Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 4242–4246, June 1988
Biochemistry

Codon choice and gene expression: Synonymous codons differ in their ability to direct aminoacylated-transfer RNA binding to ribosomes *in vitro*

(translational efficiency/*Escherichia coli*/codon bias)

LINDA K. THOMAS, DANIEL B. DIX, AND ROBERT C. THOMPSON

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309

Communicated by Bernard D. Davis, February 29, 1988

ABSTRACT **Phe-tRNA (anticodon GAA)–polypeptide-chain elongation factor Tu–GTP ternary complexes react faster with ribosomes programmed with UUC codons than with ribosomes programmed with UUU codons. A similar preference is shown by Leu-tRNA$_2$ (anticodon GAG) complexes, which react faster with ribosomes programmed with CUC than with those programmed with CUU. The difference is seen in the rate of ternary-complex binding to the ribosome; no differences are seen in peptide-bond formation. Highly expressed mRNAs in *Escherichia coli* favor codons terminating in cytosine rather than uracil when both codons are read by a single tRNA with an anticodon beginning with guanine. The results suggest that intrinsic differences between the efficiencies of synonymous codons play an important role in modulating gene expression in *E. coli*.**

The genetic code is degenerate in that most amino acids may be specified by more than one codon. In the extreme case, incorporation of serine, arginine, or leucine may be directed by six different codons. Whether apparently synonymous codons behave differently in translation has been much debated; the choice of one codon over another might then modulate gene expression. That many organisms show a significant bias in the codons used in highly expressed genes (1, 2) supports this possibility. Indeed, there is direct evidence that codon choice can affect translation: replacing a rare leucine codon by a common one in the attenuator region of the *leu* operon of *Salmonella typhimurium* prevents attenuation (3) and replacing common codons with rare ones can reduce the amount of resulting protein (4). Pederson (5) and Varenne *et al.* (6) have provided a rationale for the preferential use of codons by showing that the presence of rare codons in highly expressed mRNAs is associated with slower translational rates (5) and that pauses in the synthesis of colicin E1 *in vivo* occur in regions of the mRNA containing rare codons (6).

Preference for one codon over another may be based on the structure of the translational apparatus (ribosomes, mRNAs, and tRNAs) or on extrinsic factors like the cellular concentrations of aminoacyl-tRNAs (aa-tRNAs). The latter factors are clearly important because the rate of aa-tRNA binding to ribosomes is proportional to its concentration (7), and the aa-tRNAs used to translate the codons favored in highly expressed genes are generally present at high concentrations in the cell (8). However, it is also possible that the translational apparatus intrinsically prefers one codon over another. This sort of specificity would be very instructive concerning the nature of the translational apparatus and might even explain the evident coevolution of bias in codon usage and tRNA pools.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Evidence for some inherent specificity in the translational apparatus is provided by the finding of Parker and his colleagues (9) that Lys-tRNA is more likely to react with ribosomes programmed with the noncognate AAU codon (Asn) than with ribosomes programmed with the synonymous noncognate codon, AAC. However, demonstrating that cognate aa-tRNAs distinguish between ribosomes programmed with synonymous codons has been difficult. Recent advances in the synthesis of simple mRNAs and in techniques for measuring the kinetics of *in vitro* protein synthesis now allow us to measure *in vitro* the rate constants for the reaction of an aa-tRNA ternary complex (TC) with ribosomes programmed by synonymous codons. In this paper we show that synonymous codons do differ in the rate with which they direct ribosomes to react with the cognate aa-tRNA *in vitro* and that an intrinsic specificity for certain codons is built into the translational apparatus. In a parallel series of experiments J. Curran and M. Yarus (unpublished work) demonstrated that the same set of synonymous codons are translated at different rates *in vivo*. Therefore, the specificity we see *in vitro* is probably physiologically significant and could direct the evolution of codon bias.

MATERIALS AND METHODS

Sources of tRNAs, radioactive amino acids, and ^{32}P-labeled inorganic phosphate have been described (7, 10), except that tRNA$_2^{Leu}$ and tRNA$_1^{Leu}$ were purchased from Subriden RNA (Rollingbay, WA) and *N*-formylmethionine-specific tRNA was purchased from Boehringer Mannheim.

