Gross, M. (1979) *J. Biol. Chem. 254*, 2370–2377.
Jen, G., Detjen, B.-M., & Thach, R. E. (1980) *J. Virol. 35*, 150–156.
Kozak, M. (1980) *Cell (Cambridge, Mass.) 19*, 79–90.
Kozak, M., & Shatkin, A. J. (1978) *J. Biol. Chem. 253*, 6568–6577.
Kryostek, A., Cawthon, C., & Kabat, D. (1975) *J. Biol. Chem. 250*, 6077–6084.
Leinwand, L., & Ruddle, F. H. (1977) *Science (Washington, D.C.) 197*, 381–383.
McReynolds, L., O'Malley, B. W., Nisbet, A. D., Fothergill, J. E., Givol, D., Fields, S., Robertson, M., & Brownlee, G. G. (1978) *Nature (London) 273*, 723–728.
Nygard, O., & Hultin, T. (1977) *Eur. J. Biochem. 72*, 537–542.
Palmiter, R. D., & Schimke, R. T. (1973) *J. Biol. Chem. 248*, 1502–1512.
Rosen, J. M., Woo, S. L. C., Holder, J. W., Means, A. R., & O'Malley, B. W. O. (1975) *Biochemistry 14*, 69–78.
Safer, B., Kemper, W., & Jagus, R. (1978) *J. Biol. Chem. 253*, 3384–3386.
Sargent, T. D., & Raff, R. A. (1976) *Dev. Biol. 48*, 327–335.
Schroeder, H. W., Liarkos, C. D., Gupta, R. C., Randerrath, K., & O'Malley, B. W. O. (1979) *Biochemistry 18*, 5798–5808.
Singer, R. H., & Penman, S. (1972) *Nature (London) 240*, 100–102.
Singer, R. H., & Penman, S. (1973) *J. Mol. Biol. 78*, 321–334.
Sonenberg, N., Guertin, D., Cleveland, D., & Trachsel, H. (1981) *Cell (Cambridge, Mass.) 27*, 563–572.
Tarnowka, M. A., & Baglioni, C. (1979) *J. Cell. Physiol. 99*, 359–368.
Thomas, G. P., & Mathews, M. B. (1984) *Mol. Cell. Biol. 4*, 1063–1072.
Thompson, L. H., & Baker, R. M. (1973) *Methods Cell Biol. 6*, 210–281.
Walseth, T. F., & Johnson, R. A. (1979a) *Biochim. Biophys. Acta 526*, 11–31.

Accuracy of Natural Messenger Translation: Analysis of Codon–Anticodon Recognition in a Simplified Cell-Free System†

Didier Nègre, Alain J. Cozzone,* and Yves Cenatiempo
Laboratory of Molecular Biology, University of Lyon, 69622 Villeurbanne, France
Received February 28, 1986; Revised Manuscript Received May 27, 1986

ABSTRACT: A simplified plasmid-directed coupled system [Robakis, N., Cenatiempo, Y., Meza-Basso, L., Brot, N., & Weissbach, H. (1983) *Methods Enzymol. 101*, 690–706] was used to study the accuracy of natural messenger translation in vitro. In this system, protein synthesis is limited to the formation of the N-terminal di- or tripeptide of the gene product. Such a control is obtained by restricting the supply of aminoacyl-tRNAs in the assay medium to those corresponding specifically to the first two or three triplets in the mRNA coding sequence. We analyzed comparatively the interaction of 6 different codons with their cognate tRNAs and 18 noncognate tRNAs able to recognize triplets differing from the legitimate sequences by one base only. Special attention was paid to the single base errors occurring at the first and second codon positions during ribosomal selection of aminoacyl-tRNA molecules. The noncognate tRNAs were assayed either in the absence of the legitimate tRNAs or under competition conditions. They were chosen so that all the possibilities for misreading any particular base as each of the other three bases could be studied. First, it was mainly observed that translation mistakes can be equally detected in the first and second codon positions; there is no compelling evidence for a most or least accurate site. Second, pyrimidines seem to be read more accurately than purines. In particular, U cannot be read as either C or G, and C can hardly be mistaken for any other base. By contrast, A is easily read as U; also, G can be taken as A, U, or C, at least when the cognate tRNAs are totally missing. Finally, there is no selective misreading of any base of a given class, purine or pyrimidine, by the other base of the same class. Thus, the purine A can be read as the pyrimidine U and vice versa; similarly, G can be mistaken for U or C.

Several types of translational errors occurring in prokaryotes have been described (Hopfield & Yamane, 1980; Abraham, 1983), but there is ample evidence that misreading of the genetic information is mostly due to inaccurate aminoacyl-tRNA selection on the ribosome during protein synthesis (Gallant & Foley, 1980; Yarus & Thompson, 1983; Brakier-Gingras & Phoenix, 1984). To study the process of this selection and to characterize the missense errors that result from anomalous codon–anticodon recognition, a number of in vivo situations have been analyzed. These include, for instance, illegitimate incorporation of cysteine in *Escherichia coli* flagellin (Edelmann & Gallant, 1977), mistranslation of an ochre codon in the *phoA* gene for alkaline phosphatase (Ellis & Gallant, 1982), or modification of the total charge of proteins synthesized during amino acid starvation (O'Farrell, 1978). However, most experiments in the field have been carried out in vitro by analyzing the expression of both natural and synthetic messengers in various cell-free systems (Yarus & Thompson, 1983). Some fundamental patterns of codon–anticodon interaction have thus been elucidated such as the classical wobble rules (Crick, 1966), the more recently proposed "two-out-of-three" hypothesis (Lagerkvist, 1978), or else the notion of "extended anticodon" (Yarus, 1982). The

† This research was supported in part by the Centre National de la Recherche Scientifique (ATP Microbiologie, Contract 06931).

