ACCOUNTS
of chemical research

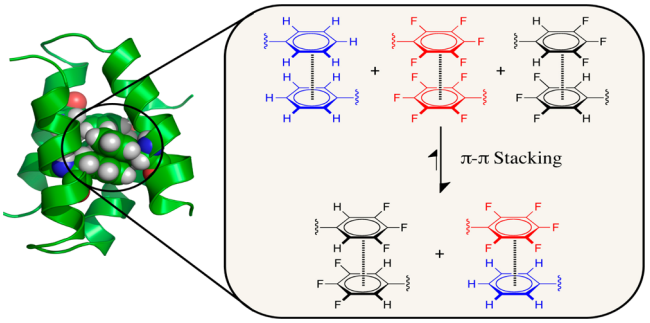
Exploring and Exploiting Polar−π Interactions with Fluorinated Aromatic Amino Acids

CHRISTOPHER J. PACE AND JIANMIN GAO*

Department of Chemistry, Merkert Chemistry Center, Boston College, 2609 Beacon Street, Chestnut Hill, Massachusetts 02467, United States

RECEIVED ON MARCH 15, 2012

CONSPECTUS



Fluorination has become an increasingly attractive strategy in protein engineering for both basic research and biomedical applications. Thus researchers would like to understand the consequences of fluorination to the structure, stability, and function of target proteins. Although a substantial amount of work has focused on understanding the properties of fluorinated aliphatic amino acids, much less is known about fluorinated aromatic residues. In addition, polar−π interactions, often referred to as aromatic interactions, may play a significant role in protein folding and protein−protein interactions. Fluorination of aromatic residues presents an ideal strategy for probing polar−π interactions in proteins.

This Account summarizes the recent studies of the incorporation of fluorinated aromatic amino acids into proteins. Herein we discuss the effects of fluorinating aromatic residues and rationalize them in the context of polar−π interactions. The results strongly support the proposal that polar−π interactions are energetically significant to protein folding and function. For example, an edge−face interaction of a pair of phenylalanines contributes as much as −1 kcal/mol to protein stability, while cation−π interactions can be much stronger. Furthermore, this new knowledge provides guidelines for protein engineering with fluorination. Importantly, incorporating perfluorinated aromatic residues into proteins enables novel mechanisms of molecular recognition that do not exist in native proteins, such as arene-perfluoroarene stacking. Such novel mechanisms can be used for programming protein folding specificity and engineering peptide-based materials.

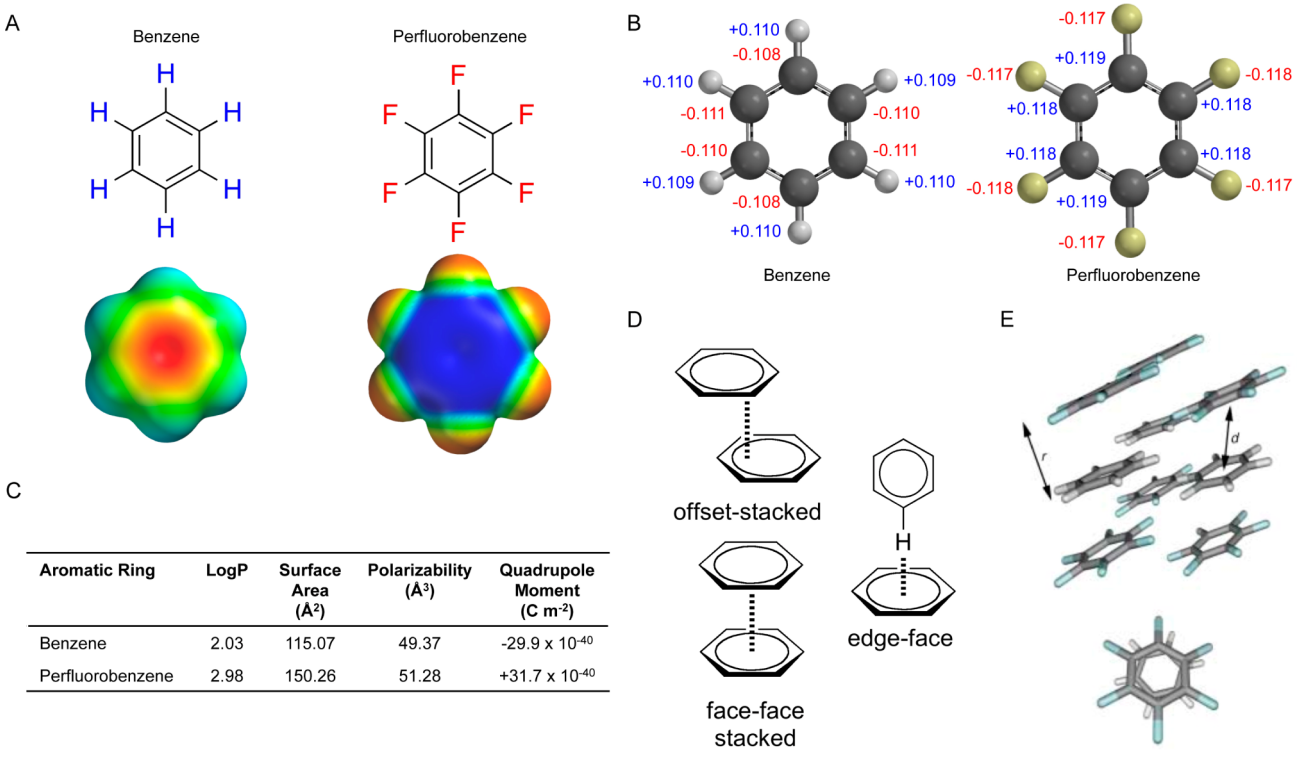
Introduction: Fluorinating Proteins

Fluorination has become an increasingly popular strategy in protein biochemistry.[1−3] Site-specific incorporation of fluorinated residues can now be accomplished through chemical synthesis and semisynthesis of proteins,[4] as well as molecular biology techniques.[5,6] Fluorinated residues serve as ideal probes of protein structure and function for several reasons: (1) Although a C−F bond (1.34 Å for an sp^2 carbon) is ∼20% longer than a C−H bond (1.09 Å), fluorination has been shown to be well tolerated by a variety of proteins without introducing much steric perturbation to the parent structure.[7,8] (2) Fluorination usually increases the hydrophobicity of the amino acid and thus favors protein folding and stability.[3] (3) Fluorine is essentially nonexistent in biology, so fluorinated residues can be tracked in complex biological systems with zero background. The stable and naturally abundant isotope of fluorine, ^{19}F, is almost as sensitive (83%) as a proton in NMR spectroscopy. It can therefore be tracked with NMR even in living cells.[9] (4) Fluorination affords novel mechanisms of molecular recognition that can be used for programming protein folding and assembly. This will be further discussed in later sections. In order to realize the full potential of fluorination in protein biochemistry, it is necessary to understand the structural and functional perturbations caused by introducing fluorinated amino acids.

