J|A|C|S COMMUNICATIONS

Published on Web 12/09/2008

Expanding the Fluorous Arsenal: Tetrafluorinated Phenylalanines for Protein Design

Hong Zheng, Kristofer Comeforo, and Jianmin Gao*

Department of Chemistry, Merkert Chemistry Center, Boston College, Chestnut Hill, Massachusetts 02467-3801

Received August 7, 2008; E-mail: jianmin.gao@bc.edu

In the search for strategies that stabilize designed protein/peptide structures, fluorinated amino acids have attracted much attention[1] because of the comparable size of fluorine to hydrogen atoms, the superhydrophobicity of fluorocarbons, and the potential "fluorous effect," which refers to the selective self-association between fluorinated moieties. Fluorinated aliphatic amino acids have been demonstrated to be effective in stabilizing protein structures[2] and functioning as recognition motifs.[3] In contrast, amino acids with fluorinated aromatic side chains are only beginning to be investigated and their energetic contribution to protein structural stability remains unclear.[4]

Statistical survey of the Protein Data Bank reveals that aromatic clusters are prevalent in protein structures, suggesting that aromatic interaction plays a significant role in defining a protein fold.[5] In contrast to the aliphatic amino acids, interaction between aromatic residues is complicated in nature and involves van der Waals, hydrophobic, and electrostatic forces.[6] The electrostatic component favors the interacting aromatic rings to adopt the geometry in which the partially positive charged aromatic hydrogens (ArHs) interact with the π-electron cloud. The ArH$\cdots\pi$ interaction is exemplified in the structure of the villin headpiece subdomain (HP35), a 35-residue α-helical protein, which has been widely used as a model system in protein folding studies.[7] HP35 structure features three interacting phenylalanines in the hydrophobic core, with the H-4 of F6 packing against the phenyl ring of F17 and the H-6 of F10 packing against the π cloud of F6 (Figure 1A). Fluorination of these phenylalanine side chains increases their hydrophobicity; however, it may potentially eliminate the favorable ArH$\cdots\pi$ interactions, resulting in protein destabilization.[4b] Based on these considerations, we propose that aromatic amino acids, with side chains *highly but not fully* fluorinated, should be able to offer enhanced hydrophobicity yet retain the key aromatic hydrogens to engage ArH$\cdots\pi$ interactions, consequentally affording a stabilized protein fold. To test the hypothesis, we synthesized two phenylalanine analogues: 2,3,4,5-tetrafluoro-L-phenylalanine (Z_o) and 2,3,5,6-tetrafluoro-L-phenylalanine (Z_p, Figure 1B). These unnatural amino acids were incorporated into HP35, and their energetic perturbation was examined in comparison to that of pentafluoro-L-phenylalanine (Z).

Figure 1B displays the surface electrostatic potential of the phenylalanine side chain and the fluorinated mimics. As expected, heavy fluorination renders the remaining ArHs on the phenyl ring more electron-deficient: H-4 of Z_p bears a partial charge of 0.202, more than double that in F (0.089); the partial charge on H-6 in Z_o also compares favorably to that in F by a large margin (0.168 vs 0.108). The increased positive charge on the ArHs should give strengthened ArH$\cdots\pi$ interactions. It is worth noting that the fluorination reduces the electron density of the π-cloud. We expect the ArH$\cdots\pi$ interaction would be weakened if the π cloud is provided by a fluorinated phenyl ring.

The tetrafluorinated phenylalanines were conveniently prepared by stereoselective benzylation of Seebach's (*S*)-Boc-BMI (BMI = 2-*tert*-butyl-3-methyl-4-imidazolidinone)[8] followed by hydrolysis of the products (Figure 1C). The amino acid analogues were then transformed into the Fmoc-protected form, which were readily incorporated into HP35 through solid phase peptide synthesis. HP35 variants F6Z_p and F10Z_o were synthesized, aiming to keep the above-mentioned ArH$\cdots\pi$ interactions (Figure 1A). As control sequences, we also prepared HP35 variants F6Z and F10Z.

