J. Mol. Biol. (1983) **164**, 73–87

Context Effects: Translation of UAG Codon by Suppressor tRNA is Affected by the Sequence Following UAG in the Message

LIONELLO BOSSI†

European Molecular Biology Laboratory
Meyerhofstrasse 1, 6900 Heidelberg
Federal Republic of Germany

(Received 1 September 1982)

The efficiency of various suppressor tRNAs in reading the UAG amber codon has been measured at 42 sites in the *lacI* gene. Results indicate that: (1) for all suppressors, efficiency is not an *a priori* value; rather, it is determined at each site by the specific reading context of the suppressed codon; (2) the degree of sensitivity to context effects differs among suppressors. Most affected is amber suppressor *supE* (su2), whose activity varies over a 20-fold range depending on context; (3) context effects are produced by residues present at the 3′ side of the UAG codon. The most important role appears to be played by the base that is immediately adjacent to the codon. When this base is a purine, the amber codon is suppressed more efficiently than when a pyrimidine is in the same position. Superimposed on this initial pattern, the influence of bases further downstream to the UAG triplet can be detected also. The possibility is discussed that context effects are produced by the whole codon following UAG in the message.

1. Introduction

The results of several independent studies on informational suppression have shown that the ability of nonsense suppressor tRNA to induce translational readthrough of a nonsense codon can be affected by the nucleotide sequence surrounding this codon in the messenger RNA. The phenomenon has been described as the effect of reading context on suppression (Salser, 1969; Salser *et al.*, 1969; Yahata *et al.*, 1970; Akaboshi *et al.*, 1976; Colby *et al.*, 1976; Fluck *et al.*, 1977; Feinstein & Altman, 1977,1978; Fluck & Epstein, 1980). More recently, Bossi & Roth (1980) provided a first indication about the molecular nature of context effects. Looking for mutations increasing the efficiency of an amber suppressor, *supE*, at a particular site, these authors isolated and characterized context mutants. In these mutants the initial amber mutation has become efficiently suppressed by *supE* due to a base-pair change in the neighborhood of the amber codon. DNA sequence analysis showed that the context mutation results

† Present address: Howard Hughes Medical Institute, Department of Biology, University of Utah, Salt Lake City, Utah 84112, U.S.A.

0022-2836/83/050073-15 $03.00/0

from the substitution of A for C in the residue adjacent to the 3′ side of the UAG codon. This change determines a tenfold increase of *supE* efficiency in reading UAG (Bossi & Roth, 1980).

In recent years, the *lacI* gene of *Escherichia coli* has been the object of a fine structure genetic dissection resulting in the characterization, at DNA sequence level, of a large number of different nonsense mutations. The existence of these nonsense sites in a wide variety of known sequence environments provides a unique opportunity for a study on context effects: one can systematically analyze possible correlations between suppressor function and the sequence surrounding the nonsense codon. This approach is greatly facilitated by the use of the genetic system described by Miller & Albertini in the accompanying paper. In this new system, a number of *lacI* nonsense mutations have been transferred onto the *lacI* portion of an *in-frame lacI-Z* gene fusion. Normally, this fusion produces a hybrid protein that retains the activity of β-galactosidase (the *lacZ* gene product: Müller-Hill & Kania, 1974). Most of the activity is lost, however, if a *lacI* nonsense mutation is present in the fusion, since translation is prevented from entering the *lacZ* coding sequence. In this case, expression of the *lacZ* portion of the fused gene depends upon suppression of the upstream nonsense site. Since the amino acid sequence of the *lacI*-encoded portion of the hybrid protein is dispensable for its β-galactosidase activity, it is expected that the nature of the amino acid inserted by the suppressor at the nonsense site should be irrelevant to β-galactosidase function. Thus, the enzymatic activity of β-galactosidase can be taken as a direct measure of the suppressor-induced readthrough of translation. (For a detailed description of the system, see Miller & Albertini, 1983). Using this system, I have measured the suppression efficiencies of four different amber suppressors in *Salmonella typhimurium* at 42 different *lacI* amber sites. In the accompanying paper, an analogous approach was taken in the study of four amber suppressors and one UGA suppressor from *E. coli*. Results of both studies show that suppression is dramatically affected by the mRNA sequence on the 3′ side of the UAG codon.

2. Materials and Methods

(a) *Media and growth conditions*

L broth (1% (w/v) Difco Bacto-tryptone, 0·5% (w/v) Difco yeast extract, 0·5% (w/v) NaCl) was used as rich medium for routine cell growth. The E medium of Vogel & Bonner (1956) supplemented with 0·2% (w/v) glucose was used as minimal medium. When required, minimal medium was supplemented with 120 μg kanamycin sulfate/ml (Sigma), 0·3 mM-amino acids and 40 μg Xgal/ml (5-bromo-4-chloro-3-indolyl-β-D-galactoside: Sigma). Solid media contained 1·5% (w/v) Difco agar.

(b) *Bacterial strains*

The genotypes and the sources of the *S. typhimurium* strains harboring different suppressors are listed in Table 1. All these strains have a Lac⁻ phenotype due to the fact that wild-type *S. typhimurium* does not carry a functional *lac* operon. Also in Table 1 is the genotype of the original *E. coli* host of the F′ *lac* episomes carrying the various $lacI^{am}$-*Z* fusions. The *lacI* nonsense mutations and their sequence locations have been described (Miller *et al.*, 1977,1978). In this study, only amber mutants were used. Some of them had

TABLE 1

List of strains

Strain	Genotype	Source
S. typhimurium LT2		
DB7136	*leuA414*^{am} *hisC527*^{am}	D. Botstein *via* J. Roth
DB7154	*leuA414*^{am} *hisC527*^{am} *supD10*	D. Botstein *via* J. Roth
DB7155	*leuA414*^{am} *hisC527*^{am} *supE20*	D. Botstein *via* J. Roth
DB7156	*leuA414*^{am} *hisC527*^{am} *supF30*	D. Botstein *via* J. Roth
DB7157	*leuA414*^{am} *hisC527*^{am} *supJ60*	D. Botstein *via* J. Roth
E. coli		
XA103	*ara (lac-pro) nalA metB argE*^{am} *rif supF*	J. H. Miller

