NIH Public Access
Author Manuscript
Chembiochem. Author manuscript; available in PMC 2011 November 9.

Published in final edited form as:
Chembiochem. 2009 January 5; 10(1): 84–86. 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

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

David A. Tirrell: tirrell@caltech.edu

[a]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA, Fax: (+1) 626-793-8472

[b]Department of Chemical and Biological Sciences, Polytechnic University, Brooklyn, NY 11201, USA, Department of Biochemistry, SUNY Downstate Medical Center, Brooklyn, NY 11203, USA

[c]Department of Chemistry, Tufts University, Medford, MA 02155, USA

[d]Cancer Center, Tufts-New England Medical Center, Boston, MA 02110, USA

Keywords

Non-canonical amino acids; protein engineering; thermostability; biosynthesis; stereochemical control

Introduction of non-canonical 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, rendering 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]

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 *Escherichia 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 non-canonical substrates, amino acid activation can be acutely sensitive to side-chain stereochemistry. For example, isoleucine contains two stereocenters, both of *S*-

Correspondence to: David A. Tirrell, tirrell@caltech.edu.
+These authors contributed equally.
Supporting information for this article is available

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

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 non-canonical 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 use in investigations of protein biosynthesis and stability.[17–20] Stereoisomers **3** and **4** were prepared (Scheme 1S, supporting information) and evaluated for incorporation into coiled-coil peptide A1[18] (Figure 1) in an *E. coli* strain auxotrophic for leucine. A1 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 and 9 ± 3 mg/L upon incubation with **3** and **4**, respectively, compared to 45 ± 6 mg/L 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, corresponding 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 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, showing 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. The relative k_{cat}/K_m values show that **3** is a slightly better substrate than **4**, 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 non-canonical amino acids.[21] Although a slight stereochemical preference is observed with respect to activation of **3** vs. **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 *2S,3R* 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 ca. 90% helical content in each peptide as judged from the molar ellipticity at 222 nm (Figure 3A).[22, 23] CD provided no evidence that fluorination affects the secondary structure of A1.

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

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 affects 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, 11°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, suggested that SR-A1 and SS-A1 form heterodimers rather than a mixture of homodimers (Table 1S, supporting information). 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] nearly identical to that observed for equimolar mixtures of SS-A1 and RR-A1.

The results reported here demonstrate that both the SR- and SS-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 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 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 via UV spectroscopy.

Protein characterization

CD data were collected on an Aviv 62DS spectropolarimeter (Lakewood, NJ) using 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 3 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 an ATP-PP_i exchange assay. The assay buffer conditions were 30 mM HEPES, pH 7.4, 10 mM $MgCl_2$, 1 mM DTT, 2 mM ATP and 2 mM [^{32}P]-PP_i (0.5 TBq/mol). A fixed concentration of 75 nM of His_6-LeuRS was used. Depending upon the activity of the enzyme toward the substrate, the following substrate concentrations ranges were used (**1**: 0.6–312.5 μM; **2**, **3**, **4**: 6.1–6250 μM). Once the reaction was completed, the reaction mixture was quenched by addition of 200 mM PP_i, 7% w/v $HClO_4$ and 3% activated charcoal. The charcoal was washed twice and measured on a scintillation counter. Kinetic data were fit using non-linear regression analysis.

NIH-PA Author Manuscript

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the NIH GM 62523 and 5FM GM67375-2 (DAT and JKM), GM65500 (KK), NSF graduate fellowship (SS), GAANN fellowship (GAC) and a NSF CAREER award (KK).

