recently reviewed elsewhere, so they are not discussed here.

Synthesis of Hexafluoroleucine

To initiate this study, we needed both a highly fluorinated amino acid and a protein structure into which it could be incorporated. A logical choice of fluorinated amino acid was L-5,5,5,5′,5′,5-hexafluoroleucine (hFLeu). Leucine is the most abundant amino acid in proteins, and fluorination of its two methyl groups results in a side chain with a very high fluorine content, ∼70% by weight, which should provide a high potential for fluorous interactions. Somewhat surprisingly, no practical synthesis for hFLeu existed when we started our work. We therefore developed an efficient route to make hFLeu,[17] which is shown in Figure 2. The key step involves coupling together the carbon skeleton of hLeu from a protected serine analogue and hexafluoroacetone. This is accomplished by first converting the serine hydroxyl to iodide and then converting this to the zincate. This compound is reacted in situ with hexafluoroacetone using copper as a catalyst. It is worth noting that ketones are normally unreactive toward alkyl-zincates, but the presence of six electron-withdrawing fluorine atoms makes hexafluoroacetone an extremely good electrophile, so the reaction proceeds in excellent yield. The resulting 4-hydroxy-hFLeu derivative can then be converted to the phenyloxalate ester and deoxygenated under radical conditions using tributyl tin hydride as the reducing agent to yield the protected hFLeu.

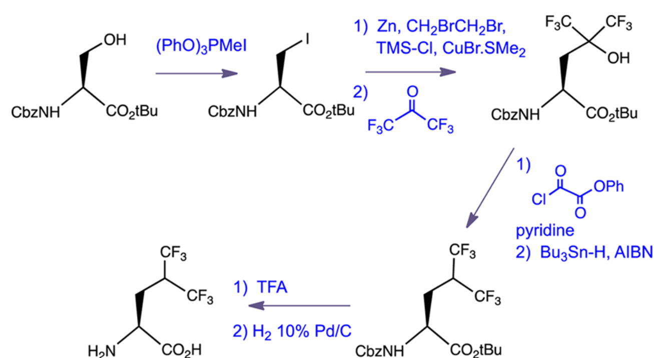


Figure 2. Scheme for the synthesis of L-5,5,5,5′,5′,5′-hexafluoroleucine from protected L-serine. For experimental details, see ref 17.

I note that in parallel with our work both Kumar and Cheng's laboratories also developed syntheses of hFLeu.[18,19]

α_4H: A Model Protein To Study the Effects of Fluorination

Highly fluorinated amino acids such as hFLeu are not efficiently incorporated into proteins biosynthetically, although Tirrell's group has made advances toward solving this problem,[20] so we focused on small de novo designed proteins that could be made by peptide synthesis. The antiparallel 4-helix bundle, or coiled-coil, motif proved to be ideal for this purpose. As shown in Figure 3, in these proteins four α-helices pack together in a regular pattern defined by a pseudorepetitive 7-residue sequence, known as a heptad repeat, that encompasses two turns of the α-helix. The most hydrophobic positions of the heptad are the a and d residues, which point toward the center of the bundle. The residues at the b and e positions form interhelix interactions with one neighboring helix, whereas those at c and g form interactions with the other neighboring helix; these interactions can be controlled by judicious choice of complementary charged residues. Lastly, the f position is solvent-exposed and can be any hydrophilic residue. This arrangement of α-helixes is present in many natural proteins, and well-characterized, de novo designed 4-helix bundles have been studied by DeGrado's laboratory.[21,22]

We designed a 27-residue peptide to form a tetrameric antiparallel 4-helix bundle in which six leucine residues at the a and d positions packed the hydrophobic core in six layers

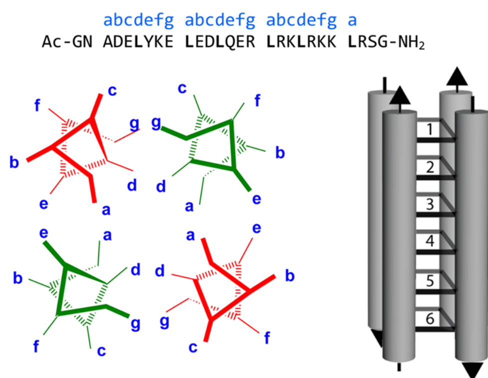


Figure 3. Design of antiparallel 4-helix bundle proteins. (Top) Sequence of α_4H. (Left) Helical wheel diagram showing interhelical interactions in the heptad repeat. (Right) Topology of α_4H; the Leu residues form six layers that pack the hydrophobic core. Adapted with permission from ref 30. Copyright 2012 National Academy of Sciences.

(Figure 3).[23] This parent 4-helix bundle protein we refer to as α_4H. Having established by techniques such CD spectroscopy, gel filtration, and analytical ultracentrifugation that this protein adopted its intended structure, we could then examine the effect of substituting one or more of the core Leu residues with hFLeu. Depending on which of the six a and d residues are substituted, different patterns of fluorination can be introduced into the protein core; for example, substituting the central a and d positions with hFLeu results in a protein in which the central two layers are packed with fluorinated residues, whereas substituting either all of the a positions or all of the d positions with hFLeu results in each layer being packed with two hFLeu side chains from opposite helices.