The set of isogenic *Salmonella* strains harboring different amber suppressors has been described (Winston *et al.*, 1979). *E. coli* strain XA103 is the original host of the F'*lac* episomes containing the different *lacI*^{am}-*Z* fusions.

been derived from wild-type by a single base change (indicated by the A prefix) or by a double base change (mutants X3 and X15; Miller *et al.*, 1978). Other mutants had been obtained by conversion of previously isolated ochre mutations into amber sites (J. Miller, unpublished data); they are indicated with the same O prefix as the original ochre site (Miller *et al.*, 1978). Recombination of the various *lacI* amber mutations into *lacI-Z* deletion *71-56-14* is described in the accompanying paper (Miller & Albertini, 1982). In all of the episomes the fused gene is expressed from the I^Q promoter (Müller-Hill *et al.*, 1968). A kanamycin-resistance Tn*5* insertion is present in the F factor DNA.

(c) *Episome transfer*

The transfer of the F' factors from their *E. coli* host (XA103) to the various suppressor-carrying strains of *Salmonella* was carried out as follows. L broth cultures (2 ml) of donor and recipient cells were separately grown to late log phase: 0·2 ml of donors was mixed with 0·5 ml of recipients and the mixtures diluted into 5 ml of L broth and incubated for 6 to 8 h at 37°C. Drops of the mating mixtures were spotted on selective plates. Selection was for kanamycin resistance on E medium plates supplemented with leucine, histidine and Xgal.

(d) *Assay of β-galactosidase*

β-Galactosidase was assayed as described by Miller (1972). *O*-Nitrophenyl-β-D-galactoside was obtained from Sigma.

3. Results

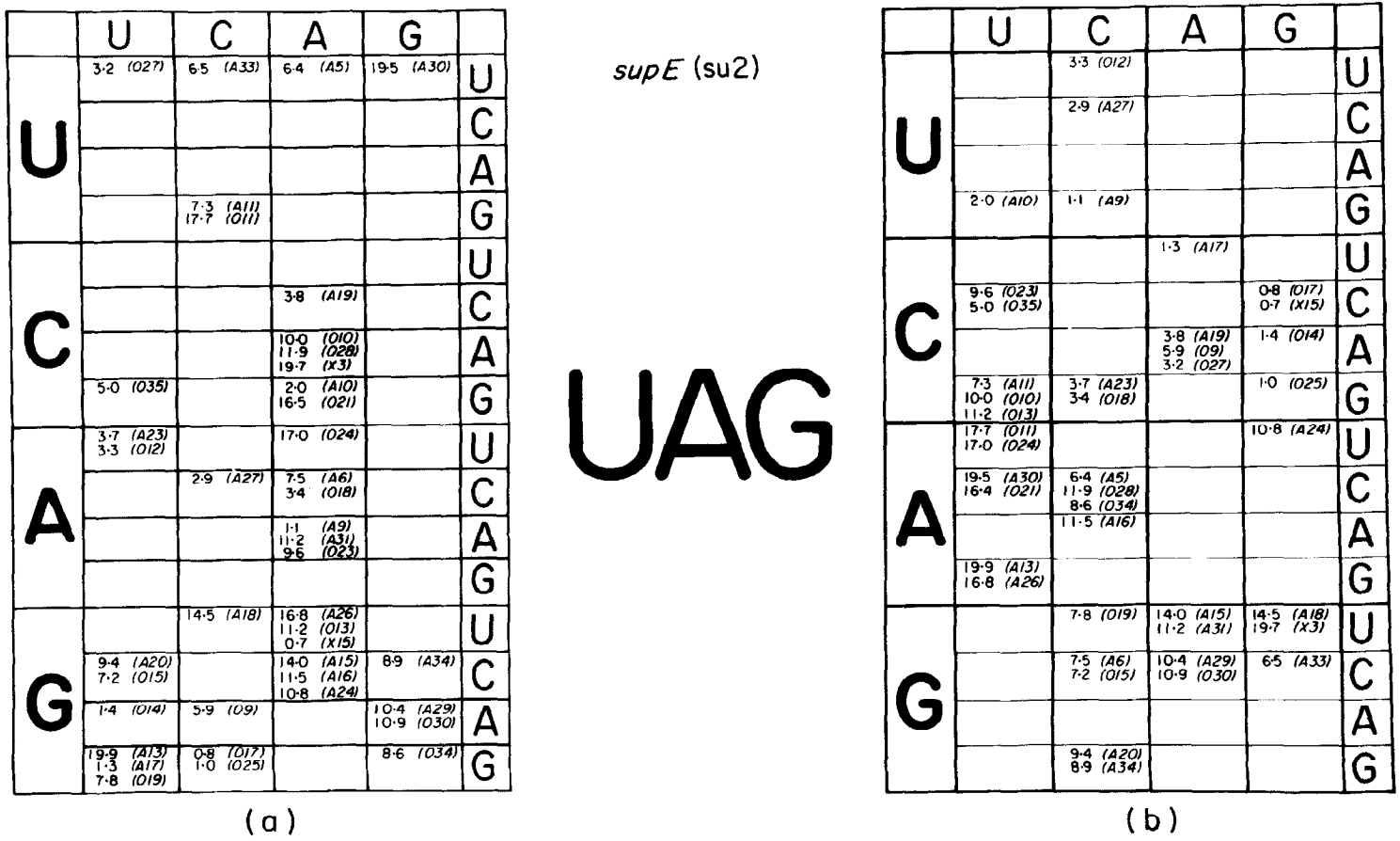
The various F' *lacI*^{am}-*Z* episomes were moved into a set of five isogenic *Salmonella* strains. Four of these strains contain a different amber suppressor, the fifth is wild-type for the suppressor alleles. This last strain was used to determine the levels of β-galactosidase produced by the various mutants in the absence of suppression. The resulting values, listed in Table 2, show a great variability in the "leakiness" of the amber mutation at the different sites. To explain this variability, one must consider that two causes are probably responsible for the leakiness: (1) low-level spontaneous readthrough at the amber codon; (2) reinitiation of translation downstream to the amber site. Context effects might cause differences

