Specificity of Translation for N-Alkyl-Amino Acids

Baolin Zhang, Zhongping Tan, Lucas Gartenmann Dickson, Madhavi N.L. Nalam,
Virginia W. Cornish, and Anthony C. Forster

Supporting Information

General Information. A, P and **hP** were purchased from Aldrich. **F**, **NMA** and **NMF** were purchased from Bachem. **NBA** and **NBF** were synthesized as published.[1] Dowex 50X8-200 ion-exchange resin was from Acros. ^{1}H NMR spectra were recorded on a Bruker DPX-400 (400MHz) spectrometer and are reported in ppm using $CDCl_3$ and D_2O (Cambridge Isotope Laboratories) as the solvent. NMR data are reported as follows: s, singlet; d, doublet; t, triplet; q, quartet; quin, quintet; m, multiplet; dd, doublet of doublets. Analytical HPLC was performed on a Waters 600 using a LiChroCART 250-4, RP18 column. Preparative HPLC was performed on a Waters 600 using a Whatman Partisil 10 ODS-2 (C18) column.

***N*-NVOC-aminoacyl-pdCpA syntheses.** The pdCpA dinucleotide was synthesized using standard nucleotide chemistry.[2] Briefly, 6-*N*, 6-*N*, 2'-*O*, 3'-*O*-tetrabenzoyl adenosine was prepared by transiently protecting the 5' hydroxyl of adenosine with dimethoxytrityl. The protected adenosine was then coupled to the 2'-deoxycytidinylphosphoramidite and oxidized under standard conditions. A phosphoramidite was then added to the 5' hydroxyl group of deoxycytidine and subsequently oxidized. Finally, the benzoyl and cyanoethyl protecting groups were removed under basic conditions, and the pdCpA was purified by HPLC and "activated" as the tetrabutylammonium salt.

The natural amino acids and their analogs were prepared as the *N*-nitroveratryloxycarbonyl (NVOC) cyanomethyl active ester.[2] Briefly, the NVOC protecting group was installed using 6-nitroveratryl chloroformate under standard conditions[3,4], and then the cyanomethyl group was introduced using chloroacetonitrile and triethylamine as the base. The cyanomethyl active ester was then used to selectively acylate pdCpA at the 2'/3' hydroxyl group using the tetrabutylammonium salt of pdCpA in anhydrous DMF at room temperature. The reaction process was monitored by comparing the pdCpA peak and the NVOC-acyl-pdCpA peak on analytical HPLC. After the ratio of these two peaks was 1:1, the reaction was terminated by adding 4:1 NH_4Ac (50 mM, pH 4.5):acetonitrile, and the NVOC-acyl-pdCpA was purified using preparative HPLC. The resulting NVOC-acyl-pdCpA was dissolved in ddH_2O and the concentration was calculated by UV absorption at 260nm and 350nm (_260 for pdCpA is 25,000 cm^{-1} M^{-1}; ε260 and _350 for 6-nitroveratryl group are 2140 cm^{-1} M^{-1} and 6336 cm^{-1} M^{-1}, respectively).[2]

***N*-NVOC-L-Ala Cyanomethyl Ester.** Yellow solid (60% yield): ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.72(s, 1H, Ar-**H**), 6.99(s, 1H, Ar-**H**), 5.54(dd, *J* = 52, 15 Hz, 2H, Ar-C**H**$_2$), 5.32(d, *J* = 7 Hz, 1H, N-**H**), 4.79(dd, *J* = 47, 16 Hz, 2H, NC-C**H**$_2$), 4.97(quin, *J* = 8 Hz, 1H, NH-C**H**), 4.01(s, 3H, Ar-OC**H**$_3$), 3.96(s, 3H, Ar-OC**H**$_3$), 1.51(d, *J* = 7 Hz, 1H, CH-C**H**$_3$).

***N*-NVOC-L-Ala-pdCpA.** Yellow solid. LRMS (FAB+), *m/z* calculated from $C_{32}H_{40}N_{10}O_{20}P_2$ 946, found $(M+H)^+$ 947, $(M+Na)^+$ 969.