References

1. a) Link AJ, Mock ML, Tirrell DA. Curr Op Biotech. 2003; 14:603–609.b) Connor RE, Tirrrell DA. Polymer Rev. 2007; 47:9–28.Budisa N. Angew Chem Int Ed. 2004; 43:6426–6463.
2. a) Hendrickson TL, de Crecy-Lagard V, Schimmel P. Ann Rev Biochem. 2004; 73:147–176. [PubMed: 15189139] b) Wang L, Schultz PG. Angew Chem Int Ed. 2005; 44:34–66.
3. a) Bilgicer B, Kumar K. J Chem Ed. 2003; 80:1275–1281.b) Jäckel C, Koksch B. Eur J Org Chem. 2005; 21:4483–4503.
4. Rennert OM, Anker HS. Biochemistry. 1963; 2:471–476. [PubMed: 14069531]
5. Marsh ENG. Chem Biol. 2000; 7:R153–R157. [PubMed: 10903940]
6. Tang Y, Ghirlanda G, Petka WA, Nakajima T, DeGrado WF, Tirrell DA. Angew Chem Int Ed. 2001; 40:1494–1496.
7. Bilgicer B, Fichera A, Kumar K. J Am Chem Soc. 2001; 123:4393–4399. [PubMed: 11457223]
8. Tang Y, Ghirlanda G, Vaidehi N, Kua J, Mainz DT, Goddard WA, DeGrado WF, Tirrell DA. Biochemistry. 2001; 40:2790–2796. [PubMed: 11258889]
9. a) Panchenko T, Zhu WW, Montclare JK. Biotech Bioeng. 2006; 94:921–930.b) Montclare JK, Tirrell DA. Angew Chem Int Ed. 2006; 45:4518–4521.c) Voloshchuk N, Lee MX, Tanrikulu IC, Montclare JK. Bioorg Med Chem Lett. 2007; 17:5907–5911. [PubMed: 17845847]
10. Cusack S, Yaremchuk A, Tukalo M. EMBO J. 2000; 19:2351–2361. [PubMed: 10811626]
11. Loftfield RB. Biochem J. 1963; 89:82–92. [PubMed: 14097371]
12. Flossdorf J, Kula MR. Eur J Biochem. 1973; 36:534–540. [PubMed: 4581277]
13. Loftfield RB, Eigner EA. Biochim Biophys Acta. 1966; 130:426–448. [PubMed: 4291467]
14. Polet H, Conrad ME. Proc Soc Exp Biol Med. 1969; 130:581–586. [PubMed: 5765302]
15. Mock ML, Michon T, van Hest JCM, Tirrell DA. ChemBioChem. 2006; 7:83–87. [PubMed: 16397872]
16. Wang P, Fichera A, Kumar K, Tirrell DA. Angew Chem Int Ed. 2004; 43:3664–3666.
17. O'Shea EK, Klemm JD, Kim PS, Alber T. Science. 1991; 254:539–344. [PubMed: 1948029]
18. Petka WA, Harden JL, McGrath KP, Wirtz D, Tirrell DA. Science. 1998; 281:389–392. [PubMed: 9665877]
19. a) Tang Y, Tirrell DA. J Am Chem Soc. 2001; 123:11089–11090. [PubMed: 11686725] b) Son S, Tanrikulu IC, Tirrell DA. ChemBioChem. 2006; 7:1251–1257. [PubMed: 16758500]
20. a) Jäckel C, Seufert W, Thust S, Koksch B. ChemBioChem. 2004; 5:717–720. [PubMed: 15122644] b) Jäckel C, Salwiczek M, Koksch B. Angew Chem Int Ed. 2006; 45:4198–4203.
21. Kiick KL, Weberskirch R, Tirrell DA. FEBS Lett. 2001; 505:25–30. [PubMed: 11478942]
22. The software CDNN was employed to calculate the % helical content from the CD wavelength data. This program is available at http://bioinformatik.biochemtech.uni-halle.de/.
23. Protein concentrations were determined by Bradford Assay (BioRad, Hercules, CA).
24. a) Lee KH, Lee HY, Slutsky MM, Anderson JT, Marsh ENG. Biochemistry. 2004; 43:16277–16284. [PubMed: 15610021] b) Lee HY, Lee KH, Al-Hashimi HM, Marsh ENG. J Am Chem Soc. 2006; 128:337–343. [PubMed: 16390163] c) Gottler LM, de la Salud-Bea R, Marsh ENG. Biochemistry. 2008; 47:4484–4490. [PubMed: 18361500]
25. Xing X, Fichera A, Kumar K. J Org Chem. 2002; 67:1722–1725. [PubMed: 11871915]
26. Weinges K, Kromm E. Liebigs Ann Chem. 1985:90–102.

27. Yamada S, Hongo C, Yoshioka R, Chibata I. J Org Chem. 1983; 48:843–846.

NIH-PA Author Manuscript

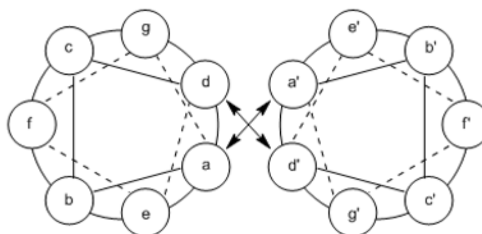
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

a b c d e f g a b c d e f g a b c d e f g
MRGSHHHHHHGSMA SGD**L**ENE VAQ**L**ERE VRS**L**EDE

a b c d e f g a b c d e f g a b c d e f g a b c d e f g
AAE**L**EQK VSR**L**KNE IED**L**KAE IGD**L**NNT SGIRRPAAKLN



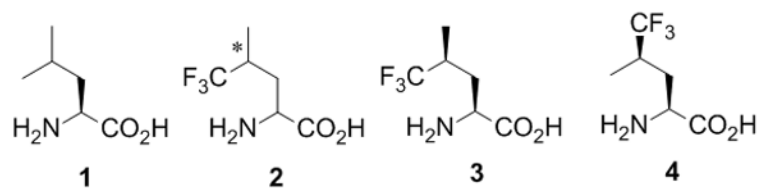


Figure 1.
A1 peptide sequence and helical wheel diagram of the dimer in which leucines are highlighted in bold. Structures of leucine (**1**) and the trifluoromethyl analogs (**2–4**). The asterisk in structure **2** denotes unresolved stereochemistry at the 4-position.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

