Communications to the Editor

Protein Biosynthesis with Conformationally Restricted Amino Acids

David Mendel,† Jonathan Ellman,‡ and Peter G. Schultz*

Department of Chemistry
University of California
Berkeley, California 94720
Center for Advanced Materials
Lawrence Berkeley Laboratory
Berkeley, California 94720

Received January 19, 1993

The incorporation of conformationally constrained amino acids into peptides is a powerful approach for generating structurally defined peptides as conformational probes and bioactive agents.[1] Amino acids with the C_α proton replaced with an alkyl group,[2] *N*-alkyl amino acids,[3] and β-branched[4] amino acids have all been used extensively in the design of conformationally restricted peptides. The ability to site-specifically introduce constrained amino acids into large polypeptide chains would provide a similar opportunity to probe the flexibility, conformation, folding and stability of proteins. To this end, we have examined the competence of the *Escherichia coli* protein biosynthetic machinery to incorporate a number of these unnatural amino acids into the 164 residue protein T4 lysozyme (T4L).

Both *E. coli* and rabbit reticulocyte in vitro protein synthesis systems have been used to site-specifically incorporate unnatural amino acids into proteins[5,6] and peptides.[7,8] In order to test the ability of a variety of constrained amino acids (Figure 1) to function in polypeptide elongation, aminoacylated suppressor tRNAs were assayed for their ability to suppress the stop codon 5′-UAG-3′ substituted in place of the alanine-specific codon 5′-GCU-3′ at position 82 of the T4L structural gene.[5b] Results are shown in the autoradiogram in Figure 2. The band at 18.7 kDa in lane 1 is [^{35}S]methionine-labeled intact wild-type T4L expressed in vitro by plasmid pHSe54,97.TA.[5b-d] pT4LA82am-directed synthesis produced less than 2% full-length T4L when suppressor $tRNA_{CUA}$ was omitted from the reaction mixture or when this tRNA did not bear an amino acid (lanes 2 and 3, respectively). These controls reveal the in vitro protein synthesis extract to be free both of endogenous suppressor $tRNA_{CUA}$s and of synthetases capable of aminoacylating the added suppressor $tRNA_{CUA}$.

L-Alanine was incorporated at position 82 of T4L with reasonable efficiency (lane 4), producing about 26% as much protein as the wild-type gene (lane 1). α-Aminoisobutyric acid (AIBA) (**3**), cyclopropylglycine (**9**), and most L-α-methyl amino acids were incorporated with efficiencies greater than or equal to that of alanine: AIBA (**3**), 27% (lane 6); cyclopropylglycine (**9**), 37% (lane 7); α-methyl-L-2-aminobutyric acid (L-isovaline) (**4**), 53% (lane 11); α-methyl-L-leucine (**7**), 13% (lane 14); and α-methyl-L-phenylalanine (**13**), 52% (lane 16). In contrast, the D-α-methyl amino acids **5** (lane 12, 7%), **8** (lane 15, 1%), and **14** (lane 17, 3%) were incorporated poorly or not at all above background levels. Suppressor tRNA charged with D-alanine (**2**) functioned poorly in polypeptide elongation (≤6.8% suppression efficiency),[9] consistent with previous studies.[5b,6] α,α-Diethylglycine (**6**) (lane 13, 2%) was a poor acceptor, whereas the cyclic constrained analogues, cyclobutylglycine (**10**) (lane 8, 35%), cyclopentylglycine (**11**) (lane 9, 50%), and cyclohexylglycine (**12**) (lane 10, 43%) integrated with good efficiencies. However, when further steric bulk was added to the cyclopentyl ring of **11** to give 2-aminoindan-2-carboxylic acid (**15**) or to the cyclohexyl ring of **12** to give TEMPO derivative **18**, suppression was totally abolished (data not shown). The structurally bulky L-*tert*-leucine (**16**) was incorporated very efficiently (lane 18, 76%), whereas its enantiomer **17** afforded little full length protein (lane 19, 3.8%).

The above results clearly demonstrate that the protein biosynthetic machinery can accommodate a wide variety of conformationally constrained amino acids. The fact that AIBA (**3**) is incorporated into T4L with the same efficiency as alanine shows that the efficiency of biosynthetic peptide bond formation does not correspond to that of the nonenzymatic reaction. Somewhat surprisingly, the C_α methyl and cyclic α,α-disubstituted amino acids were, in general, incorporated significantly more efficiently than D-alanine. One explanation for this observation is that the side chain of an L-amino acid may play a role in positioning the α-amino group for the transacylation reaction. The D-amino acid might therefore be expected to bind in a nonproductive conformation, whereas the disubstituted amino acids, which retain the requisite substituent, function in protein biosynthesis. There appears to be a limit on the size of C_α substituents that can be incorporated into L-amino acids. For example, the α-ethyl-substituted L-amino acids **5** and **6** are poorly incorporated into T4L, whereas the cyclic amino acids **11** and **12** are efficiency

* Author to whom correspondence should be addressed at the University of California, Berkeley.

† Present address: Lilly Research Laboratories, Eli Lilly and Company, Lilly Corporate Center, Indianapolis, IN 46285-1543.

