J|A|C|S
COMMUNICATIONS

Published on Web 09/18/2004

Amino Acid Backbone Specificity of the *Escherichia coli* Translation Machinery

Zhongping Tan,[†] Anthony C. Forster,[‡] Stephen C. Blacklow,[‡] and Virginia W. Cornish*,[†]

Department of Chemistry, Columbia University, 3000 Broadway, New York, New York 10027, and Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115

Received May 12, 2004; E-mail: vc114@columbia.edu

The ribosome is unique among Nature's biosynthetic machines in that the catalytic center is distinct from the substrate recognition pocket. This separation of catalysis and substrate binding suggests the ribosome may be uniquely well suited to manipulation for the synthesis of novel polymers. Indeed, in 1971, Rich and Fahnestock demonstrated that the protein biosynthetic machinery could synthesize polymers containing a random mixture of ester and amide linkages using mis-acylated hydroxy-Phe-tRNA^{Phe}.[1] With the development of an efficient method for the chemoenzymatic synthesis of aa-tRNAs,[2] several groups then used misacylated suppressor tRNAs or A/P-site substrates to test the ability of the protein biosynthetic machinery to accept backbone analogues, including α-hydroxy acids,[3a−h] *N*-methyl amino acids,[3a−c,i,j] α,α-disubstituted amino acids,[3c,k] *β*-amino acids,[3a,c,l−o] and D-amino acids.[3a−c,p−t] While many of these analogues could be incorporated in response to a stop codon, the yields for incorporation generally were low.[3] Recently, we demonstrated that a sense codon reassignment strategy with a pure translation system allowed translation of multiple, adjacent sense codons with synthetic acyl-tRNA substrates.[4] This approach and related approaches breaking codon degeneracy[5] open the possibility of using the protein biosynthetic machinery for template encoded synthesis of novel backbone polymers of defined length and composition with a pool of synthetic acyl-tRNAs. The question, however, is whether the substrate tolerance and, in particular, poor yields reported previously for backbone analogues reflect use of a suppressor tRNA, which must compete with endogenous release factors for translation of a stop codon, processing by endogenous aa-tRNA synthetases or other metabolic enzymes, such as D-aa-deacylase, or the intrinsic specificity of the protein biosynthetic machinery. Here, as a further step toward polymer synthesis with sense codon reassignment, we determined the relative yields of peptides biosynthesized using a series of backbone analogues at a sense codon in a pure *E. coli* translation system (Figure 1). For efficient incorporation of any analogue, the acyl-tRNAs must be able to bind to EF-Tu, be delivered to the ribosome, and function well in both peptide bond formation and translocation.

To isolate the effects of the protein biosynthetic machinery, a series of backbone analogues based on L-Ala and L-Phe were scored for their yield of synthesis of the tripeptide fMet-U-Glu using an *E. coli* tRNA^{Asn}-based tRNA adaptor for translation of a GUU (Val) codon (termed tRNA$^{AsnB}_{GAC}$, Figure 2). To compare with the prior literature, α-hydroxy acids, *N*-alkyl amino acids, α,α-disubstituted amino acids, *β*-amino acids, and D-amino acids were all tested (Figure 2A). Because previous reports showed that Phe analogues are incorporated with higher yields than Ala analogues,[7] both Ala and Phe variants of these analogues were included as controls for the effect of the amino acid R-group. For these analogues, we used

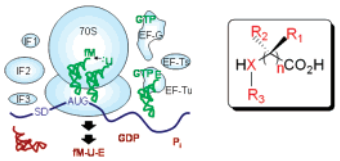

Figure 1. Sense codon translation with amino acid backbone analogues. The purified translation system[4,6] is depicted incorporating unnatural analogue (U) into tripeptide product fM-U-E. *E. coli* served as the source of all components: initiation factors, IF1, IF2, and IF3; elongation factors, EF-Tu, EF-Ts, and EF-G; ribosome and natural aa-tRNAs. SD is a Shine−Dalgarno ribosome binding sequence.

