The EMBO Journal vol.13 no.1 pp.249–257, 1994

The second to last amino acid in the nascent peptide as a codon context determinant

Salim Mottagui-Tabar, Asgeir Björnsson and Leif A.Isaksson

Department of Microbiology, Stockholm University, S-106 91 Stockholm, Sweden

Communicated by K.Nordström

Forty-two different sense codons, coding for all 20 amino acids, were placed at the ribosomal E site location, two codons upstream of a UGA or UAG codon. The influence of these variable codons on readthrough of the stop codons was measured in *Escherichia coli*. A 30-fold difference in readthrough of the UGA codon was observed. Readthrough is not related to any property of the upstream codon, its cognate tRNA or the nature of its codon–anticodon interaction. Instead, it is the amino acid corresponding to the second upstream codon, in particular the acidic/basic property of this amino acid, which seems to be a major determinant. This amino acid effect is influenced by the identity of the A site stop codon and the efficiency of its decoding tRNA, which suggests a correlation with ribosomal pausing. The magnitude of the amino acid effect is in some cases different when UGA is decoded by a wildtype form of $tRNA^{Trp}$ as compared with a suppressor form of the same tRNA. This indicates that the structure of the A site decoding tRNA is also a determinant for the amino acid effect.

Key words: amino acid/codon context/exit site/readthrough/translation termination

Introduction

Decoding efficiency *in vivo* can be studied conveniently by measuring readthrough of nonsense codons by near-cognate or suppressor tRNAs. Such studies have revealed that decoding at the ribosomal A site (aminoacyl-tRNA acceptor site) is dependent on its codon context. At the 3′ side the efficiency of UGA readthrough is dependent on the base following the UGA codon such that A > G > C > U (Kopelowitz *et al.*, 1992; Björnsson and Isaksson, 1993). The 5′ flanking base or the P site codon (peptidyl tRNA binding site) is also a determinant (Buckingham *et al.*, 1990; Björnsson and Isaksson, 1993).

Recent work indicates that the mRNA region beyond the two codons flanking the A site codon can also influence its decoding. The *prfB* mRNA, which codes for release factor 2 (RF-2) in *Escherichia coli*, has an in-frame UGA codon in its coding region which is bypassed by a frameshift event giving the mature gene product (Craigen and Caskey, 1986; Donly *et al.*, 1990). One feature of this autoregulatory mechanism is an interaction between the anti-Shine–Dalgarno region of the 16S rRNA and a six base Shine–Dalgarno element two codons upstream of the UGA codon (Weiss *et al.*, 1987, 1988; Curran and Yarus, 1988). Another example is the programmed incorporation of selenocysteine at an internal UGA codon into the protein product of the *E.coli* gene *fdhF* (Zinoni *et al.*, 1990; Heider *et al.*, 1992; Baron *et al.*, 1993). The change of identity from a UGA termination codon to a selenocysteine codon is accomplished by a downstream stem-loop mRNA region. *E.coli* ribosomes can also bypass (hop) an mRNA region with several stop codons in the bacteriophage T4 gene *60* mRNA. The mRNA sequence in the bypass region as well as a stretch of 15 amino acids located ~18 amino acids upstream of the P site amino acid in the nascent peptide are important for this programmed event to take place (Weiss *et al.*, 1990).

It is clear from these examples that codon context can affect the decoding of stop codons and that cells can use codon context to regulate gene expression. Therefore it is important to unravel the molecular interactions which produce codon context effects. The simplest case seems to be the effect of the 3′ flanking nucleotide on A site decoding. The UAG decoding efficiency of both an amber suppressor tRNA and release factor 1 (RF-1) is dependent on this nucleotide (Pedersen and Curran, 1991). The high readthrough caused by A or G at the 3′ side of the A site codon probably results from an increased potential for a stacking interaction between the decoding tRNA anticodon–codon complex and the 3′ purine base (Stromo *et al.*, 1986). RF-1 interacts more readily with UAG if uracil is the 3′ base (Pedersen and Curran, 1991). Furthermore, uracil is the base most frequently found at the 3′ side of the UAA stop codon in *E.coli* (Brown *et al.*, 1990). These observations support the notion that the *E.coli* release factors might recognize a tetranucleotide signal involving the stop codon and a 3′ uracil base (Brown *et al.*, 1990). Release factor interaction has previously been shown to be a source of codon context effects (Ryden and Isaksson, 1984; Martin *et al.*, 1988).

It is more difficult to define the molecular interactions responsible for codon context effects of the 5′ flanking codon. One report shows that an interaction can take place between a suppressor tRNA at the A site, and a mutated tRNA at the 5′ flanking P site (Smith and Yarus, 1989). Other data also suggest that a tRNA–tRNA interaction between the P and A site tRNAs can influence A site decoding (Kato *et al.*, 1990).

Ribosomal decoding activity has been shown by *in vitro* experiments to be dependent on the binding of uncharged, cognate tRNA to the second codon upstream of the A site. This location corresponds to the ribosomal E site (exit site), which is the binding site for the uncharged tRNA prior to dissociation from the ribosome, after translocation from the P site (Rheinberger *et al.*, 1981; Grajevskaja *et al.*, 1982; Krillov *et al.*, 1983; Lill *et al.*, 1984). Even though most experimental support for the E site originates from *in vitro* experiments, tRNA binding to the E site has been implicated in polysomal ribosomes in extracts from growing bacteria,

suggesting that this site also is operative *in vivo* (Remme *et al.*, 1989). The results from *in vitro* experiments show that the ribosomal E and A sites interact by an allosteric mechanism which affects tRNA binding activity at each site (Rheinberger *et al.*, 1990). An uncharged cognate tRNA at the E site induces a low affinity state of the A site with respect to EF-Tu ternary complex binding, thereby giving a more accurate selection against erroneous tRNAs (Gnirke *et al.*, 1989). Therefore, the nature of the second codon upstream of a stop codon could be a codon context determinant for nonsense codon readthrough *in vivo*, although such experiments have not yet been reported.

Here, we have addressed the question of whether the second codon upstream of a suppressible UGA or UAG stop codon is a readthrough determinant *in vivo*. Our data show that altering the codon at this location changes UGA readthrough by as much as 30-fold. The observed readthrough effects can be correlated with the nature of the amino acid donated by the E site tRNA, the length of the postulated ribosomal pause at the stop codon, and in some cases the tRNA giving the readthrough.