Fluorination and Polar−π Interactions in Proteins

Earlier work on fluorinated proteins largely concentrated on using fluorinated aliphatic amino acids to create

Published on the Web 10/24/2012 www.pubs.acs.org/accounts
10.1021/ar300086n



C

Aromatic Ring	LogP	Surface Area (Å²)	Polarizability (Å³)	Quadrupole Moment (C m⁻²)
Benzene	2.03	115.07	49.37	-29.9 x 10^{-40}
Perfluorobenzene	2.98	150.26	51.28	+31.7 x 10^{-40}

FIGURE 1. Influence of fluorination on an aromatic ring. (A) Electrostatic potential maps showing the reversed electron distribution upon perfluorination of benzene. (B) Partial charges of benzene and perfluorobenzene calculated by using the PM3 semi empirical model in Spartan '08 (Wave function, Inc.). (C) Comparison of the physical properties of benzene and perfluorobenzene. (D) Illustration of three packing modes of benzene dimers. The electrostatic potential of benzene favors edge-face and offset-stacked geometries, while disfavoring the face−face stacked geometry. (E) Face−face stacking of benzene and perfluorobenzene revealed by the cocrystal structure (image adapted from ref 16).

Teflon-like proteins.[10−13] This line of research was initially inspired by the "fluorous effect", which refers to the phase separating behavior of fluorocarbons from both aqueous and hydrocarbon solvents. The fluorous effect has been successfully utilized for product separation in organic synthesis. By fluorinating all the core residues of coiled-coil model proteins, Kumar and co-workers demonstrated self-sorting behavior when the fluorinated and nonfluorinated peptides were mixed.[14,15] *The fluorous effect, however, does not apply to fluorinated aromatics.* For example, benzene and perfluorobenzene have similar solubility in hexanes and they mix readily with each other with a favorable enthalpy.[16] The cocrystal structure of benzene and perfluorobenzene revealed an alternating arrangement with benzene and perfluorobenzene in a face−face stacked geometry (Figure 1).[17]

The distinct behavior of fluorinated aromatics from aliphatics can be attributed at least partially to the electrostatic interactions of aromatic rings.[18,19] For example, an aromatic structure such as benzene displays an uneven electron distribution even though it does not have a net dipole (Figure 1). Because an sp^2 carbon is more electronegative than hydrogen, benzene exhibits a negative potential on the π face and positive potential around the periphery. This electron distribution, referred to as a quadrupole, allows aromatic moieties to engage in electrostatic interactions with a charge (e.g., cation−π interaction), a partial charge, a dipole, or another quadrupole. Such polar−π interactions explain why benzene molecules pack in the edge-face (T-shaped) or offset-stacked geometries in crystal structures. In the case of benzene-perfluorobenzene stacking, perfluorination reverses the quadrupole of benzene to give a negative potential on the periphery and consequently a positive potential on the aromatic ring (Figure 1). Recent theoretical studies have suggested that the shift of electron density upon fluorination is more localized to the carbon skeleton than the π-clouds.[20,21] Nevertheless, the charge distribution of benzene and perfluorobenzene complement each other, driving them to associate in the face-to-face fashion.

A structural survey of the Protein Data Bank revealed a large abundance of potentially important polar−π interactions. In particular, there appears to be one energetically significant cation−π contact every 77 residues, where aromatic side chains (Trp, Tyr, Phe) pack against a lysine or an arginine.[22] On the other hand, ~60% of aromatic side chains

are found to form close contact with another aromatic residue, suggesting the significance of π−π interactions.[23] Although polar−π interactions have been a subject of investigation for several decades, their significance had not been directly examined in protein systems until relatively recently.[24] Fluorination of aromatic residues presents an ideal way of probing polar−π interactions due to the small steric yet large electronic perturbations (Figure 1). Pioneering work by Dougherty and co-workers has demonstrated the power of fluorination in investigating cation−π interactions.[25] This account will concentrate on site-specific fluorination of aromatic residues in the context of π−π interactions. Biophysical analysis of these fluorinated proteins has generated further insights into the role of polar−π interactions in protein folding and stability. Of equal importance, it has provided guidelines for using fluorination as a tool in protein engineering.

Fluorinating Aromatic Residues Involved in π−π Interactions

Experimental and theoretical studies have shown that the edge-face and offset-stacked geometries (Figure 1) are preferred between two interacting benzene molecules.[16] In contrast, face−face stacking between two benzene molecules is somewhat unfavorable and usually avoided. This again can be explained by the electrostatic potential of benzene: positive edge and negative center. Consistent with the structural preference of benzene dimers, aromatic side chains of proteins predominantly interact in the edge-face or offset-stacked fashion, while the face−face stacked geometry is rarely seen.[23,26]

a. Edge−Face or Offset−Stacked Aromatic Pairs. It is important to point out that aromatic pairs in protein structures display a continuous distribution of cross angles from 0 to 90°, instead of being strictly parallel (offset-stacked) or perpendicular (edge−face) as seen in benzene crystal structures. Therefore, we discuss these geometries of aromatic interactions together. In the early 2000s, Waters and coworkers examined the role of π−π interactions in stabilizing the secondary structures of proteins.[27] By incorporating aromatic pairs into a host Ala-Lys peptide, Waters et al. show that a phenylalanine-phenylalanine pair at *i* and *i*+4 positions (in comparison to *i* and *i*+5) stabilizes the α-helix by −0.27 to −0.8 kcal/mol. Interestingly, a phenylalanine-perfluorophenylalanine (Z) pair stabilizes the α-helix to a less extent (−0.27 to −0.55 kcal/mol). Given the side chain of Z is more hydrophobic than that of F, the F−F pair being more stabilizing than F−Z indicates that the F−F interaction is primarily electrostatic instead of hydrophobic. Molecular modeling shows that the phenylalanine side chains can interact with each other either in the edge−face or offset−stacked geometries. The face−face-stacked geometry is not feasible with aromatic pairs at the *i* and *i*+4 positions of an α-helix. It is worth noting that perfluorophenylalanine displays lower helical propensity than the native counterpart.[27,28] However, the physical origin of this effect remains poorly understood. Fortunately, by adopting the double mutant cycle analysis,[29] the enigmatic reduction of helical propensity by fluorination can be corrected, thus causing no complications to the assessment of pairwise interaction between aromatic residues. Similar to the stabilization of α-helices, the Waters group found that an F−F pair in the edge−face geometry could stabilize a β-hairpin by −0.55 kcal/mol.[30]

