J|A|C|S COMMUNICATIONS

Published on Web 08/28/2009

Effect of Highly Fluorinated Amino Acids on Protein Stability at a Solvent-Exposed Position on an Internal Strand of Protein G B1 Domain

Hsien-Po Chiu,[†] Bashkim Kokona,[‡] Robert Fairman,[‡] and Richard P. Cheng*[,§]

Department of Chemistry, University at Buffalo, The State University of New York, Buffalo, New York 14260-3000, Department of Biology, Haverford College, Haverford, Pennsylvania 19041, and Department of Chemistry, National Taiwan University, Taipei, Taiwan 10617

Received May 5, 2009; E-mail: rpcheng@ntu.edu.tw

Highly fluorinated amino acids can stabilize proteins[1] for potential application in various protein biotechnologies including therapeutics[2] and biosensors.[3] Pioneering work to enhance protein stability by substituting natural hydrocarbon amino acids with fluoro-amino acids has mostly focused on helical proteins.[1,4] However, the helicity of monomeric Ala-based peptides decreases upon replacing hydrocarbon amino acids with the corresponding fluorocarbon amino acids,[5,6] suggesting that fluoro-amino acids may be more suitable for nonhelical secondary structures such as β-sheets. Indeed, substituting Val with trifluorovaline at a mostly buried β-sheet position stabilized protein NTL9 by 1.44 kcal·mol^{-1}·residue^{-1},[7] larger than most of the fluoro-stabilizations in helices.[1a−i] Furthermore, many proteins used for therapeutics[2] and biosensors[3] involve β-sheet proteins such as antibodies. Therefore, understanding the effect of fluoro-amino acids on β-sheet stability should facilitate the use of these amino acids in protein biotechnologies and bioactive compounds.[8] As a first step, we report the effect of fluoro-amino acids at a solvent-exposed position in the β-sheet of protein G B1 domain[9] (Figure 1A).

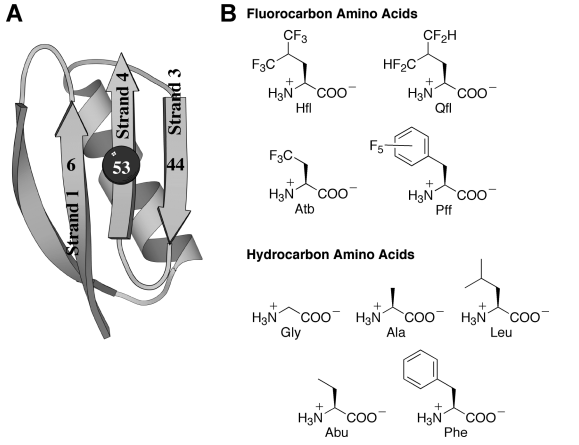

Figure 1. Panel A: ribbon diagram of protein G B1 domain[9] (PDB ID: 1PGA) generated using the program Molscript.[10] The guest position 53 (shown as black ball) on the internal strand 4, the immediate cross-strand position 6 on internal strand 1, and position 44 on edge strand 3 are labeled. Panel B: chemical structure of fluorocarbon and hydrocarbon amino acids: (*S*)-5,5,5,5′,5′,5′-hexafluoroleucine (Hfl), (*S*)-5,5,5′,5′-tetrafluoroleucine (Qfl), (*S*)-2-amino-4,4,4-trifluorobutyric acid (Atb), (*S*)-pentafluorophenylalanine (Pff), glycine (Gly), L-alanine (Ala), L-leucine (Leu), (*S*)-2-aminobutyric acid (Abu), and L-phenylalanine (Phe).