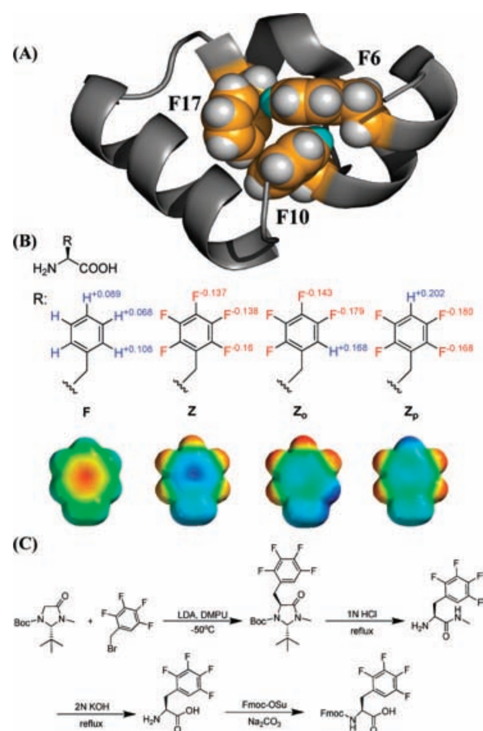

Figure 1. (A) Cartoon presentation of HP35 structure (PDB: 1YRF) highlighting the interacting aromatic residues (F6, F10, and F17) in the hydrophobic core. The main chain is shown in gray, three phenylalanines in orange, the hydrogen atoms engaging in ArH$\cdots\pi$ interactions highlighted in cyan, and other hydrogen atoms in white. (B) Structures and models of the phenylalanine side chain and its fluorinated analogues. The space-filling models are generated with Spartan, with electrostatic potential mapped on molecular surface (red for negative and blue for positive). Calculated partial charges for the relevant hydrogen and fluorine atoms are labeled on the structure. (C) Synthetic scheme of 2,3,4,5-tetrafluorophenylalanine (Z_o), the same strategy is used for the synthesis of 2,3,5,6-tetrafluorophenylalanine Z_p.

The HP35 variants were subjected to spectroscopic and thermodynamic analysis. The circular dichroism spectra of all HP35 mutants display ellipticity minima at 208 and 222 nm (Supporting Information), which are characteristic of α-helical structures. The structures of HP35 mutants are further examined by NMR spectroscopy. Shown in Figure 2A is the 1D ^{1}H NMR data. Well-dispersed resonances and nearly identical patterns were observed for all HP35 variants, indicating that the fluorinated mutants all adopt stable and native-like structures. More specifically, HP35 WT displays two upfield shifted resonances (−0.51 and 0.10 ppm), which are assigned to be the H^γ of Val9 and H^δ of Leu20, respectively.[7a] Their packing against aromatic side chains is responsible for the upfield shift. These characteristic resonances are

10.1021/ja8062309 CCC: $40.75
© 2009 American Chemical Society

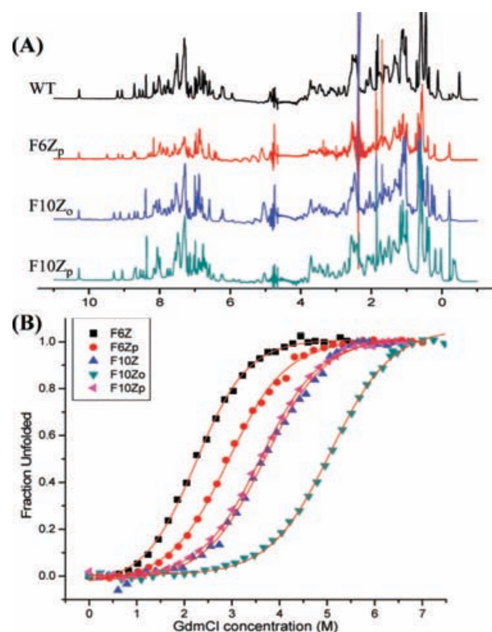

Figure 2. (A) 1D ^1H NMR spectra of the WT HP35 and variants thereof. Spectra of all mutants exhibit well-dispersed resonances and similar patterns to that of the WT. (B) Gdm•Cl denaturation curves of HP35 variants, displaying apparent two state transitions.

evidently preserved in the HP35 mutants, strongly suggesting no major structural perturbations by the fluorinated amino acids.