0006-2960/86/0425-6391$01.50/0

cell-free systems have been of great use to investigate the fidelity of translation particularly because they are sensitive to multiple effectors that modify, in one way or the other, the frequency of missense errors, e.g., temperature, cation concentrations, alcohols, and aminoglycosides (Gallant & Foley, 1980; Abraham, 1983). But their utilization has sometimes been found to be limited for conceptual or technical reasons (Laughrea, 1981). For example, even when using optimized analytical conditions (Jelenc & Kurland, 1979), it may be difficult to transpose to living cells the results obtained with some synthetic polymers, namely, the widely used polyuridylic acid, whose structure is quite different from that of natural messengers. The systems using viral RNA as template are obviously more relevant to the situation in vivo, but their handling has often been found to be technically complex (Goldman, 1982; Khazaie et al., 1984). Moreover, the amount of information that can be obtained, in each type of assay, on the possible amino acid substitutions during erroneous translation is generally restricted (Yarus, 1979).

Miscoding has been previously studied extensively at the third codon position (Yarus & Thompson, 1983) and to a lesser degree at the first and second positions, in particular with poly(U)-programmed ribosomes (Thompson & Stone, 1977; Thompson et al., 1981). In this work, we have focused our attention on the errors occurring at the first and second positions using natural messenger RNAs. Mistranslation has been studied in vitro in a simplified plasmid-directed coupled system that is based on the formation of only the amino-terminal end of the gene product instead of the completed protein (Cenatiempo et al., 1982a). By use of the appropriate aminoacyl-tRNA species, a specific di- or tripeptide can be selectively synthesized and then characterized in a rapid assay (Robakis et al., 1983). This system has been shown to correctly reflect the initiation of both transcription and translation and to be suited for studying the regulation of gene expression [for review, see Cenatiempo (1986)]. It has been utilized here to perform a comparative analysis of codon recognition by cognate and noncognate tRNAs using the messengers transcribed from various plasmids.

Materials and Methods

Plasmids. The construction of plasmids pNF1337, pNF1341, and pJEA-4 has been previously reported (Fiil et al., 1979; Erion et al., 1981; Cenatiempo et al., 1982). These plasmids were kindly provided by N. Brot, Roche Institute of Molecular Biology, Nutley, NJ.

Plasmid pND51-1 was constructed by restricting pBR322 with *Nae*I endonuclease. An 880 base pair (bp)[1] segment of the *tet* region was subsequently removed. The resulting plasmid (3480 bp) was used to transform *E. coli* RR1 competent cells (Bolivar & Backman, 1979) selecting for amp^r tet^s.

An *Hin*dIII–*Pvu*II DNA fragment (1700 bp) harboring the beginning of *rplJ* structural gene (ribosomal protein L10) was excised from pNF1337 and cloned between *Hin*dIII–*Pvu*II sites of pND51-1. The recombinant plasmid pYD1338 (3500 bp) confers amp resistance to host cells.

DNA-Directed Di- and Tripeptide Synthesis. The incubation conditions for in vitro synthesis of dipeptides were slightly modified from Cenatiempo et al. (1982a), Robakis et al. (1983), and Plumbridge et al. (1985). The incubation mixture contained in a total volume of 35 μL the following: 30 mM Tris–acetate, pH 7.5; 10 mM DMGA, pH 6.0; 35 mM ammonium acetate; 2 mM DTT; 8–10 mM magnesium acetate; 2.9 mM ATP; 0.7 mM CTP, GTP, and UTP; 28 mM phosphoenolpyruvate; 0.5 μg of pyruvate kinase; 39 mM potassium acetate; 0.8 mM spermidine; 4 units of RNasin; 4% poly(ethylene glycol) 6000; 15 μg of RSW; 0.4 μg of EF-Tu; 10 μg of RNA polymerase; 0.6 A_{260} unit of NH_4Cl-washed *E. coli* 70S ribosomes; 10 pmol of fMet-$tRNA_f^{Met}$; 7 pmol of the second aminoacyl-tRNA. The incubation medium for tripeptide synthesis was supplemented with 0.4 μg of EF-G and 7 pmol of the third aminoacyl-tRNA. In the dipeptide assay, the second aminoacyl-tRNA was labeled with ^3H amino acid, whereas in the tripeptide assay only the third aminoacyl-tRNA was labeled.

The reaction was initiated by adding 0.7 pmol of plasmid DNA. The kinetic curves of di- or tripeptide formation were linear for 90 min at 37 °C or 160 min at 30 °C and then reached a plateau. The same type of curve was obtained when using either cognate or noncognate tRNAs in the incubation medium, with a constant ratio between the corresponding levels of incorporation. Therefore, the incubation was performed throughout the linear part of kinetics so as to produce, in all cases, the maximum final amount of di- or tripeptide. It was stopped with 2.5 μL of 1 M KOH, and the mixture was incubated for an additional 10 min at 37 °C to hydrolyze any peptidyl-tRNA. The assay for the di- or tripeptide product was previously described (Weissbach et al., 1984). In this procedure, the formylated ^3H-labeled di- and tripeptides that are neutral at acidic pH flew through a minicolumn of Dowex 50WX-8 (H^+ form), whereas the positively charged free ^3H amino acids were not eluted.

Biochemicals. Purified *E. coli* tRNA isoacceptor species $tRNA_2^{Ala}$, $tRNA_3^{Ala}$, $tRNA_3^{Gly}$, $tRNA^{Ile}$, $tRNA_2^{Leu}$, $tRNA_4^{Leu}$, $tRNA_1^{Ser}$, $tRNA_3^{Ser}$, $tRNA_3^{Thr}$, $tRNA_1^{Val}$, and $tRNA_2^{Val}$ were purchased from Subriden RNA (Rolling Bay, WA), and $tRNA_f^{Met}$ and $tRNA^{Phe}$ were purchased from Boehringer Mannheim. Unfractionated *E. coli* tRNA used for $tRNA^{Asp}$ aminoacylation was from Sigma (St. Louis, MO). A 0.25 M DEAE salt eluate (Kung et al., 1975) was used as a source of enzymes to acylate tRNA species and to formylate L-methionyl-$tRNA_f^{Met}$. The amino acids used for charging tRNAs were purchased from Amersham Centre (England). Their purity was checked by applying the chromatographic procedure recommended by the manufacturer and was found, in each case, to be higher than 99%. The tRNAs were identified by a combination of controls carried out by Subriden RNA (chromatography, characterization of unique nucleotides, and sequencing). Their purity was ascertained by their chargeability with the cognate radiolabeled amino acid and their inability to be charged with other amino acids. RNA polymerase was purified as previously reported (Burgess & Jendrisak, 1975). 70S ribosomes and the RSW fraction containing initiation factors were prepared according to Hershey et al. (1979). DNA modification enzymes were from commercial sources and were used as recommended by the manufacturers or as described by Maniatis et al. (1982). Ribonuclease inhibitor (RNasin) was purchased from Genofit (Geneva). Elongation factors EF-Tu and EF-G were generous gifts of B. Antonsson and R. Leberman, EMBL, Grenoble.