To probe the energetic significance of aromatic packing inside a protein core, the Gellman group,[31] as well as our own,[32] incorporated perfluorophenylalanine into the core of the villin headpiece subdomain (HP35), one of smallest globular domains with a stable structure.[33] This 35-residue polypeptide folds into three helices, which wrap around each other to give an aromatic core (Figure 2A).[34,35] The three core Phe residues pack against each other with the edge−face geometry: the edge of F10 contacts the π-face of F6 with a dihedral angle of 61°, while the edge of F6 packs against the face of F17 with a cross angle of 92°. Interestingly, individual F-to-Z mutations at these three positions have distinct effects on the stability of HP35: while F10Z stabilizes the protein structure, both F6Z and F17Z are actually destabilizing. Because all three Phe residues in HP35 are largely buried, the hydrophobicity difference for the three F-to-Z mutations is expected to be minimal. In addition, since all three Phe residues are in helical conformation, helical propensity is also irrelevant in the stabilities of these mutants. We hypothesized that perfluorination of a Phe side chain disturbed the edge−face interaction between the aromatic residues, which resulted in the position-dependent stabilization by an F-to-Z mutation. This hypothesis was tested by using a set of tetrafluorinated phenylalanines that we named as Z_x, where *x* could be o, m or p and denotes the position of the remaining hydrogen (relative to the main chain) on the phenyl ring (Figure 2B). The tetrafluorophenylalanines were obtained in high yields and enantiopurity through a facile four-step synthesis using the Seebach or Schöllkopf chiral auxiliary.[32,43] Our results show that F10Z_o is more stable than F10Z by −1.1 kcal/mol, presumably because the F10−F6 edge−face interaction is

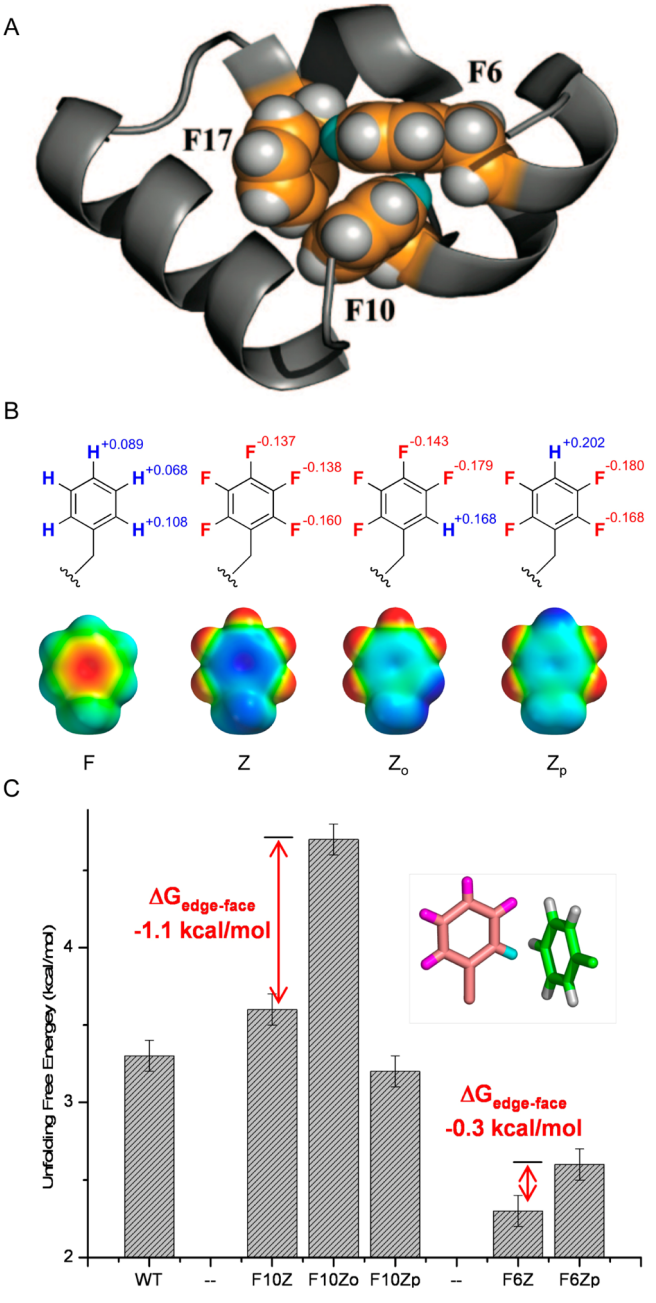


FIGURE 2. Fluorinating aromatic residues involved in edge-face interactions. (A) Cartoon representation of HP35 structure (PDB: 1YRF) showing the aromatic residues in the hydrophobic core (F6, F10, F17). Hydrogen atoms thought to be engaging in edge-face interactions are colored cyan, with all others in white. (B) Structure of the Phe side chain and the fluorinated mutants Z, Z_o, and Z_p with calculated partial charges and electrostatic potential maps (generated using Spartan). (C) Graph summarizing the influence of edge–face interactions on the unfolding free energy of HP35 variants. In comparison to F10Z, the F10Z_o mutation yields a free energy gain of −1.1 kcal/mol through retaining the edge–face contact with the π cloud of F6.

retained by the F10Z_o (Figure 2C) mutant but not by F10Z. Similarly, by retaining the F6–F17 edge–face interaction, F6Z_p affords an improved stability in comparison to F6Z, albeit by a smaller margin (−0.3 kcal/mol). The different magnitude of stabilization is perhaps not surprising given that the F6 position is more solvent exposed than F10. In other words, these results suggest that edge-face aromatic interactions can be stronger when located in a protein core than more solvent exposed positions.[27,30] They also showcase the potential of using the Z_x series to stabilize protein structures, as these novel amino acids may simultaneously recruit hydrophobic and aromatic edge–face interactions to favor protein folding. Importantly, heavy fluorination also renders the remaining protons on the Z_x side chains to be more acidic, as indicated by the increased partial charges on the aromatic hydrogens (Figure 2B). These more "activated" hydrogens are expected to strengthen the edge-face interaction between the aromatic rings.

b. Face–Face Stacked Aromatic Pairs. Previous studies using small molecule model systems have convincingly shown that face–face stacking of benzene molecules is unfavorable.[18,36] This is presumably why the face–face stacked geometry is scarce in protein structures. However, aromatic moieties of reversed quadrupoles prefer to interact with each other in face–face stacked geometries. A well-known example is the stacking of benzene and perfluorobenzene (Figure 1), which displays nearly identical magnitudes in their quadrupole moments, but with opposite signs.[37] The phenyl–perfluorophenyl pair has been utilized as a supramolecular synthon to program molecular packing in crystals,[38,39] as well as to control the stereoselectivity of polymerization reactions.[40,41] Although face–face stacked aromatic pairs are rare in native protein structures, this mode of association may be used to program protein–protein interactions and self-assembled peptide materials.