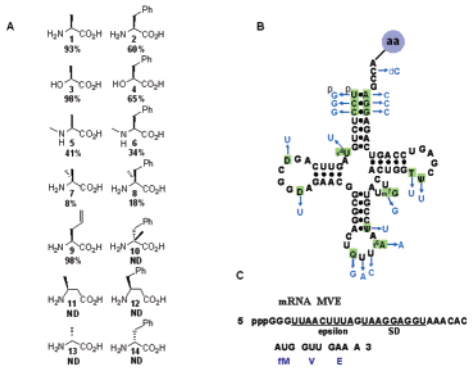

Figure 2. Backbone analogue incorporation via a synthetic tRNA$^{AsnB}_{GAC}$ adaptor at the sense codon GUU (Val). (A) Peptide yields for each analogue as determined in triplicate on the basis of incorporation of ^{3}H-Glu (ND, not detectable above background). The yields for **1**−**11** were calculated from the Dowex assay; the yields for **12**−**14** are from the HPLC assay. (B) Natural *E. coli* tRNA^{Asn} (black) and its synthetic counterpart tRNA$^{AsnB}_{GAC}$ (blue). (C) Complete mRNA sequence and encoded natural translation tripeptide product (blue). Epsilon is a translation enhancer sequence.[6] SD is a Shine−Dalgarno ribosome binding sequence.

a designed *E. coli* tRNA$^{AsnB}_{GAC}$, which gave high yields in our previous report (Figure 2B).[4] The unnatural aa-tRNA substrates were prepared as reported[2] from aa-pdCpA and tRNA^{-CA}, and the aa-tRNA ligation products were confirmed by acid urea polyacrylamide gel electrophoresis (Figure S3, Supporting Information). Natural *E. coli* aa-tRNAs were used at the *N*-terminus (fMet-tRNA^{fMet}) and *C*-terminus (^{3}H-Glu-tRNA^{Glu}). Importantly, all of the components of the pure translation system used here—the ribosome, the six recombinant his-tagged protein factors, and the tRNAs—are from *E. coli*,[6] ruling out factor incompatibility as a reason for low yields. Synthesis of a tripeptide made it possible to confirm peptide yields and product identity by comigration with authentic marker on analytical HPLC. Initially, peptide yields were scored on the basis of the incorporation of ^{3}H-Glu into full-length fMet-U-Glu. A 30 min translation reaction was terminated by the addition of NaOH, and the *N*-formylated or hydroxy peptide products were separated from free amino acids and other components by passage through a Dowex cation exchange column.[8] The

[†] Columbia University.
[‡] Harvard Medical School.

10.1021/ja0472174 CCC: $27.50

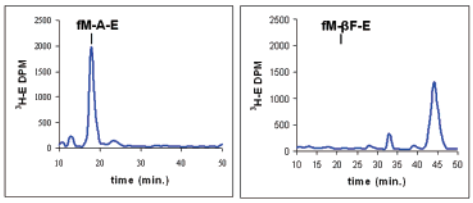

Figure 3. Analyzing of the translation products fM-A-E (93%) and fM-βF-E (ND) by HPLC comigration with authentic standard. The elution positions of the marker peptides (black) prepared by solid-phase peptide synthesis are indicated above the chromatogram of ^{3}H-radiolabeled translation products (blue).

yields for acyl-tRNA analogues are reported relative to that for Val-tRNA^{Val}, a natural substrate that has been well characterized in our purified translation system and verified to give high peptide yields. Unacylated tRNA$^{AsnB}_{GAC}$ was used to define background counts. The high-yielding amino acid "eU" (Figure 2A, **9**) was also included for comparison to our previous report. All the translation experiments were carried out in triplicate, and the standard deviations were low (Figure S4, Supporting Information). The product yields and identities for the peptide products corresponding to **1**, **2**, **5**, **7**, **8**, **9**, **12**, **13**, and **14** were verified on analytical HPLC by comparison with authentic product prepared by solid-phase peptide synthesis. Except for the β- and D-amino acids as detailed below, the peptide yields calculated from the HPLC comigration assay were in agreement with those of the Dowex assay (Figure 3; Figures S4 and S5, Supporting Information).