Results

In the transcription–translation coupled system used here to study translational fidelity, protein synthesis is limited to the formation of the initial di- or tripeptide of the gene product.

[1] Abbreviations: RSW, ribosomal salt wash; DMGA, dimethyl glutarate sodium salt; LS-RuBPCase, large subunit of ribulose-bisphosphate carboxylase; A*, unknown derivative of adenosine; bp, base pair; Tris, tris(hydroxymethyl)aminomethane; DTT, dithiothreitol; RNasin, ribonuclease inhibitor; DEAE, diethylaminoethyl; kDa, kilodalton(s); V, uridine-5-oxyacetic acid; Q, 7-(4,5-*cis*-dihydroxy-1-cyclopentenyl-3-aminomethyl)-7-deazaguanosine (quenosine).

Table I: Plasmids and Related Gene Products[a]

plasmid	gene product analyzed	5′-terminal nucleotide sequence	N-terminal amino acid sequence	ref
pND51-1	β-lactamase	AUG-AGU-AUU	fMet-Ser-Ile	this work
pJEA-4	LS-RuBPCase	AUG-UCA-CCA	fMet-Ser-Pro	Erion et al. (1981); Cenatiempo et al. (1982a)
pNF1341	ribosomal protein L12	AUG-UCU-AUU	fMet-Ser-Ile	Fiil et al. (1979)
pYD1338	ribosomal protein L10	AUG-GCU-UUA	fMet-Ala-Leu	this work

[a] The construction of the plasmids listed is described under Materials and Methods or in the indicated references.

Such a control is compelled by restricting the supply of charged tRNAs in the assay medium to those corresponding respectively to the first, second, and eventually third triplet in the mRNA coding sequence. The ability for a noncognate tRNA to read a codon and, therefore, to be used for di- or tripeptide synthesis is tested by substituting this tRNA for the legitimate one. Such in vitro system has previously been demonstrated to be highly specific and suited for analyzing accurately the initiation of the translation process (Weissbach et al., 1984).

Nature of Plasmids. In this work, four different plasmids, all derived from pBR322, were used as DNA templates (Table I). They were chosen so as to avoid the emergence of ambiguous results due to the presence of several expressible genes on the same template, which may direct the simultaneous synthesis of protein products with identical N-termini. Although pBR322 can potentially code for 12 proteins (Sutcliffe, 1979), it is known that in vivo no more than two proteins, the β-lactamase enzyme and a 37-kDa protein from the *tet* gene (Sancar et al., 1979), and in vitro only β-lactamase (Meza-Basso et al., 1981) are synthesized to a detectable level because of a limited promoter activity (Stüber & Bujard, 1981). Nevertheless, the plasmids listed in Table I were engineered, whenever possible, to eliminate the DNA sequences that theoretically might lead to misinterpretation of the experimental data. Thus, we constructed pND51-1, a pBR322 derivative with a deletion in the *tet* gene (see Materials and Methods). This plasmid was used to study the possible misreading of the second codon, AGU, of β-lactamase (dipeptide fMet-Ser). As shown below, in Figure 1, $tRNA_3^{Ser}$, which is required for reading the AGU codon, can be replaced by either $tRNA_3^{Thr}$ (dipeptide fMet-Thr) or $tRNA_3^{Gly}$ (dipeptide fMet-Gly), which recognize ACU and GGU codons, respectively. Using unmodified pBR322 as template could give dubious results since it harbors putative genes that can score for fMet-Thr and fMet-Gly in the presence of these two tRNAs (Sutcliffe, 1979). Therefore, pBR322 was restricted to yield plasmid pND51-1, which does not contain these genes. Only inaccurate reading of the second codon of β-lactamase mRNA could then produce fMet-Thr and fMet-Gly dipeptides.

We constructed also pYD1338 by using the same plasmid pND51-1 as a cloning vector for a DNA fragment coding only for the beginning of ribosomal protein L10 that we excised from another plasmid, pNF1337 (see Materials and Methods). The other two plasmids, pJEA-4 and pNF1341, used in this work did not require modification since their gene content and the nature of the tRNA substitutions assayed could hardly generate ambiguous results.

Characteristics of the in Vitro System. One of the parameters that critically influence the accuracy of translation in vitro is the magnesium concentration. In this system, it was maintained between 8 and 10 mM, i.e., in the same relatively low range as that used in other transcription–translation coupled systems (Chen & Zubay, 1983; Kung et al., 1984) to avoid abnormal translation (Pestka, 1968; Gavrilova & Spirin, 1971). In each series of assays, the precise magnesium concentration was chosen to get optimal di- or tripeptide synthesis: this value varying slightly when the incubation components were changed, e.g., different ribosomes or tRNAs (Robakis et al., 1983).

Table II: Dependency of Dipeptide and Tripeptide Synthesis on Various Factors[a]

system	assay medium	synthesis (pmol)
dipeptide (fMet-Ser)	complete	1.10
	−pND51-1 DNA	0
	−RNA polymerase	0
	−RSW	0
	−70S ribosomes	0
	−fMet-$tRNA_f^{Met}$	0
	−EF-Tu	0.80
	−Ser-$tRNA_3^{Ser}$ + Ser-$tRNA_1^{Ser}$	0
	−Ser-$tRNA_3^{Ser}$ + Thr-$tRNA_3^{Thr}$	0.24
tripeptide (fMet-Ser-Ile)	complete	0.50
	−pND51-1 DNA	0
	−fMet-$tRNA_f^{Met}$	0
	−Ser-$tRNA_3^{Ser}$	0
	−EF-G	0.36
	−Ile-$tRNA^{Ile}$ + Leu-$tRNA_4^{Leu}$	0
	−Ile-$tRNA^{Ile}$ + Phe-$tRNA^{Phe}$	0.18

[a] The composition of the complete assay medium for di- and tripeptide synthesis was as described under Materials and Methods. Plasmid pND51-1 used as DNA template contains the gene coding for β-lactamase with the initial amino acid sequence fMet-Ser-Ile (see Table I). Components were omitted (−) or added (+) as indicated. $tRNA_3^{Ser}$ is the legitimate tRNA for the second codon AGU; $tRNA_1^{Ser}$ and $tRNA_3^{Thr}$ are noncognate tRNAs. $tRNA^{Ile}$ is the legitimate tRNA for the third codon AUU; $tRNA_4^{Leu}$ and $tRNA^{Phe}$ are noncognate tRNAs.