Results

The 3A′ translational readthrough assay system

An *in vivo* assay system for readthrough of nonsense (stop) codons was used to probe for the influence of a variable codon located two codons upstream of the stop codon (Figure 1). This upstream location corresponds to the ribosomal E site when the UGA stop codon is at the A site, and is referred to here as the E site codon or the (second) upstream codon. The assay system is based on a gene which codes for three identical, engineered antibody binding B-domains of protein A from *Staphylococcus aureus*. The antibody binding domain was originally called Z (Nilsson *et al.*, 1987), but is here referred to as A′ (3A′ being the whole gene) in order to avoid confusion with *lacZ*. The sequences with the stop codon (UGA or UAG) and the different upstream codons were inserted into a linker between the segments coding for the second and third of the identical IgG binding domains. Influence of the upstream codon on readthrough of the stop codon in growing bacteria is detected as a difference in the ratio between the amounts of the three-domain (3A′) and two-domain (2A′) protein products formed. The 3A′ protein results from translational readthrough of the stop codon, and the 2A′ protein results from termination at this codon in the linker region. Protein products of the 3A′ gene were purified by IgG affinity chromatography and separated by SDS−PAGE. The protein gels are stained with CBB (Coomassie Brilliant Blue) G-250 stain, and the amounts of each protein product of the 3A′ gene were measured by laser densitometry. The molar ratio between the amounts of the 3A′ and 2A′ proteins thus reflects the relative efficiency of the tRNA and release factor reading of the stop codon. This ratio is referred to as the transmission value, and it gives quantitative information about A site decoding as a function of the variable upstream codon.

A related test gene, S3A′, which codes for a protein with a secretion signal peptide in front of 3A′, was used in a previous study on translational readthrough of UGA codons (Björnsson and Isaksson, 1993). We later found that other strains containing a plasmid with the S3A′ gene are disturbed in growth when the S3A′ gene has UAG or, in particular,

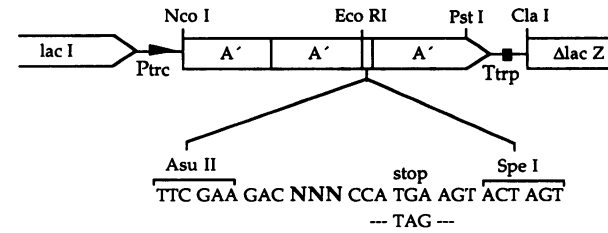

Fig. 1. The 3A′ reporter gene derivatives of the pAB93 plasmid. The synthetic 3A′ gene is transcribed from the P_{trc} promoter. This promoter is repressed by the *lacI* gene product constitutively expressed from the same plasmid. The *Asu*II and *Spe*I sites were used to clone E site sequence variants. The stop codons introduced are either UGA or UAG; E site codons are indicated by NNN. The third codon upstream of UGA was also varied using the same strategy (-TTC GAA XXX GAC CCA TGA AGT-), where XXX denotes the variable codon. T_{trp} is the 3A′ gene transcription terminator and ΔlacZ represents a *lacZ* gene with its 5′-part deleted to its first *Cla*I site.

UAA test codons in some codon contexts. This disturbance is dependent on the growth medium, and is mainly seen on Luria broth plates. It seems to be caused by a combination of several factors, in particular secretion of the gene product into the growth medium, together with the gene expression level, which is stop codon dependent. The growth disturbance is probably caused by induction of the heat shock response (Abrahmsén *et al.*, 1986). Furthermore, the S3A′ system yields partially processed signal peptide derivatives of the S3A′ protein products in addition to the correctly processed ones, which can make SDS−PAGE analysis difficult. Even though constructs with UGA do not give the indicated growth problems (Björnsson and Isaksson, 1993), we have chosen to use the 3A′ variant described here instead of S3A′, since none of the indicated problems have been observed using the 3A′ reporter gene.

Influence of the second upstream codon on readthrough of UGA

Forty-two different sense codons coding for all 20 amino acids were placed two codons upstream (E site) of a UGA stop codon and were analysed for their effects on readthrough of this stop codon. The E site codon variants were inserted into the linker between the second and third A′ coding domains of the 3A′ gene and assayed with respect to UGA readthrough. Protein products of the 3A′ gene were quantified in two *E.coli* strains, having a *trpT*$^{+}$(Su^{-}) or a *trpT*(Su9) allele coding for a wildtype or UGA suppressor form of $tRNA^{Trp}$, respectively (Coulondre and Miller, 1977). The amount of readthrough protein product 3A′ and termination product 2A′ was measured by laser scanning of the 3A′ and the 2A′ protein bands in gels; the results are shown in Table I as transmission values (3A′/2A′ protein molar ratios). These values are proportional to stop codon readthrough, and vary from 0.03 to 0.91 (30-fold) in the *trpT*$^{+}$(Su^{-}) strain. In the case of the *trpT*(Su9) strain, the range is slightly smaller (26-fold). The differences in readthrough are clearly visible by direct inspection of samples on protein gels (Figure 2) and are exemplified by the GAC (Asp), UAC (Tyr) and CGC (Arg) E site codon variants which yield the highest, intermediate and lowest UGA readthrough, respectively. A similar difference in the E site codon dependent UGA readthrough, but with systematically higher readthrough levels, is also observed in the *E.coli* strain MG1655 which represents a typical

Table I. Influence of different E site codons on UGA decoding

Plasmid	E site		Transmission		
	Codon	tRNA	*trpT*$^+$(Su$^-$)	*trpT*(Su9)	Su9/Su$^-$
pSM1	UUU	Phe	0.56	3.1	5.5
pSM2	UUC		0.52	2.2	4.3
pSM6	UCU	Ser_5	0.23	0.99	4.3
pSM9	UCC		0.22	1.3	5.4
pSM30	AGU	Ser_3	0.14	0.90	6.4
pSM31	AGC		0.26	1.1	4.4
pSM18	AGA	Arg_4	0.04	0.28	7.0
pSM17	AGG	$Arg_{4,5}$	0.04	0.26	6.5
pSM26	CGU	Arg_2	0.03	0.20	6.6
pSM27	CGC		0.03	0.18	6.0
pSM24	CGA		0.05	0.20	4.0
pSM25	CGG	Arg_3	0.04	0.32	7.3
pSM13	UAU	$Tyr_{1,2}$	0.14	1.1	8.1
pSM12	UAC		0.11	0.96	8.7
pSM14	UGU	Cys	0.18	1.0	5.8
pSM15	UGC		0.15	1.1	7.1
pSM40	UGG	Trp	0.24	1.2	5.0
pSM7	CUU	Leu_2	0.21	1.0	4.8
pSM10	CUC		0.17	1.1	6.5
pSM8	CCU	Pro_2	0.52	1.8	3.5
pSM3	CCC		0.54	1.3	2.5
pSM28	CAU	His	0.09	0.53	5.8
pSM29	CAC		0.08	0.49	6.1
pSM41	CAA	Gln_1	0.14	0.72	5.1
pSM42	CAG	$Gln_{1,2}$	0.15	1.0	6.7
pSM34	AUA	Ile_2	0.28	1.0	3.6
pSM35	AUG	Met_m	0.26	1.4	5.2
pSM37	ACA	Thr_4	0.08	0.36	4.5
pSM38	AAU	Asn	0.04	0.60	15
pSM45	AAC		0.09	0.49	5.4
pSM5	AAA	Lys	0.28	1.5	5.3
pSM46	AAG		0.20	0.81	4.0
pSM36	GUA	Val_{1B}	0.28	1.5	5.3
pSM39	GCA	Ala_1	0.50	1.4	2.8
pSM21	GAU	Asp_1	0.80	3.8	4.8
pSM11	GAC		0.91	4.6	5.1
pSM20	GAA	Glu	0.68	3.8	5.6
pSM19	GAG		0.53	2.2	4.3
pSM23	GGU	Gly_3	0.27	2.1	7.6
pSM16	GGC		0.36	2.4	6.6
pSM22	GGA	Gly_2	0.37	2.0	5.3
pSM4	GGG	$Gly_{1,2}$	0.31	1.5	5.0

