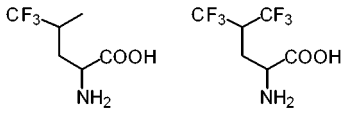
Biosynthesis of a Highly Stable Coiled-Coil Protein Containing Hexafluoroleucine in an Engineered Bacterial Host

Yi Tang and David A. Tirrell*

Division of Chemistry and Chemical Engineering
California Institute of Technology
Pasadena, California 91125

Received July 18, 2001

Incorporation of nonnatural amino acid residues allows engineering of proteins with novel chemical functionality and unusual physical properties.[1] We have shown recently that coiled-coil proteins prepared in vivo can be stabilized significantly by replacement of leucine by trifluoroleucine (**1**).[2] In the same series of experiments, however, we were unsuccessful in our attempts to incorporate the more highly fluorinated analogue hexafluoroleucine (**2**). We report here that modification of the leucyl-tRNA synthetase (LeuRS) activity of the host allows efficient incorporation of **2** into recombinant proteins prepared in *Escherichia coli*. Furthermore, the coiled-coil protein used to demonstrate incorporation of **2** exhibits enhanced stability in comparison to the same protein enriched in **1**, possibly due to the increased hydrophobic character of the additional trifluoromethyl group in the protein core.



1 **2**

Translational incorporation of an amino acid analogue in vivo requires that the analogue serve as a substrate for one of the aminoacyl-tRNA synthetases (aaRS) of the host. Several strategies have been employed to address this requirement, including directed evolution[1c] and rational design[3] of aaRS that exhibit enhanced activity toward nonnatural substrates. Fluorinated amino acids are ideal candidates for incorporation because the small size of the fluorine atom renders such analogues nearly isosteric to their hydrocarbon counterparts, and **1** is incorporated readily into recombinant proteins without alteration of the host.[2] Other fluorinated amino acids have also been introduced into proteins in conventional *E. coli* expression strains.[4] Because both the (2*S*,4*S*) and the (2*S*,4*R*) diastereomers of **1** support protein synthesis in vivo,[5] it seemed likely that **2** would be activated and charged with reasonable efficiency by the *E. coli* LeuRS. We did not anticipate editing by the synthetase because the near identity of the geometries of leucine and **2** should preclude entry of the latter into the editing site of the enzyme.[6] When we compared the efficiencies of activation of leucine, **1**, and **2** by LeuRS via an in vitro ATP-PP_i exchange assay, it became clear that the failure of **2** to support protein synthesis under normal expression conditions is due to the fact that it is activated approximately 4000-fold more slowly than leucine (Table 1).[7]

Table 1. Kinetic Parameters for Activation of Leucine, **1**, and **2** by *E. coli* LeuRS[a]

substrate	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (*rel*)
leucine	3.8 ± 0.3	17.8 ± 5.1	1
1	0.49 ± 0.05	561 ± 121	1/242
2	0.11 ± 0.01	1979 ± 804	1/4100

[a] Leucine was used as the L-isomer; **1** as a mixture of the (2*S*,4*S*), (2*S*,4*R*), (2*R*,4*S*), and (2*R*,4*R*) forms, and **2** as the D,L-mixture. K_m values are reported in terms of the concentrations of the LOLI-isomers. We have shown previously by ^{19}F NMR spectroscopy that both of the 2*S*-isomers of **1** are incorporated in vivo;[5] however, we have not evaluated the kinetic parameters for each isomer separately.

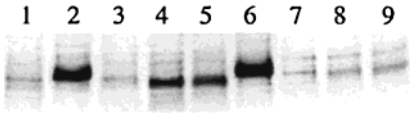


Figure 1. SDS-PAGE evidence for incorporation of **2** into recombinant protein A1. Lanes 1−5: Vector pA1EL−1, no amino acid added; 2, leucine added; 3, **2** at 40 mg/L; 4, **2** at 320 mg/L; 5, **2** at 640 mg/L. Lanes 6−9: Vector pQEA1−6, leucine added; 7, **2** at 40 mg/L; 8, **2** at 320 mg/L; 9, **2** at 640 mg/L. All concentrations are for the L-isomer.