TABLE 2

List of lacI *amber mutations*

Mutant site	Amber codon context	Affected residue	Residual β-galactosidase activity in the *lacI*^{am}-*Z* fusions
A5	UAU UAG ACC	Gln	1·4
A6	AAC UAG GCC	Gln	2·0
A9	AAA UAG UCG	Gln	1·2
A10	CAG UAG UUG	Ser	1·3
A11	UCG UAG CUG	Leu	<0·1
A13	GUG UAG AUG	Ser	0·7
A15	GAC UAG GAU	Gln	0·1
A16	GAC UAG ACA	Gln	<0·1
A17	GUG UAG CAU	Glu	<0·1
A18	GCU UAG GGU	Leu	<0·1
A19	CAC UAG CAA	Gln	<0·1
A20	GUC UAG GCG	Ser	<0·1
A23	AUU UAG CCG	Gln	<0·1
A24	GAC UAG AGU	Trp	0·1
A26	GAU UAG AUG	Gln	0·1
A27	ACC UAG UCC	Glu	<0·1
A29	GGA UAG GAC	Tyr	0·7
A30	UGU UAG AUC	Tyr	1·2
A31	AAA UAG GAU	Gln	0·1
A33	UCU UAG GGC	Gln	0·1
A34	GGC UAG GCG	Gln	0·3
O9	GCA UAG CAA	Gln	0·7
O10	CAA UAG CUG	Gln	0·7
O11	UCG UAG AUU	Gln	<0·1
O12	AUU UAG UCU	Lys	<0·1
O13	GAU UAG CUG	Gln	0·2
O14	GUA UAG CGA	Glu	<0·1
O15	GUC UAG GCC	Glu	<0·1
O17	GCG UAG CGC	Gln	<0·1
O18	ACC UAG CCG	Tyr	<0·1
O19	GUG UAG GCU	Glu	<0·1
O21	CAG UAG AUC	Gln	0·1
O23	AAA UAG CUC	Tyr	<0·1
O24	AAU UAG AUU	Gln	0·1
O25	GCG UAG CGG	Glu	<0·1
O27	UUU UAG CAA	Gln	<0·1
O28	CAA UAG ACC	Gln	<0·1
O30	GGA UAG GAC	Tyr	0·6
O34	GGG UAG ACC	Gln	0·2
O35	CUG UAG CUC	Gln	0·2
X3	CAA UAG GGU	Leu	0·2
X15	GAU UAG CGC	Phe	0·1

The various *lacI* mutant sites, their surrounding nucleotide sequences, and the amino acid specificities of the affected codons are from Miller *et al.* (1978). The amber sites with an *O* designation were obtained by conversion of previously isolated ochre mutations (Miller *et al.* 1978; and unpublished data). In all the mutants the *lacI* gene was genetically fused to the *lac* operon. The various *lacI*^{am}-*Z* fusions (present on F′ *lac* episomes) were transferred into strain DB7136 (*supE*^{wt}) as described in Materials and Methods. β-Galactosidase activity was measured in the resulting strains as described by Miller (1972). These values are expressed as the percentage of the activity measured for the *lacI-Z* parental fusion in the same genetic background. Values are the averages of 4 enzymatic measurements.



	U	C	A	G	
U	3·2 (027)	6·5 (A33)	6·4 (A5)	19·5 (A30)	U
					C
					A
		7·3 (A11) 17·7 (011)			G
C					U
			3·8 (A19)		C
			10·0 (010) 11·9 (028) 19·7 (x3)		A
	5·0 (035)		2·0 (A10) 16·5 (021)		G
A	3·7 (A23) 3·3 (012)		17·0 (024)		U
		2·9 (A27)	7·5 (A6) 3·4 (018)		C
			1·1 (A9) 11·2 (A31) 9·6 (023)		A
					G
G		14·5 (A18)	16·8 (A26) 11·2 (013) 0·7 (x15)		U
	9·4 (A20) 7·2 (015)		14·0 (A15) 11·5 (A16) 10·8 (A24)	8·9 (A34)	C
	1·4 (014)	5·9 (09)		10·4 (A29) 10·9 (030)	A
	19·9 (A13) 1·3 (A17) 7·8 (019)	0·8 (017) 1·0 (025)		8·6 (034)	G

	U	C	A	G	
U		3·3 (012)			U
		2·9 (A27)			C
					A
	2·0 (A10)	1·1 (A9)			G
C			1·3 (A17)		U
	9·6 (023) 5·0 (035)			0·8 (017) 0·7 (x15)	C
			3·8 (A19) 5·9 (09) 3·2 (027)	1·4 (014)	A
	7·3 (A11) 10·0 (010) 11·2 (013)	3·7 (A23) 3·4 (018)		1·0 (025)	G
A	17·7 (011) 17·0 (024)			10·8 (A24)	U
	19·5 (A30) 16·4 (021)	6·4 (A5) 11·9 (028) 8·6 (034)			C
		11·5 (A16)			A
	19·9 (A13) 16·8 (A26)				G
G		7·8 (019)	14·0 (A15) 11·2 (A31)	14·5 (A18) 19·7 (x3)	U
		7·5 (A6) 7·2 (015)	10·4 (A29) 10·9 (030)	6·5 (A33)	C
					A
		9·4 (A20) 8·9 (A34)			G

FIG. 1. Correlation between *supE* efficiency of suppression and the sequence surrounding the amber codon. The β-galactosidase activities produced by the different *lacI*^{am}-*Z* fusions in the presence of the *supE* suppressor were determined. These values, expressed as the percentage of the activity of the parental *lacI*-*Z* fusion (assayed in the same genetic background), were corrected for the respective background levels (measured in the absence of suppression: Table 2). The resulting figures, representing the *supE* efficiencies of suppression at the various sites, are listed. (a) Suppression efficiencies are correlated to the sequence of the triplets preceding the UAG codon: (b) the same values are shown in relation to the triplets following UAG. Values are the averages of 6 determinations.

in the readthrough values. It seems possible that the occasional suppression activity of normal tRNA (responsible for the readthrough) might be influenced by the particular context of the amber codon. On the other hand, this interesting possibility cannot be firmly established, because of the presence of a second source of variation. At some sites within the *lacI* sequence, translation can reinitiate after being terminated by a preceding nonsense mutation. The efficiency of reinitiation varies depending upon the position of the nonsense codon relative to the reinitiation site (Platt *et al.*, 1972; Files *et al.*, 1974; Miller & Albertini, 1983). In spite of their fluctuations, the β-galactosidase levels in Table 2 in most cases represent only a small percentage of the activity restored by a nonsense suppressor. These background levels were taken into account when determining the suppression efficiencies in the following experiments.