The protein products of the 3A′ gene were analysed as described in Materials and methods and Figure 2. The protein molar ratio 3A′/2A′ shows the proportion of readthrough events relative to termination events at the stop codon, and is referred to as the transmission value. The assay results were obtained in two *E.coli* strains, which differ with respect to the *trpT* allele, as indicated. The ratios of transmission values obtained in the *trpT*(Su9) and the *trpT*$^+$(Su$^-$) strains are also shown (Su9/Su$^-$). Each transmission value is the average of at least four independent experiments. The SEM (standard error of the mean) ranges from 1 to 20% for the different constructs.

wildtype strain (Bachmann, 1987), with another genetic background (data not shown).

The results from an experiment to control for proteolytic breakdown of the 3A′ protein are also shown in Figure 2. The last two lanes in the figure contain protein expressed by a 3A′ gene with a CGC (Arg) codon as the second upstream codon and a UGG (Trp) codon in place of UGA at the ribosomal A site. Decoding of UGG as well as readthrough of UGA leads to the insertion of tryptophan (see below). As can be seen, both the *trpT*$^+$(Su$^-$) and the *trpT*(Su9) strains express only the 3A′ protein when UGG is the A site codon. Therefore, the 2A′ protein found using UGA instead of UGG is not the result of degradation of the 3A′ readthrough protein, since both codons give a 3A′ protein with an identical amino acid sequence.

The effects of the second upstream codon described here could be explained in terms of a ribosome hopping event, which might allow the ribosome to bypass the stop codon (Atkins *et al.*, 1991). However, measurements of [^{3}H]tryptophan incorporation into the 3A′ readthrough protein from the test sequence -GAC CCA UGA-, which promotes the highest UGA readthrough, show that tryptophan is incorporated in stoichiometric amounts in the 3A′ protein, despite the lack of any tryptophan codon in the gene [1.1 and 0.87 mol of tryptophan/mol of 3A′ protein from the *trpT*(Su9) and *trpT*$^+$(Su$^-$) strain, respectively].

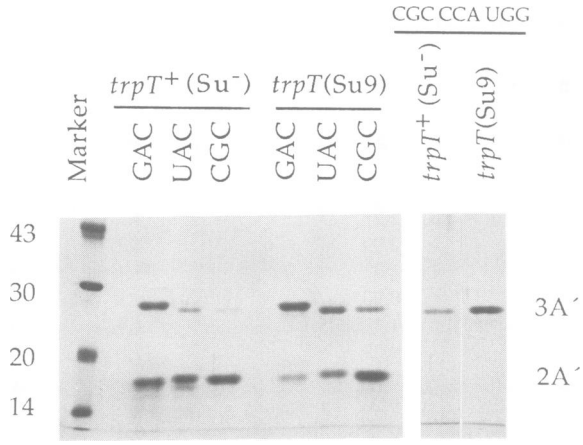

Fig. 2. SDS−PAGE analysis of protein products expressed from different 3A′ genes. The first lane contains molecular size markers with the molecular weights (kDa) indicated. Lanes 2−7 show protein products of the 3A′ gene from three representative E site codon variants [GAC (Asp), UAC (Tyr) and CGC (Arg)] with UGA two codons downstream. The different 3A′ alleles were assayed in both *trpT*+(Su−) and *trpT*(Su9) *E.coli* strains as indicated. Each E site variant yields both 3A′ protein (three-domain readthrough product) and 2A′ protein (two-domain termination product) as shown. The last two lanes show protein expressed from a 3A′ gene with UGG (Trp) as A site codon in place of UGA, with a CGC codon at the E site as indicated. Only three-domain protein (3A′) is produced by this 3A′ allele in both the *trpT*+(Su−) and *trpT*(Su9) strains. The protein products of the 3A′ gene were purified from crude cell extracts with IgG affinity chromatography, as described in Materials and methods.

For the E site codon CGC (Arg), which gives the lowest effect on the A site decoding, the corresponding value is 0.93 mol of tryptophan per mol of 3A′ protein in the *trpT*(Su9) strain. Similar results on $tRNA^{Trp}$ decoding of UGA codons have been described previously (Björnsson and Isaksson, 1993). Clearly, a major portion of the 3A′ readthrough protein originates from decoding of UGA by the wildtype or suppressor form of $tRNA^{Trp}$, and not by ribosomal hopping.

From the data in Table I, it can be seen that for the *trpT*+(Su−) strain, the E site codons UUU (Phe) and CCC (Pro) are almost identical in promoting A site readthrough. The values for AAA (Lys) and GGG (Gly) are quite similar to each other, but are significantly lower than values for the former codons. In the case of the *trpT*(Su9) strain, the codons CCC, AAA and GGG give similar results whereas UUU stimulates readthrough more significantly. From these results, as well as from some other data in Table I, it can be concluded that the number of G-C or A-U base pairs of the codon−anticodon interaction between the codon corresponding to the E site and its cognate tRNA does not influence decoding of UGA at the A site in any systematic manner.

Several of the E site codons tested are synonymous, since they are decoded by the same isoacceptor tRNA. For instance, the phenylalanine codons UUC and UUU are translated by the same $tRNA^{Phe}$ species (Komine *et al.*, 1990). The decoding of these codons involves a G-C interaction in the first case and a G-U wobble base pairing in the second case. Several other analogous examples were also analysed. The data in Table I show that the synonymous codons at the E site have an equal influence on UGA readthrough in the *trpT*+(Su−) strain. The effects of the synonymous E site codons are similar also in the case of the *trpT*(Su9) strain, with the possible exception of the GAA/GAG (Glu) pair and a couple of others. Thus, in cases when one particular tRNA has a cognate anticodon to a synonymous codon pair at the E site, there is no significant difference in the effect on A site decoding. It does not matter if the E site codon−anticodon recognition is via Watson−Crick base pairing, or whether it comprises a wobble interaction.