Previous work with methionine analogues has shown that elevation of the MetRS activity of the host facilitates incorporation of analogues that are poor substrates for the synthetase.[8] The *E. coli leuS* gene[9] and its endogenous promoter were amplified from genomic DNA and ligated into the expression vector pQEA1 to yield pA1EL.[10] The LeuRS activity of strains carrying pA1EL was approximately 8-fold higher than that of otherwise identical strains carrying pQEA1, confirming overexpression of LeuRS.

The synthetic leucine zipper protein A1[11] was used to assay protein expression in *E. coli* cultures depleted of leucine and supplemented with **2**.[12] Results are shown in Figure 1. Only the strain bearing pA1EL supports protein synthesis in cultures supplemented with **2** (lanes 4 and 5); the recombinant protein is not detected in similar cultures transformed with pQEA1 (lanes 7−9). The enhanced electrophoretic mobility of the fluorinated

(1) For example see: (a) Sharma, N.; Furter, R.; Kast, P.; Tirrell, D. A. *FEBS Lett.* **2000**, *467*, 37−40. (b) van Hest, J. C. M.; Kiick, K. L.; Tirrell, D. A. *J. Am. Chem. Soc.* **2000**, *122*, 1282−1288. (c) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498−501. (d) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 666−667.

(2) (a) Tang, Y.; Ghirlanda, G.; Petka, W. A.; Nakajima, T.; DeGrado, W. F.; Tirrell, D. A. *Angew. Chem., Int. Ed.* **2001**, *40*, 1494−1497. (b) Similar stabilization has been reported by Kumar and co-workers (Bilgicer, B.; Fichera, A.; Kumar, K. *J. Am. Chem. Soc.* **2001**, *123*, 4393−4399) and by our laboratory (Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, W. A.; DeGrado, W. F.; Tirrell, D. A. *Biochemistry* **2001**, *40*, 2790−2796) for coiled-coil peptides prepared by solid-phase peptide synthesis.

(3) Kast, P.; Hennecke, H. *J. Mol. Biol.* **1991**, *222*, 99−124.

(4) (a) Feeney, J.; McCormick, J. E.; Bauer, C. J.; Birdsall, B.; Moody, C. M.; Starkmann, B. A.; Young, D. W.; Francis P.; Havlin, R. H.; Arnold, W. D.; Oldfield E. *J. Am. Chem. Soc.* **1996**, *118*, 8700−8706. (b) Renner, C.; Alefelder, S.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L. *Angew. Chem., Int. Ed.* **2001**, *40*, 923−925.

(5) Kothakota, S. Ph.D. Thesis, University of Massachusetts, Amherst, 1993.

(6) (a) Chen, J.-F.; Li, T.; Wang, E. D.; Wang, Y. L. *Biochemistry* **2001**, *40*, 1144−1149. (b) Murisinna, R. S.; Lincecum, T. L.; Martinis, S. A. *Biochemistry* **2001**, *40*, 5376−5381.

(7) LeuRS was expressed and purified according to procedures described previously. The kinetic parameters reported here for leucine agree well with literature values (Li, T.; Wang, E. D.; Wang, Y. L. *Acta Biochim. Biophys. Sinica* **1997**, *29*, 591−596. Li, T.; Guo, N. N.; Xia, X.; Wang, E. D.; Wang, Y. L. *Biochemistry* **1999**, *38*, 13063−13069).

(8) Kiick, K. L.; van Hest, J. C. M.; Tirrell, D. A. *Angew. Chem., Intl. Ed.* **2000**, *39*, 2148−2151.

(9) Hartlein, M.; Madern, D. *Nucl. Acid. Res.* **1987**, *15*, 10199−10210.

(10) For primer sequence, see Supporting Information. The *leuS* gene was cloned into the *NheI* site of the pQEA1. *E. coli* genomic DNA was purified by using the DNEasy Tissue Kit from Qiagen (Chatsworth, CA).