(a) *The* supE *suppressor*

The β-galactosidase activities produced by the action of the *supE* suppressor at the different amber sites were determined. The values of suppression are given in Figure 1, both in relation to the base triplet preceding UAG (Fig. 1(a)) and to the triplet following the amber codon (Fig. 1(b)). Examination of the data in Figure 1(a) shows that there is no significant correlation between *supE* efficiency and the sequence at the 5′ side of UAG. For example, the role of the base immediately preceding UAG can be ruled out by comparing the values in the bottom row of (a). At all the sites in this row, UAG is preceded by codons ending with the same G residue; yet *supE* efficiency can be as high as 20% in one case (*A13*) and as low as 1% at other sites (*O17*, *O25*). The same bottom part of (a) reveals that, even when the entire sequence of the preceding triplet is the same, as it is for mutations *A13* and *A17* (both preceded by GUG), *supE* activity can vary over one order of magnitude. A similar situation is observed at sites preceded by triplets AAA (efficiency is 1·1% at *A9* and 11·2% at *A31*) and GAU (where efficiency varies from 0·7% to 16·8%). In light of these observations, it is likely that the few cases in which a correlation seems to exist between the β-galactosidase values and the sequence preceding the amber codon are either fortuitous, or due to the presence of similar sequences at the 3′ side of UAG (e.g. *O17* and *O25*; see below).

A clearcut pattern emerges if one examines *supE* efficiency in relation to the sequence of the triplet that follows the amber codon (Fig. 1(b)). A predominant effect on suppression can be attributed to the base immediately adjacent to the 3′ side of the amber codon. The highest values of *supE* efficiency are seen at sites where UAG is followed by triplets starting with A, and to a somewhat lesser extent, with G residues. In contrast, at most locations where UAG is followed by U or C, *supE* activity is low. Some deviations from this initial pattern appear upon closer examination. The most relevant exception is represented by the relatively high suppression values at sites followed by CUN triplets. This is particularly evident by comparing the sites followed by CUN to those followed by CGN triplets. In Figure 1(b), five sites are listed in which UAG is followed by either CUG or CUC, while in four cases UAG is followed by CGC, CGG or CGA. Despite the fact that in all these cases C is the base adjacent to the 3′ side of the amber codon, the *supE*

suppressor appears to distinguish between the two sets of triplets: the average value of suppression efficiency for the CUN sites is 8%, but as low as 1% for the CGN sites. Similar, though less dramatic, differences can be seen in the suppression of sites followed by triplets starting with A. Although all these sites are, in general, efficiently suppressed by *supE*, it is clear that the highest efficiency values (18% on the average) correlate with the presence of an AUN triplet on the 3′ side of the amber codon. At sites where the amber codon is followed by ACN triplets, the values of suppression are consistently lower (9% on the average).

In summary, the above results indicate that the first two residues in the sequence adjacent to the 3′ side of UAG have the property of affecting suppression. Depending on whether these residues act *individually* or as components of the *following codon*, two alternative explanations for the production of context effects can be entertained (see Discussion).

(b) *The other suppressors*

The effects of context were analyzed for three other amber suppressors: *supD* (sul), *supF* (su3) and *supJ* (Winston *et al.*, 1979). The values of the suppression efficiencies are listed in Figures 2, 3 and 4, tabulated with respect to the sequence present on the 3′ side of the amber codon. (Listing of the data in relation to the sequence preceding UAG has been omitted here since, as in the case of *supE*, no significant correlation could be found.) These data indicate that all three suppressors are affected by context, although to a lesser extent than *supE*. The efficiency of *supD* shows a sevenfold variation (ranging from 8·4% to 57·3%; Fig. 2), while that of *supF*, ranging from 9·5% to 79·4% (Fig. 3), varies over an eightfold range. *supJ* is the least affected by context changes; its efficiency varies about threefold (from 14·8% to 51·9%; Fig. 4). Despite their reduced sensitivity to context effects, these three suppressors exhibit the same qualitative behavior as *supE*: at sites where *supE* is inefficient, the other suppressors also show a lower efficiency; the same holds true for the more efficiently suppressed sites as well. In this respect, these findings are somewhat at odds with the original observation reported by Bossi & Roth (1980). These authors compared various suppressors for their ability to read the amber codon in two different contexts: UAG followed by CGU or AGU. While the behavior of *supE* and *supF* observed in that case is consistent with the results presented here, the *supD* suppressor showed a lower efficiency at the UAG-AGU site. In light of the data in this paper (see *supD* suppression of *A24* in Fig. 2), it seems likely that the previous result could be in error, due perhaps to a mutation impairing *supD* activity that was selected accidentally in the construction of the strain containing the UAG-AGU site. This seems possible in view of the "sickness" caused in that strain by the constitutive expression of the *his* operon (Murray & Hartman, 1972). This sickness is retained in efficiently suppressed *his* nonsense mutants due to suppressor-induced release of polarity.