Even though there is no suggestive correlation between readthrough and the interaction between the upstream codon and its cognate tRNA, the data in Table I reveal an interesting pattern. Transmission values for codons within a family for a given E site amino acid are quite similar for the *trpT*+(Su−) strain. This is true even in cases when different tRNA isoacceptors are used for a particular amino acid (serine, arginine and glycine). The correlation also holds for the *trpT*(Su9) strain with its suppressor form of the same tRNA species with the possible exception of GAA/G (Glu) and AAA/G (Lys). Thus, readthrough of UGA at the A site is correlated to the nature of the amino acid corresponding to the E site codon more than to any property of its cognate tRNA or codon−anticodon interaction. A particularly interesting example is provided by the arginine codons, which yield the lowest transmission values of all E site codons. The six arginine codon transmission values are similar, despite the fact that AGA/G and CGG codons are minor and their decoding tRNAs are rare relative to the CGA/C/U codons and their decoding tRNA (Ikemura, 1981; Emilsson *et al.*, 1993).

The amino acids arginine, asparagine, histidine and threonine are associated with uniquely low transmission values. The highest readthrough is promoted by the most acidic amino acids—aspartic and glutamic acid—but phenylalanine, proline and glycine also give a rather high readthrough. If the readthrough value for each amino acid is plotted as a function of its pI (isoelectric point), a suggestive correlation between the acidic/basic nature of the amino acids and their effects on readthrough is found (Figure 3). This is true of both $tRNA^{Trp}$ and its UGA suppressor form encoded by *trpT*(Su9). Acidic amino acids seem to be associated with high readthrough and basic amino acids with low readthrough, whereas for the neutral amino acids the transmission values vary considerably.

Table I also shows that the effect associated with the second upstream codon is not determined by the nature of the first two bases in the codons. For instance, AAY (Asn) and AAR (Lys) have significantly different effects on readthrough, as do AGY (Ser) and AGR (Arg). Other similar examples, although less dramatic, can also be found in the table.

It can be also seen in Table I that proline gives a readthrough value which is 2.5 times higher in the *trpT*(Su9) strain than in the *trpT*+(Su−) strain. For others, for instance tyrosine, the corresponding difference is significantly higher. Thus, the amino acid corresponding to the second upstream codon affects readthrough of UGA in a manner which is dependent on the nature of the tRNA reading this stop codon.

Influence of the second upstream codon on readthrough of UAG

Some of the E site codon variants were altered to contain UAG as the A site codon, instead of UGA. The GAC (Asp), UAC (Tyr) and CGC (Arg) codons were chosen as E site codons because they are respectively associated with the

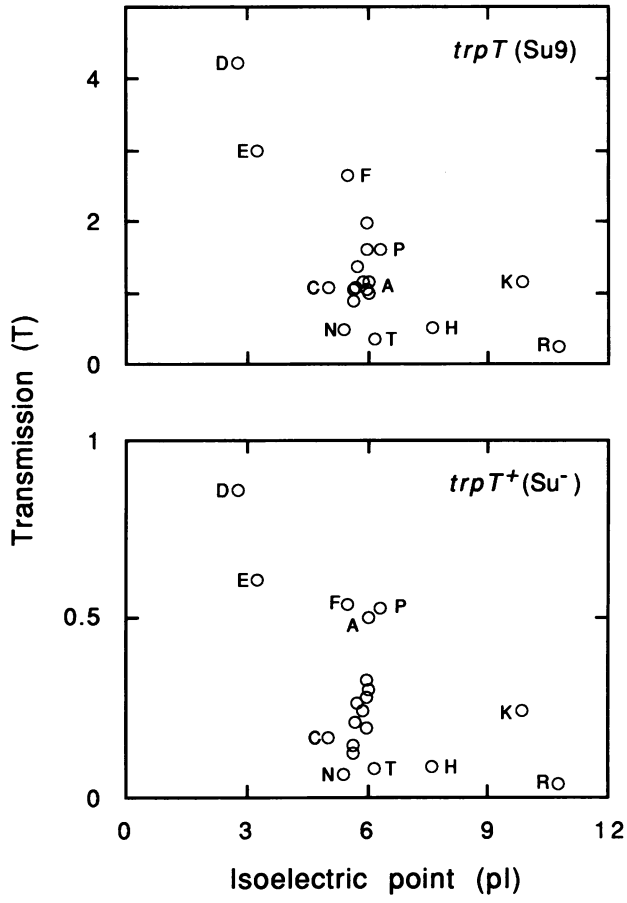

Fig. 3. UGA transmission values (T) obtained for different E site codon variants (Table I) plotted as a function of the isoelectric point (pI) of the corresponding amino acids (Hardy, 1985). One plot (**Top**) shows the results of UGA decoding by the *trpT*(Su9) tRNA and the other (**Bottom**) those for the *trpT*$^+$(Su$^-$) tRNA, as indicated. The amino acids indicated on the plots are D, E, F, C, N, T, P, A, H, K and R. Nine amino acids (Q, G, I, L, M, S, W, Y and V) are not indicated on the plots since they tend to plot in the same area. Values for all of the 20 amino acids are plotted. Average transmission values are used for amino acids representing more than one E site codon.

highest, an intermediate and the lowest UGA readthrough values. The transmission values, measured in strains with two different amber suppressor tRNAs, are presented in Figure 4 and Table II. As can be seen, changing the second upstream codon also influences readthrough of UAG, but the difference between the highest and lowest readthrough values is much smaller than that between the corresponding UGA codon contexts. The ranking of the effects by the upstream codons GAC, UAC and CGC seems to be similar for UAG, as for UGA at the A site. In addition, the range of the amino acid effects is smaller for the strong *tyrT*(Su3) amber suppressor (1.6-fold) than for the weak *lysT*(Su5) ochre suppressor (2.5-fold).

Mutant aminoacyl tRNA ligases and readthrough effects

The influence of the *tyrS565* mutation on decoding of UAG and UGA codons was analysed. This mutation lowers the charging level of tyrosine tRNAs significantly, giving slow growth at 37°C and no growth at 42°C (L.Isaksson, unpublished results). The results are shown in Table III. The readthrough of UAG by the *tyrT*(Su3) suppressor is decreased 7-fold for the GAC (Asp) codon, 13-fold for the

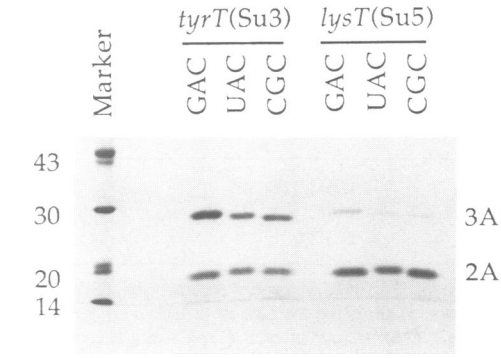

Fig. 4. SDS−PAGE analysis of protein products expressed from three representative E site codon variants (GAC, UAC and CGC) with a UAG codon two codons downstream. The 3A′ genes were assayed in both *tyrT*(Su3) and *lysT*(Su5) *E.coli* strains, as indicated. Processing of samples and identities of proteins are as described in the legend to Figure 2.