Formylmethionyl tRNA (fMet-tRNA) was prepared by the method of Dubnoff and Maitra (11). Purified initiation factors (IF-1, IF-2, and IF-3) were prepared according to Hershey *et al.* (12). Free ribosomes and ribosomes programmed with poly(U) and with *N*-acetylphenylalanine-tRNA in the peptidyl (P) site were prepared and assayed according to Thompson and Dix (10). TCs of [^{3}H]aa-tRNA, polypeptide-chain elongation factor of Tu (EFTu), and [γ-^{32}P]GTP were prepared as described (13).

To prepare plasmids encoding the required mRNAs the following synthetic oligodeoxyribonucleotides were made on the Applied Biosystems model 380A synthesizer: 1, GGAGGATTTAATCATGTTTAAGAGCT; 2, CTTAAACATGATTAAATCCTCCTGCA; 3, GGAGGATTTAATCATGTTCAAGAGCT; 4, CTTGAACATGATTAAATCCTCCTGCA; 5, GGAGGATTTAATCATGCTTAAGAGCT; 6, GGAGGATTTAATCATGCTCAAGAGCT; and 7, GGAGGATT-

Abbreviations: EFTu, polypeptide-chain elongation factor Tu; IF-1, -2, and -3, polypeptide-chain initiation factors 1, 2, and 3; fMet-tRNA, formylmethionyl tRNA; k_{GTP} and k_{PEP}, apparent rate constants for GTP hydrolysis and peptide bond formation, respectively; RS, ribosomes programmed with mRNA and with fMet-tRNA in the peptidyl site; aa-tRNA, aminoacyl-tRNA; TC, ternary complexes of aa-tRNA·EFTu·GTP.

TAATCATGCTGAAGAGCT. Pairs of these were hybridized, and the resulting duplexes were ligated into pGEM-2 (Promega Biotec, Madison, WI) previously cut with *Sac* I and *Pst* I according to standard methods (14). Oligonucleotides 1 and 2 were used to prepare the plasmid encoding RNA containing the codon UUU, oligonucleotides 3 and 4 were used for the codon UUC, oligonucleotides 5 and 2 were used for the codon CUU, oligonucleotides 6 and 4 were used for the codon CUC, and oligonucleotides 7 and 2 were used for the codon CUG. As expected, the presence of up to two mismatches did not hinder the hybridizing and cloning. The ligation reaction mixture was used directly to transform competent JM107 cells as described by Mandel and Higa (15), and colonies containing the plasmid of interest were selected by their ability to hybridize to ^{32}P end-labeled oligodeoxyribonucleotide encoding the appropriate RNA. This determination was by the method of Woo (16) using tetraethylammonium chloride wash solutions as described by Wood *et al.* (17).

mRNA was prepared by *in vitro* transcription of the appropriate plasmid, which had been previously linearized by incubation with *Eco*RI (1600 units per 300 μg of DNA). The reaction mixture was extracted with phenol, with phenol/chloroform (1:1), and then with chloroform. The DNA was then ethanol precipitated and resuspended in 50 μl of 10 mM Tris·HCl, pH 8.0/1 mM EDTA. The transcription reaction contained 58 μg of linearized plasmid, 150 units of RNasin, 150 μg of bovine serum albumin, 1000 units of T7 RNA polymerase (Promega Biotec), and 150 nmol of each ribonucleotide in 1.5 ml of 40 mM Tris·HCl, pH 7.5/6 mM $MgCl_2$/2 mM spermidine/10 mM NaCl/10 mM dithiothreitol. After 1 hr at 37°C the RNA was purified by being extracted first with phenol and then with ether, as described for aa-tRNA (13). The RNA was precipitated with 2.5 volumes of ethanol from 2 M ammonium acetate. The RNA pellet was dissolved in 100 μl of water and purified from any remaining unincorporated nucleotide using a 0.5-ml Bio-Gel (Bio-Rad) P-6 spin column equilibrated in 2 mM KOAc, pH 4.5/50 mM KCl. Typically, 80–120 μg of RNA was recovered based upon one A_{260} unit being equivalent to 40 μg of RNA. The A_{260}/A_{280} ratios were 2.0 ± 0.1.

Ribosomes were programmed with these RNAs by incubating 1–2 μM ribosomes, 2–4 μM mRNA, f[^{14}C]Met-tRNA at 2 pmol/pmol of ribosomes, 2 μM IF-1, 1.6 μM IF-2, 1.6 μM IF-3, and 1 mM GTP in the buffer to be used for the subsequent reaction with TC (see below) for 10 min at 37°C. Typically 40–45% of the ribosomes bound f[^{14}C]Met-tRNA as determined by nitrocellulose filter binding of the ^{14}C radiolabel. The concentration of active initiation complexes was determined from the endpoint of both GTP hydrolysis and aa-tRNA binding reactions in the presence of a 2- to 3-fold excess of cognate TC. Typically 35–40% of the ribosomes were active by this assay.