Effect of Fluorination on Protein Stability

In principle, there are 64 differently fluorinated proteins that could be produced on the basis of this simple, binary substitution strategy; in practice, during the course of our studies we have synthesized and characterized 10 of the possible Leu/hFLeu combinations.[23−25] Although a relatively small number, it is sufficient to draw some useful conclusions. First, fluorination appears to be generally nonperturbing, at least to the overall structure of the protein; all of the α_4H variants that we have synthesized remain extensively α-helical, as judged by CD spectroscopy, and tetrameric, as judged by analytical ultracentrifugation. Second, we found that fluorination invariably stabilizes the structure against unfolding by denaturants such as guanidinium chloride (GuHCl) (Figure 4). By fitting the unfolding of the proteins as a function of GuHCl concentration to appropriate equations (the details are described elsewhere[23]), one can determine ΔG_{unfold} for the proteins and thereby quantify the stabilizing effect of the fluorinated side chain. We found that increasing the number of fluorinated residues increases ΔG_{unfold} in a roughly monotonic fashion (Figure 4), as has been observed in other proteins in which fluorinated side chains have been introduced. The stabilizing effect of fluorine can be quite large when multiple hFLeu residues are incorporated; for example, replacing all Leu with hFLeu increases ΔG_{unfold} by ~14 kcal/mol.[25,26] This stability allows fluorinated proteins to resist unfolding by organic solvents and degradation by proteases.[27]

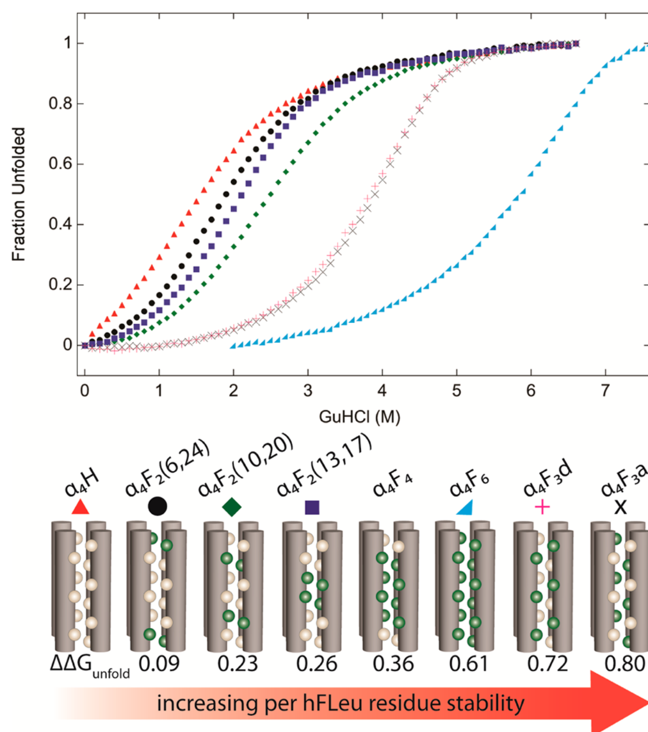


Figure 4. Fluorination increases protein stability of α_4 proteins. (Top) GuHCl-induced unfolding of α_4 proteins followed by CD at 222 nm; protein identities are listed in the center. (Bottom) Hydrophobic core packing arrangement of various α_4 proteins; Leu = tan spheres, hFLeu = green spheres. The per-residue increase in stability, $\Delta\Delta G_{unfold}$ (kcal/mol/hFLeu residue), is shown below each structure. Reproduced with permission from ref 11. Copyright 2009 Royal Society of Chemistry.

Searching for Elusive "Fluorous" Interactions

We were especially interested in whether particular patterns of fluorinated residues conferred greater stabilization than others. At the time, there was speculation that packing configurations that maximized fluorine−fluorine contacts would be especially stable because of favorable "fluorophilic" interactions. This idea was derived from the self-segregating properties of fluorocarbons, which have been effectively exploited and popularized in organic synthesis to purify molecules from reaction mixtures: the organic molecules can be purified by attaching a long perfluoroalkane "tail", which allows it to be selectively extracted from ordinary organic solvents by perfluorocarbon solvents.[28] It therefore seemed reasonable that "fluorous" interactions could be exploited in the design of proteins (Figure 1); after all, phase segregation, in the form of partitioning of hydrophobic side chains out of water, plays an important role in protein folding.

However, after calculating the stability of the fluorinated α_4H variants on a per-residue basis, we failed to find evidence that arrangements that maximized fluorine−fluorine contacts conferred additional thermodynamic stability.[24] For example, we compared the stabilities of two α_4 proteins, one in which two adjacent layers (layers 3 and 4) were packed with hFLeu and the other in which the layers were separated (layers 2 and 5). If stabilizing fluorous interactions were occurring, then we would expect the former to be more stable than the latter. In practice, we found that these proteins differed only marginally in their stabilities, by ~0.12 kcal/mol/residue, a difference that was within the limits of experimental error. Furthermore, we found that the most stabilizing pattern of fluorination involved placing hFLeu at either all a or all d positions, which results in