UAC (Tyr) codon and 17-fold for the CGC (Arg) codon by the *tyrS565* mutation. Thus, as expected, the *tyrS565* mutation lowers readthrough of UAG in all cases. It should be noted that the influence of the tRNA ligase mutation (Table III) seems to be strongest on UAG, E site codon variants which have a low readthrough in the *tyrS*$^+$ strain (UAC and CGC in Table II). In the *tyrS565* strain, the difference between the transmission value obtained with the *tyrT*(Su3) suppressor for GAC and CGC E site codons is 4-fold (Table III), but it is only 1.6-fold in the *tyrS*$^+$ strain (Table II). Thus, the range of the E site effect is larger in the former strain with its reduced readthrough ability relative to the latter control strain. These results are in line with the data discussed above that UAG readthrough by the weak *lysT*(Su5) ochre suppressor gives a larger range of transmission values than the stronger *tyrT*(Su3) amber suppressor.

Table III also shows that there is almost no influence on UGA decoding by wildtype tRNA^{Trp} in a suppressor-free strain when the GAC (Asp) and the UAC (Tyr) E site codon variants are analysed in a *tyrS565* strain, compared with results from a *tyrS*$^+$ strain. As discussed above, the efficiency of the *tyrT*(Su3) suppressor is decreased by about one order of magnitude (Tables II and III) by the tyrosine charging deficiency associated with the *tyrS565* mutation. This charging defect should lead to a correspondingly increased intracellular pool of uncharged tRNA^{Tyr}. These data suggest that an increased pool of the uncharged tRNA^{Tyr} matching UAC (Tyr), two codons upstream of the A site, has no specific effect on decoding of UGA.

It is observed that the E site codon CGC (Arg), which is associated with an extremely low readthrough, shows a significant decrease in UGA readthrough (>10-fold) together with the *tyrS565* allele (Table III). Therefore, the CGC upstream codon, as well as GAC (Asp) and GGC (Gly), were investigated for their effects on UGA readthrough in a *glyS709* strain. This mutation codes for a temperature sensitive glycine tRNA ligase (L.Isaksson, unpublished results). The data in Table III show that the *glyS709* mutation halves UGA readthrough with GAC (Asp) and GGC (Gly) as E site codons, while the readthrough of the CGC (Arg) codon is >10-fold lower than measurements in the strain with the wildtype glycyl-tRNA ligase (Table I). The data therefore show that even if the *glyS709* allele lowers

Table II. Influence of different E site codons on UAG decoding

Plasmid	E site		Transmission		
	Codon	tRNA	*tyrT*(Su3)	*lysT*(Su5)	Su3/Su5
pAG53	GAC	Asp_1	0.80	0.10	8
pAG54	UAC	$Tyr_{1,2}$	0.52	0.06	9
pAG55	CGC	Arg_2	0.51	0.04	13

The protein products of the 3A′ gene were analysed as described in Materials and methods and Figure 2; transmission values are defined in Table I. Each transmission value is the average of at least four independent experiments. The SEM ranges from 5 to 17% for the different constructs. The tRNA suppressor gene present in each *E.coli* strain is indicated. The ratios of transmission values obtained in the *tyrT*(Su3) and the *lysT*(Su5) strains are also shown (Su3/Su5).

Table III. Influence of mutant aminoacyl tRNA ligases on decoding of UAG and UGA

Test codons E/P/A sites	tRNA ligase	Transmission			
		trpT$^+$(Su$^-$)	*tyrT*(Su3)	wildtype/mutant ligase	
				trpT$^+$(Su$^-$)	*tyrT*(Su3)
GAC CCA UAG	*tyrS565*	ND	0.12		7
UAC CCA UAG	*tyrS565*	ND	0.04		13
CGC CCA UAG	*tyrS565*	ND	0.03		17
GAC CCA UGA	*tyrS565*	0.66	ND	1.4	
UAC CCA UGA	*tyrS565*	0.10	ND	1.1	
CGC CCA UGA	*tyrS565*	<0.003	ND	>10	
GAC CCA UGA	*glyS709*	0.36	ND	2.5	
GGC CCA UGA	*glyS709*	0.19	ND	1.9	
CGC CCA UGA	*glyS709*	<0.003	ND	>10	

The protein products of the 3A′ gene were analysed as described in Materials and methods and Figure 2; transmission values, which are defined in Table I, are given for the indicated tRNA ligase mutant strains. The ratio wildtype/mutant refers to transmission ratios which were obtained by dividing transmission values from the tRNA ligase wildtype strains (from Tables I and II) by the corresponding values obtained in the tRNA ligase mutant strains. The measurements of the pSM27 (-CGC CCA UGA-) protein products of the 3A′ gene, expressed in the mutant tRNA ligase strains, are the result of two independent experiments in each case, with severe overloading of samples on the SDS gels so that a laser scanner reading of a 3A′ readthrough protein could be obtained. The SEM ranges from 5 to 15% for the different constructs. ND, not determined.

Table IV. Influence of the third upstream codon/amino acid on UGA decoding

Plasmids	-XXX GAC CCA UGA-		Transmission		
	Codon	Amino acid	*trpT*$^+$(Su$^-$)	*trpT*(Su9)	Su9/Su$^-$
pSM43	UAC	Tyr	0.50	2.5	5.0
pSM32	CGC	Arg	0.55	3.7	6.7
pSM11	GAC	Asp	0.91	4.6	5.1
pAB94[a]	GAA	Glu	0.79	3.6	4.5
pSM44	AAC	Asn	0.37	2.3	6.3

The protein products of the 3A′ gene were analysed as described in Materials and methods and Figure 2; transmission values are defined in Table I. The location of the variable codon is denoted by XXX. Each transmission value is the average of at least four independent experiments. The SEM ranges from 6 to 17% for the different constructs. Data for pSM11 are also given in Table I.
[a]The 3A′ gene in pAB94 lacks a flanking GAC codon upstream its E site codon, resulting in the following sequence: -TTC GAA GAC CCA TGA-.

readthrough of UGA, this effect is not specific for the construct with GGC (Gly) as the E site codon. The very low readthrough of a UGA context with CGC (Arg) as the E site codon seems to be non-specifically lowered by tRNA charging deficiency as exemplified by tyrosine and glycine.

Influence of the third codon upstream of UGA on readthrough

From the data discussed above it appears that the amino acid coded by the second upstream codon is important for readthrough of both UGA and UAG. In order to investigate the effect of an altered third upstream codon, constructs with different codons at this location were analysed for readthrough of UGA.

It can be seen from Table IV that varying the third codon upstream of the UGA codon has a significant, but not profound effect on readthrough of the stop codon with given E site (GAC) and P site (CCA) codons. The amino acids chosen for this analysis at the third upstream position are associated with the highest (aspartic and glutamic acids), an

intermediate (tyrosine) and with the lowest (arginine and asparagine) readthrough when tested as the second upstream codon (see Table I).

