DOI: 10.1002/cbic.200800164

Biosynthesis and Stability of Coiled-Coil Peptides Containing (2*S*,4*R*)-5,5,5-Trifluoroleucine and (2*S*,4*S*)-5,5,5-Trifluoroleucine

Jin Kim Montclare,[b] Soojin Son,[a] Ginevra A. Clark,[c] Krishna Kumar,[c, d] and David A. Tirrell*[a]

Introduction of noncanonical amino acids into proteins is a powerful method for the creation of macromolecules with novel properties.[1–3] In particular, 5,5,5-trifluoroleucine (TFL, **2**) has been utilized as a more hydrophobic surrogate of leucine in various contexts (Figure 1).[4,5] When incorporated into the hydrophobic cores of certain coiled-coil proteins, TFL triggers an increase in stability and renders proteins more resistant to thermal and chemical denaturation.[6,7] Furthermore, despite the larger volume of CF_3 compared to CH_3, protein structure and activity can be retained upon replacement of leucine (Leu, **1**) by TFL.[8,9]

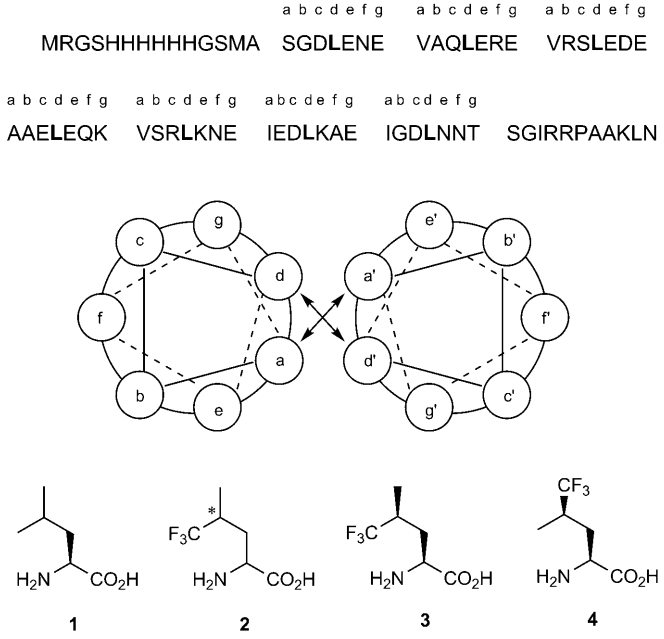

Figure 1. The A1 peptide sequence and helical wheel diagram of the dimer in which leucines are highlighted in bold. Structures of leucine (**1**) and the trifluoromethyl analogues (**2**–**4**) are also shown. The asterisk in structure **2** denotes unresolved stereochemistry at the 4-position.

Leu contains one stereocenter (C_α), which has the *S* configuration. Replacement of a methyl group by a trifluoromethyl group at C_γ introduces an additional stereocenter and yields the two diastereoisomers (2*S*,4*S*)-5,5,5-trifluoroleucine (**3**) and (2*S*,4*R*)-5,5,5-trifluoroleucine (**4**; Figure 1). Here we report the effects of TFL stereochemistry on coiled-coil peptide biosynthesis and stability. We demonstrate that both **3** and **4** are activated and incorporated into recombinant peptides prepared in *E. coli*. Coiled-coil homodimers of peptides bearing **3** or **4** exhibit increased stability when compared to dimers of the Leu form of the peptide. An equimolar mixture of the two fluorinated peptides forms a heterodimer of modestly enhanced thermal stability relative to the homodimers.

The fidelity of translation is governed in large part by the activation of amino acids by their cognate aminoacyl–tRNA synthetases (AARS).[10] Although some AARSs are known to tolerate noncanonical substrates, amino acid activation can be acutely sensitive to side-chain stereochemistry. For example, isoleucine contains two stereocenters, both of *S* configuration, one at C_α and another at C_β. The 2*S*,3*R* isomer (*allo*-isoleucine) is not incorporated into proteins, although there is evidence that it is bound and activated by isoleucyl–tRNA synthetase (IleRS).[11–14] Stereochemistry can also determine the fate of noncanonical amino acids as possible substrates for protein synthesis in *E. coli*. For instance, the isoleucine analogue, 2-amino-3-methyl-4-pentenoic acid, is accepted only in its 2*S*,3*S* form, while the valine (or isoleucine) analogue, 4,4,4-trifluorovaline (TFV), is active only in its 2*S*,3*R* form.[15,16] We sought to explore whether leucyl–tRNA synthetase (LeuRS) exhibits a stereochemical preference with respect to activation of TFL.