Reactions between TC and the mRNA-programmed ribosomes were conducted under single-turnover conditions in the rapid-mixing apparatus described by Eccleston *et al.* (18) at either 5°C or 25°C in buffers of final composition: 50 mM Tris·HCl, pH 7.2/55 mM KCl/<10 mM NH_4Cl/1 mM dithiothreitol/either 5 mM $MgCl_2$ or 3 mM $MgCl_2$/2 mM spermidine/10 mM putrescine as listed. At 5 mM $MgCl_2$ and 5°C, 5 μl of enzymatically initiated ribosomes was added by a Hamilton syringe to 15 μl of a stirred solution of TC, EFTu·GTP·aa-tRNA. The ribosome complex was typically between 0.025 and 0.06 μM, and the TC concentration was between 0.08 and 0.17 μM. At 5 mM $MgCl_2$ and 25°C, 5 μl of ribosomes and 5 μl of TC were added simultaneously, but in separate Hamilton syringes, to 90 μl of buffer, with final concentrations between 0.006 and 0.02 μM and between 0.035 and 0.05 μM, respectively. At 3 mM $MgCl_2$ and 5°C, 5 μl of ribosomes was added to 15 μl of TC with final concentrations between 0.02 and 0.05 μM and between 0.07 and 0.13 μM, respectively. At 3 mM $MgCl_2$ and 25°C, 5 μl of ribosomes and 5 μl of TC were added to 140 μl of buffer with final concentrations of 0.003 and 0.01 μM, respectively. The reactions were stopped 0.2–30 sec later by addition of 10 μl of EDTA (500 mM). Analysis of the reaction mixture has been described (10).

RESULTS

The mRNAs used in this study are 52 nucleotides long and are identical apart from the three bases after the translational initiation codon. The mRNAs were made by *in vitro* transcription of synthetic DNA cloned between the *Pst* I and *Sac* I sites of plasmid pGEM2 (Promega Biotec). All mRNAs were designed to have the general sequence shown below, where XYZ is the codon to be tested, AUG is the initiation codon, and GGAGGA is the Shine–Dalgarno sequence:

5′ GGGAGACCGGAAGCUUGGGCUGCA-

$\overline{\text{GGAGGA}}$UUUAAUC $\overline{\text{AUG}}$ $\overline{\text{XYZ}}$ AAG AUC UCG 3′

Polyacrylamide gel electrophoresis of the purified transcripts showed the RNAs to be homogeneous and of the expected size (Fig. 1).

We tested the ability of these synthetic mRNAs to properly initiate protein synthesis by incubating them with ribosomes, purified initiation factors, fMet-tRNA, and GTP. As shown in Table 1, the ribosomes bound fMet-tRNA, and this binding was dependent on initiation factors and on mRNA. When ribosomes initiated on mRNAs containing the codons UUU or UUC were mixed with TCs of Phe-tRNA·EFTu·GTP, this aa-tRNA also bound to the ribosomes. An approximate molar equivalent of GTP was hydrolyzed with a rate (see below) close to that expected (7). A large fraction of the phenylalanine bound to the ribosomes was found in the form of a peptide (Table 2). The same ribosomes mixed with TCs containing Leu-$tRNA_2$ hydrolyze GTP more slowly and form peptide less efficiently. These findings point to the ability of the synthetic mRNA to form a true initiation complex and to selectively direct the binding of a cognate aa-tRNA to the vacant ribosomal A site (ribosomal aa-tRNA binding site).

To determine the rate at which aa-tRNA TCs react with the vacant ribosomal A site we mixed ribosome·mRNA·fMet-