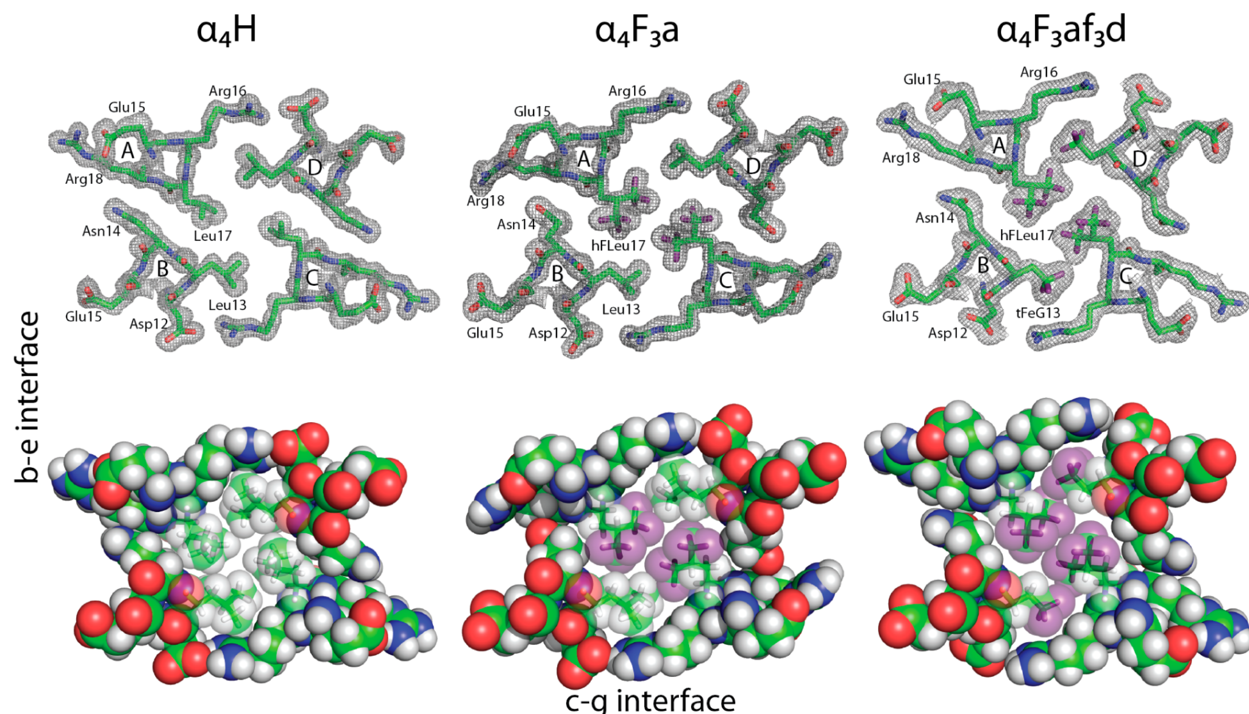


Figure 5. How fluorine is accommodated in the hydrophobic cores of a protein. (Top) Electron density maps for α_4H, α_4F_3a, and $\alpha_4F_3af_3$d ($2F_o - F_c$ maps; residues contoured at 1.0σ). (Bottom) Space-filling representations of the hydrophobic core illustrating how fluorination conserves the tight packing of side chains. Fluorine atoms are colored purple. Reproduced with permission from ref 30. Copyright 2012 National Academy of Sciences.

an interdigitating arrangement of Leu and hFLeu side chains throughout the hydrophobic core. This suggested that efficient packing of residues within the protein core likely trumps any stabilizing contribution that might arise from fluorous interactions.

As a further test of whether fluorous effects could operate in proteins, we examined whether α_4H and its fully fluorinated variant, α_4F_6, exhibited self-segregating properties akin to those of perfluorocarbon solvents.[27] For this, we exploited ^{19}F NMR. The chemical shift of the ^{19}F nucleus is extremely sensitive to its chemical environment, which allowed changes in the interaction of α_4F_6 with α_4H to be monitored. If the proteins were truly self-segregating, then the ^{19}F NMR spectrum of α_4F_6 should, of course, be unchanged by adding α_4H. However, progressive changes were observed in both the number and chemical shifts of most peaks in the spectrum of α_4F_6 as the ratio of α_4H to α_4F_6 was increased. Control experiments confirmed that the protein mixtures remained both α-helical and tetrameric. The most likely explanation, therefore, is that upon mixing α_4H and α_4F_6 exchange helices to from mixtures of 4-helix bundles, with α-helices contributed from both peptides.

From the above experiments, it seemed unlikely that fluorous effects, as envisaged by the phase-segregating properties of perfluorinated small molecules, could be exploited in the design of bio-orthogonal protein−protein interactions. To understand why this may be, it is necessary to consider the nature of the phase-segregating properties of perfluorocarbons effect in more detail.

Although the phase separation of fluorocarbon−hydrocarbon solvent mixtures are often ascribed to an attractive "fluorous" interaction between fluorocarbon molecules, this is not strictly correct. The phenomenon arises because the cohesive dispersion forces between two hydrocarbon molecules are greater than those between two fluorocarbon molecules, or between a fluorocarbon and a hydrocarbon molecule (this is because hydrocarbons are more polarizable than fluorocarbons), and thus fluorocarbons are excluded from the hydrocarbons. In fact, the origin of the fluorous effect was described quantitatively by eq 1 as long ago as 1948,[29] which relates the mutual solubility of a mixture of two nonpolar liquids to the difference in the solubility parameter, δ, which depends on the energy of vaporization, ΔE^V, and the molal volume of the pure liquid, V:

$$\delta = (\Delta E^V/V)^{1/2} \tag{1}$$

As the difference in δ increases, the heat of mixing becomes more unfavorable until the two liquids are no longer miscible. Fluorocarbons have low δ values because they have both lower boiling points and larger molal volumes than hydrocarbons.

Protein interfaces, of course, are highly structured and formed by specific interactions between side chains. In contrast, solvent−solute interactions are transient, nonspecific, and dynamic. So, it is perhaps in retrospect not surprising that the principles that underlie the segregating tendency of small fluorocarbon molecules cannot be applied to protein−protein interactions. We were admittedly somewhat disappointed by this realization, but this did raise the interesting question of why fluorocarbon side chains are so effective at stabilizing proteins, an effect that has been found for almost all proteins where this has been examined.[4−8,10]

To answer this question, we needed more detailed structural and thermodynamic information. Therefore, we set about determining crystal structures for a number of fluorinated α_4 variants to examine how fluorinated residues are accommodated within the hydrophobic core, and we undertook a detailed thermodynamic analysis of all of the proteins to

uncover the contributions that changes in enthalpy, entropy, and heat capacity make to ΔG_{unfold}.

Structures of Highly Fluorinated Proteins

Although non-canonical amino acids are increasingly being incorporated into proteins, there is very little structural data on how non-canonical side chains are accommodated into proteins, and before our studies, no structures for extensively fluorinated proteins had been determined. We were initially able to solve high-resolution crystal structures for both the parent protein α_4H (at 1.36 Å) and a variant, α_4F_3a, (at 1.54 Å) which contained hFLeu at all of the a positions so that 50% of the hydrophobic core was fluorocarbon.[30] α_4F_3a was one of the most stable fluorinated proteins that we had designed (each hFLeu contributes an additional ∼0.8 kcal/mol/residue to ΔG_{unfold}), so we were particularly interested in its structure. Comparison of the structures (Figure 5) showed that incorporating hFLeu, which has ∼30% larger side chain volume than Leu, caused remarkably little perturbation of the structure: the helices moved slightly further apart, displacing the backbone atoms of α_4F_3a by an rmsd of only 0.95 Å from the coordinates of α_4H. Except for one position, the "knobs into holes" packing arrangement that characterizes the hydrophobic core of coiled-coil structures was identical between the two proteins. We also carefully examined these structures for any evidence of fluorous interactions between residues (i.e., any preference for the fluorinated side chains to adopt conformations that maximized fluorine−fluorine contacts or minimized fluorine−hydrocarbon contacts), but we found none.