(11) Petka, W. A.; Harden, J. L.; McGrath, K. P.; Wirtz, D.; Tirrell, D. A. *Science* **1998**, *281*, 389−392. A1 is a 74-residue protein containing eight leucine residues and six heptad repeats in its coiled-coil region. (For descriptions of the coiled-coil motif, see, for example: (a) McClain, D. L.; Woods, H. L.; Oakley, M. G. *J. Am. Chem. Soc.* **2001** *123*, 3151−3152. (b) Tripet, B.; Wagschal, K.; Lavigne, P.; Mant, C. T.; Hodges, R. S. *J. Mol. Biol.* **2000** *30*, 377−402 and references therein.

(12) The procedure used to assay protein expression has been described previously.[2a] pA1EL carries the gene encoding A1 under control of a *lac* promoter. The plasmid pREP4 (Qiagen, Chatsworth, CA) was used to co-transform the leucine auxotrophic strain LAM1000 in order to reduce the basal level of protein expression.

10.1021/ja016652k CCC: $20.00
Published on Web 10/11/2001

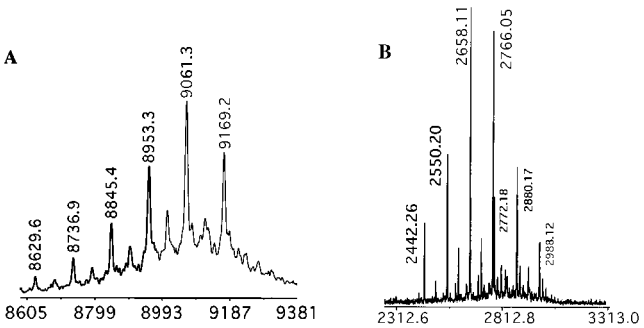


Figure 2. Replacement of leucine by **2** in HA1 can be detected in MALDI mass spectra of the intact protein (A) and of tryptic fragments (B). (A) The mass of the wild-type protein is 8307 (not observed). Each substitution of **2** for leucine leads to a mass increase of 108 units. The strongest signal (at *m*/*z* 9061.3) corresponds to replacement of 7 out of 8 leucine positions. (B) The tryptic fragment **L**KNEIED**L**KAEIGD**L**-NNTSGIR contains three leucine sites. Fragments that correspond to 0, 1, 2, and 3 sites of replacement are observed at masses of 2442.26, 2550.20, 2658.11, and 2766.05, respectively. A different set of tryptic fragments containing two sites of substitution is apparent at slightly higher masses (0, 2772.18; 1, 2880.17; 2, 2988.12).[13]

protein is consistent with earlier observations of the behavior of A1 prepared in cultures supplemented with **1**.[2a]

To analyze the composition and physical properties of the protein (HA1) prepared in cultures supplemented with **2**, a sample was purified on a Ni-NTA affinity column via the His-tag leader sequence of the protein. The yield of HA1 was 8 mg/L, compared to 40 mg/L for A1 prepared under similar conditions.

Amino acid analysis of HA1 indicates 74% replacement of leucine by **2**. Incomplete substitution is likely due to release of leucine through turnover of cellular proteins; given the large difference in the efficiencies of activation of leucine and **2**, even a severely depleted leucine pool should give rise to competitive incorporation of the two amino acids.

Replacement of leucine by **2** causes the expected increase in the mass of the protein (Figure 2A). The predominant signals in the MALDI mass spectrum correspond to occupancy of 6 (8953.3), 7 (9061.3), and 8 (9169.2) leucine sites, respectively (of a total of 8 such sites in A1). Incorporation of **2** is also evident from the MALDI-MS of tryptic fragments of the substituted protein (Figure 2B).

The secondary structure of HA1 was determined by circular dichroism spectroscopy to be >90% α-helical. Ultracentrifugation shows the protein to be predominantly dimeric at concentrations of 10−100 μM, with a dissociation constant for the dimer-to-tetramer equilibrium of 188 ± 4 μM (as compared to a value of 80.5 μM previously reported for A1[2a]). These results suggest minimal disruption of the folding and assembly of the coiled-coil architecture of A1 upon near-complete fluorination of the leucine sites.