An ideal host system−guest position combination for exploring sheet propensity should have the following two characteristics: (1) high sensitivity to mutation at the guest position and (2) minimal interference from tertiary interactions including lateral and diagonal cross-strand interactions. Sheet formation energetics of natural amino acids have been determined in two hosts: zinc finger[11] and protein G B1 domain (GB1).[12−14] The internal strand guest position 53[12,14] of GB1 is more sensitive than both the edge strand guest position 44[13] of GB1 and the zinc finger host−guest system.[11] Apparently, cross-strand interactions may not be significant in GB1 based on phage display studies.[15] Many studies have focused on β-hairpins;[16] however the stability of such motifs is determined by turn stability, intrinsic sheet propensity, and lateral and diagonal cross-strand interactions, making deciphering the intrinsic sheet propensity difficult in β-hairpins. Accordingly, we chose to investigate the GB1 system using the I6A T44A double mutant to minimize possible cross-strand interactions (Figure 1A).[12,14,15] The solvent-exposed guest position 53[12,14] on internal strand 4 was systematically changed to the residues in Figure 1B.[17] All the proteins were monomeric in solution by sedimentation equilibrium;[18] therefore intermolecular interactions should not affect the stability of the proteins.

Thermal denaturation of the proteins was monitored by circular dichroism spectroscopy (CD) at 222 nm (Figure S1).[18] The high cooperativity of GB1[9a] has enabled the unfolding of the helix (monitored at 222 nm) to represent the unfolding of the overall structure and thus β-sheet.[14,19] There was some variation in the CD signal near 4 °C for the GB1 mutants, especially for GB1-Atb and GB1-Hfl.[18] Such differences have been reported with minimal effect on GB1 bioactivity[12] or structure,[12,14,19] and the variations were attributed to differences in aromatic contributions.[12] To confirm the structural integrity of GB1-Atb and GB1-Hfl, these two proteins along with GB1-Ala and GB1-Qfl were investigated by NMR.[18] The sheet structure near the guest site for all four mutants was consistent with the native GB1 fold[9] based on chemical shift deviations,[18,20a] sequential HCα(i)-HN(i+1) NOEs,[18,20b] and interstrand NOEs.[18] Furthermore, the structure of the helix for all four proteins was also consistent with the native fold[9] based on chemical shift deviations,[18,20a] sequential HN(i)-HN(i+1) NOEs,[18,20b] and sequential HCα(i)-HCβ(i+3) NOEs.[18,20b] Thus, the different CD signals for GB1-Atb and GB1-Hfl may be due to different contributions from aromatic side chains,[12] but not the lack of sheet formation near the guest site or helix formation.

The thermal unfolding and folding of all the proteins were reversible (Figure S1).[18] The CD data were converted to fraction unfolded protein (Figure 2). Data near the 50% unfolded for each protein were used to obtain the T_m (Table 1)[18] and van't Hoff unfolding enthalpy and entropy,[18] which were used to derive the relative unfolding free energy at 60 °C ($\Delta\Delta G_{unfold\ 60°C}$, Table 1).[18] This temperature was chosen to minimize extrapolation of the data from the T_m for each protein[12] and to enable direct comparison with literature values at 60 °C.[14] The T_m and $\Delta\Delta G_{unfold\ 60°C}$ for proteins with natural amino acids in the guest position were similar to literature values.[14]

[†] The State University of New York.
[‡] Haverford College.
[§] National Taiwan University.

10.1021/ja903631h CCC: $40.75 © 2009 American Chemical Society

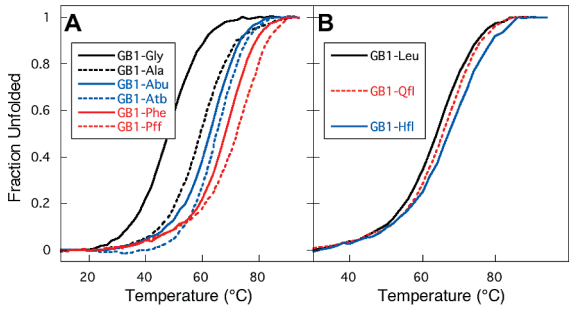

Figure 2. Fraction unfolded plotted against temperature for GB1-based mutants. Panel A: GB1-Gly, GB1-Ala, GB1-Abu, GB1-Atb, GB1-Phe, and GB1-Pff. Panel B: GB1-Leu, GB1-Qfl, and GB1-Hfl.