Coiled-coil peptides constitute simple model systems for investigation of protein biosynthesis and stability.[17–20] Stereoisomers **3** and **4** were prepared (Scheme S1 in the Supporting Information) and evaluated for incorporation into coiled-coil peptide A1[18] (Figure 1) in an *E. coli* strain that is auxotrophic for leucine. The A1 peptide was also expressed in media supplemented with **1** or **2**. Following purification on Ni–nitrilotriacetic acid affinity columns, protein yields were determined to be (18 ± 4) mg $L^{-1}$and (9 ± 3) mg L^{-1} upon incubation with **3** and **4**, respectively, compared to (45 ± 6) mg L^{-1} for A1 prepared with **1**. Peptides containing **3** and **4** were designated *SS*-A1 and *SR*-A1, respectively.

Matrix-assisted laser desorption/ionization (MALDI) mass spectrometry analysis of A1 fragments obtained by trypsin digestion was performed to assess the extent of substitution by **3** or **4**. Fragment **L**KNEIED**L**KAEIGD**L**NNTSGIR, which corre-

[a] *Dr. S. Son,+ Prof. D. A. Tirrell*
Division of Chemistry and Chemical Engineering
California Institute of Technology, Pasadena, CA 91125 (USA)
Fax: (+1) 626-793-8472
E-mail: tirrell@caltech.edu

[b] *Prof. J. K. Montclare+*
Department of Chemical and Biological Sciences
Polytechnic University, Department of Biochemistry
SUNY Downstate Medical Center, Brooklyn, NY 11203 (USA)

[c] *G. A. Clark, Prof. K. Kumar*
Department of Chemistry, Tufts University
Medford, MA 02155 (USA)

[d] *Prof. K. Kumar*
Cancer Center, Tufts-New England Medical Center
Boston, MA 02110 (USA)

[+] *These authors contributed equally to this work.*

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.

sponds to residues 46–67 in A1, contains three leucine residues (shown in bold type). Fragments that correspond to replacement at 0, 1, 2, and 3 sites by either **3** or **4** were observed (Figure 2). The expected mass increment of 54 Da was visible for each leucine residue that was replaced by TFL. Incomplete replacement of **1** most likely reflects a persistent pool of the natural amino acid replenished by cellular protein turnover.[19] The distribution of peak intensities (though not simply related to the relative abundances of fragments) is roughly consistent with unbiased substitution of **3** or **4**, and suggests a slight preference for incorporation of **3** (90% replacement of Leu in *SS*-A1) versus **4** (82% in *SR*-A1). Quantitative amino acid analysis was consistent with the MALDI results and showed 91% replacement in *SS*-A1 and 80% in *SR*-A1.

The relative rates of activation of **1** and fluorinated analogues **2**–**4** by LeuRS were determined by an in vitro ATP–PP_i exchange assay. The kinetic parameters are shown in Table 1.

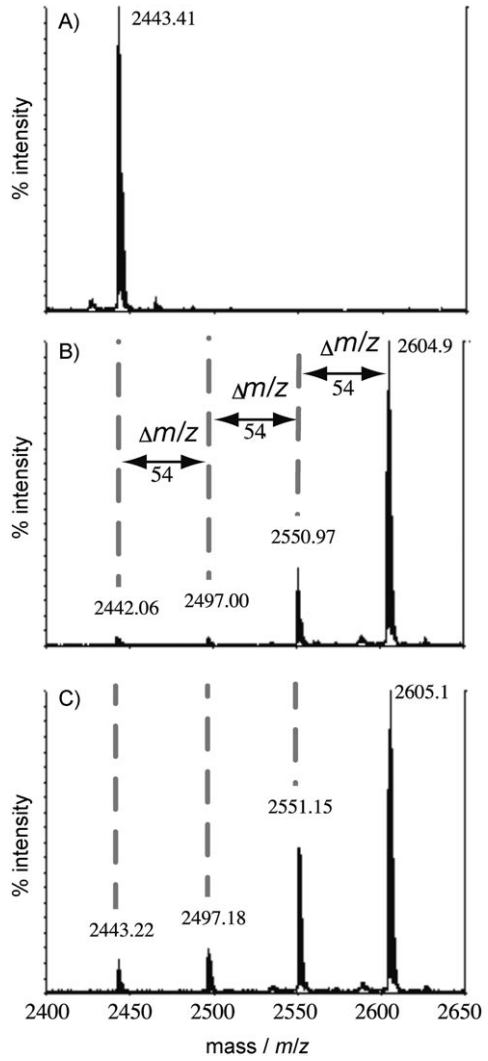

Figure 2. MALDI mass spectra of a tryptic fragment of A1 (residues 46–67) containing three leucine positions. A1 was expressed in media supplemented with either A) **1**, B) **3**, or C) **4**. Fragments corresponding to 0, 1, 2 and 3 sites of substitution are represented as ~2442, 2497, 2551 and 2605 Da, respectively.