The thermal stabilities of HP35 variants were evaluated through thermal melting experiments on a circular dichroism spectrometer monitoring the ellipticity at 222 nm. The (un)folding is completely reversible (Supporting Information), and the melting curves were fitted to a two-state model to give the melting temperature (T_m) values (Table 1). The folding free energies (ΔG_f) were determined through analyzing

Table 1. Thermodynamic Parameters of the HP35 Variants

protein	T_m[a] (°C)	C_m[b] (M)	m value[b] (kcal/(mol·M))	ΔG_f[b] (kcal/mol)	$\Delta\Delta G_f$[c] (kcal/mol)
WT	66	3.7	0.88	−3.3 ± 0.1	0
F6Z	51	2.3	1.02	−2.3 ± 0.1	+1.0 ± 0.1
F6Z_p	62	3.0	0.89	−2.6 ± 0.1	+0.7 ± 0.1
F10Z	68	3.6	0.99	−3.6 ± 0.1	−0.3 ± 0.1
F10Z_o	80	5.0	0.93	−4.7 ± 0.1	−1.4 ± 0.1
F10Z_p	64	3.6	0.89	−3.2 ± 0.1	+0.1 ± 0.1

[a] Calculated by fitting thermal denaturation curves. [b] Calculated by fitting guanidine hydrochloride denaturation curves. [c] $\Delta\Delta G_f = \Delta G_{f,mutant} - \Delta G_{f,wt}$.

guanidinium chloride denaturation data (Figure 2B), and the results are summarized in Table 1. Interestingly, F to Z_o mutation at residue 10 enhances HP35 stability by a large margin; F10Z_o is more stable than the WT by 14 °C in T_m and −1.4 kcal/mol in ΔG_f. More interestingly, the tetrafluorophenylalanine mutant F10Z_o exhibits higher stability than the pentafluorophenylalanine mutant F10Z (12 °C in T_m and −1.1 kcal/mol in ΔG_f). Given that there is only one atom difference between Z_o and Z (hydrogen vs fluorine), we attribute the superior stability of F10Z_o to the ArH···π interaction, which is retained in F10Z_o and eliminated in F10Z.

To further explore the physical origin of the hyperstability of F10Z_o, we synthesized and examined the HP35 variant F10Z_p, in which the positively charged H-4 on the Z_p side chain does not engage in ArH···π interactions. F10Z_p is less stable than F10Z_o by −1.5 kcal/mol, which strongly argues the hyperstability of F10Z_o is largely attributed to the ArH···π interaction. In addition, the folding free energy of F10Z_p is comparable to that of F10Z, indicating that the slightly larger size of Z (130 Å^3 vs 125 Å^3 for Z_p or Z_o) is easily accommodated by the protein structure.

Compared to the pentafluorophenylalanine variant F6Z, the tetrafluorophenylalanine variant F6Z_p also exhibits improved stability (11 °C in T_m and −0.3 kcal/mol in ΔG_f), indicating that the ArH···π interaction is an important energetic factor. However, both F6Z_p and F6Z are destabilized in comparison to the WT protein. This can be explained by the fact that F6 engages in two ArH···π interactions: it provides its H-4 to interact with the phenyl ring of F17; it also provides its π-cloud to interact with H-6 of F10. Heavy fluorination of F6 yields an electron-deficient aromatic ring, which weakens the ArH···π interaction between the face of F6 and the edge of F10. It is interesting to note that the ArH···π interaction contributes −1.1 kcal/mol to protein stability at residue 10 (F10Z_o vs F10Z), while it only contributes to −0.3 kcal/mol at residue 6 (F6Z_p vs F6Z). This may be explained by the fact that residue 6 is partially solvent-exposed (solvent accessible surface area for CH-4 of F6 is 4.6 Å^2) while residue 10 is completely buried. Solvent exposure weakens the ArH···π interaction that is electrostatic in nature.