To quantitatively assess aromatic stacking in proteins, our group has been utilizing the model system α_2D, a de novo designed polypeptide that folds into a dimeric helix bundle (Figure 3A).[42] Similar to HP35, the core of α_2D primarily consists of aromatic residues. In the folded α_2D dimer, the F10 side chains of each monomer stack with those of F29 from the other in the face-to-face fashion; the potential repulsion between the stacked Phe side chains is compensated by the folding free energy of α_2D. This unique structure makes α_2D an ideal system to probe face–face stacking of aromatic residues. For ease of discussion, we named each α_2D variant based on residues 10 and 29. For instance, the wild type α_2D has a phenylalanine at both positions; we named it as (***F***, ***F***). To probe the physical origin of aromatic stacking, we have incorporated a series of fluorinated Phe analogues into the α_2D core structure

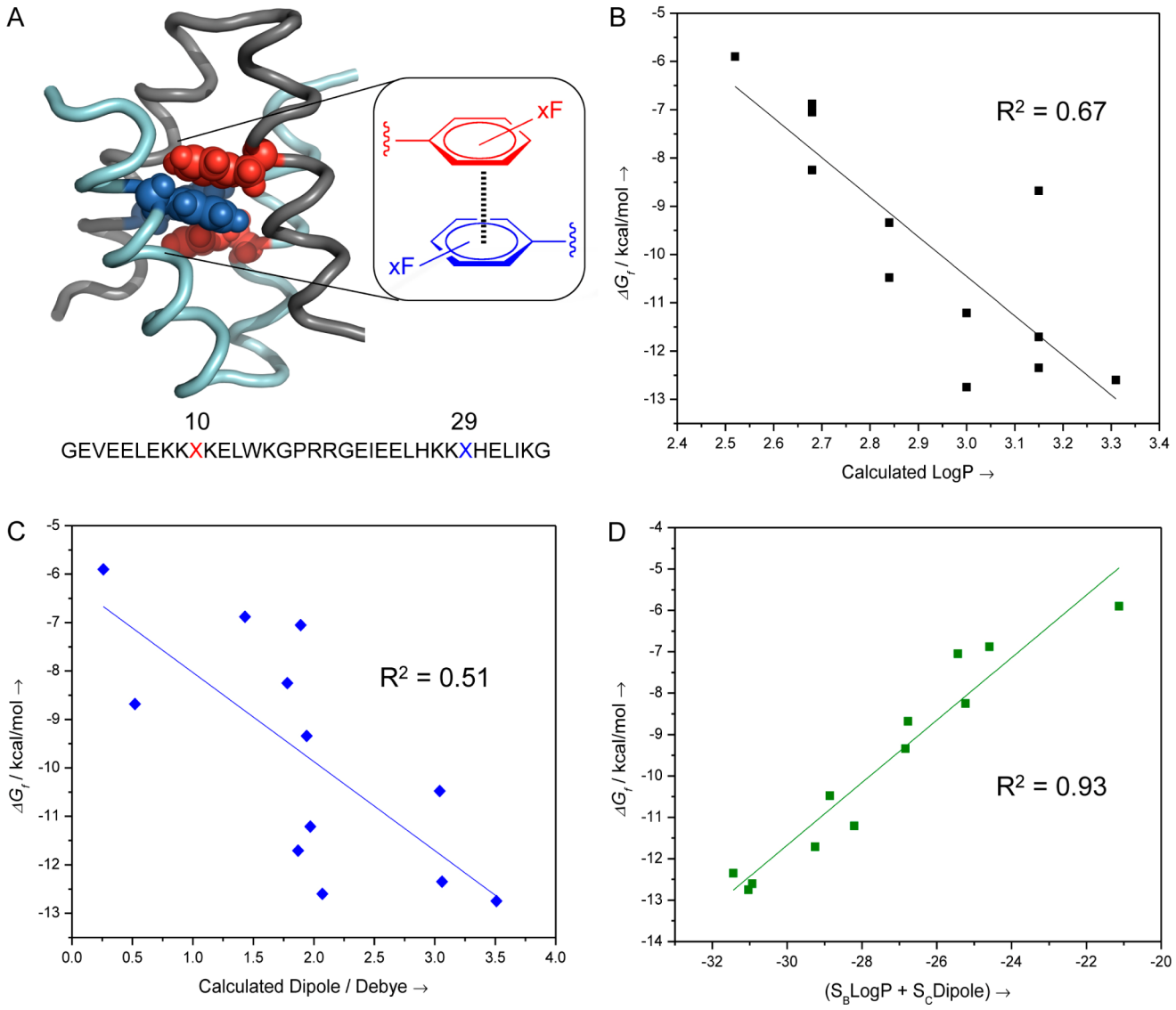


FIGURE 3. Investigating the energetics of face−face stacking using fluorinated Phe analogues. (A) Cartoon representation of the α_2D dimer (PDB 1QP6) highlighting the two face−face stacked pairs in the hydrophobic core (inset). The monomers are shown in gray and cyan. Residues 10 and 29 are highlighted in red and blue, respectively, in the structure and sequence. (B) Plot of all α_2D double mutants showing correlation between ΔG_f and hydrophobicity (calculated log *P*) of the core residues. (C) Plot of all α_2D double mutants showing correlation between ΔG_f and net dipole moment. (D) Plot of all α_2D double mutants showing excellent correlation between ΔG_f and a synthetic parameter that takes into account both hydrophobicity and dipole moment.

(Figure 3A) at positions 10 and/or 29.[43] Our experiments show that, despite the size increase, fluorination of the core Phe residues is well accommodated by the α_2D fold. In fact, fluorination of all core Phe residues dramatically improved the α_2D structural stability (by ~ -6 kcal/mol), with the melting temperatures increasing from 29 °C for (***F***, ***F***) to 80 °C for (***Z***, ***Z***). This is perhaps not surprising considering the enhanced hydrophobicity and the small steric perturbation caused by fluorination. An F-to-Z mutation appears to contribute ~ -1.5 kcal/mol to protein stability by the enhanced hydrophobicity alone. The hydrophobic contribution of an F-to-Z mutation has been further validated in other model protein systems (Pace and Gao, unpublished results).