Table 1. Kinetic parameters for activation of **1**–**4** by *E. coli* LeuRS.[a]

Substrate	K_m [μM]	k_{cat} [s^{-1}]	k_{cat}/K_m [rel]
1	16.9 ± 4.5	4.22 ± 0.35	1
2	659 ± 103	0.40 ± 0.03	1/412
3	252 ± 92	0.59 ± 0.05	1/107
4	708 ± 280	0.19 ± 0.02	1/933

[a] Substrate **1** was used as the L-isomer; **2** as a mixture of 2*S*,4*S*, 2*S*,4*R*, 2*R*,4*S* and 2*R*,4*R* forms; **3** as the 2*S*,4*S* form and **4** as the 2*S*,4*R* form.

The relative k_{cat}/K_m values show that **3** is a slightly better substrate than **4**; this is consistent with the modest differences in yield and incorporation levels described above. As expected, the apparent k_{cat}/K_m for **2** fell between the values for **3** and **4** (Table 1). The activation rates for both **3** and **4** are within the range of rates that have been shown to support protein synthesis in conventional hosts cultured with other noncanonical amino acids.[21] Although a slight stereochemical preference is observed with respect to activation of **3** versus **4**, this result stands in sharp contrast to the absolute selectivity imposed by IleRS and ValRS in the activation of TFV, in which only the 2*S*,3*R* isomer is tolerated.[15,16] Furthermore, the fact that both **3** and **4** can be incorporated into proteins in *E. coli* is consistent with previous work showing that hexafluoroleucine is activated by LeuRS.[21] All of these results indicate that fluorination at either of the C_δ positions is tolerated by LeuRS.

The secondary structures of all four A1 peptides were examined by circular dichroism (CD) spectroscopy. Because the peptides can form dimers, Leu-A1, *SR*-A1 and *SS*-A1 represent homodimers whereas the equimolar mixture of *SR*-A1 and *SS*-A1 can form either homo- or heterodimers. All four spectra were nearly identical at 1 °C, and indicated about 90% helical content in each peptide as judged from the molar ellipticity at 222 nm (Figure 3A).[22,23] The CD spectra showed no evidence that fluorination affects the secondary structure of A1.

Previous studies have shown that incorporation of fluorinated amino acids into coiled-coil peptides and α-helical bundles results in enhanced stability;[7,8,19,20,24] the extent of stabilization varies depending on the identity of the fluorinated analogue.[19,24] To determine whether the stereochemistry of TFL effects the extent of stabilization of A1, thermal denaturation was monitored by CD spectroscopy (Figure 3B). For both *SR*-A1 and *SS*-A1, the thermal melting temperature (T_m) was 65 °C, 10 °C higher than that of Leu-A1 (T_m = 55 °C). The equimolar mixture of *SR*-A1 and *SS*-A1 exhibited a T_m of 68 °C. This additional 3 °C increase in T_m suggests that *SR*-A1 and *SS*-A1 form heterodimers rather than a mixture of homodimers (Table S1). When a mixture of the 2*S*,4*S* and 2*S*,4*R* forms of TFL was used for expression of A1, ΔT_m was 13 °C,[6] which is nearly identical to that observed for equimolar mixtures of *SS*-A1 and *RR*-A1.

The results reported here demonstrate that both the *S*,*R* and *S*,*S* isomers of TFL are incorporated into proteins expressed in *E. coli*; the *S* isomer is activated at a slightly higher rate by LeuRS. The higher activation rate leads to higher protein yields for *SS*-A1 relative to *SR*-A1 and to increased levels of incorporation of the fluorinated analogue. Neither stereoisomer appears