The structure of α_4F_3a also suggested why we had been unable, despite considerable effort, to obtain good crystals of α_4F_6, in which all of the core residues are fluorinated: whereas an alternating arrangement of Leu and hFLeu packs one layer of the core very efficiently, there did not seem to be room for four hFleu residues to pack a layer without disrupting other interhelical interactions. We therefore designed a protein that incorporated a smaller fluorinated residue trifluoroethylglycine (tFeG) at the d positions together with hFLeu at the a positions.[30] We initially predicted that this protein, $\alpha_4F_3$af$_3$d, would be more stable than α_4F_3a because of its high fluorine content and enhanced potential for supposedly favorable fluorine−fluorine contacts. As we had hoped, this protein gave well-diffracting crystals, and when we solved its structure, it revealed that the hFLeu and tFeG residues fit very efficiently into the core in an arrangement essentially identical to that of α_4H and α_4F_3a (Figure 5). However, to our surprise, $\alpha_4F_3$af$_3$d was actually no more stable than α_4H (ΔG_{unfold} = 17.8 and 18 kcal/mol, respectively). Clearly designing a well-folded, extensively fluorinated protein was not per se a recipe for stability!

Further analysis of the structures of α_4H, α_4F_3a, and $\alpha_4F_3$af$_3$d revealed that ΔG_{unfold} correlated very well with changes in buried hydrophobic surface area, as has been well-established for natural proteins.[30] Replacing Leu with hFLeu increases the buried hydrophobic surface area by ∼20 Å^2/residue; this would be expected to increase ΔG_{unfold} by ∼7 kcal/mol for α_4F_3a, which is in reasonable agreement with what we found. In contrast, replacing Leu with tFeG decreases the buried hydrophobic surface area by ∼20 Å^2/residue such that the buried surface areas in α_4H and $\alpha_4F_3$af$_3$d are almost identical. We should note here that although fluorocarbons are often described as being intrinsically more hydrophobic than hydrocarbons, the larger volume and surface area of fluorocarbons is often overlooked in such comparisons; when accounting for these factors, fluorocarbons and hydrocarbons exhibit similar hydrophobicities.

The Importance of Being Fluorine

If simple changes in hydrophobicity are responsible for the increased stability, then why are fluorinated amino acids are so effective at stabilizing protein structure? The reason, we contend, is that fluorination, while increasing size and hydrophobicity, closely preserves the shape of side chains, which is important for correctly packing the hydrophobic core. This allows the fluorinated residue to be introduced with minimal adjustment of the surrounding structure, as we saw in the structure of α_4F_3a. The conventional approach to increasing side chain hydrophobicity would be to add extra carbon atoms to the side chain, for example, by changing a valine to isoleucine. However, this also changes the side chain's shape, often significantly enough to perturb the protein's structure and compromise its stability and/or biological activity.

Support for this hypothesis comes from analysis of other structures of α_4H variants that we solved later.[31] Two of these proteins contained different arrangements of hFLeu, but arrangements that still result in each layer of the core being packed with alternating Leu−hFLeu residues. In one, α_4F_3d, hFLeu is present at all of the d positions; its structure is analogous to that of α_4F_3a. In the other, α_4F_3(6−13), the first two a and the first d positions of each helix contain hFLeu; this arrangement also results in each layer of the core being packed with two Leu and two hFLeu, although in this case, the hFLeu residues alternate between occupying the a and d positions. As shown in Figure 6, structural analysis of both of these proteins revealed that the hFLeu residues fit into the core almost seamlessly.

To investigate the shape-preserving properties of fluorinated residues, we designed a non-fluorinated α_4 variant with a core volume very close to that of the α_4F_3 series of proteins. This protein, $\alpha_4$tbA$_6$, incorporated *β-t*-butylalanine (tBAla) at all a and d positions; this has the effect of adding an extra methyl group at C_γ to each Leu of the original α_4H protein. We

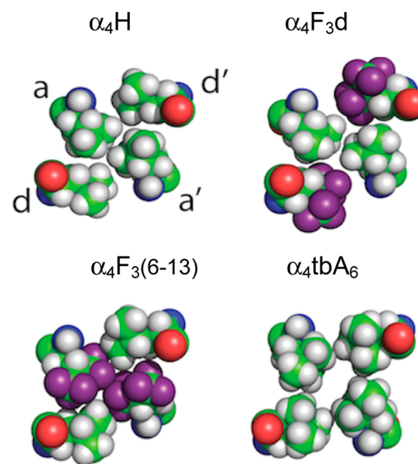


Figure 6. Packing of one layer of the hydrophobic core of α_4H. The larger hFLeu residue is accommodated at either the a (α_4F_3(6−13)) or d (α_4F_3d) positions with minimal disruption. In contrast, the similarly sized tbAla residue ($\alpha_4$tbA$_6$) does not fit well, causing a destabilizing void in the center of the core. Adapted with permission from ref 31. Copyright 2012 The Protein Society.