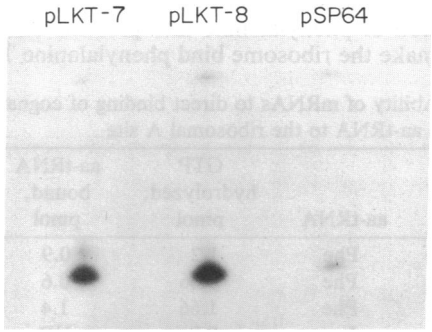

FIG. 1. Analysis of RNAs transcribed from pLKT-7 (UUU) and pLKT-8 (UUC). *Eco*RI-linearized plasmid (3.8 μg) was mixed with 50 nmol each of ATP, GTP, CTP, and (α-^{32}P)UTP (7.5 μCi/μmol; 1 Ci = 37 GBq), and 125 units of T7 RNA polymerase in 100 μl of transcription buffer at 37°C. Twenty microliters was removed and purified as described. Five microliters of loading buffer (0.25% bromophenol blue/0.25% xylene cyanol/40% sucrose) was added to the sample and loaded onto a 12% polyacrylamide/7 M urea gel. pLKT-7 and pLKT-8 produce a 52-nucleotide transcript, whereas the parental pSP64 (Promega Biotec) should yield a 55-nucleotide transcript.

Table 1. Dependence of initiation complex formation on initiation factors and mRNA

Reactants	Addition	[^{3}H]fMet-tRNA bound, pmol
RS, fMet-tRNA, GTP		0.2
RS, fMet-tRNA, GTP +IFs		0.06
RS, fMet-tRNA, GTP +IFs	Poly(UG)	2.2
RS, fMet-tRNA, GTP +IFs	AUGUUU	1.4
RS, fMet-tRNA, GTP +IFs	AUGUUC	1.1

Ten picomoles of 70S ribosomes and 120 pmol of [^{3}H]fMet-tRNA were mixed, where indicated, with 18 pmol each of IF-1, IF-2, and IF-3 and with either 0.55 μg of poly(UG), 20 pmol of RNA containing the sequence AUGUUU, or 20 pmol of RNA containing the sequence AUGUUC and incubated for 10 min at 37°C in 20 μl of the following buffer: 50 mM Tris·HCl, pH 7.2/50 mM KCl/100 mM NH_4Cl/1 mM GTP/1 mM dithiothreitol/4 mM $MgCl_2$. Ten microliters of the reaction mixture was analyzed for bound fMet-tRNA by filtration through nitrocellulose.

tRNA complexes (RS), made as described above, with excess TCs in a rapid mixing device (18) and stopped the reaction between 0.2 and 30 sec later with a large excess of EDTA. Analysis of the reaction mixture for [^{32}P]P_i and ^{3}H-labeled peptide allowed us to reconstruct the progress of the reaction. Simulation of the rate of product accumulation according to the following simplified mechanism, enabled us to determine k_{GTP} and k_{PEP}, apparent rate constants for GTP hydrolysis and peptide-bond formation for each aa-tRNA–codon pair.

$$RS + TC \xrightarrow{k_{GTP}} P_i + RS\cdot aa\text{-}tRNA\cdot EFTu\cdot GDP \xrightarrow{k_{PEP}} RS\cdot peptide\text{-}tRNA$$

The course of a typical reaction and a simulation of the results are shown in Fig. 2.

We measured the rate constants for the reaction of Phe-tRNA TCs with UUU and UUC-programmed ribosomes under a variety of conditions with the results shown in Table 3. With ribosomes programmed with UUC the rate constant for GTP hydrolysis, $k_{GTP(UUC)}$, is consistently greater than that for the ribosomes programmed with UUU, $k_{GTP(UUU)}$. Although the ratio between the rate constants depends to some extent on the reaction conditions, the k_{GTP} ratio is always between 1.6 and 1.9. In contrast, k_{PEP} values on ribosomes programmed with UUU and UUC codons are very similar. We conclude that the two codons used to direct the incorporation of phenylalanine into proteins differ in their ability to make the ribosome bind phenylalanine TCs but do not affect the ability of that ribosome to subsequently incorporate bound Phe-tRNA into protein.

Table 2. Ability of mRNAs to direct binding of cognate but not noncognate aa-tRNA to the ribosomal A site

mRNA	aa-tRNA	GTP hydrolyzed, pmol	aa-tRNA bound, pmol	Peptide formed, pmol
Poly(U)	Phe	1.2	0.9	0.4
AUGUUU	Phe	0.66	0.6	0.3
AUGUUC	Phe	1.66	1.4	0.8
Poly(U)	Leu_2	0.9	ND	0.01
AUGUUU	Leu_2	0.03	ND	<0.001
AUGUUC	Leu_2	0.35	ND	0.002

Twenty microliters of 0.3 μM ribosomes programmed with poly(U) or 0.5 μM ribosomes programmed with RNA containing the sequence AUGUUU or AUGUUC were mixed with 30 μl of 0.5 μM TC containing the indicated aa-tRNA at 0°C. After 30 s 20 μl from the cognate reaction was filtered through nitrocellulose to determine the extent of Phe-tRNA binding, and 20 μl was added to 20 μl of 0.1 M EDTA. After 2 min, 20 μl from the near-cognate reaction was added to 20 μl of 0.1 M EDTA. The reactions stopped with EDTA were analyzed for GTP hydrolysis and peptide formation. ND, not done.