The 3A′ gene in pAB94 has one of the two GAC codons at the second and third upstream location in plasmid pSM11 deleted, so that a different 5′ sequence, relative to the other constructs used in this study, is covered by the ribosome when it is decoding UGA. The transmission values obtained for these two plasmids are only slightly different from each other (Table IV). This further indicates that a sequence 5′ to the E site codon is relatively unimportant for UGA readthrough.

Discussion

The amino acid corresponding to the second upstream codon influences UGA readthrough

We have measured the influence of 42 different sense codons, coding for all 20 amino acids, on translational readthrough of a UGA codon positioned two codons downstream in the mRNA. The location of these sense codons corresponds to the ribosomal E site when UGA is at the A site. The observed 30-fold difference in readthrough of this stop codon between the different constructs shows that the upstream E site codon directly or indirectly has an effect on decoding of the A site codon by *E.coli* ribosomes *in vivo* during exponential cellular growth (Table I).

Our UGA readthrough data show a very suggestive dependence on the nature of the amino acid which corresponds to the E site codon. For any such given amino acid, the effect on A site readthrough is very similar irrespective of the nature of its codon or its decoding tRNA. The highest readthrough is promoted by aspartic acid and glutamic acid, and the lowest readthrough is associated with arginine and asparagine. A plot of the pI values for the amino acids versus readthrough data suggests a correlation between the isoelectric point of the amino acid and the effect on readthrough (Figure 3). Even though the neutral amino acids give widely different effects it appears that high readthrough is promoted by acidic, and low readthrough is promoted by basic amino acids.

The second upstream codon or its cognate tRNA is not directly affecting UGA readthrough

We have tried to correlate the effects on UGA readthrough obtained by changing the second upstream codon to other parameters besides the amino acid. We cannot correlate the GC content of a codon–anticodon interaction between the upstream codon and its cognate tRNA with the observed effects on readthrough. However, it must be kept in mind that even though a G-C interaction between the strands in an RNA stem is stronger than other base interactions such as A-U or G-U, this is not necessarily true for codon–anticodon interactions (Potapov, 1982). Furthermore, it has been found that GC richness in a codon–anticodon interaction does not influence translation speed of homopolymers *in vitro* (Andersson *et al.*, 1984). Also, anticodon–anticodon binding experiments by Grosjean and collaborators did not indicate any stronger association between tRNAs of complementing anticodons which are rich in G-C interactions (Grosjean *et al.*, 1978). Taken together, these results suggest that the importance of the GC content for the strength of the A site codon–anticodon interaction is uncertain, which makes it even more difficult to speculate about its effect on an E site interaction.

Our data show no correlation between the genomic frequency of the E site codon and its context effect. Furthermore, sequence comparisons of the different E site tRNAs did not reveal any obvious correlation between their primary or secondary structure and the associated A site readthrough values (data not shown). Also, we did not find any correlation between the intracellular abundance of the E site tRNA for bacteria grown in the medium used here and the observed E site codon effects. An interesting example is provided by the arginine codon family. The arginine codons CGU, CGC and CGA are decoded by the abundant species $tRNA^{Arg2}$ (6%), whereas CGG is read by the less common $tRNA^{Arg3}$ (0.2%). The rare arginine codon AGA is decoded exclusively by $tRNA^{Arg4}$ (1.5%), whereas the other rare codon (AGG) is read by this tRNA plus $tRNA^{Arg5}$ (0.3%). All these codons give similar UGA readthrough values when placed two codons upstream of UGA. Also, the different species of isoacceptor tRNAs for the glycine codons occur at different intracellular concentrations (Emilsson *et al.*, 1993), but this does not seem to have any effect on UGA readthrough.

It is conceivable that the E site codons in translating *E.coli* ribosomes could bind reversibly to their cognate deacylated tRNA species *in vivo*, even though it has been found that E site tRNA is tightly bound to ribosomes during centrifugation sedimentation (Remme *et al.*, 1989). An empty E site could affect A site decoding, in line with the allosteric interaction model put forward by Nierhaus and collaborators (Gnirke *et al.*, 1989). The *tyrS565* mutation used here decreases the charging level of tyrosine tRNAs significantly, resulting in a decrease in efficiency of the *tyrT*(Su3) suppressor of about one order of magnitude. The intracellular level of uncharged $tRNA^{Tyr}$ should be correspondingly increased. However, no influence from such decreased charging was observed on UGA readthrough by a UAC (Tyr), E site codon construct in the *tyrS565* strain (Table III). A *glyS709* mutation which codes for a temperature sensitive glycine tRNA ligase was also tested. No specific influence on UGA decoding was observed when a GGC (Gly) codon was located at the E site position. These data indicate that the intracellular level of the uncharged tRNA corresponding to the E site codon has no specific effect on A site decoding of UGA.

Ribosomal pausing and the E site codon effect

Most of our experiments are based on the stop codon UGA as the A site codon. However, as a comparison we also measured the E site codon influence on UAG readthrough and found that it was much lower than that for UGA. This observation correlates to a fundamental difference in the speed of RF-1 and RF-2 mediated termination reactions at UAG codons and UGA codons, respectively, since the RF-2 reaction seems to be slower than the RF-1 reaction in some codon contexts (A.Björnsson and L.Isaksson, unpublished results). The slower recognition of UGA by RF-2 at the A site selection step would increase the probability of a tRNA decoding event in the case of UGA relative to UAG or UAA codons. We also observe an indication that the magnitude of the amino acid effect is larger in a strain with the *trpT*$^{+}$(Su^{-}) wildtype $tRNA^{Trp}$, than in a *trpT*(Su9) strain with its more efficient UGA reading suppressor tRNA.

Similarly, the effect associated with the second upstream codon is larger for the weak *lysT*(Su5) ochre suppressor than for the more efficient *tyrT*(Su3) amber suppressor. In addition, the E site effect on readthrough by the *tyrT*(Su3) suppressor is larger in a *tyrS565* strain, with its deficient charging of the suppressor tRNA giving low readthrough, than in a $tyrS^+$ strain, with its high readthrough. These three cases suggest a correlation of the readthrough effect associated with the upstream codon and the efficiency of the A site decoding tRNA. Such findings, together with the observed correlation of the E site effect with the efficiency of RF-2 versus RF-1 (UGA or UAG readthrough), indicate that the magnitude of the observed amino acid effect is correlated with ribosomal pausing. One implication of this would be that the much faster translation of sense codons instead of stop codons at the A site should be considerably less sensitive to such amino acid effects.