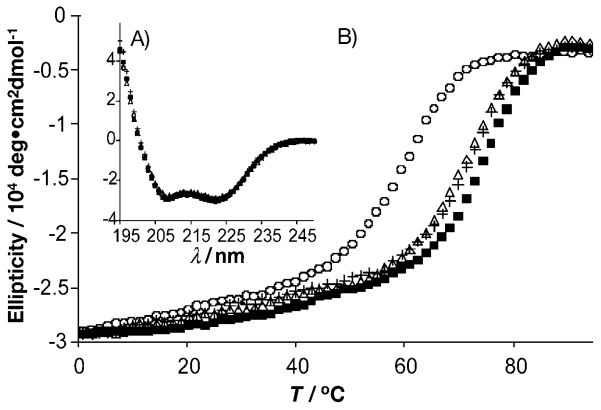

Figure 3. CD spectra of Leu-A1 (○), *SR*-A1 (△), *SS*-A1 (+), and an equimolar mixture of *SR*-A1 and *SS*-A1 (■). A) Wavelength scan at 1 °C, 10 μM protein concentration, PBS buffer, pH 7.4. B) Thermal denaturation (1.5 °C interval, 1 min equilibration time, 10 s averaging time) at 10 μM protein concentration, PBS buffer, pH 7.4.

to alter the coiled-coil structure of A1. Replacement of Leu by either isomer enhances the thermostability of A1; the heterodimer of *SS*- and *SR*-A1 shows an additional modest increase in stability. Experiments are underway to extend these findings and explore more fully the influence of side-chain fluorination on protein stability.

Experimental Section

Synthesis of 3 and 4: Amino acids **3** and **4** were prepared as described previously (see the Supporting Information).[25–27]

Protein biosynthesis and purification: Leucine auxotrophic strain LAM1000 transformed with pREP4 (Qiagen) was used as the *E. coli* host to express A1, which was encoded within pQEA1 under the control of a *lac* promoter. Protein expression and purification were performed as described previously.[7] Protein concentrations were determined by UV spectroscopy.

Protein characterization: CD data were collected by using an Aviv 62DS spectropolarimeter (Lakewood, NJ, USA) with a 1 mm pathlength cell. Wavelength scans were taken from 195 to 250 nm, with points taken every 1 nm. Temperature scans were performed from 0–95 °C in 1.5 °C steps. Each plot represents an average of three scans.

Activation kinetics: An N-terminal His_6–LeuRS fusion was expressed and purified as previously reported.[21] Measurement of the rates of activation of leucine and analogues was performed by using an ATP–PP_i exchange assay. The assay buffer conditions were HEPES (30 mM, pH 7.4), $MgCl_2$ (10 mM), DTT (1 mM), ATP (2 mM) and [^{32}P]-PP_i (2 mM; 0.5 TBq mol^{-1}). A fixed concentration of His_6–LeuRS (75 nM) was used. Depending on the activity of the enzyme toward the substrate, the following substrate concentration ranges were used: **1**: 0.6–312.5 μM; **2**, **3**, **4**: 6.1–6250 μM. Once the reaction was complete, the reaction mixture was quenched by the addition of inorganic pyrophosphate (PP_i; 200 mM), $HClO_4$ (7%, *w/v*) and activated charcoal (3%). The charcoal was washed twice and measured by using a scintillation counter. Kinetic data were fitted by using nonlinear regression analysis.

Acknowledgements

This work was supported by the NIH grants GM62523 and 5FM GM67375-2 (D.A.T. and J.K.M.), and GM65500 (K.K.), NSF graduate fellowship (S.S.), GAANN fellowship (G.A.C.), and a NSF CAREER award (K.K.).

Keywords: biosynthesis · noncanonical amino acids · protein engineering · stereochemical control · thermostability

[1] a) A. J. Link, M. L. Mock, D. A. Tirrell, *Curr. Opin. Biotechnol.* **2003**, *14*, 603–609; b) R. E. Connor, D. A. Tirrrell, *Polym. Rev.* **2007**, *47*, 9–28; N. Budisa, *Angew. Chem.* **2004**, *116*, 6586–6624; *Angew. Chem. Int. Ed.* **2004**, *43*, 6426–6463.

[2] a) T. L. Hendrickson, V. de Crecy-Lagard, P. Schimmel, *Annu. Rev. Biochem.* **2004**, *73*, 147–176; b) L. Wang, P. G. Schultz, *Angew. Chem.* **2005**, *117*, 34–68; *Angew. Chem. Int. Ed.* **2005**, *44*, 34–66.

[3] a) B. Bilgicer, K. Kumar, *J. Chem. Educ.* **2003**, *80*, 1275–1281; b) C. Jäckel, B. Koksch, *Eur. J. Org. Chem.* **2005**, 4483–4503.

[4] O. M. Rennert, H. S. Anker, *Biochemistry* **1963**, *2*, 471–476.

[5] E. Neil, G. Marsh, *Chem. Biol.* **2000**, *7*, R153–R157.