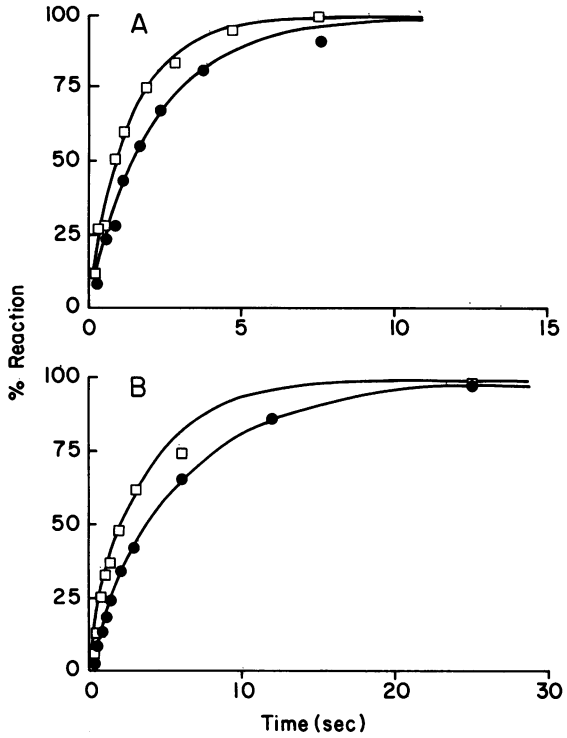

FIG. 2. Time course of GTP hydrolysis. (*A*) The reaction of 0.17 μM phenylalanine TC with 0.04 μM UUC-fMet-tRNA-programmed ribosomes (□) and 0.04 μM UUU-fMet-tRNA-programmed ribosomes (●). (*B*) The reaction of 0.09 μM $tRNA_2^{Leu}$ TC with 0.04 μM CUC-fMet-tRNA-programmed ribosomes (□) and 0.04 μM CUU-fMet-tRNA-programmed ribosomes (●). Lines show the theoretical amount of (^{32}P)P_i predicted by computer simulation of the reactions with $k_{GTP(UUC)} = 4.5 \times 10^6\ M^{-1}\cdot sec^{-1}$, $k_{GTP(UUU)} = 2.8 \times 10^6\ M^{-1}\cdot sec^{-1}$, $k_{GTP(CUC)} = 4.5 \times 10^6\ M^{-1}\cdot sec^{-1}$, and $k_{GTP(CUU)} = 2.5 \times 10^6\ M^{-1}\cdot sec^{-1}$ (k values all listed in Table 3).

To determine whether ribosomes programmed with codons terminating in cytosine might generally react more rapidly than ribosomes programmed with codons terminating in uracil we did similar experiments on the reaction of CUC- and CUU-programmed ribosomes with TCs of Leu-$tRNA_2$. The results, also shown in Table 3, indicate that ribosomes programmed with the CUC codon do react faster than ribosomes programmed with the CUU codon. Again, the k_{PEP} values of these ribosomes do not differ significantly.

The k_{GTP} for UUC-programmed ribosomes reacting with a phenylalanine TC is similar to the k_{GTP} for CUC-programmed ribosomes reacting with a Leu-$tRNA_2$ TC, even though UUC is the major codon for phenylalanine and CUC is used comparatively rarely for leucine. To determine whether CUG, the major codon for leucine, is any more effective than CUC in selecting Leu-tRNA TCs, we measured the rate of reaction of CUG-programmed ribosomes with Leu-$tRNA_1$ TCs. The result (Table 3) indicates that this rate constant is actually lower than that for ribosomes programmed by the minor codon.