‡ Current position: Assistant Professor of Chemistry, University of California, Berkeley, CA 94720.

(1) (a) DeGrado, W. F. *Adv. Protein Chem.* **1988**, *39*, 51. (b) Hruby, V. J. *Life Sciences* **1982**, *31*, 189–199. (c) *Biochemistry of Peptide Antibiotics*; Kleinkauf, H., von Döhren, H., Eds.; Walter de Gruyter: Berlin, New York, 1990. (d) Gupta, S.; Krasnoff, S. B.; Roberts, D. W.; Renwick, J. A. A.; Brinen, L. S.; Clardy, J. *J. Am. Chem. Soc.* **1991**, *113*, 707–709. (e) *The Peptides: Analysis, Synthesis, Biology*; Hruby, V. J., Ed.; Academic Press: Orlando, FL, 1985; Vol. 7, pp 1–14.

(2) (a) Prasad, B. V.; Balaram, P. *CRC Crit. Rev. Biochem.* **1984**, *16*, 307. (b) Bosch, R.; Jung, G.; Schmitt, H.; Sheldrick, G. M.; Winter, W. *Angew. Chem., Int. Ed. Engl.* **1984**, *23*, 450. (c) Schiller, P. W.; Weltrowska, G.; Nguyen, T. M.-D.; Lemieux, C.; Chung, N. N.; Marsden, B. J.; Wilkes, B. C. *J. Med. Chem.* **1991**, *34*, 3125–3132. (d) Toniolo, C.; Crisma, M.; Bonora, G. M.; Benedetti, E.; Di Blasio, B.; Pavone, V.; Pedone, C.; Santini, A. *Biopolymers* **1991**, *31*, 129–138. (e) Paterson, Y.; Rumsey, S. M.; Benedetti, E.; Nemethy, G.; Scheraga, H. A. *J. Am. Chem. Soc.* **1981**, *103*, 2947–2955.

(3) (a) Vitoux, B.; Aubry, A.; Chung, M. T.; Marrand, M. *Int. J. Pept. Protein Res.* **1986**, 17, 469. (b) Manavalan, P.; Momany, F. A. *Biopolymers* **1980**, *19*, 1943.

(4) (a) Lyu, P. C.; Sherman, J. C.; Chen, A.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 5317. (b) Patterson, Y.; Leach, S. J. *Macromolecules* **1978**, *11*, 409. (c) Stokrova, S.; Pospisek, J.; Sponer, J.; Blaha, K. *Collect. Czech. Chem. Commun.* **1985**, *50*, 2925.

(5) (a) Chung, H. H.; Benson, D. R.; Schultz, P. G. *Science* **1993**, *259*, 806–809. (b) Ellman, J. A.; Mendel, D.; Schultz, P. G. *Science* **1992**, *255*, 197–200. (c) Mendel, D.; Ellman, J. A.; Schultz, P. G. *J. Am. Chem. Soc.* **1991**, *113*, 2758–2760. (d) Mendel, D.; Ellman, J. A.; Chang, Z.; Veenstra, D. L.; Kollman, P. A.; Schultz, P. G. *Science* **1992**, *256*, 1798–1802. (e) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182–188. (f) Judice, J. K.; Murphy, E.; Schultz, P. G. *Science*, submitted for publication.

(6) (a) Ellman, J.; Mendel, D.; Anthony-Cahill, S.; Noren, C. J.; Schultz, P. G. *Methods Enzymol.* **1991**, *202*, 301–336. (b) Noren, C. J.; Anthony-Cahill, S. J.; Suich, D. J.; Noren, K. A.; Griffith, M. C.; Schultz, P. G. *Nucleic Acids Research* **1990**, *18*, 83–88. (c) Robertson, S. A.; Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Nucleic Acids Res.* **1989**, *17*, 9649–9660. (d) Robertson, S. A.; Ellman, J. A.; Schultz, P. G. *J. Am. Chem. Soc.* **1991**, *113*, 2722–2729.

(7) Hecht, S. M. *Acc. Chem. Res.* **1992**, *25*, 545–552.

(8) (a) Bain, J. D.; Glabe, C. G.; Dix, T. A.; Chamberlin, A. R.; Diala, E. S. *J. Am. Chem. Soc.* **1989**, *111*, 8013–8014. (b) Bain, J. D.; Diala, E. S.; Glabe, C. G.; Wacker, D. A.; Lyttle, M. H.; Dix, T. A.; Chamberlin, A. R. *Biochemistry* **1991**, *30*, 5411–5421. (c) Bain, J. D.; Wacker, D. A.; Kuo, E. E.; Chamberlin, A. R. *Tetrahedron* **1991**, *47*, 2389–2400.

(9) Some racemization of L-alanine may have occurred during the aminoacylation/translation process.