In order to show that the initiation of translation was taking place at the predicted physiological sites but not randomly within the plasmid-encoded genes, the dependency of the synthesis of the N-terminal di- or tripeptide of β-lactamase on various factors was analyzed. First, the data presented in Table II show that the synthesis of the dipeptide fMet-Ser is totally dependent on the presence of the DNA template (plasmid pND51-1), as well as on the RNA polymerase that catalyzes the formation of the corresponding mRNA. Second, the system exhibits a strict dependency on all the components necessary for the formation of the 70S initiation complex: initiation factors (present in the RSW fraction),fMet-$tRNA_f^{Met}$, and 70S ribosomes. In addition, the first dipeptide bond is generated only if the legitimate charged tRNA, in this case Ser-$tRNA_3^{Ser}$, is added together with elongation factor EF-Tu, but it seems that the dependency on EF-Tu is merely partial. This finding may be explained by considering that only trace amounts of EF-Tu are required for di- and tripeptide synthesis (Cenatiempo et al., 1982b). Consequently, the concentration of this factor in the RSW fraction used in the system is likely sufficient to sustain a significant synthesis of dipeptide. The same type of situation is encountered for elongation factor EF-G in tripeptide synthesis, and a similar explanation can be proposed. In any case, it has clearly been shown in previous experiments that this system is totally dependent on factor EF-Tu for dipeptide synthesis (Weissbach et al., 1984) and on both factors EF-Tu and EF-G for tripeptide synthesis (Cenatiempo et al., 1982b). This observation therefore excludes the event of an artifactual factor-inde-

Table III: Nature of Codons and Cognate/Noncognate tRNAs Assayed[a]

	cognate		noncognate		
codon	tRNA species	anticodon	tRNA species	anticodon	recognized codon
AGU	$tRNA_3^{Ser}$	GCU	$tRNA_3^{Gly}$	GCC	*G*GU
			$tRNA_3^{Thr}$	GGU	A*C*U
UCA	$tRNA_1^{Ser}$	VGA	$tRNA_2^{Ala}$	VGC	*G*CA
UCU	$tRNA_1^{Ser}$	VGA	$tRNA_2^{Ala}$	VGC	*G*CU
			$tRNA_3^{Ala}$	GGC	*G*CU
			$tRNA^{Phe}$	GAA	U*U*U
			$tRNA_3^{Thr}$	GGU	*A*CU
GCU	$tRNA_3^{Ala}$	GGC	$tRNA^{Asp}$	QUC	G*A*U
			$tRNA_3^{Gly}$	GCC	G*G*U
			$tRNA_1^{Ser}$	VGA	*U*CU
			$tRNA_3^{Thr}$	GGU	*A*CU
			$tRNA^{Val}$	GAC	G*U*U
AUU	$tRNA^{Ile}$	GAU	$tRNA_2^{Leu}$	GAG	*C*UU
			$tRNA_3^{Thr}$	GGU	A*C*U
			$tRNA^{Phe}$	GAA	*U*UU
			$tRNA_2^{Val}$	GAC	*G*UU
UUA	$tRNA_4^{Leu}$	A*AA	$tRNA_1^{Ser}$	VGA	U*C*A
			$tRNA_1^{Val}$	VAC	*G*UA

[a] The interaction of six different codons with their respective cognate tRNAs and a variety of noncognate tRNAs was assayed as indicated, taking into account the patterns of codon–anticodon recognition published by Ikemura (1981). The noncognate tRNAs were chosen so that they could recognize a codon sequence differing by only one base (italicized) from that of the corresponding legitimate triplet.

pendent translation. Concerning the dependency on tRNA, when a cognate tRNA-isoaccepting species is substituted by an illegitimate tRNA whose anticodon sequence is differing by two or three nucleotides, no peptide synthesis occurs. This is observed, for example, in the dipeptide system where the cognate $tRNA_3^{Ser}$ that reads AGU and AGC codons cannot be replaced by $tRNA_1^{Ser}$, which recognizes UCA, UCG, and UCU. This applies also to the tripeptide system in which the cognate $tRNA^{Ile}$ specific for AUU and AUC codons cannot be replaced by $tRNA_4^{Leu}$, which reads UUA and UUG. However, when the difference between cognate and noncognate tRNAs concerns only one nucleotide in the recognized codon sequences, substitution may take place and significant peptide synthesis can then be measured. Examples are given in Table II for $tRNA_3^{Thr}$, which reads ACU codon and can partially replace $tRNA_3^{Ser}$ in dipeptide formation, and $tRNA^{Phe}$ specific for UUU, which can substitute to $tRNA^{Ile}$ in tripeptide synthesis.

All together, these observations indicated that the in vitro system was quite efficient in tRNA selection and suited for studying single base errors during translation. Therefore, it was further used to analyze the recognition of 6 different codons by, comparatively, their cognate tRNAs and 18 distinct noncognate tRNAs able to recognize codons differing from the legitimate sequences by only one base, in either the first or second position (Table III).

Inaccurate Dipeptide Synthesis Produced by tRNA Substitutions. A series of experiments were performed to analyze, qualitatively and quantitatively, the possible replacement of a cognate tRNA by near-cognate tRNAs at the level of the second codon of mRNA. For this purpose, dipeptide synthesis was measured in the presence of either the legitimate tRNA as a control or various illegitimate tRNAs assayed individually. The codons and tRNAs that were studied are described in Tables I and III. Incubation was carried out at 30 or 37 °C. Only the results obtained at 30 °C are presented in Figure 1. Similar qualitative results were obtained at 37 °C, but the degree of inaccurate translation was then lower (data not shown).