The picture of the substrate tolerance of the protein biosynthetic machinery obtained here from translating a sense codon without competitors is very similar to that from the suppressor tRNA and other prior literature. The hydroxy acids and *N*-methyl amino acids are incorporated with good yield, and α,α-disubstituted amino acids are incorporated with poor yield, while the β- and D-amino acids are not detectable. Control translation reactions with two other mRNA templates suggest that the results with these substrates are representative and general (Figures S6 and S7, Supporting Information), but further studies are underway to understand other factors that influence analogue incorporation. Product characterization by analytical HPLC shed light on previous conflicting reports about β-amino acids[3a,c,l−o] and D-amino acids.[3a−c,p−t] Initial Dowex ^{3}H counts suggested β-Phe incorporated with 15% yield and D-Ala with 37% yield (Figure S4). However, further analysis by HPLC comigration with authentic product and related standards showed no β-Phe and D-Ala product, but instead a major product of unknown structure for β-Phe and what appears to be L-Ala incorporation based on co-migration with fMet-(L)-Ala-Glu for D-Ala (Figure S5). A caveat is that the sensitivities of these translation assays are only ca. $\geq$5% yield and so cannot rule out that an analogue is just a very poor substrate. Encouragingly, Hecht and co-workers recently showed that we may not be limited by the natural substrate specificity of the ribosomal machinery, demonstrating that mutation of the ribosome in the peptidyl transferase center and helix 89 regions allowed incorporation of D-amino acids.[3s]

The rank ordering and product yields of the good substrates, however, differ significantly from those observed with suppressor tRNAs. With suppressor tRNAs, hydroxy acids generally gave yields much lower than those for their amino acid counterparts in vitro, with yields in the low 20% region.[3a,d−g] Here, with translation of a sense codon in a purified system, the yields for hydroxy acids are greatly improved, comparable to or greater than those for the corresponding amino acids. The yields for *N*-methyl amino acids are also improved, which is in agreement with recent results for *N*-methyl amino acids.[3i,j] Interestingly, the yields for the Ala analogues generally are higher than those for the Phe analogues, in contrast to the higher yields for Phe analogues observed with the yeast tRNA$^{Phe}_{CUA}$ suppressor.[7]

These results further support the substrate plasticity of the protein biosynthetic machinery. They confirm hydroxy acids and *N*-methyl amino acids as immediate candidates for ribosomally encoded synthesis of backbone polymers, and encouragingly show good yields for these analogues with sense translation. Currently, we are working toward ribosome-catalyzed synthesis of *N*-alkyl amino acid polymers. These results also speak to the role of the amino acid in the aa-tRNA recognition process. The ranking order of good substrates and Ala vs Phe analogues differs here with a tRNA$^{AsnB}_{GAC}$ adaptor from that seen previously with a tRNA$^{Phe}_{CUA}$ suppressor. This observation provides further support for the intriguing possibility that analogue yields can be improved by tRNA paring and that there is cross-talk between the amino acid and tRNA body/anticodon in aa-tRNA decoding on the ribosome.[9]

Acknowledgment. This work is supported by a fellowship from the Sloan Foundation to V.W.C. and an Army Idea Award from the Department of Defense (DAMD17-00-1-0163 to S.C.B.). V.W.C. is a Sloan Foundation Fellow, S.C.B. is a Pew Scholar and an Established Investigator of the American Heart Association, and A.C.F. is supported by a K08 grant from the NIH.