Table 1. T_m and Relative Unfolding Free Energy at 60 °C ($\Delta\Delta G_{unfold\ 60°C}$) of GB1-Based Proteins

protein[17]	T_m (°C)	$\Delta\Delta G_{unfold\ 60°C}$ (kcal·mol^{-1})[a]	T_m (°C)[14]	$\Delta\Delta G_{unfold\ 60°C}$ (kcal·mol^{-1})[14]
GB1-Gly	47.7 ± 0.2	−1.207	45.95	−1.21
GB1-Ala	59.2 ± 0.2	0	57.05	0
GB1-Abu	62.1 ± 0.6	0.387		
GB1-Leu	63.3 ± 0.7	0.513	62.47	0.45
GB1-Phe	67.5 ± 0.6	1.073	67.68	1.08
GB1-Atb	64.9 ± 0.9	0.737		
GB1-Qfl	64.9 ± 1.1	0.722		
GB1-Hfl	65.9 ± 0.9	0.806		
GB1-Pff	71.1 ± 0.8	1.410		

[a] $\Delta\Delta G_{unfold\ 60°C}$(GB1-Xaa) = $\Delta G_{unfold\ 60°C}$(GB1-Xaa) − $\Delta G_{unfold\ 60°C}$(GB1-Ala). $\Delta G_{unfold\ 60°C}$(GB1-Ala) = −0.086 kcal·mol^{-1}.

Introducing fluorines onto the amino acids at the solvent-exposed position 53 on internal strand 4 appears to stabilize GB1 based on T_m and $\Delta\Delta G_{unfold\ 60°C}$ (Table 1). Replacing Phe with Pff stabilizes GB1 by 0.34 kcal·mol^{-1}, whereas replacing Abu with Atb increases the stability by 0.35 kcal·mol^{-1}. Furthermore, replacing Leu with Qfl and Hfl stabilizes GB1 by 0.21 and 0.29 kcal·mol^{-1}, respectively. The increased stability upon introducing the fluorines may be due to hydrophobics,[21a] sterics,[21b,c] or both,[21d,e] because linear correlations of $\Delta\Delta G_{unfold\ 60°C}$ with hydrophobicity (log P, $R = 0.812$) and size (volume, $R = 0.891$) were similar (Figure S6).[18] Hydrophobic side chains can facilitate backbone desolvation.[21f] In contrast, large side chains can limit available backbone conformations to favor sheet formation,[21b] obstruct backbone−solvent interaction,[21c] and shield cross-strand hydrogen bonds in the folded form.[21f] The stabilization observed is less than that observed by Raleigh upon introducing trifluorovaline at a largely buried sheet position,[7] most likely because the current study involves a solvent-exposed position, which cannot take full advantage of burying the highly hydrophobic fluorous side chains. Furthermore, energetics in the current study are reported at a higher temperature (60 °C versus 25°), which attenuates the values. The apparent discrepancy may also be due to the difference in the shape of the fluoro-amino acids investigated (i.e., β-branched versus non-β-branched). Nevertheless, the stability of GB1 increases upon introducing fluorines onto the amino acids at the solvent-exposed guest position 53 on internal strand 4. This is in sharp contrast to helix formation energetics (in Ala-based peptides), which become less favorable upon introducing fluorines by up to 1.72 kcal·mol^{-1}.[5,6] Overall, fluoro-amino acids may be worthwhile building blocks to explore for stabilizing β-sheet proteins, which are especially important for biotechnologies such as therapeutics[2] and biosensors.[3]

Acknowledgment. This work was supported by NYSTAR Watson Young Investigator Program, ACS Petroleum Research Fund (R.P.C., #44532-G4), Kapoor fund (R.P.C.), the State University of New York at Buffalo (R.P.C.), National Science Council in Taiwan (R.P.C., NSC-97-2113-M-002-019-MY2), National Taiwan University (R.P.C.), and National Science Foundation (R.F., MCB-0211754).