The stability of HA1 with respect to thermal and chemical denaturation is significantly improved compared to the wild-type protein (Figure 3). The melting temperature of HA1 is elevated from 54 °C (for A1) to 76 °C; for comparison, the same protein outfitted with **1** melts at 67 °C.[2a] The free energy of unfolding at 25 °C was determined by thermodynamic fitting of the melting

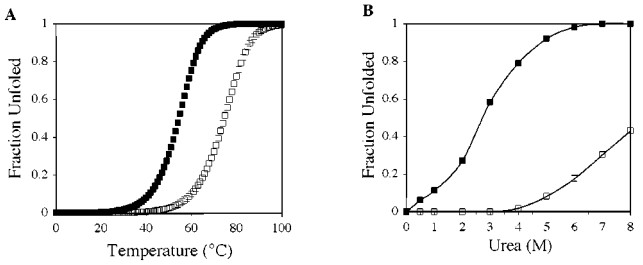


Figure 3. HA1 (□) exhibits significant elevation in stability compared to A1 (■). CD spectra were recorded at a protein concentration of 10 μM in PBS buffer, pH 7.4. (A) Thermal denaturation profiles. The melting temperature is increased from 54 to 76 °C upon replacement of leucine by **2**. (B) Urea denaturation profiles. CD reports less than 50% denaturation of HA1 at 8 M urea.

curve to be 14.4 kcal/mol,[14] an elevation of 3.7 kcal/mol compared to A1. HA1 also exhibits remarkable resistance to denaturation by urea; the protein is nearly fully folded (as reported by circular dichroism spectroscopy) at 4 M urea, and unfolding remains far from complete even in 8 M solutions of the denaturant. Juxtaposition of the aliphatic side chains of the leucine residues contributes significantly to the stability of coiled-coil proteins.[15] Replacement of these residues with the highly hydrophobic, fluorinated analogue **2** increases the driving force for protein folding and assembly.

The results described here illustrate the synthetic potential of microbial hosts equipped with elevated aaRS activities. Nonnatural amino acids that do not support protein synthesis in conventional hosts can be incorporated efficiently into recombinant proteins in such engineered hosts. Introduction of **2** into the coiled-coil protein A1 is accompanied by significant increases in the thermal and chemical stability of the folded form of the protein. We believe that these results can be extended to the design of highly stable proteins and to the engineering of protein−protein interactions.

Acknowledgment. This work was supported by the NSF MRSEC program at Caltech. Y.T. is supported by a fellowship from the Whitaker Foundation. We thank Dr. Giovanna Ghirlanda and Prof. William DeGrado for ultracentrifugation studies. We also thank Kent Kirshenbaum, Isaac Carrico, and Pin Wang for insightful discussions.

Supporting Information Available: Amino acid synthesis, plasmid construction, protein expression, and protein analysis (PDF). This material is available free of charge via the Internet at http://pubs.acs.org. *Address correspondence to this author. Phone: (626)395-3140. Fax: (626)793-8472. E-mail: tirrell@caltech.edu.

JA016652K

(13) In both spectra, signals of reduced intensity are detected between the major peaks. We attribute these weaker signals to proteins that contain **1**, which was present at a level of <0.2% in the sample of **2** used to prepare HA1 (see Supporting Information). The ratio of **1** to **2** in the protein was determined to be approximately 0.04 by amino acid analysis. Enrichment of **1** in the protein as compared to the feed is expected on the basis of its more efficient activation by LeuRS. The intensities of the intermediate signals are enhanced by combinatorial effects associated with the occupancy of the leucine sites by three different amino acid variants.

(14) Schneider, J. P.; Lear, J. D.; DeGrado, W. F. *J. Am. Chem. Soc.* **1997**, *119*, 5742−5743.

(15) Moitra, J.; Szilak, L.; Krylov, D.; Vinson, C. *Biochemistry* **1997**, *36*, 12567−12573.