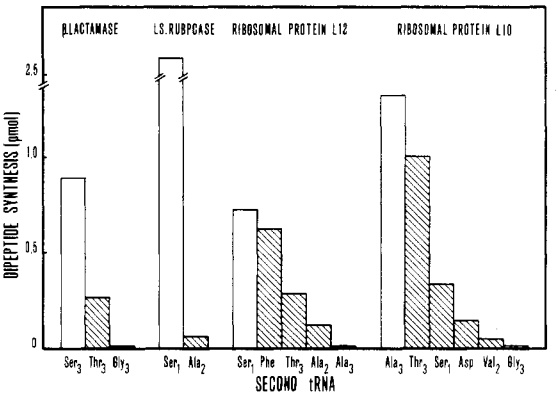

FIGURE 1: Dipeptide formation in the presence of cognate or noncognate tRNAs. Dipeptide synthesis was measured as described under Materials and Methods in the presence of the initiator $tRNA_f^{Met}$ and varying radioactive second aminoacyl-tRNAs, either legitimate (empty bars) or illegitimate (hatched bars) as indicated. Plasmids coding respectively for β-lactamase (fMet-Ser), LS-RuBPCase (fMet-Ser), and ribosomal proteins L12 (fMet-Ser) and L10 (fMet-Ala) were used as templates (see Table I).

Although all the illegitimate tRNAs assayed shared the common property of reading codons, differing in only one base from those present in the 5′ coding region of mRNAs, they appeared to behave quite differently depending on the nature of the interchange. In fact, only a limited number of substitutions, 5 out of 11 tested, were found to occur to a significant extent (in our experimental conditions, values of the amount of dipeptide synthesized lower than 0.1 pmol were not taken as significant). The data in Figure 1 show namely that $tRNA^{Phe}$ is almost as efficient as $tRNA_1^{Ser}$ in reading the UCU codon of the mRNA of ribosomal protein L12 and, similarly, $tRNA_3^{Thr}$ reads effectively the GCU codon of the mRNA of protein L10, which is normally decoded by $tRNA_3^{Ala}$; to a lower extent, $tRNA_1^{Ser}$ can be substituted to $tRNA_3^{Ala}$ to read the same GCU codon, and $tRNA_3^{Thr}$ can be used in the place of $tRNA_3^{Ser}$ to translate the AGU triplet of β-lactamase mRNA. Misreading of the L12-UCU (Ser) codon by $tRNA_3^{Thr}$ may be questionable (see below) since the plasmid used is harboring a truncated β-lactamase gene with an intact 5′ coding terminus and, therefore, fMet-Thr synthesis can arise as well from mistranslation of the β-lactamase-AGU (Ser) codon.

All the other substitutions assayed were found to be inoperative: $tRNA_3^{Gly}$ cannot read the β-lactamase-AGU codon nor recognize the L10-GCU codon, which, on the other hand, cannot be misread by $tRNA_2^{Val}$ or $tRNA^{Asp}$; $tRNA_2^{Ala}$ does not read LS-RuBPCase-UCA and L10-UCU triplets, the latter being not recognized either by $tRNA_3^{Ala}$.

Mistranslation in the Tripeptide System. Experiments similar to those described above for dipeptide synthesis were carried out in the case of tripeptide formation by measuring the replacement of the cognate tRNA species for the third codon of mRNA by near-cognate tRNAs. The codons analyzed were those present in the third position of β-lactamase and protein L10 messengers, which are read legitimately by $tRNA^{Ile}$ and $tRNA_4^{Leu}$, respectively (see Tables I and III). Among the six different substitutions checked (Figure 2), only one gave a positive response: $tRNA^{Phe}$ reads significantly the β-lactamase-AUU codon whereas $tRNA_3^{Thr}$, $tRNA_2^{Val}$, and $tRNA_2^{Leu}$ are all three unable to read it. On the other hand, both $tRNA_1^{Ser}$ and $tRNA_1^{Val}$ assayed for mistranslation of the L10-UUA triplet gave unsignificant results.

Competition between Cognate and Near-Cognate tRNAs. In the preceding assays, di- or tripeptide synthesis was mea-

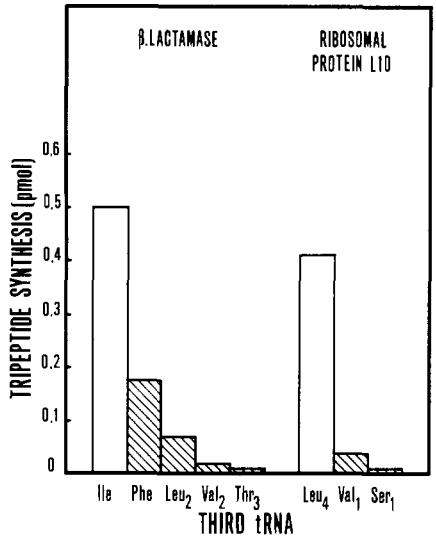

FIGURE 2: Tripeptide synthesis in the presence of cognate or noncognate tRNAs. The incubation medium contained all components required for tripeptide synthesis (see Materials and Methods) including the isoacceptor species for the second amino acid of the gene products: $tRNA_3^{Ser}$ for β-lactamase (fMet-Ser-Ile) and $tRNA_3^{Ala}$ for protein L10 (fMet-Ala-Leu). The amount of tripeptide formed was measured from the incorporation of the radioactive third aminoacyl-tRNAs, either legitimate (empty bars) or illegitimate (hatched bars). Background synthesis in the absence of second tRNA was subtracted from the experimental data.