Varying the degree and pattern of Phe fluorination affords a wide distribution in the physical properties of the side chains, including hydrophobicity, dipole and quadrupole moments. A double mutant of α_2D, (***X***, ***X***), would have two stacked X−X pairs in the protein core, where X denotes Phe or its fluorinated analogues (Figure 3A). Thermodynamic analysis of the (***X***, ***X***) series allowed us to correlate the folding free energy of the α_2D variants with the physical properties of each Phe analogue. Interestingly, the

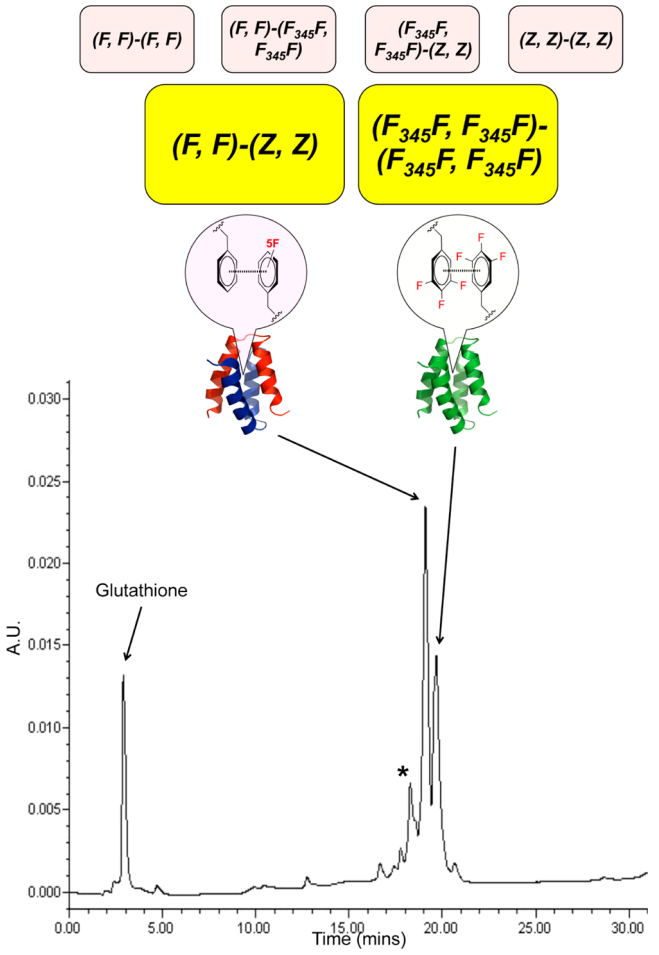


FIGURE 4. Self-sorting of the α_2D variants induced by aromatic stacking. Shown is an LC trace of a disulfide cross-linking experiment with a three-component mixture of (***F***, ***F***), (***Z***, ***Z***), and (***F*$_{345}$*F***, ***F*$_{345}$*F***). Peptides are mixed as reduced monomer and allowed to cross-link in an oxidized glutathione buffer. The six possible dimers are shown in boxes, and the two primary products, the (***F***, ***F***)-(***Z***, ***Z***) heterodimer and the (***F*$_{345}$*F***, ***F*$_{345}$*F***)-(***F*$_{345}$*F***, ***F*$_{345}$*F***) homodimer are highlighted in yellow. The third observable peak, denoted by an asterisk (*), is composed of monomer−glutathione adducts for (***Z***, ***Z***) and (***F*$_{345}$*F***, ***F*$_{345}$*F***), which are side products of the cross-linking reaction.

homodimer stability displays a positive but moderate correlation ($R^2 = 0.67$) with the log *P* value of the Phe analogues, indicating that factors other than hydrophobicity must play an important role in stabilizing the α_2D fold (Figure 3B). Indeed, a positive correlation ($R^2 = 0.51$) was also observed between the protein stability and the dipole moment of the side chain analogs, indicating that dipole−dipole interactions contribute significantly to α_2D folding (Figure 3C). Consequently, (***F*$_{345}$*F***, ***F*$_{345}$*F***) exhibits the most favorable folding free energy possibly because $F_{345}F$ has the largest dipole (3.5 D; see structure in Figure 4) among this group of Phe analogues. More interestingly, a synthetic parameter combining hydrophobicity and dipole moment yielded an excellent correlation with the stability of the α_2D variants ($R^2 = 0.93$, Figure 3D). On the other hand, the experiments with the single mutant (***F***, ***X***) series show that dipole−induced dipole interactions can be quite significant as well. Among the single mutant series, (***F***, ***Z*$_o$**) gives the most stable homodimer; it is even more stable than (***F***, ***Z***), which carries a more hydrophobic core residue. This is presumably because (***F***, ***Z*$_o$**) benefits from both hydrophobicity of fluorination and the larger dipole of Z_o (3.1 D) in comparison to Z (2.1 D). However, as previously noted, these highly fluorinated aromatic rings (e.g., $F_{345}F$ and Z_o) exhibit more acidic protons (Figure 2B), which could potentially serve as stronger hydrogen bond donors. The structural and energetic significance of such hydrogen bonds remains to be examined in greater detail. Nevertheless, the series of α_2D mutants presented here highlight the inherent complexities of aromatic interactions.

We further evaluated the dimerization specificity of the α_2D variants, to which the hydrophobicity of the core residues (Phe or its fluorinated analogues) is no longer a significant contributor because they are completely buried in both homo and heterodimers of α_2D. Rather, the electrostatic interactions of aromatic stacking are expected to dictate the dimerization preference of the α_2D variants. For example, (***F***, ***F***) and (***Z***, ***Z***) both fold into stable homodimers, yet upon mixing they exchange completely to give heterodimers.[44] Since the homodimer−heterodimer equilibrium is not driven by the hydrophobic effect, the preference of dimerization is presumably a result of the favorable face−face stacking between the quadrupoles of F and Z. Using double mutant cycle analysis, a useful strategy for estimating the energetic scale of pairwise interactions,[29] we estimated the F−Z cross affinity to be ~ -1.0 kcal/mol, which is comparable to the strength of weak hydrogen bonds. Furthermore, by capitalizing on both quadrupole stacking (***F***−***Z***) and dipole−dipole interactions (***F*$_{345}$*F***−***F*$_{345}$*F***), we demonstrate that self-sorting behavior can be afforded by stacked aromatic pairs (Figure 4).[43] Specifically, equilibration of a mixture of (***F***, ***F***), (***Z***, ***Z***) and (***F*$_{345}$*F***, ***F*$_{345}$*F***) gives primarily two dimeric species: the heterodimer (***F***, ***F***)−(***Z***, ***Z***) and the homodimer of (***F*$_{345}$*F***, ***F*$_{345}$*F***). These two dimers emerge exclusively out of six possible dimers presumably due to the synergy of the two favorable stacking pairs (F−Z and $F_{345}F$−$F_{345}F$).