In a few cases, the analogies between the behavior of the three suppressors and that of *supE* are somewhat questionable. This can be seen in the suppression of mutations *A13* and *A18*. At both sites the efficiency of *supE* is quite high (relative

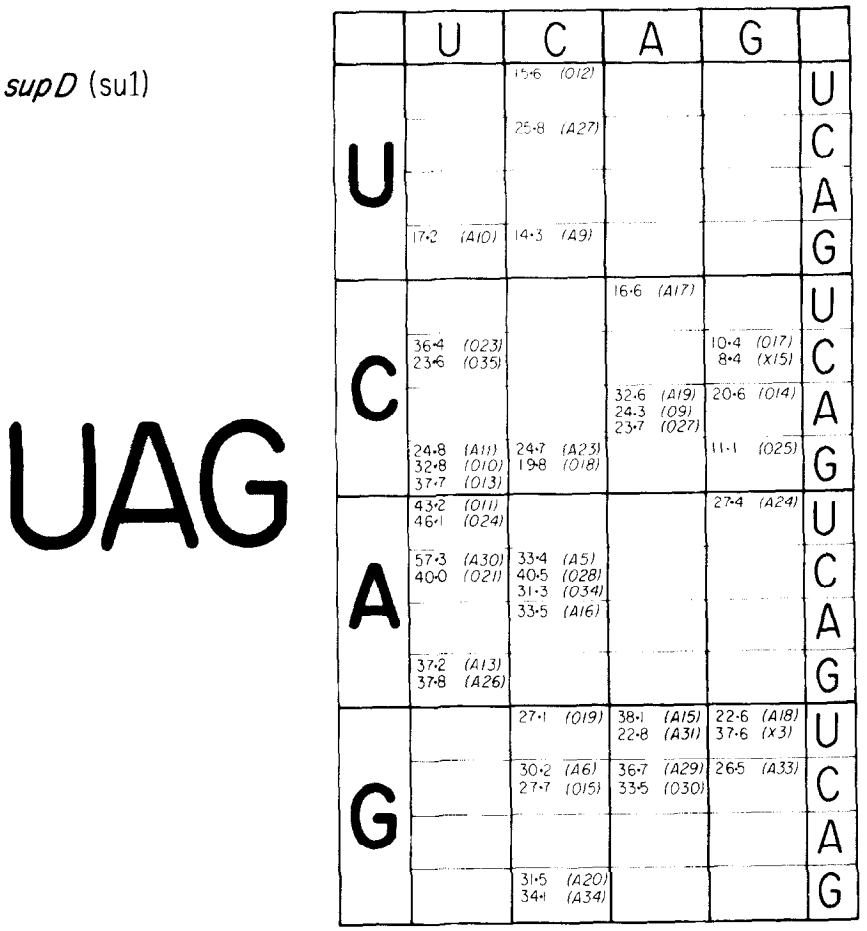


FIG. 2. Correlation between the efficiency of *supD* and the sequence at the 3′ side of the amber codon. Suppression efficiencies were determined as described for *supE* (see the legend to Fig. 1). Values are the averages of 4 enzymatic measurements.

to this suppressor range of efficiencies; Fig. 1). On the contrary, the other suppressors *(supD* in particular) elicit, at the same sites, low relative efficiencies (Figs 2 to 4). If, for example, we compare suppression of *A18* to that of a site like *A19*, we see that, while *supE* efficiencies are 14·5% and 3·8%, respectively, *supD* is respectively 22·6% and 32·6% efficient (Figs 1 and 2).

4. Discussion

The data presented in this paper indicate that a correlation exists between efficiency of suppression and the sequence immediately adjacent to the nonsense codon at the 3′ side. In the particular system where these effects were studied (amber mutations in the *lacI* portion of a *lacI-Z* gene fusion), the assumption was made that the enzymatic activity of the *lacZ* gene product could be considered a

supF (su3)

UAG

	U	C	A	G	
U		16·6 (012)			U
		25·4 (A27)			C
					A
	20·7 (A10)	23·4 (A9)			G
C			18·6 (A17)		U
	34·8 (023) 26·6 (035)			9·5 (017) 15·4 (X15)	C
			31·9 (A19) 29·9 (09) 26·6 (027)	16·1 (014)	A
	21·1 (A11) 32·2 (010) 42·9 (013)	26·1 (A23) 24·5 (018)		11·7 (025)	G
A	47·1 (011) 47·0 (024)			35·2 (A24)	U
	55·7 (A30) 39·6 (021)	31·3 (A5) 44·9 (028) 34·4 (034)			C
		34·8 (A16)			A
	49·7 (A13) 79·4 (A26)				G
G		29·9 (019)	36·9 (A15) 31·8 (A31)	38·2 (A18) 42·7 (X3)	U
		26·1 (A6) 32·1 (015)	41·5 (A29) 42·7 (030)	31·6 (A33)	C
					A
		36·1 (A20) 32·7 (A34)			G

FIG. 3. Correlation between the efficiency of *supF* and the sequence at the 3′ side of the amber codon. Suppression efficiencies were determined as described in the legend to Fig. 1. Values are the averages of 2 determinations.

faithful measure of the suppressor-induced readthrough at the amber codon. This assumption neglects the possibility that, in some cases, particular amino acid substitutions in the *lacI* portion of the hybrid protein (resulting from suppressor action) could affect the activity of the *lacZ* moiety (perhaps by interfering with correct folding or by affecting stability of the hybrid polypeptide). Although this possibility has not been ruled out, it is also clear that it does not invalidate the conclusions of the work described here. The conclusions are based on the analysis of suppression at a large number of different sites. The result is an overall consistent picture with no significant exceptions. Moreover, in the case of the *supE* suppressor, in principle one could limit the analysis of suppression to the subsets of amber mutations that are derived from glutamine codons. Since glutamine is the amino acid being inserted by *supE*, the suppressor action at these sites regenerates a wild-type *lacI* polypeptide sequence. There are as many as 22 such sites in the

supJ (su6)