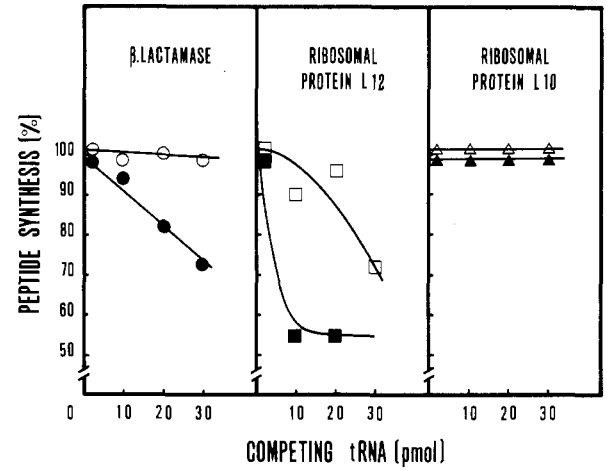

FIGURE 3: Competition between cognate and noncognate tRNAs for accurate peptide synthesis. A fixed amount of legitimate tRNA (7 pmol) charged with ^3H amino acid was incubated in the presence of competing unlabeled illegitimate aminoacyl-tRNA at various concentrations as indicated. Results are expressed as a percentage of the di- or tripeptide synthesis measured in the absence of competing tRNA. β-Lactamase di- and tripeptide and L12 and L10 dipeptide synthesis was directed respectively by pND51-1, pNF1341, and pYD1338 (see Table I). (○) fMet-Ser ([^{3}H]Ser-$tRNA_3^{Ser}$ vs. Thr-$tRNA_3^{Thr}$); (●) fMet-Ser-Ile ([^{3}H]Ile-$tRNA^{Ile}$ vs. Phe-$tRNA^{Phe}$); (□) fMet-Ser ([^{3}H]Ser-$tRNA_3^{Ser}$ vs. Phe-$tRNA^{Phe}$); (■) fMet-Ser ([^{3}H]Ser-$tRNA_1^{Ser}$ vs. Thr-$tRNA_3^{Thr}$); (△) fMet-Ala ([^{3}H]Ala-$tRNA_3^{Ala}$ vs. Thr-$tRNA_3^{Thr}$); (▲) fMet-Ala ([^{3}H]Ala-$tRNA_3^{Ala}$ vs. Ser-$tRNA_1^{Ser}$).

sured after substituting in the incubation medium a near-cognate tRNA for a legitimate tRNA. To analyze further the process of codon–anticodon recognition, competition experiments were carried out by incubating simultaneously a varying amount of nonradioactive aminoacylated near-cognate tRNA (up to 30 pmol/assay) with a fixed amount of legitimate tRNA aminoacylated with a labeled amino acid (7 pmol/assay). In this procedure, any decrease in the synthesis of the labeled di- or tripeptide could be interpreted as a consequence of mistranslation by the illegitimate tRNA. Only those six tRNAs previously found to be efficient in substitution experiments (Figures 1 and 2) were assayed.

Figure 3 shows that, in our experimental conditions, three near-cognate tRNAs were able to partially outcompete cognate tRNAs: $tRNA^{Phe}$ vs. $tRNA^{Ile}$, $tRNA_3^{Thr}$ vs. $tRNA_1^{Ser}$, and $tRNA^{Phe}$ vs. $tRNA_1^{Ser}$. The other three competitions tested, i.e., $tRNA_3^{Thr}$ vs. $tRNA_3^{Ser}$, $tRNA_3^{Thr}$ vs. $tRNA_3^{Ala}$, and $tRNA_1^{Ser}$ vs. $tRNA_3^{Ala}$, did not result in any decrease of legitimate amino acid incorporation in fMet-peptides. The possibility that a limiting concentration of factor EF-Tu in the reaction medium might generate a competition between cognate and near-cognate tRNAs for this factor was analyzed. For this, the same type of experiments was performed in the presence of a fivefold higher concentration of EF-Tu (57.1 μg/mL). No significant difference in the degree of replacement of cognate tRNAs by near-cognate tRNAs was observed (data not shown) in comparison to the data obtained with the lower EF-Tu concentration (Figure 3).

We stressed above that some ambiguity might exist between the misreading of L12-UCU codon and that of β-lactamase-AGU codon by the same $tRNA_3^{Thr}$ when using plasmid pNF1341. In fact, these two possibilities could not be confused with each other in the type of experiments described in Figure 3 because any direct competition between $tRNA_3^{Thr}$ and $tRNA_1^{Ser}$ could take place only for reading the L12-UCU codon but not the β-lactamase-AGU codon.

By comparison of the results obtained in substitution and competition experiments, it can be noted that some tRNAs were equally efficient in both types of assay while others were not. For instance, $tRNA^{Phe}$ could read the protein L12-UCU triplet in substitution experiments (Figure 1) and could compete as well with the cognate isoacceptor $tRNA_1^{Ser}$ (Figure 3). By contrast, $tRNA_3^{Thr}$, which could read the protein L10-GCU codon in the absence of the legitimate $tRNA_3^{Ala}$, was unable to compete with this same tRNA.

Table IV: Summary of Results Obtained in Mistranslation Experiments[a]

nature of base misread	position in codon	type of assay	
		substitution	competition
A as C	1	–	nd
G	1	–	nd
U	1	+	+
C as A	2	–	nd
G	2	–	nd
U	2	+, –	+
G as A	1	+	–
C	2	+	–
U	1	+	–
U as A	1	+	+
C	2	–	nd
G	1	–	nd

[a] The possibility for each base in first or second codon position to be read as any of the other three bases was assayed by substituting various noncognate tRNAs to the legitimate tRNA (Figures 1 and 2). Substitution and competition were found either to occur (+) or not (–). Only noncognate tRNAs that yielded significant misreading were further tested in competition experiments (Figure 3). The effect of the other noncognate tRNAs, presumably inoperative, was not determined (nd).

All the results described in this study are summarized in Table IV. Only the nature of the base misread was considered since, in each experiment, mistranslation could be due to an error in codon–anticodon recognition exclusively at the level of a single base. It must be pointed out that all the possibilities for misreading any particular base as any of the other three bases were studied in substitution experiments. In some cases, several different types of tRNAs could be investigated. Thus,

to study the misreading of A as G, two distinct assays were performed, at the level of dipeptide or tripeptide synthesis, both leading to negative results. The possible misreading of U as G was tested in four different situations (three dipeptide and one tripeptide experiments); all of them giving also negative results. The only result apparently conflictual was obtained in studying the ability of C to be read as U: $tRNA^{Phe}$, which normally reads codon UUU, recognized the protein L12-UCU codon, whereas $tRNA_2^{Val}$, which reads GUU, could not recognize the protein L10-GCU codon.