Exploiting Polar−π Interactions in Peptide Design

The arene-fluoroarene stacking interaction is beginning to be exploited in the design of peptide-based materials. For

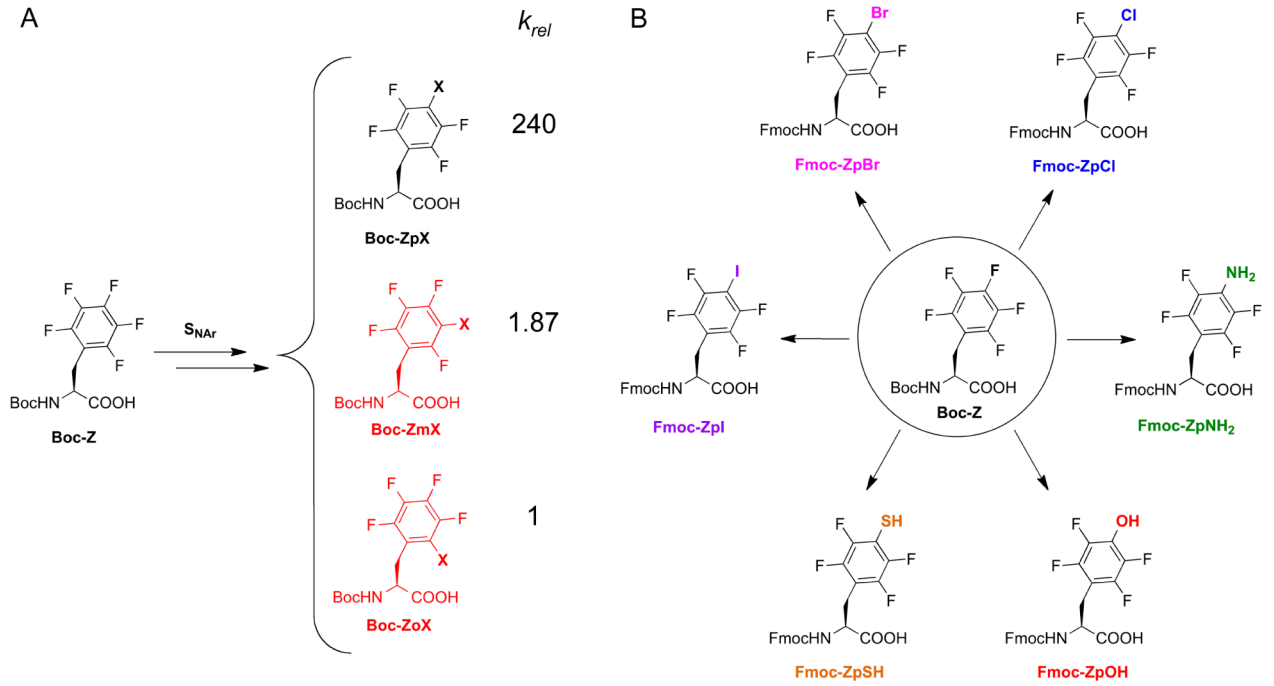


FIGURE 5. Expanding the repertoire of fluorinated aromatic amino acids via S_{NAr} chemistry. (A) Regioselectivity of the S_{NAr} reaction on perfluorophenylalanine. (B) Scope of the S_{NAr} reaction of Boc-Z to obtain ZpX derivatives with diverse functionality. These novel fluorinated amino acids may serve as probes of protein structure and function; they may also introduce novel mechanisms of molecular recognition to protein designers.

example, conjugation of a phenylalanine (F) and a perfluorophenylalanine (Z) onto the N- and C-termini of a short 30-residue collagen peptide afforded self-assembly to give micrometer-scale fibrils, which induced platelet aggregation as effectively as natural collagen does.[45,46] Computational modeling suggested a critical role of F−Z stacking in the organized assembly of this system. Specifically, intermolecular F−Z stacking at the assembly interface gave an association energy more favorable (by −6 kcal/mol) than that of F−F stacking of the control peptide. Similarly, Fmoc-protected perfluorophenylalanine (Fmoc-Z) was reported by Nilsson et al. to undergo rapid self-assembly to give rigid supramolecular gels.[47] In contrast, the nonfluorinated counterpart (Fmoc-F) failed to form hydrogels under the same conditions. This comparison indicates the importance of the reversed quadrupole of Z, which could potentially stack with the Fmoc moiety. Interestingly, the Nilsson group also found that Fmoc-F and Fmoc-Z readily coassembled to form two-component fibrils under conditions that Fmoc-F alone did not form fibrils at all.[48] However, Fmoc-F was found to coassemble with several side-chain monohalogenated (F, Cl, Br) derivatives as well. Without detailed structural information of these fibrillization products, it is difficult to determine whether F−Z stacking contributes to the specificity of the coassembly or not.

Synthesis of Fluorinated Aromatic Amino Acids

As discussed above, fluorinated aromatic amino acids serve as powerful tools for exploring polar−π interactions in proteins and enabling novel recognition mechanisms for protein design. Synthesis of fluorinated aromatic amino acids is nontrivial, with only a handful available from commercial sources. It has been particularly difficult to obtain heavily fluorinated aromatic residues; yet such residues are most desirable to probe aromatic interactions as multifluorination gives maximum electronic perturbation of aromatic rings. Recent work from our group has expanded the scope of heavily fluorinated aromatic amino acids. Due to the electron withdrawing nature of fluorine substitution, heavily fluorinated aromatic rings are subject to nucleophilic aromatic substitution (S_{NAr}) reactions.[49] Because of the counterbalance between the inductive and conjugation effect of fluorine substitution, the S_{NAr} reaction on an aromatic ring is dramatically accelerated by *meta*-fluorination, slightly less so by *ortho*-fluorination. A fluorine substituent on the *para*-position actually decelerates the S_{NAr} reaction. Collectively, for a pentafluorophenyl moiety (like the side chain of Z), the S_{NAr} reaction exhibits remarkable regioselectivity, with the *para*-fluorine substituted exclusively (Figure 5A). Importantly, the reactivity of the resulting tetrafluorophenyl

moiety is greatly reduced, preventing further substitution. Earlier work from the Ghadiri group shows that racemic *para*-azido-tetrafluorophenylalanine can be synthesized through the S_{NAr} reaction of perfluorophenylalanine in its properly protected form.[50] Recent research from our group has shown that S_{NAr} chemistry can be carried out with enantiomerically pure starting material and full retention of the stereochemistry.[51,52] Further exploration of the S_{NAr} chemistry of L-perfluorophenylalanine with Boc protection (Boc-Z) has resulted in a series of *para*-substituted tetrafluorophenylalanines. We name a member of this class of unnatural amino acids as ZpX, with X denoting the *para* substituent. Specifically, by using an alcohol or thiol as a nucleophile, one can synthesize tetrafluorotyrosine (ZpOH) or tetrafluorothiotyrosine (ZpSH), respectively. Chemoselective reduction of ZpN_3 readily gives $ZpNH_2$, which can be further converted into *para*-halogenated tetrafluorophenylalanines (Figure 5B). Preliminary studies of this expanded repertoire of fluoroaromatic building blocks have already shown appealing characteristics such as their highly dispersed ^{19}F-NMR resonances.[52] Furthermore, we expect novel mechanisms of polar−π interactions to be enabled with these new fluoroaromatic residues. For instance, the Zp-halide series may allow halogen bonding[53,54] to be used as a specificity element to control peptide assemblies.