The nascent peptide as a codon context determinant

The amino acid corresponding to the E site codon could influence A site decoding by several means. The effect probably arises after the amino acid has been incorporated into the nascent peptide chain as the penultimate residue. If so, the nature of the neighbouring amino acids could also be of importance. For instance, if the effect arises in connection with passage of the nascent peptide through a ribosomal tunnel (Picking *et al.*, 1992), it is possible that the temporarily exposed sequence of the nascent peptide is of importance for the kinetics of such passage. This process could have an indirect effect on A site activity. However, in the few cases studied here where the third amino acid from the C-terminal end of the nascent peptide has been varied, the results suggest that the nature of this neighbouring amino acid does not dramatically affect the E site amino acid effect (Table IV). Furthermore, deletion of the GAC (Asp), E site codon (plasmid pAB94), which brings the third upstream codon, also GAC, into the E site position, should lead to a phase shift of the preceding amino acid sequence in the nascent peptide. However, a comparison of the results obtained with plasmids pSM11 and pAB94 (Table IV) does not reveal any major difference in the readthrough values. This indicates that there is no demand for a fixed distance to a given amino acid sequence in the nascent peptide chain upstream of the aspartic acid residue, in this case, to promote the effect on UGA readthrough.

We note that the upstream amino acids proline and tyrosine influence UGA readthrough differently when results from a $trpT^+$(Su$^-$) and a *trpT*(Su9) strain are compared. This suggests that the tRNA causing the readthrough is directly responding to these amino acids in the nascent peptide. As discussed above, we have reason to believe that the nature of the release factor is also important since the stronger amino acid effect observed in UGA decoding relative to UAG seems to correlate with a difference in ribosomal pausing at these stop codons. The penultimate amino acid in the nascent peptide might therefore affect ribosomal interactions at the A site with both the tRNA and the release factor. We are currently further investigating these effects by altering the P site codon together with the E site codon. The P site tRNA has been implicated in codon context effects which manifest themselves through tRNA–tRNA interactions with the A site decoding tRNA (Smith and Yarus, 1989; Kato *et al.*, 1990). The E site amino acid might influence the conformation of the P site tRNA so that it interacts differently with the A site decoding tRNA and/or the release factor.

Materials and methods

Materials

L-[5-^{3}H]tryptophan and [^{35}S]dATP were purchased from Amersham, restriction enzymes and T4 DNA ligase from Promega, and Sequenase from United States Biochemical Corporation. Deoxyoligonucleotides were made in a Clone Plus DNA Synthesizer (Millipore). A Midget electrophoresis system (Pharmacia) was used for gel electrophoresis.

Growth media and bacterial strains

LB and M9 media were made as previously described (Miller, 1972). M9 medium supplemented with glucose, thiamine and necessary amino acids at recommended concentrations was used for growth of bacteria at 37°C for readthrough determinations (Neidhardt *et al.*, 1974). Plasmids were selected using ampicillin at concentrations of 100 μg/ml in plates and 200 μg/ml in liquid medium. IPTG was used for induction of the P_{trc} promoter (Figure 1) at a concentration of 0.4 mM in M9-based medium. The *E.coli* strain MC1061 was used for primary cloning (Sambrook *et al.*, 1989). The *E.coli* strains XA100 (wildtype with respect to tRNA genes), CDJ64 [with a *trpT*(Su9) suppressor gene], XA103 [with a *tyrT*(Su3) suppressor] and XA105 [with a *lysT*(Su5) suppressor] were used for assay purposes (Coulondre and Miller, 1977). Translational readthrough of UGA stop codons in the XA100 strain results from decoding by wildtype $tRNA^{Trp}$ coded by the $trpT^+$(Su$^-$) gene (Hirsh and Gold, 1971; Weiner and Weber, 1971; Björnsson and Isaksson, 1993). The CDJ64 strain carries a mutation in this gene [*trpT*(Su9)] which gives an A24→G alteration in the $tRNA^{Trp}$ and increases its UGA decoding ability. The strain MG1655 was also used as a representative for wildtype *E.coli* (Bachmann, 1987). The *E.coli* strains UT152 and UT133 are XA100 derivatives, which contain a Tn*10* (T152) insertion linked to a *tyrS565* mutation and Tn*10* (T133) insertion linked to a *glyS709* mutation, respectively. The *tyrT*(Su3) strain UI413 is a P1 transduction derivative of XA103 containing the Tn*10* (T152) insertion together with the *tyrS565* mutation. The *tyrS565* and *glyS709* derivatives of XA100 and XA103 were assayed in the same manner as the parental strains, because they can grow at a reduced growth rate at 37°C, even though they do not grow at 42°C.

Plasmid constructs

Standard recombination techniques were applied (Sambrook *et al.*, 1989) and all synthetic sequences cloned were sequenced with a Sequenase kit. The sequence primer ABPO2 which is complementary to a unique sequence at the 3′-end of the S3A′ gene was used for sequencing (Björnsson and Isaksson, 1993). A 'magic miniprep' protocol (Promega) was used for plasmid preparations. A Geneclean Kit (Bio 101) was used to purify restriction fragments from 1% agarose gels.

UUU and UUC E site codon variants were first cloned into pAB7, a plasmid expressing an S3A′ gene (Björnsson and Isaksson, 1988, 1993). Since S3A′ plasmids sometimes cause growth disturbance (see Results) the UUU (pAB82) and UUC (pAB83) E site variants were, for the studies described here, subcloned from these plasmids into the 3A′ expression system (pAB93) using the *Eco*RI and *Cla*I sites.

All second (E site) and third upstream codon variants of the 3A′ gene, except for the UUU and UUC codon contexts, were made by direct cloning of deoxyoligonucleotides into pAB93 replacing its *Asu*II–*Spe*I restriction fragment. Plasmid pAB93 is a derivative of pAB26 and originates from subcloning of a *Eco*RI–*Cla*I restriction fragment from a S3A′ plasmid (pAB75) into pAB26. The last A′ domain of the S3A′ gene in pAB75 together with the upstream test codons (-AGC TAG TGT-) and a downstream T_{trp} element, as in pAB7, was introduced into pAB26 thereby replacing the last A′ domain of its 3A′ gene and deleting a 5′ portion of its *lacZ* gene. This cloning step disrupts the 3A′-*lacZ* translational fusion in pAB26 (Faxén *et al.*, 1991). pAB75 is identical to pAB7 except for the stop codon which is UGA in pAB7.

Analysis of protein products of the 3A′ gene

Fresh colonies from M9-based agar plates, supplemented with glucose, thiamine, amino acids and ampicillin (100 μg/ml) were used to inoculate 10 ml of liquid medium of an identical composition, except that it contained IPTG (0.4 mM) and a greater concentration of ampicillin (200 μg/ml). Bacteria in balanced growth (mid-log phase) were harvested at an A_{540} of 0.5, put on ice and pelleted by centrifugation. The resulting pellets were resuspended in 1 ml of 0.05 M Tris–HCl (pH 8.0) and sonicated for 30 s

using an A350G model (Ultrasonics Ltd). The supernatant obtained after centrifugation was applied to a column containing IgG-Sepharose 6 Fast Flow (Pharmacia) and the eluted 3A′ and 2A′ proteins were analysed by SDS–PAGE, CBB G-250 staining and laser densitometry as previously described (Björnsson and Isaksson, 1993).