Discussion

The DNA-directed coupled protein-synthesizing system (Robakis et al., 1983) used in this work provides a simple as well as a specific procedure to study the accuracy of translation in vitro. Indeed, it is highly defined and requires only a limited number of protein factors. The di- and tripeptide gene products are synthesized quite selectively since only fMet-tRNA and the appropriate aminoacyl-tRNA species for the second and third amino acids are utilized. In particular, such utilization of purified isoacceptors makes it possible to distinguish between two different triplets coding for the same amino acid. In addition, the gene products can be rapidly and easily assayed, and precise quantitative data can be obtained. Moreover, although only plasmids have been used here to direct protein formation, it is known that other types of molecules bearing genetic information can also serve as templates, such as phage DNA, restriction fragments, or mRNA of prokaryotic or eukaryotic origin (Weissbach et al., 1984). The applicability of this system is therefore wide, and it is conceivable that the expression of every codon could thus be analyzed, provided that the corresponding individual isoacceptor species be available.

Considering the experimental procedure used, the type of error that we have studied deals essentially with the ribosomal selection of the aminoacyl-tRNA molecules. The availability of the aminoacyl-tRNA cognate to a codon at the A site of the ribosome is an important determinant of specificity. Any decrease in the supply of legitimate tRNA perturbs the selection process and, thereby, enhances the frequency of ambiguities in the decoding of mRNA (Abraham, 1983). However, even in the complete absence of cognate aminoacyl-tRNA, the translation apparatus still remains very selective with regard to the acceptance of noncognate tRNAs (Yarus, 1979). Such selectivity is confirmed by our data showing that among the 18 substitutions assayed, only 6 gave a positive response. It is further demonstrated by the finding that only three illegitimate tRNAs could compete effectively with cognate tRNAs. Taken together, the various base replacements that we have analyzed in substitution experiments lend themselves to some generalizations: (i) Translation mistakes can be equally detected in the first and second codon positions. There is no compelling evidence for a most or least accurate position. By referring to some previous results that demonstrated also third position misreading under aminoacyl-tRNA limitation (Parker et al., 1978; Parker & Friesen, 1980), it can be concluded that mistakes occur in all three positions of codons. (ii) Pyrimidines are read more accurately that purines. The most striking observation is that U cannot be read as either C or G; similarly, C can hardly be mistaken for any other base. By contrast, A is easily read as U, and G can be taken as A, U, or C. However, the latter finding is not confimred by the results from competition experiments, which instead indicate that G cannot be mistaken for any other base. It therefore appears that, when it is present in a given codon, this purine can be misread only in the extreme situation where the corresponding cognate tRNA is totally missing. Such behavior seems to be specific to G because in all other cases investigated, the base replacements detected in the substitution assays have actually been confirmed by competition experiments. (iii) There is no selective misreading of any base of a given class, purine or pyrimidine, by the other base of the same class. Thus, the purine A can be read as the pyrimidine U and vice versa; also, G can be mistaken for U or C (but not the reverse). Interestingly, when analyzing the misreading of pyrimidines for each other, it appears that C can be read, or not read, as U depending on the nature of the codon. Indeed, the UCU codon can be mistaken for UUU, while the GCU codon cannot be read as GUU. The nature of the bases within a codon sequence seems therefore to influence the extent of miscoding, as already suggested by Yarus (1979) from the observation that no translation error can be detected when a central pyrimidine is flanked by two purines. From the same point of view, the nature of the base adjacent to the 5′ codon position has been shown to influence the strength and accurateness of codon–anticodon interaction (Weissenbach & Grosjean, 1981), an effect known as the "context effect", which could be analyzed also in the present cell-free system.

Obviously, to check the plausibility of the patterns of mistranslation inferred from this work, further experiments are now needed, using a variety of codons and cognate/near-cognate anticodons. It would also be of interest to analyze in detail the dependency of translational fidelity on various effectors such as guanosine polyphosphates, antibiotics, or ribosomal mutations.

Registry No. A, 73-24-5; C, 71-30-7; G, 73-40-5; U, 66-22-8.

References

Abraham, K. A. (1983) *Prog. Nucleic Acid Res. Mol. Biol. 28*, 81–100.

Bolivar, F., & Backman, K. (1979) *Methods Enzymol. 68*, 245–267.

Brakier-Gingras, L., & Phoenix, P. (1984) *Can. J. Biochem. Cell Biol. 62*, 231–244.

Burgess, R. R., & Jendrisak, J. J. (1975) *Biochemistry 14*, 4634–4638.

Cenatiempo, Y. (1986) *Biochimie 68*, 505–515.

Cenatiempo, Y., Robakis, N., Meza-Basso, L., Brot, N., Weissbach, H., & Reid, B. R. (1982a) *Proc. Natl. Acad. Sci. U.S.A. 79*, 1466–1468.

Cenatiempo, Y., Robakis, N., Reid, B., Weissbach, H., & Brot, N. (1982b) *Arch. Biochem. Biophys. 218*, 572–578.

Chen, H. Z., & Zubay, G. (1983) *Methods Enzymol. 101*, 674–690.

Crick, F. H. C. (1966) *J. Mol. Biol. 19*, 548–555.

Edelmann, P., & Gallant, J. A. (1977) *Cell (Cambridge, Mass.) 10*, 131–137.

Ellis, N., & Gallant, J. A. (1982) *MGG, Mol. Gen. Genet. 188*, 169–172.

Erion, J. L., Tarnowski, J., Weissbach, H., & Brot, N. (1981) *Proc. Natl. Acad. Sci. U.S. A. 78*, 3459–3463.

Fiil, N. P., Bendiak, D., Collins, J., & Friesen, J. D. (1979) *MGG, Mol. Gen. Genet. 173*, 39–50.

Gallant, J. A., & Foley, D. (1980) in *Ribosomes: Structure, Function, and Genetics* (Chambliss, G., Craven, G. R., Davies, J., Davis, K., Kahan, L., & Nomura, M., Eds.) pp 615–638, University Park Press, Baltimore, MD.