Conclusion and Future Perspectives

The research summarized here on fluorinated proteins demonstrates the energetic significance of polar−π interactions in protein folding, assembly, and ligand recognition. The strength of polar−π interactions estimated in the context of proteins appears to be quite consistent with the results from some small molecule model systems.[18,36,55] By studying incrementally fluorinated phenylalanines, we have found aromatic stacking to be mechanistically complex, embodying a combination of quadrupole, dipole−dipole, and dipole−induced dipole interactions, with no single factor overwhelming others. This improved understanding enables strategic application of fluoroaromatic amino acids in protein design endeavors. For example, tetrafluorophenylalanines (Z_x) can be advantageous over their perfluororinated counterpart (Z) because they allow retention of favorable edge−face interactions. On the other hand, cross affinity between arenes and fluoroarenes allows one to program protein dimerization and assembly into higher order structures. With the expanding pool of fluoroaromatic amino acids, we expect additional novel mechanisms of polar−π interactions (e.g., halogen bonding) to be realized to enrich protein chemistry.

We thank Boston College and the Smith Family Foundation for the financial support of our research. We are grateful to Professors Xin Chen and Feng Wang at Boston University for their valuable input during the preparation of this manuscript. We also extend our thanks to all the participants of the Mesilla Workshop on Aromatic Interactions for helpful discussions.

BIOGRAPHICAL INFORMATION

Christopher J. Pace is currently pursuing his Ph.D. at Boston College under the direction of Prof. Jianmin Gao. His research focuses on the use of fluorinated aromatic amino acids as tools for examining the significance of polar−π interactions in protein stability and molecular recognition.

Jianmin Gao is currently an assistant professor of chemistry at Boston College. He obtained his Ph.D. in Chemistry at Stanford University with Eric Kool and performed postdoctoral research at Scripps with Jeff Kelly. His research interest lies in understanding molecular recognition and design of functional peptides.

FOOTNOTES

*To whom correspondence should be addressed. E-mail: jianmin.gao@bc.edu.
The authors declare no competing financial interest.

REFERENCES

1 Akcay, G.; Kumar, K. A new paradigm for protein design and biological self-assembly. *J. Fluorine Chem.* **2009**, *130*, 1178–1182.
2 Marsh, E. N.; Buer, B. C.; Ramamoorthy, A. Fluorine--a new element in the design of membrane-active peptides. *Mol. BioSyst.* **2009**, *5*, 1143–1147.
3 Salwiczek, M.; Nyakatura, E. K.; Gerling, U. I.; Ye, S.; Koksch, B. Fluorinated amino acids: compatibility with native protein structures and effects on protein-protein interactions. *Chem. Soc. Rev.* **2012**, *41*, 2135–2171.
4 Muralidharan, V.; Muir, T. W. Protein ligation: an enabling technology for the biophysical analysis of proteins. *Nat. Methods* **2006**, *3*, 429–438.
5 Liu, C. C.; Schultz, P. G. Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **2010**, *79*, 413–444.
6 Minnihan, E. C.; Young, D. D.; Schultz, P. G.; Stubbe, J. Incorporation of fluorotyrosines into ribonucleotide reductase using an evolved, polyspecific aminoacyl-tRNA synthetase. *J. Am. Chem. Soc.* **2011**, *133*, 15942–15945.
7 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.
8 Mortenson, D. E.; Satyshur, K. A.; Guzei, I. A.; Forest, K. T.; Gellman, S. H. Quasiracemic crystallization as a tool to assess the accommodation of noncanonical amino residues in nativelike protein conformations. *J. Am. Chem. Soc.* **2012**, *134*, 2473–2476.
9 Li, C.; Wang, G. F.; Wang, Y.; Creager-Allen, R.; Lutz, E. A.; Scronce, H.; Slade, K. M.; Ruf, R. A.; Mehl, R. A.; Pielak, G. J. Protein ^{19}F NMR in Escherichia coli. *J. Am. Chem. Soc.* **2010**, *132*, 321–327.
10 Tang, Y.; Ghirlanda, G.; Petka, W. A.; Nakajima, T.; DeGrado, W. F.; Tirrell, D. A. Fluorinated coiled-coil proteins prepared in vivo display enhanced thermal and chemical stability. *Angew. Chem., Int. Ed.* **2001**, *40*, 1494–1496.
11 Bilgicer, B.; Fichera, A.; Kumar, K. A coiled-coil with a fluorous core. *J. Am. Chem. Soc.* **2001**, *123*, 4393–4399.
12 Lee, K. H.; Lee, H. Y.; Slutsky, M. M.; Anderson, J. T.; Marsh, E. N. Fluorous effect in proteins: de novo design and characterization of a four-alpha-helix bundle protein containing hexafluoroleucine. *Biochemistry* **2004**, *43*, 16277–16284.
13 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.

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

15 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.

16 Meyer, E. A.; Castellano, R. K.; Diederich, F. Interactions with aromatic rings in chemical and biological recognition. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210–1250.

17 Williams, J. H.; Cockcroft, J. K.; Fitch, A. N. Structure of the lowest temperature phase of the solid benzene hexafluorobenzene adduct. *Angew. Chem., Int. Ed.* **1992**, *31*, 1655–1657.

18 Cozzi, F.; Ponzini, F.; Annunziata, R.; Cinquini, M.; Siegel, J. S. Polar interactions between stacked π-systems in fluorinated 1,8-Diarylnaphthalenes - Importance of quadrupole-moments in molecular recognition. *Angew. Chem., Int. Ed.* **1995**, *34*, 1019–1020.

19 Hunter, C. A.; Sanders, J. K. M. The nature of π-π interactions. *J. Am. Chem. Soc.* **1990**, *112*, 5525–5534.

20 Wheeler, S. E.; Houk, K. N. Substituent effects in cation/π interactions and electrostatic potentials above the centers of substituted benzenes are due primarily to through-space effects of the substituents. *J. Am. Chem. Soc.* **2009**, *131*, 3126–3127.

21 Wheeler, S. E.; Houk, K. N. Through-space effects of substituents dominate molecular electrostatic potentials of substituted arenes. *J. Chem. Theory Comput.* **2009**, *5*, 2301–2312.

22 Gallivan, J. P.; Dougherty, D. A. Cation-π interactions in structural biology. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9459–9464.

23 Burley, S. K.; Petsko, G. A. Aromatic-aromatic interaction - A mechanism of protein-structure stabilization. *Science* **1985**, *229*, 23–28.

24 Salonen, L. M.; Ellermann, M.; Diederich, F. Aromatic rings in chemical and biological recognition: Energetics and structures. *Angew. Chem., Int. Ed.* **2011**, *50*, 4808–4842.

25 For a recent review: Dougherty, D. A. Physical organic chemistry on the brain. *J. Org. Chem.* **2008**, *73*, 3667–3673.