Supporting Information Available: Preparation of acyl-tRNA$^{AsnB}_{GAC}$'s, in vitro translation reactions, product characterization by analytical HPLC, and mRNA dependency experiments. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Fahnestock, S.; Rich, A. *Science* **1971**, *173*, 340−343.
(2) (a) Heckler, T. G.; Chang, L. H.; Zama, Y.; Naka, T.; Chorghade, M. S.; Hecht, S. M. *Biochemistry* **1984**, *23*, 1468−1473. (b) Baldini, G.; Martoglio, B.; Schachenmann, A.; Zugliani, C.; Brunner, J. *Biochemistry* **1988**, *27*, 7951−7959. (c) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182−188. (d) Bain, J. D.; Diala, E. S.; Glabe, C. G.; Dix, T. A.; Chamberlin, A. R. *J. Am. Chem. Soc.* **1989**, *111*, 8013−8014. (e) Robertson, S. A.; Ellman, J. A.; Schultz, P. G. *J. Am. Chem. Soc.* **1991**, *113*, 2722−2729.
(3) (a) Bain, J. D.; Wacker, D. A.; Kuo, E. E.; Chamberlin, A. R. *Tetrahedron* **1991**, *47*, 2389−2400. (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) Ellman, J. A.; Mendel, D.; Schultz, P. G. *Science* **1992**, *255*, 197−200. (d) Koh, J. T.; Cornish, V. W.; Schultz, P. G. *Biochemistry* **1997**, *36*, 11314−11322. (e) England, P. M.; Zhang, Y.; Dougherty, D. A.; Lester, H. A. *Cell* **1999**, *96*, 89−98. (f) England, P. M.; Lester, H. A.; Dougherty, D. A. *Tetrahedron Lett.* **1999**, *40*, 6189−6192. (g) England, P. M.; Lester, H. A.; Dougherty, D. A. *Biochemistry* **1999**, *38*, 14409−14415. (h) Lu, T.; Ting, A. Y.; Mainland, J.; Jan, L. Y.; Schultz, P. G.; Yang, J. *Nat. Neurosci.* **2001**, *4*, 239−246. (i) Frankel, A.; Millward, S. W.; Roberts, R. W. *Chem. Biol.* **2003**, *10*, 1043−1050. (j) Merryman, C.; Green, R. *Chem. Biol.* **2004**, *11*, 575−582. (k) Mendel, D.; Ellman, J.; Schultz, P. G. *J. Am. Chem. Soc.* **1993**, *115*, 4359−4360. (l) Heckler, T. G.; Zama, Y.; Naka, T.; Hecht, S. M. *J. Biol. Chem.* **1983**, *258*, 4492−4495. (m) Roesser, J. R.; Xu, C.; Payne, R. C.; Surratt, C. K.; Hecht, S. M. *Biochemistry* **1989**, *28*, 5185−5195. (n) Killian, J. A.; Van Cleve, M. D.; Shayo, Y. F.; Hecht, S. M. *J. Am. Chem. Soc.* **1998**, *120*, 3032−3034. (o) Eisenhauer, B. M.; Hecht, S. M. *Biochemistry* **2002**, *41*, 11472−11478. (p) Calendar, R.; Berg, P. *J. Mol. Biol.* **1967**, *26*, 39−54. (q) Yamane, T.; Miller, D. L.; Hopfield, J. J. *Biochemistry* **1981**, *20*, 7059−7064. (r) Roesser, J. R.; Chorghade, M. S.; Hecht, S. M. *Biochemistry* **1986**, *25*, 6361−6365. (s) Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M. *J. Am. Chem. Soc.* **2003**, *125*, 6616−6617. (t) Starck, S. R.; Qi, X.; Olsen, B. N.; Roberts, R. W. *J. Am. Chem. Soc.* **2003**, *125*, 8090−8091.
(4) Forster, A. C.; Tan, Z.; Nalam, M. N. L.; Lin, H.; Qu, H.; Cornish, V. W.; Blacklow, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6353−6357.
(5) (a) Kwon, I.; Kirshenbaum, K.; Tirrell, D. A. *J. Am. Chem. Soc.* **2003**, *125*, 7512−7513. (b) Frankel, A.; Roberts, R. W. *RNA* **2003**, *9*, 780−786.
(6) Forster, A. C.; Weissbach, H.; Blacklow, S. C. *Anal. Biochem.* **2001**, *297*, 60−70.
(7) Cornish, V. W.; Mendel, D.; Schultz, P. G. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 621−633.
(8) Peacock, S.; Weissbach, H.; Nash, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 6009−6013.
(9) (a) Cload, S. T.; Liu, D. R.; Froland, W. A.; Schultz, P. G. *Chem. Biol.* **1996**, *3*, 1033−1038. (b) LaRiviere, F. J.; Wolfson, A. D.; Uhlenbeck, O. C. *Science* **2001**, *294*, 165−168.

JA0472174