UAG

	U	C	A	G	
U		17·4 (012)			U
		27·3 (A27)			C
					A
	18·5 (A10)	26· (A9)			G
C			20·3 (A17)		U
	34·8 (023) 31·5 (035)			16·4 (017) 15·6 (x15)	C
			39·0 (A19) 32·4 (09) 28·4 (027)	18·3 (014)	A
	29·1 (A11) 33·9 (010) 41·2 (013)	30·8 (A23) 23·5 (018)		14·8 (025)	G
A	48·0 (011) 45·4 (024)			34·6 (A24)	U
	51·9 (A30) 39·8 (021)	35·8 (A5) 46·1 (028) 33·1 (034)			C
		38·0 (A16)			A
	45·2 (A13) 42·0 (A26)				G
G		26·9 (019)	39·7 (A15) 26·1 (A31)	31·7 (A18) 36·0 (x3)	U
		36·7 (A6) 31·5 (015)	30·9 (A29) 30·7 (030)	33·6 (A33)	C
					A
		39·8 (A20) 32·3 (A34)			G

FIG. 4. Correlation between the efficiency of *supJ* and the sequence at the 3′ side of the amber codon. Suppression efficiencies were determined as described in the legend to Fig. 1. Values are the averages of 2 determinations.

collection of mutants studied here (Table 2). The sample is, therefore, sufficiently large to represent a wide variety of different contexts. Examination of the *supE* efficiency values at these sites, shows the presence of the familiar pattern of context sensitivity (Fig. 1).

(a) *Efficiency of suppression: a reconsideration*

Suppression efficiency is a parameter that is often used to indicate the strength of a given suppressor. Implicitly, therefore, efficiency of suppression is considered a tRNA attribute. The results presented in this paper challenge this conclusion. Rather, suppression efficiency seems to depend on a combination of tRNA and mRNA features. With the *supE* suppressor, the reading context (i.e. the message component), plays a major role in determining efficiency of suppression. Experiments described in the accompanying paper on a UGA suppressor indicate that in some cases, the message component practically dictates whether tRNA can

perform its function. On the other hand, the results presented here show that suppressors such as *supD*, *supF* and *supJ* are much less affected by context and the overall efficiency of reading resides mainly in the tRNA component. It appears, therefore, that some features in the suppressor tRNA structure determine the susceptibility of the suppressor to context effects. The different behavior of *supE* relative to the other suppressors is reminiscent of other idiosyncrasies of this particular suppressor. It has been known for some time that *supE* is, among the amber suppressors, the most dramatically "restricted" by some *strA* alleles (Strigini & Gorini, 1970). More recently, it has also been shown that the mutation *hisT* (altering post-transcriptional modification of several tRNAs) dramatically affects the function of *supE* (Bossi & Roth, 1980). It seems possible that all these characteristics, i.e. sensitivity to context effects, sensitivity to *strA* and *hisT* mutations, are different phenotypes related to a unique feature in *supE* structure. This hypothesis is supported by the existence of mutations, mapping in the *supE* genetic locus, which simultaneously render the suppressor insensitive to *strA*, *hisT* and to the effects of poor reading contexts (Bossi & Roth, unpublished results).

(b) *Nature of context effects*

The molecular mechanism determining context effects is still largely unknown. The starting point in developing models is the notion that nonsense suppression is the outcome of a competition between the suppressor tRNA and the factor(s) responsible for termination of translation. It follows that one must make a basic separation between two alternative sets of hypotheses.

(1) Context affects the reaction of peptide-chain termination.
(2) Context affects the function of the suppressor tRNA.

Several aspects in the process of peptide chain termination are still unclear, including the mechanism of action of the release factors in their recognition of the nonsense triplet (reviewed by Caskey & Campbell, 1979). One can speculate that the mechanism of translation termination could operate on an mRNA sequence-target longer than the nonsense triplet. As a result, the efficiency of termination could be affected by the sequence immediately following the nonsense codon. This hypothesis was originally proposed to explain the context effects observed in the suppression of bacteriophage T4 amber mutants (Salser, 1969; Salser *et al.*, 1969; Fluck *et al.*, 1977; Fluck & Epstein, 1980). According to this model, one would predict that different suppressors should behave similarly with respect to context; that is, to be efficient at sites where the termination mechanism is poorly active and *vice versa*. In general, the results presented in this paper are consistent with this prediction. However, a possible objection is raised by previous studies on the behavior of ochre suppressors. Although ochre suppressors appear to follow the same general rules when decoding ochre codons (Fluck & Epstein, 1980), they behave differently in the suppression of UAG amber sites (by wobble pairing). At these sites, ochre suppressors seem to follow a different pattern of context sensitivity (Akaboshi *et al.* 1976; Feinstein & Altman, 1978; Bossi & Roth, 1980).

The second hypothesis noted above is that codon context might influence the

function of the suppressor tRNA. In turn, this hypothesis can be developed along two different lines depending on the following possibilities: (1) the bases at the 3′ side of the amber codon are involved directly and individually in the creation of context effects; (2) the role of the bases at the 3′ side of UAG is the indirect consequence of the fact that they are part of the adjacent codon.

(i) *The "swollen" codon hypothesis*

It seems possible that the tRNA–mRNA interaction at the ribosomal A site could involve a larger number of base-pairs than the conventional three: namely, four or perhaps five base-pairs. This possibility has been suggested in the past. Data suggesting that a fourth base-pair could be formed between fMet-tRNA and the residue adjacent to the 3′ side of the AUG initiator codon were obtained by Taniguchi & Weissmann (1978). The hypothesis of a five base-pair tRNA–mRNA interaction was originally proposed by Crick *et al.* (1976) as a model for a primordial translation mechanism and, more recently, has been adapted by Pieczenik (1980) to contemporary translation. The hypothesis suggests that, during the interaction with the message, the anticodon loop of the aminoacyl tRNA is in a configuration (5 bases stacked on top of each other) that allows formation of two additional base-pairs involving the two residues adjacent to the 3′ side of the codon. How do these hypotheses fit with the results obtained in the present study? All known elongator tRNAs contain an invariant uracil residue, U33, at the position adjacent to the anticodon on the 5′ side. If U33 can be involved in a fourth base-pair with the residue adjacent to the 3′ side of the codon, one would expect that the presence of A or G in the mRNA would strengthen the interaction; in the case of suppression this would presumably result in an increased efficiency. It is clear that the results shown here support this prediction. At sites where the amber codon is followed by a purine residue, suppression efficiency is consistently high. In contrast, the presence of a pyrimidine at the 3′ side of an amber codon usually causes it to be poorly read. As previously pointed out, however, the residue immediately adjacent to UAG is not sufficient to account for all the data. In particular, the efficient *supE* suppression of sites followed by CUN triplets (discussed in the text) might suggest a role for the U present in the second position of this triplet. One could be tempted to rationalize these results on the basis of the five base-pair interaction model. Accordingly, the U in the CUN triplets should be able to base-pair with the corresponding residue on the 5′ side of the suppressor tRNA anticodon. However, inspection of the *supE* tRNA sequence (Inoguchi *et al.*, 1979), shows that no such base-pairing is possible: the base on the 5′ side of U33 in the *supE* anticodon loop is also a U.