***N*-NVOC-L-Phe Cyanomethyl Ester.** Yellow solid (84% yield): ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.69(s, 1H, Ar-**H**), 7.24-7.31 (m, 3H, Ph-**H**), 7.13(d, *J* = 7 Hz, 2H, Ph-**H**), 6.90(s, 1H, Ar-**H**), 5.49(dd, *J* = 34,

15 Hz, 2H, Ar-C**H**$_2$), 5.24(d, *J* = 8 Hz, 1H, N-**H**), 4.66-4.77(m, 3H, NC-C**H**$_2$ and NH-C**H**), 3.93(s, 3H, Ar-OC**H**$_3$), 3.92(s, 3H, Ar-OC**H**$_3$), 3.12-3.16(m, 2H, Ph-C**H**$_2$).

***N*-NVOC-L-Phe-pdCpA.** Yellow solid. MALDI-TOF, *m/z* calculated from $C_{38}H_{44}N_{10}O_{20}P_2$ 1022, found $(M+Na)^+$ 1045.

***N*-NVOC-L-Pro Cyanomethyl Ester.** Yellow solid (55% yield): ^{1}H NMR (400 MHz, CDCl3) δ 7.24+7.72(s, 1H, Ar-**H**), 7.03+6.99(s, 1H, Ar-**H**), 5.70-5.58(m, 1H, Ar-C**H**), 5.51-5.42(m, 1H, Ar-C**H**), 4.90-4.68(m, 2H, NC-C**H**$_2$), 4.52-4.45(m, 1H, N-C**H**), 4.04(s, 3H, Ar-OC**H**$_3$), 3.98(s, 3H, Ar-**OCH**$_3$), 3.70-3.58(m, 2H, N-C**H**$_2$), 2.38-2.32(m, 1H, N-CH-C**H**), 2.13-2.03(m, 3H, N-CH-C**H** and N-CH-CH$_2$-C**H**$_2$).

***N*-NVOC-L-Pro-pdCpA.** Yellow solid. LRMS (FAB+), *m/z* calculated from $C_{34}H_{42}N_{10}O_{20}P_2$ 972, found $(M+Na)^+$ 996.

N-NVOC-(4R)-4-OTBDMS-L-Pro Cyanomethyl Ester. Yellow foam (47% yield): ^{1}H NMR (400 MHz, CDCl3) δ 7.69+7.68(s, 1H, Ar-**H**), 6.96+6.95(s, 1H, Ar-**H**), 5.64-5.41(m, 2H, Ar-C**H**$_2$), 4.87-4.65(m, 2H, NC-C**H**$_2$), 4.54-4.41(m, 2H, N-C**H** and TBDMSO-C**H**), 4.00-3.93(m, 6H, Ar-OC**H**$_3$), 3.68-3.51(m, 2H, N-**CH**$_2$), 2.27-2.05(m, 2H, N-CH-C**H**$_2$), 0.85(s, 9H, Si-C-C**H**$_3$), 0.08-0.05(m, 6H, Si-C**H**$_3$).
Note: The two conformers of the tertiary amine exchange slowly and are diastereomers. The ^{1}H NMR data shown here is the spectra of the mixture of two diastereomers. The amino acid was carried forward as a mixture of conformers.

N-NVOC-(4R)-4-OH-L-Pro-pdCpA. Yellow solid. LRMS (FAB+), m/z calculated from $C_{35}H_{43}N_9O_{21}P_2$ 988, found $(M+H)^+$ 989, $(M+Na)^+$ 1011.