DISCUSSION

Table 3 shows that significant differences exist between the abilities of synonymous codons to direct incorporation of their cognate amino acid into nascent protein. All differences are in the rate of the GTP-hydrolysis reaction that accompanies aa-tRNA TC binding to the ribosome. The rates at which bound aa-tRNAs are incorporated into peptide do not

Table 3. Rate constants for the reaction of programmed ribosomes with cognate aa-tRNA TCs

Codon	aa-tRNA	Ionic condition	Temperature, °C	$k_{GTP} \times 10^{-4}$, $M^{-1}\cdot sec^{-1}$	k_{PEP}, sec^{-1}
UUU	Phe	M5	5	2.8 ± 0.2	1.1 ± 0.1
UUC	Phe	M5	5	4.6 ± 0.2	1.1 ± 0.2
UUU	Phe	M5	25	14 ± 1	ND
UUC	Phe	M5	25	27 ± 1	ND
UUU	Phe	M3/S2/P10	5	11 ± 1	6 ± 1
UUC	Phe	M3/S2/P10	5	17.5 ± 1	5.5 ± 1
UUU	Phe	M3/S2/P10	25	95 ± 5	ND
UUC	Phe	M3/S2/P10	25	150 ± 5	ND
CUU	Leu2	M5	5	2.5 ± 0.3	1.0 ± 0.1
CUC	Leu2	M5	5	4.5 ± 0.2	1.0 ± 0.1
CUU	Leu2	M5	25	17 ± 4	ND
CUC	Leu2	M5	25	27 ± 3	6 ± 2
CUU	Leu2	M3/S2/P10	5	14 ± 1	ND
CUC	Leu2	M3/S2/P10	5	22 ± 1	ND
CUG	Leu1	M5	5	1.9 ± 0.1	1.1 ± 0.1

M = mM $MgCl_2$, S = mM spermidine, and P = mM putrescine. ND, not done. Leu2 and Leu1 represent isoacceptors for leucine.

differ significantly. Our results indicate that intrinsic differences between codons, like differences in the concentrations of aa-tRNAs, can influence the efficiency of translation.

Physiological Significance of the Results. Differences in the rate at which aa-tRNA binds to ribosomes will be physiologically significant only if the search for an aa-tRNA *in vivo* occupies a significant fraction of the time taken to incorporate an amino acid into protein; some evidence suggests this to be so. For example, the intermediates detected by Varenne *et al.* in colicin A synthesis (19) and the results obtained by Pederson (5) imply that a step whose rate depends on aa-tRNA concentration can become the rate-limiting step in protein synthesis (19). In addition, a simple calculation of time taken for a ribosome to find a cognate aa-tRNA using the known concentrations of reactants and the rate constants measured here supports the idea that this time is a significant fraction of the total time needed to incorporate an amino acid. We assume that there are about 300,000 tRNAs (20) in a cell of average volume 2.5×10^{-15} liter (21) and that 75% of total tRNA is charged with an amino acid. Concentration of the Phe-tRNA, which represents ≈2% of the total, will then be about 3 μM. When ≈70% of this charged aa-tRNA is in the TC form and rate constants for TC selection *in vivo* resemble those measured in 5 mM Mg^{2+} at 25°C, half the UUC-programmed ribosomes bind a Phe-tRNA complex in ≈14 msec, whereas UUU-programmed ribosomes require about 26 msec. The difference between these times is ≈20% of the average time of 60 msec taken for addition of a single amino acid to the nascent polypeptide chain under such conditions (20). However, the most convincing evidence that synonymous codons differ in their ability to direct the incorporation of an amino acid *in vivo* comes from the unpublished work of J. Curran and M. Yarus. These authors have measured the relative rates of translation of UUU, UUC, CUU, CUC, and CUG codons by taking advantage of the fact that the probability of the ribosome losing frame at these codons inversely relates to the rate at which they are translated. Their results show that a UUC codon is translated 1.3 times as fast as UUU and that CUC is translated 1.3 times as fast as CUU. They find that CUG is translated faster than CUU or CUC, but when the rates are adjusted for the relative concentration of $tRNA_1^{Leu}$ and $tRNA_2^{Leu}$ they find, as do we, that the CUC codon is translated more rapidly.