[6] Y. Tang, G. Ghirlanda, W. A. Petka, T. Nakajima, W. F. DeGrado, D. A. Tirrell, *Angew. Chem.* **2001**, *113*, 1542–1544; *Angew. Chem. Int. Ed.* **2001**, *40*, 1494–1496.

[7] B. Bilgicer, A. Fichera, K. Kumar, *J. Am. Chem. Soc.* **2001**, *123*, 4393–4399.

[8] Y. Tang, G. Ghirlanda, N. Vaidehi, J. Kua, D. T. Mainz, W. A. Goddard, W. F. DeGrado, D. A. Tirrell, *Biochemistry* **2001**, *40*, 2790–2796.

[9] a) T. Panchenko, W. W. Zhu, J. K. Montclare, *Biotechnol. Bioeng.* **2006**, *94*, 921–930; b) J. K. Montclare, D. A. Tirrell, *Angew. Chem.* **2006**, *118*, 4630–4633; *Angew. Chem. Int. Ed.* **2006**, *45*, 4518–4521; c) N. Voloshchuk, M. X. Lee, I. C. Tanrikulu, J. K. Montclare, *Bioorg. Med. Chem. Lett.* **2007**, *17*, 5907–5911.

[10] S. Cusack, A. Yaremchuk, M. Tukalo, *EMBO J.* **2000**, *19*, 2351–2361.

[11] R. B. Loftfield, *Biochem. J.* **1963**, *89*, 82–92.

[12] J. Flossdorf, M. R. Kula, *Eur. J. Biochem.* **1973**, *36*, 534–540.

[13] R. B. Loftfield, E. A. Eigner, *Biochim. Biophys. Acta* **1966**, *130*, 426–448.

[14] H. Polet, M. E. Conrad, *Proc. Soc. Exp. Biol. Med.* **1969**, *130*, 581–586.

[15] M. L. Mock, T. Michon, J. C. M. van Hest, D. A. Tirrell, *ChemBioChem* **2006**, *7*, 83–87.

[16] P. Wang, A. Fichera, K. Kumar, D. A. Tirrell, *Angew. Chem.* **2004**, *116*, 3750–3752; *Angew. Chem. Int. Ed.* **2004**, *43*, 3664–3666.

[17] E. K. O'Shea, J. D. Klemm, P. S. Kim, T. Alber, *Science* **1991**, *254*, 539–544.

[18] W. A. Petka, J. L. Harden, K. P. McGrath, D. Wirtz, D. A. Tirrell, *Science* **1998**, *281*, 389–392.

[19] a) Y. Tang, D. A. Tirrell, *J. Am. Chem. Soc.* **2001**, *123*, 11089–11090; b) S. Son, I. C. Tanrikulu, D. A. Tirrell, *ChemBioChem* **2006**, *7*, 1251–1257.

[20] a) C. Jäckel, W. Seufert, S. Thust, B. Koksch, *ChemBioChem* **2004**, *5*, 717–720; b) C. Jäckel, M. Salwiczek, B. Koksch, *Angew. Chem.* **2006**, *118*, 4305–4309; *Angew. Chem. Int. Ed.* **2006**, *45*, 4198–4203.

[21] K. L. Kiick, R. Weberskirch, D. A. Tirrell, *FEBS Lett.* **2001**, *502*, 25–30.

[22] The software CDNN was employed to calculate the percent helical content from the CD wavelength data. This program is available at: http://bioinformatik.biochemtech.uni-halle.de/

[23] Protein concentrations were determined by using the Bradford assay (BioRad, Hercules, CA, USA).

[24] a) K. H. Lee, H. Y. Lee, M. M. Slutsky, J. T. Anderson, E. N. G. Marsh, *Biochemistry* **2004**, *43*, 16277–16284; b) H. Y. Lee, K. H. Lee, H. M. Al-Hashimi, E. N. G. Marsh, *J. Am. Chem. Soc.* **2006**, *128*, 337–343; c) L. M. Gottler, R. de La Salud-Bea, E. N. G. Marsh, *Biochemistry* **2008**, *47*, 4484–4490.

[25] X. Xing, A. Fichera, K. Kumar, *J. Org. Chem.* **2002**, *67*, 1722–1725.

[26] K. Weinges, E. Kromm, *Liebigs Ann. Chem.* **1985**, 90–102.

[27] S. Yamada, C. Hongo, R. Yoshioka, I. Chibata, *J. Org. Chem.* **1983**, *48*, 843–846.

Received: March 14, 2008