***N*-NVOC-*N*-Me-L-Ala Cyanomethyl Ester.** Yellow solid (83% yield): ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.69+7.71(s, 1H, Ar-**H**), 6.94+7.00(s, 1H, Ar-**H**), 5.51+5.57(dd, *J* = 47, 15 Hz, and *J* = 72, 14, 2H, Ar-C**H**$_2$), 4.69-4.90(m, 3H, NC-C**H**$_2$ and N-C**H**), 3.95+4.00(s, 6H, Ar-OC**H**$_3$), 3.01+2.97(s, 3H, N-C**H**$_3$), 1.51+1.52(s, 3H, CH-C**H**$_3$).
Note: The two conformers of the tertiary amine exchange slowly and are diastereomers. The ^{1}H NMR data shown here is the spectra of the mixture of two diastereomers. The amino acid was carried forward as a mixture of conformers.

***N*-NVOC-*N*-Me-L-Ala-pdCpA.** Yellow solid. LRMS (FAB+), *m/z* calculated from $C_{33}H_{42}N_{10}O_{20}P_2$ 960, found $(M+H)^+$ 961, $(M+Na)^+$ 983.

***N*-NVOC-*N*-Me-L-Phe Cyanomethyl Ester.** Yellow solid (62% yield): ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.70(s, 1H, Ar-**H**), 7.15-7.29 (m, 5H, Ph-**H**), 6.89+6.93(s, 1H, Ar-**H**), 5.32-5.57(m, 2H, Ar-C**H**$_2$), 4.74-4.94(m, 3H, NC-C**H**$_2$ and N-C**H**), 3.90+3.95(s, 6H, Ar-OC**H**$_3$), 3.09-3.41(m, 2H, Ph-C**H**$_2$), 2.84+2.87(s, 3H, N-C**H**$_3$).
Note: The two conformers of the tertiary amine exchange slowly and are diastereomers. The ^{1}H NMR data shown here is the spectra of the mixture of two diastereomers. The amino acid was carried forward as a mixture of conformers.

***N*-NVOC-*N*-Me-L-Phe-pdCpA.** Yellow solid. LRMS (FAB+), *m/z* calculated from $C_{39}H_{46}N_{10}O_{20}P_2$ 1036, found $(M+H)^+$ 1037, $(M+Na)^+$ 1059.

***N*-NVOC-*N*-Bu-L-Ala Cyanomethyl Ester.** Yellow foam (41% yield): ^{1}H NMR (400 MHz, CDCl3) δ 7.72+7.70(s, 1H, Ar-**H**), 6.98(s, 1H, Ar-**H**), 5.61-5.40(m, 2H, Ar-C**H**$_2$), 4.80-4.70(m, 2H, NC-C**H**$_2$), 4.37-4.25(m, 1H, N-C**H**), 4.02-3.96(m, 6H, Ar-OC**H**$_3$), 3.48-3.43+3.26-3.22(m, 2H, N-**C**H$_2$), 1.64-1.53(m, 5H, N-CH$_2$-C**H**$_2$+ N-CH-C**H**$_3$), 1.38-1.26(m, 2H, N-CH$_2$-CH$_2$-C**H**$_2$), 0.95-0.92(m, 3H, N-CH$_2$-CH$_2$-CH$_2$-C**H**$_3$).
Note: The two conformers of the tertiary amine exchange slowly and are diastereomers. The ^{1}H NMR data shown here is the spectra of the mixture of two diastereomers. The amino acid was carried forward as a mixture of conformers.

***N*-NVOC-*N*-Bu-L-Ala-pdCpA.** Yellow solid. LRMS (FAB+), *m/z* calculated from $C_{36}H_{48}N_{10}O_{20}P_2$ 1002, found $(M+H)^+$ 1003, $(M+Na)^+$ 1025.