One interesting feature of rate differences intrinsic to the structure of the translational apparatus is that these differences could explain why codon choice and tRNA pools have evolved to their present bias. However, to attribute this bias entirely to the greater translational efficiency of codons terminating in cytosine would be premature. An examination of the codon-bias pattern in highly expressed genes in *E. coli* reveals another possibility. A preference for codons ending in cytosine over codons ending in uracil is seen in only eight of the sixteen boxes of the genetic code. These are as follows: UUU/UUC (Phe), UAU/UAC (Tyr), UGU/UGC (Cys), CAU/CAC (His), AUU/AUC (Ile), AAU/AAC (Asn), AGU/AGC (Ser), and GAU/GAC (Asp), where the cytosine-ending codon is favored over the uracil-ending codon by 4×, 4.2×, 1.6×, 3.8×, 5.2×, 13×, 9×, and 2.5×, respectively (data from table 1 of ref. 22). A preference for codons ending in cytosine is not found in the other eight boxes of the code. For codon pairs UCU/UCC (Ser), CUU/CUC (Leu), CCU/CCC (Pro), CGU/CGC (Arg), ACU/ACC (Thr), GUU/GUC (Val), GCU/GCC (Ala), and GGU/GGC (Gly), the ratio of cytosine-ending to uracil-ending codon is 0.9×, 0.9×, 0.1×, 0.4×, 0.9×, 0.1×, 0.1×, and 0.8× in highly expressed genes.

The former group of codons is distinguished from the latter by two factors. (*i*) They are all in split boxes, where misreading of the 3′ pyrimidine of the codon by a pyrimidine in the anticodon would lead to an error in incorporation. As shown by Parker *et al.* (9), a codon terminating in cytosine may be less prone to this type of error. (*ii*) The former pairs of codons are all read exclusively by a single aa-tRNA the anticodon of which begins with guanine which, as we show above, reacts more rapidly with ribosomes programmed with codons terminating in cytosine. Thus, codon bias in favor of codons terminating in cytosine could reflect either a need for accuracy or a need for efficiency.

A somewhat different perspective is provided by considering the bias in favor of codons terminating in uracil in some full boxes of the genetic code. The specificity of the translational apparatus is probably not different in full boxes because $tRNA_2^{Leu}$ shows the same specificity as $tRNA^{Phe}$. However, this bias could result if codons ending in uracil react more readily than codons ending in cytosine with aa-tRNAs that have anticodons beginning with pyrimidines. This would shorten the time required to translate uracil-ending codons in full boxes. It would also explain the lower accuracy of uracil-ending codons in split boxes.

Mechanistic Significance of the Results. The mechanism of aa-tRNA binding to ribosomes is complex (10) (see Fig. 3). However, because the *elementary* rate constant for GTP cleavage (k_2) is much faster than that for dissociation of cognate TCs (k_{-1}) from the ribosome, the *observed* rate constant k_{GTP} equals the rate constant for TC binding to ribosomes (k_1, refs. 10 and 23). The finding that this rate constant depends on the codon is further evidence that the

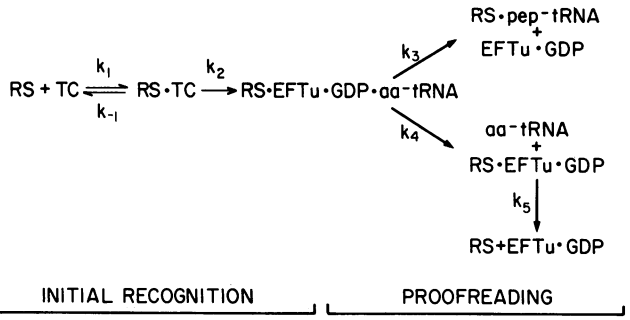

FIG. 3. Mechanism for aa-tRNA binding to ribosomes. RS, ribosomes programmed with mRNA and with fMet-tRNA in the P site.

binding reaction is not simply the diffusion-controlled encounter of a TC with the ribosome (7). In a sense, this finding complements our observation that base substitutions in the anticodon loop of an aa-tRNA change k_1 for its reaction with ribosomes (24).

Our results may indicate that both the aa-tRNA and the mRNA exist in several alternative states that have different properties and that interconvert with a rate comparable to the rate of TC binding to ribosomes. Alternatively, the binding reaction must be regarded as a complex process involving a mutual adaptation between codon and anticodon at a rate dependent on structure of the tRNA anticodon loop and the mRNA codon. Our observation that the rate constant for TC binding to ribosomes is first order in ribosomes over a 100-fold range of ribosome concentrations (7, 25) favors the latter explanation.

Our findings may relate to the greater affinity of $tRNA^{Phe}$ (yeast) for UUC over UUU trinucleotides noted by Labuda and Porschke. These workers also reported that UUC is more able than UUU to promote a conformational change in the $tRNA^{Phe}$(yeast), which is manifested as a tendency to dimerize (26, 27). Although the tRNAs used are from different organisms, the parallel between relative ability of the two trinucleotides to bind to and change the tRNA and the ability of the corresponding codons to promote binding of TCs to the ribosome supports the idea that a codon-induced conformational change in the tRNA is important in the translational process.