(ii) *The tRNA–tRNA interaction hypothesis*

An alternative hypothesis is based on the observation that context effects can be correlated with the nature of the codon adjacent to the 3′ side of the amber site. Since a codon is a real entity to the extent that it is actually translated by tRNA, one can speculate that the tRNA reading the adjacent codon is the primary source of context effects. This could occur if the activity of the suppressor can be affected

by some sort of interaction with the tRNA reading the adjacent codon on the ribosome. To entertain this model, one must challenge the initial dogma that the role of context ultimately is to affect the competition between suppressor and release factor(s). Perhaps even after translocating at the ribosomal P site, the peptidyl suppressor still has a finite probability of falling off the ribosome. (Production of abortive peptidyl-tRNA molecules has been shown to be a frequent event during translation in bacteria (Menninger, 1976).) An interaction with the following tRNA could affect the "stickiness" of the peptidyl suppressor to the translation complex.

Alternatively, a tRNA–tRNA interaction could influence the rate of peptide chain elongation. In both cases codon context would be the expression of the degree of "fitness" between the A-site tRNA and the P-site peptidyl suppressor. By applying the model to the data shown in this paper, one could conclude that most tRNAs that read codons starting with A or G might have a high degree of fitness with all the amber suppressor tRNAs, while tRNAs reading codons starting with C or U interfere with suppressor activity (with the possible exception of CUN-reading tRNAs). If this is the case, we might expect to find a correlation between some structural features of tRNA and the first residue of the codon the tRNA recognizes. Interestingly, in some cases such a correlation has been seen. Most tRNAs that read codons starting with A carry the same modified adenine at the position adjacent to the 3′ side of the anticodon, while a different adenine modification is present at the same position of most tRNAs that read codons starting with U (Nishimura, 1972). It is intriguing that the modified base is at a position of the tRNA that may end up in close proximity to the invariant U33 in the anticodon loop of the tRNA reading the preceding codon on the ribosome (see the diagram in Fig. 5). In this respect it is significant that in order for the hypothesis of a tRNA–tRNA interaction to be consistent with the data, one must conclude that the interaction involves a *constant domain* of the P-site tRNA and a *variable domain* of the A-site rRNA. Otherwise, one could not explain the observed lack of context effects due to residues at the 5′ side of the amber codon. (The interaction should also be expected between the

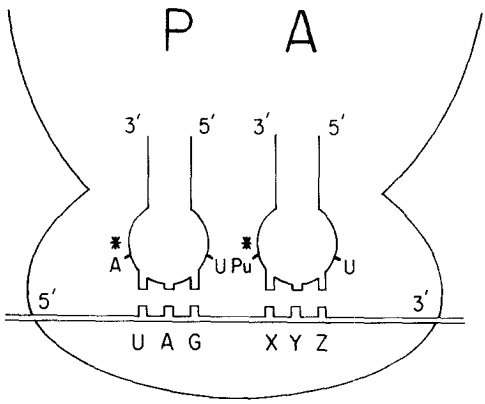


FIG. 5. Anticodon loops of tRNAs reading adjacent codons on the mRNA. The invariant U33 in the anticodon of the suppressor tRNA (at the ribosomal P site) is near to the modified purine in the anticodon of the tRNA reading the following codon (at the ribosomal A site).

tRNA reading the codon preceding UAG and the suppressor tRNA.) If the effects of context reflect the interaction between U33 in the P-site tRNA and the modified purine in the A-site tRNA, one would detect the effects only when the interaction involves the suppressor and the following tRNA (whose modified base varies from one tRNA to another); the interaction between the suppressor and the preceding tRNA (whose U33 is invariant) is expected to be constant at all sites.

5. Conclusions

The major implication of the phenomenon described in this paper is independent of the particular mechanism responsible for the production of context effects. This is the notion that the actual signals for translation termination could include sequences downstream to the nonsense triplet. The hypothesis finds support in the analysis of the sequences surrounding natural termination codons. A recent compilation of the translation termination sequences for many eukaryotic and prokaryotic mRNAs reveals a statistically significant bias in the usage of bases at the 3′ side of the termination codons (Kohli & Grosjean, 1981). In prokaryotes, U is the base most frequently found following the termination codon. In a sense these data suggest the existence of a consensus sequence as the ideal signal for translation termination. One can think of this sequence as being composed of two domains: a relatively constant region essential for termination (the nonsense triplet) and a second, variable region (the 3′-side sequence) affecting the efficiency of termination. This model permits the solution of an old puzzle: namely, the existence of leaky terminators. Proteins resulting from the readthrough of the UGA terminator are synthesized by bacteriophage Qβ (Horiuchi *et al.*, 1971; Moore *et al.*, 1971; Weiner & Weber, 1971), by bacteriophage λ (Yates *et al.*, 1977), and by filamentous phages (Engelberg-Kulka *et al.*, 1979). Amber (UAG) readthrough proteins have been identified among the translation products of MuL and MS virus RNAs (Philipson *et al.*, 1978) and of tobacco mosaic virus RNA (Pelham, 1978). How can the same termination signal be tight at some sites and leaky at others? It seems possible that in all cases in which readthrough proteins might have become functionally important, the sequence at the 3′ side of the termination triplet could have evolved so as to allow a degree of suppression by normal tRNAs. Engelberg-Kulka (1981) has surveyed the sequences present at the 3′ side of the leaky UGA terminators. In all cases, an A residue is present adjacent to the 3′ side of the UGA codon.