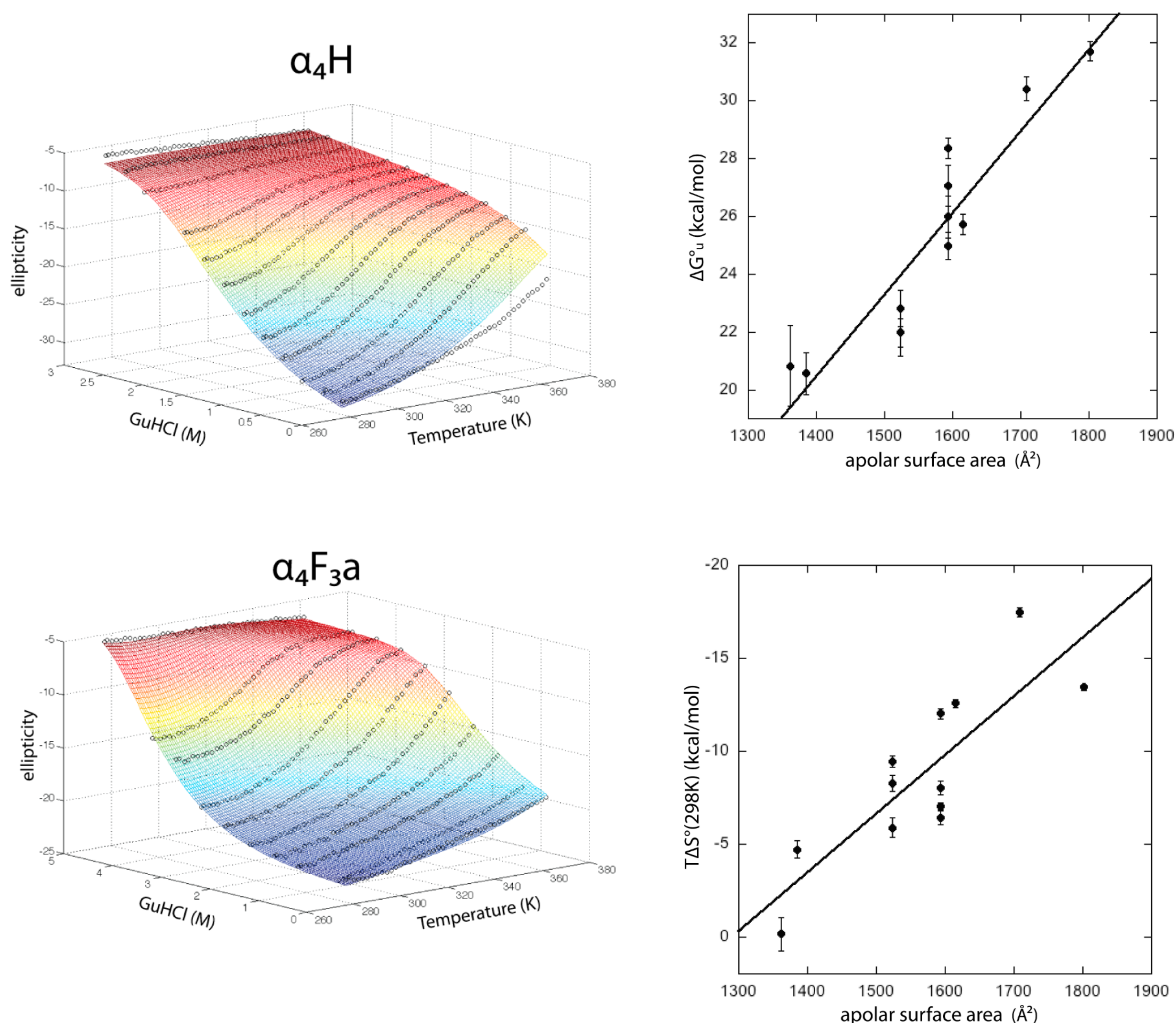


Figure 7. Thermodynamic analysis of fluorinated proteins unfolding. (Left) Unfolding surfaces for α_4H and α_4F_3a as a function of [GuHCl] and temperature. (Right) Correlation between buried apolar surface area and ΔG°_{unfold} (top) and buried apolar surface area and $T\Delta S^\circ$ (bottom). Entropy changes account for most of the additional stability of the fluorinated proteins. Adapted from ref 26. Copyright 2012 American Chemical Society.

considered this to be the least intrusive to increase side chain volume without fluorination. The volume and surface area of tBAla is intermediate between that of Leu and hFleu; thus, when incorporated at both the a and d positions, the core volume and buried surface area of $\alpha_4$tbA$_6$ would almost exactly match those of α_4F_3a, α_4F_3d, and α_4F_3(6−13).

When we solved the structure of $\alpha_4$tbA$_6$, we found that, although it resembled that of α_4H overall, the introduction of the tBAla group led to an expansion of the b−e interface to accommodate this bulkier side chain.[31] This caused the packing of the hydrophobic core to differ significantly from that of α_4H, α_4F_3d, and α_4F_3(6−13): in the central layers of the core, the *t*-butyl side chains abutted each other across the b−e and c−g interfaces, creating a cavity that ran through the center of the protein, as shown in Figure 6. This observation nicely illustrates the change in structure that even a small change in the shape of a hydrophobic residue can exert. Thus, it appears that fluorination is uniquely able to increase side chain hydrophobicity while maintaining side chain shape.

I note that our results contrast with an interesting investigation from Kumar's laboratory in which incorporation of hFLeu into a peptide designed to form a parallel, two-stranded coiled-coil did lead to the expected self-segregation of the fluorinated and non-fluorinated peptides.[32] However, in this case, introducing hFLeu also caused the peptide to convert from a two-stranded to a four-stranded coiled-coil, complicating interpretation of this result. It may be that the increased size of hFLeu relative to Leu is incompatible with the more sterically constraining two-stranded structure, leading to a four-stranded structure that is better able to accommodate the larger hFLeu side chain. In any event, the experiment shows that in some cases even fluorinated amino acids can cause significant structural changes.

Thermodynamics of α_4 Proteins Folding

To gain a better understanding of the physicochemical basis by which fluorination stabilizes protein structure, we undertook a detailed thermodynamic analysis of the unfolding of 12 α_4 proteins that we had characterized previously that incorporated different numbers and patterns of fluorination.[26] Whereas there had been many studies demonstrating, through different measurements, that fluorination stabilizes proteins against unfolding, there was very little data on how the enthalpic and

entropic contributions to protein folding change as a consequence of fluorination. Most studies had either focused on increases in T_m as a measure of protein stability or compared changes to the overall free energy of unfolding, $\Delta\Delta G^\circ_{unfold}$, resulting from fluorination.