The physicochemical basis for the greater efficiency of codons terminating in cytosine is most obviously attributed to their forming three perfect base pairs with the anticodon, whereas the synonymous uracil-ending codons form one wobble pair. However, an alternate possibility is that when the codon is stacked imperfectly by the ribosome, the cytosine-ending codon more easily adopts a productive conformation upon encountering a TC.

We thank Jim Curran and Mike Yarus for valuable discussions of the problem of codon bias and for a preprint of their paper on codon translation rates *in vivo*. We also thank John Eccleston for the gift of EFTu·GDP and John Hershey for guiding us in the isolation of translational initiation factors from *E. coli*. This work was funded by Grant GM32584 from the National Institutes of Health.

1. Post, L. E., Strycharz, G. D., Nomura, M., Lewis, H. & Dennis, P. P. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1697–1701.
2. Bennetzen, J. L. & Hall, B. D. (1982) *J. Biol. Chem.* **257**, 3026–3031.
3. Carter, P. W., Bartkus, J. M. & Calvo, J. M. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 8127–8131.
4. Robinson, M., Lilley, R., Little, S., Emtage, J. S., Yarranton, G., Stephens, P., Millican, A., Eaton, M. & Humphreys, G. (1984) *Nucleic Acids Res.* **12**, 6663–6671.
5. Pederson, S. (1984) *EMBO J.* **3**, 2895–2898.
6. Varenne, S., Knibiehler, M., Cavard, D., Morlon, J. & Lazdunski, C. (1982) *J. Mol. Biol.* **159**, 57–70.
7. Thompson, R. C., Dix, D. B. & Eccleston, J. F. (1980) *J. Biol. Chem.* **255**, 11088–11090.
8. Ikemura, T. (1981) *J. Mol. Biol.* **146**, 1–21.
9. Parker, J., Johnston, T. C., Borgia, P. T., Holtz, G., Remaut, E. & Fiers, W. (1983) *J. Biol. Chem.* **258**, 10007–10012.
10. Thompson, R. C. & Dix, D. B. (1982) *J. Biol. Chem.* **257**, 6677–6682.
11. Dubnoff, J. S. & Maitra, U. (1971) *Methods Enzymol.* **20**, 248–261.
12. Hershey, J. W. B, Yanov, J., Johnston, K. & Faukunding, J. L. (1977) *Arch. Biochem. Biophys.* **182**, 626–638.
13. Thompson, R. C., Dix, D. B., Gerson, R. B. & Karim, A. M. (1981) *J. Biol. Chem.* **256**, 81–86.
14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 98–106.
15. Mandel, M. & Higa, A. (1970) *J. Mol. Biol.* **53**, 154–164.
16. Woo, S. (1979) *Methods Enzymol.* **68**, 389–395.
17. Wood, W. I., Gitschier, J., Laskey, L. A. & Lawn, R. M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 1585–1588.
18. Eccleston, J. F., Messerschmidt, R. G. & Yates, D. W. (1980) *Anal. Biochem.* **106**, 73–77.
19. Varenne, S., Buc, J., Lloubes, R. & Lazdunski, C. (1984) *J. Mol. Biol.* **180**, 549–576.
20. Dennis, P. P. & Bremer, H. (1974) *J. Bacteriol.* **19**, 270–281.
21. Pierucci, O. (1978) *J. Bacteriol.* **135**, 559–574.
22. de Boer, H. A. & Kastelein, R. A. (1986) in *Maximizing Gene Expression*, eds. Reznikoff, W. & Gold, L. (Butterworth, Stoneham, MA), pp. 225–285.
23. Karim, A. M. & Thompson, R. C. (1986) *J. Biol. Chem.* **261**, 3238–3243.
24. Dix, D. B., Wittenberg, W. L., Uhlenbeck, O. C. & Thompson, R. C. (1986) *J. Biol. Chem.* **261**, 10112–10118.
25. Eccleston, J. F., Dix, D. B. & Thompson, R. C. (1985) *J. Biol. Chem.* **260**, 16237–16241.
26. Labuda, D. & Porschke, D. (1983) *J. Mol. Biol.* **167**, 205–209.
27. Labuda, D., Striker, G. & Porschke, D. (1984) *J. Mol. Biol.* **174**, 587–604.