***N*-NVOC-*N*-Bu-L-Phe Cyanomethyl Ester.** Yellow foam (47% yield): ^{1}H NMR (400 MHz, CDCl3) δ 7.72 (s, 1H, Ar-**H**), 7.29-7.06(m, 5H, Ar-**H**), 6.96(s, 1H, Ar-**H**), 5.66-5.45(m, 2H, Ar-C**H**$_2$), 4.78-4.70(m, 2H, NC-C**H**$_2$), 4.22-4.18(m, 1H, N-C**H**), 4.02-3.96(m, 6H, Ar-OC**H**$_3$), 3.40-3.28(m, 2H, N-**C**H$_2$), 3.26-3.19+2.71-2.61(m, 2H, N-CH-C**H**$_2$), 1.33-1.24(m, 2H, N-CH$_2$-C**H**$_2$), 1.17-1.12(m, 2H, N-CH$_2$-CH$_2$-C**H**$_2$), 0.83-0.78(m, 3H, N-CH$_2$-CH$_2$-CH$_2$-C**H**$_3$).
Note: The two conformers of the tertiary amine exchange slowly and are diastereomers. The ^{1}H NMR data shown here is the spectra of the mixture of two diastereomers. The amino acid was carried forward as a mixture of conformers.

***N*-NVOC-*N*-Bu-L-Phe-pdCpA.** Yellow solid. LRMS (FAB+), *m/z* calculated from $C_{42}H_{52}N_{10}O_{20}P_2$ 1078, found $(M+H)^+$ 1079, $(M+Na)^+$ 1101.

Synthetic tRNAs. Templates for tRNAs were constructed by cloning oligodeoxyribonucleotides or PCR-amplified oligos such that the sense strands of $tRNA^{PheB}$, $tRNA^{AsnB}$ and $tRNA^{AlaB}$ (sequences shown in Fig. 2 left) were flanked by the T7 RNA polymerase promoter and HindIII, BstNI, FokI and EcoRI sites precisely as follows:

...AAGCTTAATACGACTCACTATA**-tRNA-**GGTGATCCATCCGAATTC**...**

The inserts cut at the underlined sites were ligated into HindIII/EcoRI cut pUC19 (AlaB), pUC18 (PheB) or both (AsnB). pUC18 was the vector of choice because this force-cloned the tRNA into the orientation that is not transcribed by *E. coli* RNA polymerase (a precautionary measure aimed at preventing expression of mutant tRNA precursor RNA, which might be processed by tRNA processing enzymes into toxic tRNA). Templates for tRNAs lacking their 3'-terminal CA sequence were prepared by digestion with FokI, while control templates for production of full-length tRNAs were prepared by digestion of the same plasmids with BstNI. Transcriptions included 20 mM neutralized GMP to ensure that virtually all transcripts had the native 5' monophosphate.

Synthetic amino-/N-alkyl-amino-acyl-tRNAs. $tRNA^{minusCA}$ species were ligated to an NVOC-aminoacyl-pdCA by T4 RNA ligase. Concentrations of all 24 unnatural NVOC-aminoacyl-tRNA

substrates were estimated by urea polyacrylamide gel electrophoresis at pH 5. Only efficient ligations were kept for later photodeprotection.

mRNAs. Oligodeoxyribonucleotide templates differed at the bold sequences:

For mMTFuucV:
AATTCAAC**GAA**GGTCATGTGTTTTACCTCCTTACTAAAGTTAACCCTATAGTGAGTCGTATTA

For mMTNaacV:
AATTCAAC**GTT**GGTCATGTGTTTTACCTCCTTACTAAAGTTAACCCTATAGTGAGTCGTATTA

For mMTAgcaV:
AATTCAAC**TGC**GGTCATGTGTTTTACCTCCTTACTAAAGTTAACCCTATAGTGAGTCGTATTA

They were hybridized to 18-mer TAATACGACTCACTATAG and transcribed with T7 RNA polymerase.

Translation assays.[5] Translations contained 0.5 uM each of initiation factors 1-3 and elongation factors Ts and G, 2.5 uM elongation factor Tu, 0.25 uM purified ribosomes, 1 uM appropriate mRNA, 0.2 uM (limiting) fMet-$tRNA_i^{fMet}$, 0.5 uM each of Thr-$tRNA_3^{Thr}$ and ^{3}H-Val-$tRNA_1^{Val}$ (14,600 d.p.m/pmol), and 1.5 uM test, photodeprotected, chemoenzymatic amino-/N-alkyl-amino-acyl-tRNA (or 0.5 uM control natural Phe-$tRNA^{Phe}$). They were performed in 5 ul volumes at 37C for 40 min. without preincubation, then terminated by the addition of NaOH. N-formylated peptide products were separated from free amino acids and other components by passage through a Dowex 50X8-200 cation-exchange mini-column. All other materials and methods were as described.[5] Maximal yields corresponded typically to half of the limiting 0.2 uM fMet-$tRNA^{fMet}$ incorporated within 40 minutes into tetrapeptide d.p.m. (i.e. 0.5 pmol per 5 ul translation).