We used a Van't Hoff analysis to determine the thermodynamic parameters, ΔH°, ΔS°, and ΔC_p°, associated with the unfolding of each protein. By studying the unfolding as a function of both GuHCl concentration and temperature, we were able to generate a 2D unfolding surface for each protein (Figure 7). The addition of GuHCl perturbs the unfolding temperature, allowing measurements to be made over a wider temperature range so that ΔC_p° can be reliably determined. This allowed the data to be globally and robustly fitted to the modified version of the Gibbs−Helmholtz equation (eq 2), which includes the effect of GuHCl concentration on ΔG°_{unfold}, to determine the enthalpy, ΔH°, and entropy, ΔS°, of protein folding together with the change in heat capacity, ΔC_p°, that accompanies this transition.

$$\Delta G^\circ(T,[\mathrm{GuHCl}]) = \Delta H^\circ - T\Delta S^\circ + \Delta C_p^\circ \times \left(T - T_0 + T\ln\frac{T_0}{T}\right) - m \times [\mathrm{GuHCl}] \quad (2)$$

For natural proteins, the hydrophobic effect is generally considered to be the major force driving protein folding, and the free energy of unfolding correlates with the change in buried hydrophobic surface area. With the detailed structural information that we had obtained for several of our fluorinated proteins, we could now reliably calculate the change in buried hydrophobic surface area and look for correlations with ΔG°_{unfold}, ΔH°, ΔS°, and ΔC_p°. We saw that the increase in ΔG°_{unfold} afforded by increasing the number of fluorinated residues correlated well with the increase in buried hydrophobic surface area (Figure 7), with an average value for $\Delta\Delta G^\circ_{unfold}$ of 28.3 cal/mol/Å^2. This value is close to the generally accepted energetic contribution of the hydrophobic effect to protein folding of ΔG°_{unfold} = 25−30 cal/mol/Å^2.[33,34]

Plots of $T\Delta S^\circ$ against buried hydrophobic surface area also correlated well (Figure 7). At 298 K, the area coefficient for the entropy change, $T\Delta S^\circ$, was −31.6 cal/mol/Å^2, implying that most of the increase in the free energy of unfolding is due to entropic effects. This is another hallmark of the hydrophobic effect in which the increase in entropy of folding is arises from the release of water molecules that form an ordered clatherate around the hydrophobic side chains in the unfolded state. In contrast, we found no correlation between ΔH° and buried hydrophobic surface area nor between ΔH° and the number of trifluoromethyl groups in the protein. A correlation might have been expected if electrostatic interactions arising from the permanent dipole moments of the trifluoromethyl groups and dipole moments of hydrogen-bonding moieties on the protein were contributing to the greater thermodynamic stability. Such interactions, sometimes described as the "polar hydrophobic effect", have been observed when certain fluorinated small molecules bind to proteins.[35]

The change in heat capacity, ΔC_p°, that accompanies the transition from folded to unfolded protein can be particularly informative. A positive value indicates that unfolding is dominated by solvation of hydrophobic side chains, whereas a negative ΔC_p° indicates that solvation of polar residues dominates unfolding.[36] For the α_4 series of proteins, the ΔC_p° values were all positive and ranged from 2−6 cal/mol/K/residue. These values are much lower than those typically measured for natural, well-folded proteins, for which per-residue ΔC_p° generally lies with the range of 10−15 cal/mol/K/residue.[36] The small value of is ΔC_p° is mainly responsible for the very high thermal stability of these proteins (i.e., ΔH° and ΔS° change only slowly as a function of temperature). Indeed, the T_m's of the most stable α_4 proteins are estimated to be above 220 °C, far more stable than most natural proteins!

The most likely explanation for the low ΔC_p° values is that the proteins retain some residual structure in the unfolded state.[37] Natural proteins from thermophilic organisms also tend to exhibit low per-residue ΔC_p° values, which would provide an evolutionary strategy to maintain the folded state at high temperature. Various lines of evidence point to such proteins retaining a compact structure in the unfolded state, possibly through nonspecific hydrophobic contacts, and thereby reducing the amount of solvent-exposed area that is buried upon folding. We do not know how structured the α_4 proteins are in the unfolded state: whereas their CD spectra indicate that they lack secondary structure in the unfolded state, this does not rule out a compact structure lacking regular secondary structure.

From this quite exhaustive analysis, it seems fairly certain that the hydrophobic effect is the major driving force contributing to the stability of fluorinated proteins, just as it drives the folding of most natural proteins folding. The α_4 proteins do exhibit some atypical properties, notably a low ΔC_p° and consequently a very high thermal stability, but this is probably the result of the de novo designed 4-helix coiled-coil scaffold rather than an effect of fluorination per se. Many simple de novo designed proteins have unusually high stabilities associated with their somewhat artificial core packing arrangements, in contrast to that of natural proteins, which are not generally selected for high thermal stability and exhibit more complex core packing arrangements.

■ CONCLUSIONS

The studies described above were inspired by academic curiosity, and the α_4 proteins were designed as simple models with no biological function. However, superstable fluorinated proteins and enzymes may have practical applications, for example, by retaining activity in the presence of organic solvents and/or high temperatures. For example, we have synthesized fluorinated versions of antimicrobial peptides, which kill bacteria by disrupting their membranes, by substituting hFLeu in place of isoleucine, leucine, or valine residues. For the α-helical antimicrobial peptide, MSI-78, fluorination increased its potency against some bacterial strains and protected the peptide against proteolytic degradation when bound to lipid bilayers,[38] whereas for the β-sheet antimicrobial peptide, protegrin, we demonstrated that fluorination could be used to modulate its oligomerization state upon insertion into model membranes.[39]

A variety of methods are now available that allow for the introduction of non-canonical amino acids into large proteins,[40,41] with fluorinated amino acids being well-represented among the diverse range of amino acids that have been incorporated. As these techniques become more commonplace and easier to apply, we expect that the use of non-canonical amino acids in protein engineering will continue to expand. We hope that our studies on model proteins, aimed at exploring the effects of fluorination on protein structure and stability, will prove to be useful for those seeking to incorporate

fluorinated amino acids into more complex natural proteins. Our observations suggest that fluorinated analogues of canonical hydrophobic amino acids can generally be inserted into proteins with very little perturbation to their structures. Although there will obviously be cases where even the substitution of fluorine for hydrogen is not tolerated, by carefully choosing the fluorinated amino acid with regard to the protein's structure, it should be possible to fluorinate most proteins while retaining biological function.