Supporting Information Available: Experimental details for the synthesis and characterization of the proteins, sedimentation equilibrium, thermal denaturation, and NMR experiments. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, W. A., III.; DeGrado, W. F.; Tirrell, D. A. *Biochemistry* **2001**, *40*, 2790. (b) Tang, Y.; Ghirlanda, G.; Petka, W. A.; Nakajima, T.; DeGrado, W. F.; Tirrell, D. A. *Angew. Chem., Int. Ed.* **2001**, *40*, 1494. (c) Tang, Y.; Tirrell, D. A. *J. Am. Chem. Soc.* **2001**, *123*, 11089. (d) Montclare, J. K.; Son, S.; Clark, G. A.; Kumar, K.; Tirrell, D. A. *ChemBioChem* **2009**, *10*, 84. (e) Bilgiçer, B.; Fichera, A.; Kumar, K. *J. Am. Chem. Soc.* **2001**, *123*, 4393. (f) Bilgiçer, B.; Kumar, K. *Tetrahedron* **2002**, *58*, 4105. (g) Lee, K.-H.; Lee, H.-Y.; Slutsky, M. M.; Anderson, J. T.; Marsh, E. N. G. *Biochemistry* **2004**, *43*, 16277. (h) Lee, H.-Y.; Lee, K.-H.; Al-Hashimi, H. M.; Marsh, E. N. G. *J. Am. Chem. Soc.* **2006**, *128*, 337. (i) Woll, M. G.; Hadley, E. B.; Mecozzi, S.; Gellman, S. H. *J. Am. Chem. Soc.* **2006**, *128*, 15932. (j) Zheng, H.; Comeforo, K.; Gao, J. *J. Am. Chem. Soc.* **2009**, *131*, 18.

(2) (a) Kurtzman, A. L.; Govindarajan, S.; Vahle, K.; Jones, J. T.; Heinrichs, V.; Patten, P. A. *Curr. Opin. Biotechnol.* **2001**, *12*, 361. (b) Vasserot, A. P.; Dickinson, C. D.; Tang, Y.; Huse, W. D.; Manchester, K. S.; Watkins, J. D. *Drug Discov. Today* **2003**, *8*, 118. (c) Tang, L.; Persky, A. M.; Hochhaus, G.; Meibohm, B. *J. Pharm. Sci.* **2004**, *93*, 2184.

(3) (a) Ramsden, J. J. *J. Mol. Recognit.* **1997**, *10*, 109. (b) Baird, C. L.; Myszka, D. G. *J. Mol. Recognit.* **2001**, *14*, 261. (c) Elia, G.; Silacci, M.; Scheurer, S.; Scheuermann, J.; Neri, D. *Trends Biotechnol.* **2002**, *20*, S19. (d) Cooper, M. A. *Nat. Rev. Drug Discov.* **2002**, *1*, 515. (e) Shah, J.; Wilkins, E. *Electroanalysis* **2003**, *15*, 157.

(4) (a) Jäckel, C.; Seufert, W.; Thust, S.; Koksch, B. *ChemBioChem* **2004**, *5*, 717. (b) Jäckel, C.; Salwiczek, M.; Koksch, B. *Angew. Chem., Int. Ed* **2006**, *45*, 4198.

(5) Butterfield, S. M.; Patel, P. R.; Waters, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 9751.

(6) (a) Chiu, H.-P.; Suzuki, Y.; Gullickson, D.; Ahmad, R.; Kokona, B.; Fairman, R.; Cheng, R. P. *J. Am. Chem. Soc.* **2006**, *128*, 15556. (b) Chiu, H.-P.; Cheng, R. P. *Org. Lett.* **2007**, *9*, 5517.