26 McGaughey, G. B.; Gagne, M.; Rappe, A. K. π-stacking interactions - Alive and well in proteins. *J. Biol. Chem.* **1998**, *273*, 15458–15463.

27 Butterfield, S. M.; Patel, P. R.; Waters, M. L. Contribution of aromatic interactions to alpha-helix stability. *J. Am. Chem. Soc.* **2002**, *124*, 9751–9755.

28 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.

29 Cockroft, S. L.; Hunter, C. A. Chemical double-mutant cycles: dissecting non-covalent interactions. *Chem. Soc. Rev.* **2007**, *36*, 172–188.

30 Tatko, C. D.; Waters, M. L. Selective aromatic interactions in beta-hairpin peptides. *J. Am. Chem. Soc.* **2002**, *124*, 9372–9373.

31 Woll, M. G.; Hadley, E. B.; Mecozzi, S.; Gellman, S. H. Stabilizing and destabilizing effects of phenylalanine → F-5-phenylalanine mutations on the folding of a small protein. *J. Am. Chem. Soc.* **2006**, *128*, 15932–15933.

32 Zheng, H.; Comeforo, K.; Gao, J. Expanding the fluorous arsenal: tetrafluorinated phenylalanines for protein design. *J. Am. Chem. Soc.* **2009**, *131*, 18–19.

33 McKnight, C. J.; Doering, D. S.; Matsudaira, P. T.; Kim, P. S. A thermostable 35-residue subdomain within villin headpiece. *J. Mol. Biol.* **1996**, *260*, 126–134.

34 McKnight, C. J.; Matsudaira, P. T.; Kim, P. S. NMR structure of the 35-residue villin headpiece subdomain. *Nat. Struct. Biol.* **1997**, *4*, 180–184.

35 Chiu, T. K.; Kubelka, J.; Herbst-Irmer, R.; Eaton, W. A.; Hofrichter, J.; Davies, D. R. High-resolution X-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7517–7522.

36 Cockroft, S. L.; Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. Electrostatic control of aromatic stacking interactions. *J. Am. Chem. Soc.* **2005**, *127*, 8594–8595.

37 Battaglia, M. R.; Buckingham, A. D.; Williams, J. H. The electric quadrupole moments of benzene and hexafluorobenzene. *Chem. Phys. Lett.* **1981**, *78*, 421–423.

38 Williams, J. H. The molecular electric quadrupole-moment and solid-state architecture. *Acc. Chem. Res.* **1993**, *26*, 593–598.

39 Ponzini, F.; Zagha, R.; Hardcastle, K.; Siegel, J. S. Phenyl/pentafluorophenyl interactions and the generation of ordered mixed crystals: sym-triphenethynylbenzene and sym-tris(perfluorophenethynyl)benzene. *Angew. Chem., Int. Ed.* **2000**, *39*, 2323–2325.

40 Coates, G. W.; Dunn, A. R.; Henling, L. M.; Dougherty, D. A.; Grubbs, R. H. Phenyl-perfluorophenyl stacking interactions: A new strategy for supermolecule construction. *Angew. Chem., Int. Ed.* **1997**, *36*, 248–251.

41 Coates, G. W.; Dunn, A. R.; Henling, L. M.; Ziller, J. W.; Lobkovsky, E. B.; Grubbs, R. H. Phenyl-perfluorophenyl stacking interactions: Topochemical 2 + 2 photodimerization and photopolymerization of olefinic compounds. *J. Am. Chem. Soc.* **1998**, *120*, 3641–3649.

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

43 Pace, C. J.; Zheng, H.; Mylvaganam, R.; Kim, D.; Gao, J. Stacked fluoroaromatics as supramolecular synthons for programming protein dimerization specificity. *Angew. Chem. Int. Ed* **2012**, *51*, 103–107.

44 Zheng, H.; Gao, J. Highly specific heterodimerization mediated by quadrupole interactions. *Angew. Chem., Int. Ed.* **2010**, *49*, 8635–8639.

45 Cejas, M. A.; Kinney, W. A.; Chen, C.; Vinter, J. G.; Almond, H. R., Jr.; Balss, K. M.; Maryanoff, C. A.; Schmidt, U.; Breslav, M.; Mahan, A.; Lacy, E.; Maryanoff, B. E. Thrombogenic collagen-mimetic peptides: Self-assembly of triple helix-based fibrils driven by hydrophobic interactions. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 8513–8518.

46 Cejas, M. A.; Kinney, W. A.; Chen, C.; Leo, G. C.; Tounge, B. A.; Vinter, J. G.; Joshi, P. P.; Maryanoff, B. E. Collagen-related peptides: self-assembly of short, single strands into a functional biomaterial of micrometer scale. *J. Am. Chem. Soc.* **2007**, *129*, 2202–2203.

47 Ryan, D. M.; Anderson, S. B.; Senguen, F. T.; Youngman, R. E.; Nilsson, B. L. Self-assembly and hydrogelation promoted by F(5)-phenylalanine. *Soft Matter* **2010**, *6*, 475–479.

48 Ryan, D. M.; Doran, T. M.; Nilsson, B. L. Complementary π-π interactions induce multicomponent coassembly into functional fibrils. *Langmuir* **2011**, *27*, 11145–11156.

49 Rodionov, P. P.; Furin, G. G. Kinetics of nucleophilic-substitution reactions of polyfluoroaromatic compounds. *J. Fluorine Chem.* **1990**, *47*, 361–434.

50 Redman, J. E.; Ghadiri, M. R. Synthesis of photoactive p-azidotetrafluorophenylalanine containing peptide by solid-phase Fmoc methodology. *Org. Lett.* **2002**, *4*, 4467–4469.

51 Wang, F.; Qin, L.; Wong, P.; Gao, J. Facile synthesis of tetrafluorotyrosine and its application in pH triggered membrane lysis. *Org. Lett.* **2011**, *13*, 236–239.

52 Qin, L.; Sheridan, C.; Gao, J. Synthesis of tetrafluorinated aromatic amino acids with distinct signatures in ^{19}F NMR. *Org. Lett.* **2012**, *14*, 528–531.

53 Metrangolo, P.; Meyer, F.; Pilati, T.; Resnati, G.; Terraneo, G. Halogen bonding in supramolecular chemistry. *Angew. Chem., Int. Ed.* **2008**, *47*, 6114–6127.

54 Voth, A. R.; Khuu, P.; Oishi, K.; Ho, P. S. Halogen bonds as orthogonal molecular interactions to hydrogen bonds. *Nat. Chem.* **2009**, *1*, 74–79.

55 Carroll, W. R.; Pellechia, P.; Shimizu, K. D. A rigid molecular balance for measuring face-to-face arene−arene interactions. *Org. Lett.* **2008**, *10*, 3547–3550.