The amount of tryptophan in the 3A′ readthrough protein was determined as previously described (Björnsson and Isaksson, 1993). The same medium was used as for other readthrough determinations reported here. Furthermore, a 3A′ gene (pAB95) was used, which has a unique UGG (Trp) codon in place of the UGA codon in the test codon region (-GAA GAC AGC UGG UGU ACU-) as a reference for 100% incorporation of L-[5-^{3}H]tryptophan.

Acknowledgements

We thank Steve Muir and Mitch Dushey for comments on the manuscript. This work was supported by grants from the Swedish Natural Science Foundation (NFR).

References

Abrahmsén,L., Moks,T., Nilsson,B. and Uhlén,M. (1986) *Nucleic Acids Res.*, **14**, 7487–7500.

Andersson,S.G.E., Buckingham,R.H. and Kurland,C.G. (1984) *EMBO J.*, **3**, 91–94.

Atkins,J.F., Weiss,R.B., Thompson,S. and Gesteland,R.F. (1991) *Annu. Rev. Genet.*, **25**, 201–228.

Bachmann,B. (1987) In Neidhardt,F.C. (ed.), *Escherichia coli and Salmonella typhimurium. Cellular and Molecular Biology*. American Society for Microbiology, Washington DC, pp. 1190–1219.

Baron,C., Heider,J. and Böck,A. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 4181–4185.

Björnsson,A. and Isaksson,L.A. (1988) *Nucleosid. Nucleotid.*, **7**, 565–569.

Björnsson,A. and Isaksson,L.A. (1993) *J. Mol. Biol.*, **232**, in press.

Brown,C.M., Stockwell,P.A., Trotman,C.N.A. and Tate,W.P. (1990) *Nucleic Acids Res.*, **18**, 2079–2086.

Buckingham,R.H., Sörensen,P., Pagel,F.T., Hijazi,K.A., Mims,B.H., Brechemier-Baey,D. and Murgola,E.J. (1990) *Biochim. Biophys. Acta*, **1050**, 259–260.

Coulondre,C. and Miller,J.H. (1977) *J. Mol. Biol.*, **117**, 525–576.

Craigen,W.T. and Caskey,C.T. (1986) *Nature*, **322**, 273–275.

Curran,J.F. and Yarus,M. (1988) *J. Mol. Biol.*, **203**, 75–83.

Donly,B.C., Edgar,C.D., Adamski,F.M. and Tate,W.P. (1990) *Nucleic Acids Res.*, **18**, 6517–6522.

Emilsson,V., Näslund,A.K. and Kurland,C.G. (1993) *J. Mol. Biol.*, **230**, 483–491.

Faxén,M., Plumbridge,J. and Isaksson,L.A. (1991) *Nucleic Acids Res.*, **19**, 5247–5251.

Gnirke,A., Geigenmüller,U., Rheinberger,H.-J. and Nierhaus,K.H. (1989) *J. Biol. Chem.*, **264**, 7291–7301.

Grajevskaja,R.J., Ivanov,Y.V. and Saminsky,E.M. (1982) *Eur. J. Biochem.*, **128**, 47–52.

Grosjean,H.J., Henau,S.D. and Crothers,D.M. (1978) *Proc. Natl Acad. Sci. USA*, **75**, 610–614.

Hardy,P.M. (1985) In Barrett,C.G. (ed.), *Chemistry and Biochemistry of the Amino Acids*. Chapman and Hall Ltd, London, pp. 7–24.

Heider,J., Baron,C. and Bock,A. (1992) *EMBO J.*, **11**, 3759–3766.

Hirsh,D. and Gold,L. (1971) *J. Mol. Biol.*, **58**, 459–468.

Ikemura,T. (1981) *J. Mol. Biol.*, **151**, 389–409.

Kato,M., Nishikawa,K., Uritani,M., Miyazaki,M. and Takemura,S. (1990) *J. Biochem.*, **107**, 242–247.

Komine,Y., Adachi,T., Inokuchi,H. and Ozeki,H. (1990) *J. Mol. Biol.*, **212**, 579–598.

Kopelowitz,J., Hampe,C., Goldman,R., Reches,M. and Engelberg-Kulka,H. (1992) *J. Mol. Biol.*, **225**, 261–269.

Krillov,S.V., Makarov,E.M. and Semenkov,Y.P. (1983) *FEBS Lett.*, **157**, 91–94.

Lill,R., Robertson,J.M. and Wintermeyer,W. (1984) *Biochemistry*, **23**, 6710–6717.

Martin,R., Weiner,M. and Gallant,J. (1988) *J. Bacteriol.*, **170**, 4714–4717.

Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Neidhardt,F.C., Bloch,P.L. and Smith,D.F. (1974) *J. Bacteriol.*, **119**, 736–747.

Nilsson,B., Moks,T., Jansson,B., Abrahmsén,L., Elmblad,A., Holmgren,E., Henrichson,C., Jones,T.A. and Uhlen,M. (1987) *Protein Engng*, **1**, 107.

Pedersen,W.T. and Curran,J.F. (1991) *J. Mol. Biol.*, **219**, 231–241.

Picking,W.D., Picking,W.L., Odom,O.W. and Hardesty,B. (1992) *Biochemistry*, **31**, 2368–2375.

Potapov,A.P. (1982) *FEBS Lett.*, **146**, 5–8.

Remme,J., Margus,T., Villems,R. and Nierhaus,K.H. (1989) *Eur. J. Biochem.*, **183**, 281–284.

Rheinberger,H., Sternbach,H. and Nierhaus,K.H. (1981) *Proc. Natl Acad. Sci. USA*, **78**, 5310–5314.

Rheinberger,H.-J., Geigenmüller,U., Gnirke,A., Hausner,T.-P., Remme,J., Saruyana,H. and Nierhaus,K.H. (1990) In Hill,W.E., Dalberg,A., Garrett,R.A., Moore,P.B., Schlessinger,D. and Warner,J.R. (eds), *The Ribosome: Structure, Function & Evolution*. American Society for Microbiology, Washington DC, pp. 318–330.

Ryden,S.M. and Isaksson,L.A. (1984) *Mol. Gen. Genet.*, **193**, 38–45.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Smith,D. and Yarus,M. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 4397–4401.

Stromo,G.D., Schneider,T.D. and Gold,L. (1986) *Nucleic Acids Res.*, **14**, 6661–6679.

Weiner,A.M. and Weber,K. (1971) *Nature New Biol.*, **234**, 206–209.

Weiss,R.B., Dunn,D.M., Atkins,J.F. and Gesteland,R.F. (1987) *Cold Spring Harbor Symp. Quant. Biol.*, **52**, 687–693.

Weiss,R.B., Dunn,D.M., Dahlberg,A.E., Atkins,J.F. and Gesteland,R.F. (1988) *EMBO J.*, **7**, 1503–1507.

Weiss,R.B., Huang,W.M. and Dunn,D.M. (1990) *Cell*, **62**, 117–126.

Zinoni,F., Heider,J. and Böck,A. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 4660–4664.

Received on September 17, 1993; revised on October 13, 1993