In summary, we have synthesized two tetrafluoro-phenylalanines and shown that they can contribute more favorably to protein stability than pentafluoro-phenylalanines. Our data strongly suggest the enhanced stability is due to the ArH···π interactions retained and strengthened by these *highly*, *but not fully*, fluorinated aromatic molecules. Considering the prevalence of the ArH···π interaction in protein structures, we expect molecules of this category should find a wide array of applications in designing stable protein structures and mediating protein−protein interactions.

Acknowledgment. We thank the reviewers for their comments and Dr. John Boylan for NMR assistance. We gratefully acknowledge financial support from the Smith Family Foundation and the start-up fund from Boston College. We also acknowledge the NSF for financial support of the BC Mass Spectrometry Center (Grant No. DBI-0619576).

Supporting Information Available: Procedures for amino acids and protein synthesis; Far-UV CD of the HP35 variants; and thermal/chaotrope denaturation curve acquisition. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Yoder, N. C.; Kumar, K. *Chem. Soc. Rev.* **2002**, *31*, 335–41.
(2) (a) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 666–7. (b) Horng, J. C.; Raleigh, D. P. *J. Am. Chem. Soc.* **2003**, *125*, 9286–7. (c) Jackel, C.; Salwiczek, M.; Koksch, B. *Angew. Chem., Int. Ed.* **2006**, *45*, 4198–203. (d) Lee, H. Y.; Lee, K. H.; Al-Hashimi, H. M.; Marsh, E. N. *J. Am. Chem. Soc.* **2006**, *128*, 337–43. (e) Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, I. W.; DeGrado, W. F.; Tirrell, D. A. *Biochemistry* **2001**, *40*, 2790–6.
(3) (a) Bilgicer, B.; Kumar, K. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15324–9. (b) Bilgicer, B.; Xing, X.; Kumar, K. *J. Am. Chem. Soc.* **2001**, *123*, 11815–6.
(4) (a) Chiu, H. P.; Suzuki, Y.; Gullickson, D.; Ahmad, R.; Kokona, B.; Fairman, R.; Cheng, R. P. *J. Am. Chem. Soc.* **2006**, *128*, 15556–7. (b) Woll, M. G.; Hadley, E. B.; Mecozzi, S.; Gellman, S. H. *J. Am. Chem. Soc.* **2006**, *128*, 15932–3. (c) Butterfield, S. M.; Patel, P. R.; Waters, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 9751–5. (d) Thorson, J. S.; Chapman, E.; Murphy, E. C.; Schultz, P. G. *J. Am. Chem. Soc.* **1995**, *117*, 1157–8.
(5) (a) Burley, S. K.; Petsko, G. A. *Science* **1985**, *229*, 23–8. (b) Burley, S. K.; Petsko, G. A. *Adv. Protein Chem.* **1988**, *39*, 125–89.
(6) (a) Waters, M. L. *Biopolymers* **2004**, *76*, 435–45. (b) Waters, M. L. *Curr. Opin. Chem. Biol.* **2002**, *6*, 736–41.
(7) (a) McKnight, C. J.; Matsudaira, P. T.; Kim, P. S. *Nat. Struct. Biol.* **1997**, *4*, 180–4. (b) Chiu, T. K.; Kubelka, J.; Herbst-Irmer, R.; Eaton, W. A.; Hofrichter, J.; Davies, D. R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7517–22.
(8) Gander-coquoz, M.; Seebach, D. *Helv. Chim. Acta* **1988**, *71*, 224–36.

JA8062309