One cannot distinguish whether context effects are produced by the translation termination mechanism or whether they reflect a property of tRNA. The latter possibility is intriguing for its general implications in genetic decoding: context effects need not be limited to translation termination, but could in principle influence the translation of any codon. This would have profound consequences on the structure and evolution of the genetic code. Rules may exist, built into the structure of the code, that govern the assortment of code-words relative to each other. Optimal translation of the genetic message would require each codon to be placed in a favorable environment. The degeneracy of the code could permit this mechanism to operate without putting excessive restrictions on the evolution of protein sequences.

I am very grateful to Jeffrey Miller. Without his generous gift of all the *lacI* mutants described in the paper, this work would not have been possible. I am also grateful to Hans Lehrach, Anne-Marie Frischauf and Riccardo Cortese for the hospitality in their laboratories during the course of this work. I thank Molly Schmid for suggestions, and Ray Gesteland, Sandy Parkinson and John Atkins for comments on the manuscript. I acknowledge the European Molecular Biology Organization for the support of a long-term fellowship.

REFERENCES

Akaboshi, E., Inouye, M. & Tsugita, A. (1976). *Mol. Gen. Genet.* **149**, 1–4.
Bossi, L. & Roth, J. R. (1980). *Nature (London)*, **286**, 123–127.
Caskey, C. T. & Campbell (1979). In *Nonsense Mutations and tRNA Suppressors* (Celis, J. E. & Smith, J. D., eds), pp. 81–96, Academic Press, London.
Colby, D. S., Schedl, P. & Guthrie, C. (1976). *Cell*, **9**, 449–463.
Crick, F. H. C., Brenner, S., Klug, A. & Pieczenik, G. (1976). *Orig. Life*, **7**, 389–397.
Engelberg-Kulka, H., Dekel, L., Israeli-Reches, M. & Belfort, M. (1979). *Mol. Gen. Genet.* **170**, 155–159.
Engelberg-Kulka, H. (1981). *Nucl. Acids Res.* **9**, 983–991.
Feinstein, S. I. & Altman, S. (1977). *J. Mol. Biol.* **112**, 453–470.
Feinstein, S. I. & Altman, S. (1978). *Genetics*, **88**, 201–219.
Files, J. G., Weber, K. & Miller, J. H. (1974). *Proc. Nat. Acad. Sci., U.S.A.* **71**, 667–670.
Fluck, M. M. & Epstein, R. H. (1980). *Mol. Gen. Genet.* **177**, 615–627.
Fluck, M. M., Salser, W. & Epstein, R. H. (1977). *Mol. Gen. Genet.* **151**, 137–149.
Horiuchi, K., Webster, R. E. & Matsuhashi, S. (1971). *Virology*, **45**, 429–439.
Inoguchi, H., Yamao, F., Sakano, H. & Ozeki, H. (1979). *J. Mol. Biol.* **132**, 649–662.
Kohli, J. & Grosjean, H. (1981). *Mol. Gen. Genet.* **182**, 430–439.
Menninger, J. R. (1976). *J. Biol. Chem.* **251**, 3392–3398.
Miller, J. H. (1972). In *Experiments in Molecular Genetics*, pp. 352–355, Cold Spring Harbor Laboratory, Cold Spring Harbor.
Miller, J. H. & Albertini, A. M. (1983). *J. Mol. Biol.* **164**, 59–71.
Miller, J. H., Ganem, D., Lu, P. & Schmitz, A. (1977). *J. Mol. Biol.* **109**, 275–301.
Miller, J. H., Coulondre, C. & Farabaugh, P. J. (1978). *Nature (London)*, **274**, 770–775.
Moore, C. H., Farron, F., Bohnert, D. & Weissmann, C. (1971). *Nature New Biol.* **234**, 204–206.
Müller-Hill, B. & Kania, J. (1974). *Nature (London)*, **249**, 561–562.
Müller-Hill, B., Crapo, L. & Gilbert, W. (1968). *Proc. Nat. Acad. Sci., U.S.A.* **59**, 1259–1264.
Murray, M. L. & Hartman, P. E. (1972). *Can. J. Microbiol.* **18**, 671–681.
Nishimura, S. (1972). *Progr. Nucl. Acid Res. Mol. Biol.* **12**, 49–85.
Pelham, H. R. B. (1978). *Nature (London)*. **272**, 469–471.
Philipson, L., Andersson, P., Olshevsky, U., Weinberg, R., Baltimore, D. & Gesteland, R. (1978). *Cell*, **13**, 189–199.
Pieczenik, G. (1980). *Proc. Nat. Acad. Sci., U.S.A.* **77**, 3539–3543.
Platt, T., Weber, K., Ganem, D. & Miller, J. H. (1972). *Proc. Nat. Acad. Sci., U.S.A.* **69**, 897–901.
Salser, W. (1969). *Mol. Gen. Genet.* **105**, 125–130.
Salser, W., Fluck, M. & Epstein, R. (1969). *Cold Spring Harbor Symp. Quant. Biol.* **34**, 513–520.
Strigini, P. & Gorini, L. (1970). *J. Mol. Biol.* **47**, 517–530.
Taniguchi, T. & Weissmann, C. (1978). *J. Mol. Biol.* **119**, 533–565.
Vogel, H. J. & Bonner, D. M. (1956). *J. Biol. Chem.* **218**, 97–106.
Weiner, A. M. & Weber, K. (1971). *Nature New Biol.* **234**, 206–209.
Winston, F., Botstein, D. & Miller, J. H. (1979). *J. Bacteriol.* **137**, 433–439.
Yahata, H., Ocada, Y. & Tsugita, A. (1970). *Mol. Gen. Genet.* **106**, 208–212.
Yates, J. L., Gette, W. R., Furth, M. E. & Nomura, M. (1977). *Proc. Nat. Acad. Sci., U.S.A.* **74**, 689–693.

Edited by S. Brenner