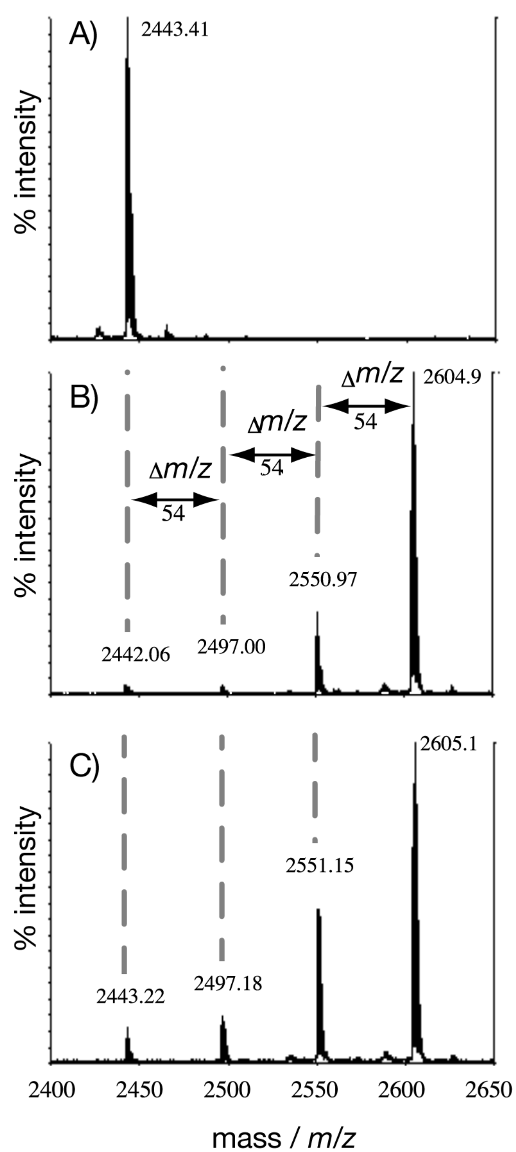


Figure 2.

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

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

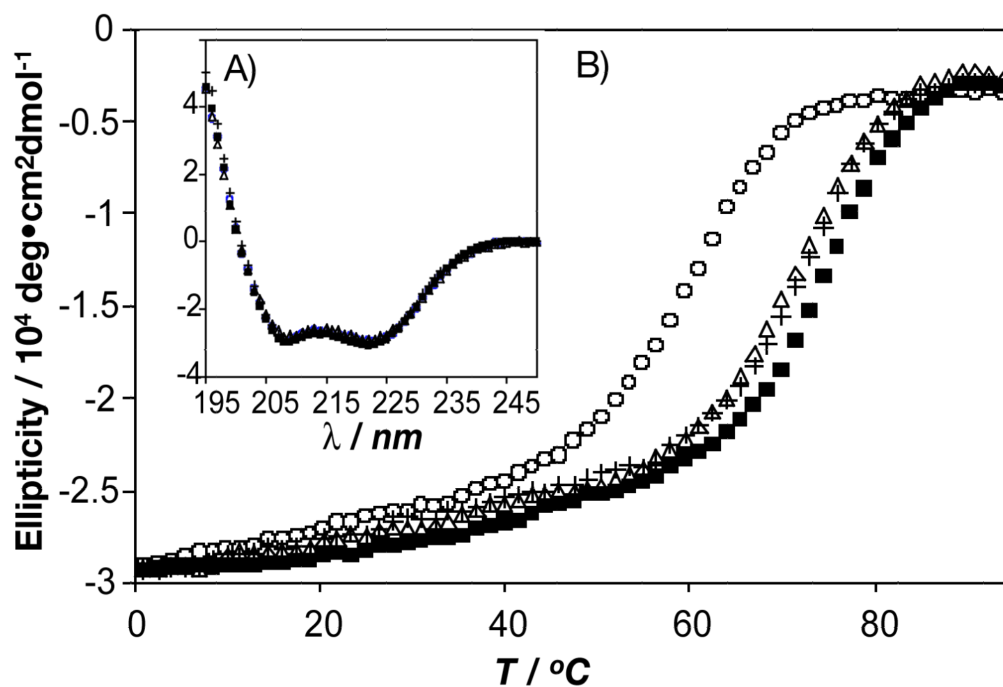


Figure 3.
CD spectra of Leu-A1 (○), SR-A1 (△), SS-A1 (+), and an equimolar mixture of SR-A1 and SS-A1 (■). A) Wavelength scan at 1 °C, 10 μM protein concentration, PBS buffer, pH 7.4. B) Thermal denaturation (1.5 °C interval, 1 minute equilibration time, 10 second averaging time) at 10 μM protein concentration,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

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

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.