Ribosome binding assays. The method was similar to that described[6]. Ribosomes (0.5 uM final) that had been pre-loaded at the P site with deacylated $tRNA^{fMet}$ (0.3 uM) using mRNAs MI or MF[7] (1.2 uM of respective non-cognate AUG **AUC** UUC UAG or cognate AUG **UUU** ACG AUU UAG coding sequences) were mixed with the amino-/N-alkyl-amino-acyl-$tRNA^{PheB}$s shown (5 nM; pre-labeled with ^{33}P using polynucleotide kinase) that had been pre-bound to EF-Tu (2 uM) in the presence of EF-Ts (0.1 uM) in polymix buffer containing pyruvate kinase and myokinase.[7] After incubation at 37C for 3 min., ribosomes were collected by filtration through nitrocellulose, washed, and the filters counted.

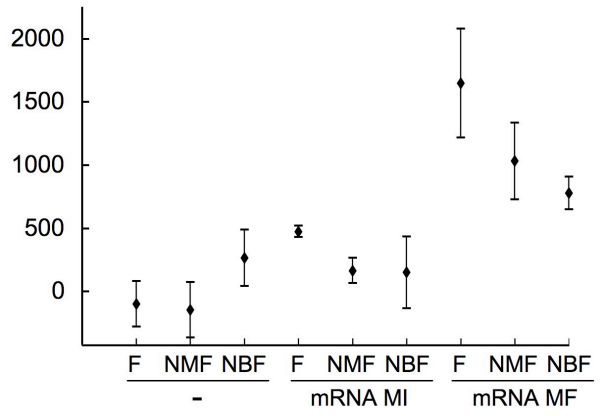

Figure S1. Relative efficiencies of binding to the ribosomal A site of amino-/N-alkyl-amino-acyl $tRNA^{PheB}$s. Ribosomes pre-loaded at the P site with deacylated $tRNA^{fMet}$ using the mRNAs indicated were incubated for three minutes with EF-Tu and the radiolabeled amino-/N-alkyl-amino-acyl $tRNA^{PheB}$s indicated. The ribosomes were then collected by filtration through nitrocellulose and the bound d.p.m. counted. -, Background d.p.m. defined by omitting mRNA. Non-cognate binding d.p.m. were defined using mRNA MI. Standard deviations of triplicate experiments are shown.

References

(1) Verardo, G.; Geatti, P.; Pol, E.; Giumanini, G. *Can. J. Chem.* **2002**, *80*, 779-788.
(2) Ellman, J.; Mendel, D.; Anthony-Cahill, S.; Noren, C. J.; Schultz, P. G. *Methods Enzymol.* **1991**, *202*, 301-336.
(3) Robertson, S. A.; Ellman, J. A.; Schultz, P. G. *J. Am. Chem. Soc.* **1991**, *113*, 2722-2729.
(4) Koh, J. T.; Cornish, V. W.; Schultz, P. G. *Biochemistry* **1997**, *36*, 11314 -11322.
(5) Tan, Z.; Blacklow, S. C.; Cornish, V. W.; Forster, A. C. *Methods* **2005**, *36*, 279-290.
(6) Dale, T.; Uhlenbeck, O. C. *RNA* **2005**, *11*, 1610-1615.
(7) Ayman, A.; Pavlov, M. Y., Tenson, T.; Ehrenberg, M. *Biol. Proced. Online* **2004**, 6, 35-54.