(7) Horng, J.-C.; Raleigh, D. P. *J. Am. Chem. Soc.* **2003**, *125*, 9286.

(8) (a) Meng, H.; Kumar, K. *J. Am. Chem. Soc.* **2007**, *129*, 15615. (b) Meng, H.; Krishnaji, S. T.; Beinborn, M.; Kumar, K. *J. Med. Chem.* **2008**, *51*, 7303. (c) Gottler, L. M.; Lee, H.-Y.; Shelburne, C. E.; Ramamoorthy, A.; Marsh, E. N. G. *ChemBioChem* **2008**, *9*, 370. (d) Gottler, L. M.; de la Salud Bea, R.; Shelburne, C. E.; Ramamoorthy, A.; Marsh, E. N. G. *Biochemistry* **2008**, *47*, 9243.

(9) (a) Gronenborn, A. M.; Filpula, D. R.; Essig, N. Z.; Achari, A.; Whitlow, M.; Wingfield, P. T.; Clore, G. M. *Science* **1991**, *253*, 657. (b) Gallagher, T.; Alexander, P.; Bryan, P.; Gilliland, G. L. *Biochemistry* **1994**, *33*, 4721.

(10) Kraulis, P. J. *J. Appl. Crystallogr.* **1991**, *24*, 946.

(11) Kim, C. A.; Berg, J. M. *Nature* **1993**, *362*, 267.

(12) Minor, D. L., Jr.; Kim, P. S. *Nature* **1994**, *367*, 660.

(13) Minor, D. L., Jr.; Kim, P. S. *Nature* **1994**, *371*, 264.

(14) Smith, C. K.; Withka, J. M.; Regan, L. *Biochemistry* **1994**, *33*, 5510.

(15) Distefano, M. D.; Zhong, A.; Cochran, A. G. *J. Mol. Biol.* **2002**, *322*, 179.

(16) (a) de Alba, E.; Rico, M.; Jiménez, M. A. *Protein Sci.* **1997**, *6*, 2548. (b) Griffiths-Jones, S. R.; Maynard, A. J.; Searle, M. S. *J. Mol. Biol.* **1999**, *292*, 1051. (c) Ramírez-Alvarado, M.; Kortemme, T.; Blanco, F. J.; Serrano, L. *Bioorg. Med. Chem.* **1999**, *7*, 93. (d) Syud, F. A.; Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* **2001**, *123*, 8667. (e) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. *Protein Sci.* **2002**, *11*, 1492. (f) Russell, S. J.; Blandl, T.; Skelton, N. J.; Cochran, A. G. *J. Am. Chem. Soc.* **2003**, *125*, 388. (g) Tatko, C. D.; Waters, M. L. *J. Am. Chem. Soc.* **2004**, *126*, 2028.

(17) The protein sequences were based on the protein G B1 domain I6A T44A mutant with the guest position at residue 53. The proteins were named by prefixing the three-letter code of the amino acid at the guest position with "GB1-"; protein GB1-Ala would have an Ala at the 53 position of the protein G B1 domain I6A T44A mutant.

(18) See Supporting Information.

(19) Merkel, J. S.; Sturtevant, J. M.; Regan, L. *Structure* **1999**, *7*, 1333.

(20) (a) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *Biochemistry* **1992**, *31*, 1647. (b) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, NY, 1986.

(21) (a) Yang, A.-S.; Honig, B. *J. Mol. Biol.* **1995**, *252*, 366. (b) Street, A. G.; Mayo, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9074. (c) Bai, Y.; Englander, S. W. *Proteins: Struct., Funct., Genet.* **1994**, *18*, 262. (d) Otzen, D. E.; Fersht, A. R. *Biochemistry* **1995**, *34*, 5718. (e) Koehl, P.; Levitt, M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 12524. (f) Avbelj, F.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 1309.

JA903631H