0002-7863/93/1515-4359$04.00/0

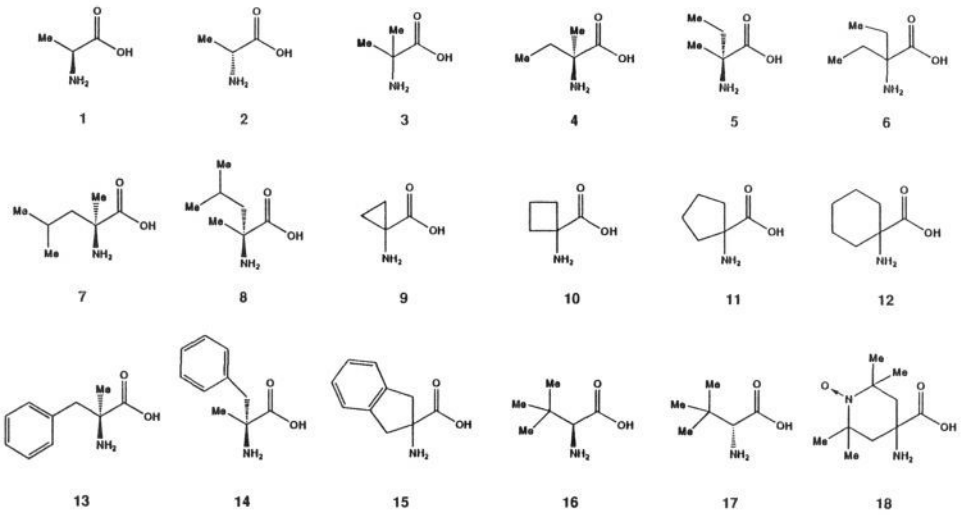

Figure 1. Amino acids[12,13] tested for incorporation in response to the termination codon 5′-UAG-3′ at position 82 in the structural gene of T4L.

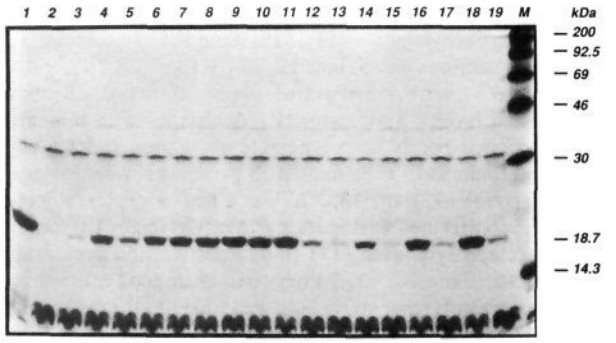

Figure 2. Autoradiogram of in vitro suppression reactions labeled with [^{35}S]-L-methionine. Lane 1 shows expression of wild-type T4L by plasmid pHSe54,97.TA without $tRNA_{CUA}$; T4L synthesis in lanes 2–19 was directed by plasmid pT4LA82am with 10-μg additions of aminoacyl suppressor tRNAs bearing the following amino acids: lane 2, no added suppressor; lane 3, 10-μg full-length unacyl suppressor; lane 4, **1**; lane 5, **2**; land 6, **3**; lane 7, **9**; lane 8, **10**; lane 9, **11**; lane 10, **12**; lane 11, **4**; lane 12, **5**; lane 13, **6**; lane 14, **7**; lane 15, **8**; lane 16, **13**; lane 17, **14**; lane 18, **16**; lane 19, **17**. Lane M contains a ^{14}C-methylated molecular mass standard set. The other major product of pT4LA82am-directed T4L synthesis is truncated T4L, residues 1–81, which appears as a band near the bottom of the gel.

incorporated. β-Branching, however, appears to be readily accommodated by the translational machinery since L-*tert*-leucine was incorporated with the highest efficiency of any amino acid assayed.[10] The extent to which peptidyl transferase, elongation factor, and context effects (either peptide or nucleotide sequence) influence suppression efficiencies remains unknown. However, AIBA (**3**) was incorporated with poor efficiency *relative* to L-alanine at position 15 of ras protein,[11] suggesting that context effects will play an important role in determining relative suppression efficiencies of amino acids at a given site.

The expansion of structural motifs that can be biosynthetically incorporated into proteins to include a large number of conformationally constrained amino acids significantly increases the power of mutagenesis methods as probes of protein structure and function and provides additional insights into the steric requirements of the translational machinery.

Acknowledgment. We are grateful for financial support for this work from the National Institutes of Health (Grant No. R01 GM41679) and by the Director, Office of Energy Research, Office of Basic Energy Sciences, Division of Material Sciences, and also by the Division of Energy Biosciences of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098 (D.M.). D.M. was supported by American Cancer Society Postdoctoral fellowship PF-4014A. We are grateful to Peter W. Schiller for the generous gift of 2-aminoindan-2-carboxylic acid and to Nariman Naber and Roger Cooke (University of California, San Francisco) for the TEMPO derivative **18**.

(10) L-Valine has also been observed to suppress more efficiently than alanine at a number of sites in T4L (data not shown).

(11) Danley, L.; Schultz, P. G., unpublished results.

(12) (a) Naef, R.; Seebach, D. *Helv. Chim. Acta* **1985**, *68*, 135–143. (b) Seebach, D.; Aebi, J. D.; Naef, R.; Weber, T. *Helv. Chim. Acta* **1986**, *68*, 144–154.

(13) Mendel, D.; Chung, H.-H.; Schultz, P. G., manuscript in preparation.