AUTHOR INFORMATION

Corresponding Author

*Tel.: 734 763 6096. Fax: 734 615 3790. E-mail: nmarsh@umich.edu.

Notes

The author declares no competing financial interest.

Biography

E. Neil G. Marsh received his undergraduate and doctoral degrees from Cambridge University. He was a postdoctoral fellow at Johns Hopkins University and then a Royal Society University Research Fellow in biochemistry at Cambridge University. He joined the faculty of the University of Michigan in 1995, where he is currently Professor of Chemistry and Biological Chemistry. He is a Fellow of the Royal Society of Chemistry.

ACKNOWLEDGMENTS

I have been fortunate to have a succession talented of postdoctoral scientists, graduate students, and undergraduates work on this project over the years. These include, in roughly chronological order, James Anderson, Morris Slutsky, Kyung-hoon Lee, Hwang-yeol Lee, Lindsey Gottler, Roberto de la Salud Bea, Ben Buer, Yuta Suzuki, and Ben Levin. This project has also benefitted from productive collaborations and discussions with my colleagues, A. Ramamoorthy, Vincent Pecoraro, Hashim Al Hashimi, and Jeanne Stuckey. Aspects of this work have been supported by grants from the National Science Foundation, CHE 0640934, and the Army Research Office, W911NF-11-1-0251.

REFERENCES

(1) Marsh, E. N. G. Towards the Non-Stick Egg: Designing Fluorous Proteins. *Chem. Biol.* **2000**, *7*, R153−R157.

(2) Sandford, G. Organofluorine Chemistry. *Philos. Trans. R. Soc., A* **2000**, *358*, 455−471.

(3) Horvath, I. T. Fluorous Biphase Chemistry. *Acc. Chem. Res.* **1998**, *31*, 641−650.

(4) Niemz, A.; Tirrell, D. A. Self-Association and Membrane-Binding Behavior of Melittins Containing Trifluoroleucine. *J. Am. Chem. Soc.* **2001**, *123*, 7407−7413.

(5) Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, W. A.; DeGrado, W. F.; Tirrell, D. A. Stabilization of Coiled-Coil Peptide Domains by Introduction of Trifluoroleucine. *Biochemistry* **2001**, *40*, 2790−2796.

(6) Bilgicer, B.; Kumar, K. De Novo Design of Defined Helical Bundles in Membrane Environments. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15324−15329.

(7) Jackel, C.; Salwiczek, M.; Koksch, B. Fluorine in a Native Protein Environment − How the Spatial Demand and Polarity of Fluoroalkyl Groups Affect Protein Folding. *Angew. Chem., Int. Ed.* **2006**, *45*, 4198−4203.

(8) Cornilescu, G.; Hadley, E. B.; Woll, M. G.; Markley, J. L.; Gellman, S. H.; Cornilescu, C. C. Solution Structure of a Small Protein Containing a Fluorinated Side Chain in the Core. *Protein Sci.* **2007**, *16*, 14−19.

(9) Hodges, J. A.; Raines, R. T. Stereoelectronic Effects on Collagen Stability: The Dichotomy of 4-Fluoroproline Diastereomers. *J. Am Chem Soc.* **2003**, *125*, 9262−9263.

(10) Chiu, H. P.; Suzuki, Y.; Gullickson, D.; Ahmad, R.; Kokona, B.; Fairman, R.; Cheng, R. P. Helix Propensity of Highly Fluorinated Amino Acids. *J. Am. Chem. Soc.* **2006**, *128*, 15556−15557.

(11) Marsh, E. N. G.; Buer, B. C.; Ramamoorthy, A. Fluorine−A New Element in the Design of Membrane-Active Peptides. *Mol. BioSyst.* **2009**, *5*, 1143−1147.

(12) Suzuki, Y.; Buer, B. C.; Al-Hashimi, H. M.; Marsh, E. N. G. Using Fluorine Nuclear Magnetic Resonance To Probe Changes in the Structure and Dynamics of Membrane-Active Peptides Interacting with Lipid Bilayers. *Biochemistry* **2011**, *50*, 5979−5987.

(13) Buer, B. C.; Chugh, J.; Al-Hashimi, H. M.; Marsh, E. N. G. Using Fluorine Nuclear Magnetic Resonance to Probe the Interaction of Membrane-Active Peptides with the Lipid Bilayer. *Biochemistry* **2010**, *49*, 5760−5765.

(14) Buer, B. C.; Levin, B. J.; Marsh, E. N. G. Perfluoro-*tert*-butyl-homoserine as a Sensitive ^{19}F NMR Reporter for Peptide−Membrane Interactions in Solution. *J. Pept. Sci.* **2013**, *19*, 308−314.

(15) Suzuki, Y.; Brender, J. R.; Hartman, K.; Ramamoorthy, A.; Marsh, E. N. G. Alternative Pathways of Human Islet Amyloid Polypeptide Aggregation Distinguished by ^{19}F Nuclear Magnetic Resonance-Detected Kinetics of Monomer Consumption. *Biochemistry* **2012**, *41*, 8154−8162.

(16) Suzuki, Y.; Brender, J. R.; Soper, M. T.; Krishnamoorthy, J.; Zhou, Y.; Ruotolo, B. T.; Kotov, N. A.; Ramamoorthy, A.; Marsh, E. N. G. Resolution of Oligomeric Species During the Aggregation of A-beta(1−40) Using ^{19}F NMR. *Biochemistry* **2013**, *52*, 1903−1912.

(17) Anderson, J. T.; Toogood, P. L.; Marsh, E. N. G. A Short and Efficient Synthesis of L-5,5,5,5′,5′,5′-Hexafluoroleucine from *N*-Cbz-L-Serine. *Org. Lett.* **2002**, *4*, 4281−4283.

(18) Xing, X. C.; Fichera, A.; Kumar, K. A Novel Synthesis of Enantiomerically Pure 5,5,5,5 ′,5 ′,5 ′-Hexafluoroleucine. *Org. Lett.* **2001**, *3*, 1285−1286.