Gavrilova, L. P., & Spirin, A. S. (1971) *FEBS Lett. 17*, 324–326.

Goldman, E. (1982) *J. Mol. Biol. 158*, 619–636.

Hershey, J. W. B., Yanov, J., & Fakunding, J. L. (1979) *Methods Enzymol. 60*, 3–11.
Hopfield, J. J., & Yamane, T. (1980) in *Ribosomes: Structure, Function, and Genetics* (Chambliss, G., Craven, G. R., Davies, J., Davis, K., Kahan, L., & Nomura, M., Eds.) pp 585–596, University Park Press, Baltimore, MD.
Ikemura, T. (1981) *J. Mol. Biol. 146*, 1–21.
Jelenc, P. C., & Kurland, C. G. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 3174–3178.
Khazaie, K., Buchanan, J. H., & Rosenberger, R. F. (1984) *Eur. J. Biochem. 144*, 491–495.
Kung, H. F., Spears, C., & Weissbach, H. (1975) *J. Biol. Chem. 250*, 1556–1562.
Kung, H. F., Bekesi, E., Guterman, S. K., Gray, J. E., Traub, L., & Calhoun, D. H. (1984) *MGG, Mol. Gen. Genet. 193*, 210–213.
Lagerkvist, U. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 1759–1762.
Laughrea, M. (1981) *Biochimie 63*, 145–168.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Meza-Basso, L., Robakis, N., Cenatiempo, Y., Weissbach, H., & Brot, N. (1981) *Biochem. Biophys. Res. Commun. 101*, 459–463.
O'Farrell, P. H. (1978) *Cell (Cambridge, Mass.) 14*, 545–557.
Parker, J., & Friesen, J. D. (1980) *MGG, Mol. Gen. Genet. 177*, 439–445.
Parker, J., Pollard, J. W., Friesen, J. D., & Stanners, C. P. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 1091–1095.
Pestka, S. (1968) *J. Biol. Chem. 243*, 2810–2820.
Plumbridge, J. A., Deville, F., Sacerdot, C., Petersen, H. U., Cenatiempo, Y., Cozzone, A. J., Grunberg-Manago, M., & Hershey, J. W. B. (1985) *EMBO J. 4*, 223–229.
Robakis, N., Cenatiempo, Y., Meza-Basso, L., Brot, N., & Weissbach, H. (1983) *Methods Enzymol. 101*, 690–706.
Sancar, A., Hack, A. M., & Rupp, W. D. (1979) *J. Bacteriol. 137*, 692–693.
Stüber, D., & Bujard, H. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 167–171.
Sutcliffe, J. G. (1979) *Cold Spring Harbor Symp. Quant. Biol. 43*, 77–90.
Thompson, R. C., & Stone, P. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 198–202.
Thompson, R. C., Dix, D. B., Gerson, R. B., & Karim, A. M. (1981) *J. Biol. Chem. 256*, 6676–6681.
Weissbach, H., Robakis, N., Cenatiempo, Y., & Brot, N. (1984) *Biotechniques 2*, 16–22.
Weissenbach, J., & Grosjean, H. (1981) *Eur. J. Biochem. 116*, 207–213.
Yarus, M. (1979) *Prog. Nucleic Acid Res. Mol. Biol. 23*, 195–225.
Yarus, M. (1982) *Science (Washington, D.C.) 218*, 646–652.
Yarus, M., & Thompson, R. C. (1983) *Cold Spring Harbor Monogr. Ser. No. 15*, 23–63.

Accessibility of the Leading End of Ribonucleic Acid in Transcription Complexes†

Susan L. Bernhard and Claude F. Meares*
Chemistry Department, University of California, Davis, California 95616
Received February 10, 1986; Revised Manuscript Received May 27, 1986

ABSTRACT: A photoaffinity-protection technique has been developed to study the accessibility of the leading (5′) end of nascent RNA as it passes through the transcription complex formed by *Escherichia coli* RNA polymerase and phage T7 DNA. The macromolecules contacted by the leading (5′) end of the growing RNA chain in the transcription complex have been determined previously by photoaffinity labeling experiments using aryl azides attached to the leading end of nascent RNA [Hanna, M. M., & Meares, C. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4238–4242]. By using thiols to reduce accessible photoprobes, we have modified the photoaffinity technique so that it tests the accessibility of the leading end of nascent RNA to small molecules in solution, as a function of RNA chain length. We examined in detail RNA molecules containing 11–50 nucleotides, whose 5′ ends label the β and β' enzyme subunits with good yield. The thiol's accessibility to the leading end of each transcript was determined by comparing the RNAs cross-linked to $\beta\beta'$ in thiol-treated samples to controls not treated with thiol. Incubation with 1 mM dithiothreitol for 5 min reduced approximately 36% of the 5′-azides on RNAs 11–13 bases long and approximately 43% on RNAs 28–37 bases long but practically none of the 5′-azides on RNAs 40–43 bases long. Also notable was the reduction of 34 ± 1% of the 5′-azides on RNA 12 bases long but only 14 ± 2% on the 14-base RNA; on the T7 A1 promoter, the leading end of the transcript diverges from the DNA template when the chain is between 12 and 14 bases long.

DNA-dependent RNA polymerases catalyze the synthesis of ribonucleic acid, using deoxyribonucleic acid as a template (Losick & Chamberlin, 1976; von Hippel et al., 1984; McClure, 1985). These enzymes are usually large, multisubunit assemblies; *Escherichia coli* RNA polymerase contains five major subunits, with a total M_r of 449K. The "core" enzyme consists of subunits β' (M_r 155 162; Ovchinnikov et al., 1982) and β (M_r 150 619; Ovchinnikov et al., 1981) and two α subunits (M_r 36 512; Ovchinnikov et al., 1977). The core enzyme is capable of elongating RNA but does not specifically initiate transcription at promoter sites on DNA. The RNA polymerase holoenzyme contains the core plus the

† Supported by Research Grant GM 25909 to C.F.M. from the National Institute of General Medical Sciences, National Institutes of Health.

0006-2960/86/0425-6397$01.50/0