(19) Chiu, H.-P.; Cheng, R. P. Chemoenzymatic Synthesis of (S)-Hexafluoroleucine and (S)-Tetrafluoroleucine. *Org. Lett.* **2007**, *9*, 5517−5520.

(20) Tang, Y.; Tirrell, D. A. Attenuation of the Editing Activity of the *Escherichia coli* Leucyl-tRNA Synthetase Allows Incorporation of Novel Amino Acids into Proteins in Vivo. *Biochemistry* **2002**, *41*, 10635−10645.

(21) Betz, S. F.; DeGrado, W. F. Controlling Topology and Native-Like Behavior of De Novo- Designed Peptides: Design and Characterization of Antiparallel Four-Stranded Coiled Coils. *Biochemistry* **1996**, *35*, 6955−6962.

(22) Hill, R. B.; Raleigh, D. P.; Lombardi, A.; Degrado, W. F. De Novo Design of Helical Bundles as Models for Understanding Protein Folding and Function. *Acc. Chem. Res.* **2000**, *33*, 745−754.

(23) Lee, K.-H.; Lee, H.-Y.; Slutsky, M. S.; Anderson, J. T.; Marsh, E. N. G. Fluorous Effect in Proteins: *De Novo* Design and Characterization of a Four-Alpha-Helix Bundle Protein Containing Hexafluoroleucine. *Biochemistry* **2004**, *43*, 16277−16284.

(24) Buer, B. C.; de la Salud-Bea, R.; Al Hashimi, H. M.; Marsh, E. N. G. Engineering Protein Stability and Specificity Using Fluorous Amino Acids: The Importance of Packing Effects. *Biochemistry* **2009**, *48*, 10810−10817.

(25) Lee, H. Y.; Lee, K. H.; Al-Hashimi, H. M.; Marsh, E. N. G. Modulating Protein Structure with Fluorous Amino Acids: Increased Stability and Native-Like Structure Conferred on a 4-Helix Bundle Protein by Hexafluoroleucine. *J. Am. Chem. Soc.* **2006**, *128*, 337−343.

(26) Buer, B. C.; Levin, B. J.; Marsh, E. N. G. Influence of Fluorination on the Thermodynamics of Protein Folding. *J. Am. Chem. Soc.* **2012**, *134*, 13027−13034.

(27) Gottler, L. M.; de la Salud-Bea, R.; Marsh, E. N. G. The Fluorous Effect in Proteins: Properties of α_4F_6, a 4-α-Helix Bundle Protein with a Fluorocarbon Core. *Biochemistry* **2008**, *47*, 4484−4490.

(28) Studer, A.; Hadida, S.; Ferritto, R.; Kim, S. Y.; Jeger, P.; Wipf, P.; Curran, D. P. Fluorous Synthesis: A Fluorous-Phase Strategy for Improving Separation Efficiency in Organic Synthesis. *Science* **1997**, *275*, 823−826.
(29) Scott, R. L. The Solubility of Fluorocarbons. *J. Am. Chem. Soc.* **1948**, *70*, 4090−4093.
(30) Buer, B. C.; Meagher, J. L.; Stuckey, J. A.; Marsh, E. N. G. Structural Basis for the Enhanced Stability of Highly Fluorinated Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 4810−4815.
(31) Buer, B. C.; Meagher, J. L.; Stuckey, J. A.; Marsh, E. N. G. Comparison of the Structures and Stabilities of Coiled-Coil Proteins Containing Hexafluoroleucine and *t*-Butylalanine Provides Insight into the Stabilizing Effects of Highly Fluorinated Amino Acid Side-Chains. *Protein Sci.* **2012**, *21*, 1705−1715.
(32) Bilgicer, B.; Xing, X. C.; Kumar, K. Programmed Self-Sorting of Coiled Coils with Leucine and Hexafluoroleucine Cores. *J. Am. Chem. Soc.* **2001**, *123*, 11815−11816.
(33) Eriksson, A. E.; Baase, W. A.; Zhang, X. J.; Heinz, D. W.; Blaber, M.; Baldwin, E. P.; Matthews, B. W. Response of a Protein-Structure to Cavity-Creating Mutations and Its Relation to the Hydrophobic Effect. *Science* **1992**, *255*, 178−183.
(34) Lins, L.; Brasseur, R. The Hydrophobic Effect in Protein-Folding. *FASEB J.* **1995**, *9*, 535−540.
(35) Biffinger, J. C.; Kim, H. W.; DiMagno, S. G. The Polar Hydrophobicity of Fluorinated Compounds. *ChemBioChem.* **2004**, *5*, 622−627.
(36) Prabhu, N. V.; Sharp, K. A. Heat Capacity in Proteins. *Annu. Rev. Phys. Chem.* **2005**, *56*, 521−548.
(37) Robic, S.; Guzman-Casado, M.; Sanchez-Ruiz, J. M.; Marqusee, S. Role of Residual Structure in the Unfolded State of a Thermophilic Protein. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11345−11349.
(38) Gottler, L. M.; Lee, H.-Y.; Shelburne, C. E.; Ramamoorthy, A.; Marsh, E. N. G. Using Fluorous Amino Acids To Modulate the Biological Activity of an Antimicrobial Peptide. *ChemBioChem.* **2008**, *9*, 370−373.
(39) Gottler, L. M.; de la Salud Bea, R.; Shelburne, C. E.; Ramamoorthy, A.; Marsh, E. N. Using Fluorous Amino Acids To Probe the Effects of Changing Hydrophobicity on the Physical and Biological Properties of the Beta-Hairpin Antimicrobial Peptide Protegrin-1. *Biochemistry* **2008**, *47*, 9243−9250.
(40) Link, A. J.; Mock, M. L.; Tirrell, D. A. Non-Canonical Amino Acids in Protein Engineering. *Curr. Opin. Biotechnol.* **2003**, *14*, 603−609.
(41) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z. W.; Schultz, P. G. An Expanded Eukaryotic Genetic Code. *Science* **2003**, *301*, 964−967.