The EMBO Journal vol.15 no.7 pp.1696–1704, 1996

Structure of the C-terminal end of the nascent peptide influences translation termination

Asgeir Björnsson[1], Salim Mottagui-Tabar and Leif A.Isaksson[2]

Department of Microbiology, Stockholm University, S-106 91 Stockholm, Sweden and [1]Department of Chemistry, Aarhus University, DK-8000 Aarhus C, Denmark

[2]Corresponding author

The efficiency of translation termination at NNN NNN UGA A stop codon contexts has been determined in *Escherichia coli*. No general effects are found which can be attributed directly to the mRNA sequence itself. Instead, termination is influenced primarily by the amino acids at the C-terminal end of the nascent peptide, which are specified by the two codons at the 5′ side of UGA. For the penultimate amino acid (–2 location), charge and hydrophobicity are important. For the last amino acid (–1 location), α-helical, β-strand and reverse turn propensities are determining factors. The van der Waals volume of the last amino acid can affect the relative efficiency of stop codon readthrough by the wild-type and suppressor forms of $tRNA^{Trp}$ (CAA). The influence of the –1 and –2 amino acids is cooperative. Accumulation of an mRNA degradation intermediate indicates mRNA protection by pausing ribosomes at contexts which give inefficient UGA termination. Highly expressed *E.coli* genes with the UGA A termination signal encode C-terminal amino acids which favour efficient termination. This restriction is not found for poorly expressed genes.

Keywords: amino acid/codon context/nascent peptide/protein structure/ribosome pausing

Introduction

Messenger RNA translation is terminated in *Escherichia coli* at the UAG stop codon by release factor one (RF-1), at UGA by release factor 2 (RF-2) and at UAA by both termination factors (Scolnik *et al.*, 1968). Recently, a third factor, RF-3 (release factor 3), has been cloned (Grentzmann *et al.*, 1994; Mikuni *et al.*, 1994) which stimulates the *in vitro* activity of both RF-1 and RF-2 (Goldstein and Caskey, 1970). However, it has been suggested that this factor preferentially stimulates RF-2-mediated termination (Grentzmann *et al.*, 1995).

The efficiency of translation termination is sensitive to both the 5′ and 3′ sequence contexts of stop codons (Salser, 1969; Bossi and Roth, 1980; Bossi, 1983; Miller and Albertini, 1983; Smith and Yarus, 1989; Buckingham *et al.*, 1990; Pedersen and Curran, 1991; Kopelowitz *et al.*, 1992; Björnsson and Isaksson, 1993; Mottagui-Tabar *et al.*, 1994; Poole *et al.*, 1995). A strong bias in the relative frequencies of the 3′-flanking bases is found in highly expressed *E.coli* genes. For UAA the 3′-flanking base preference is U>>G>A>C, for UGA it is U>>A>G>C, and for the rare stop codon UAG it is U>C = G>A (Brown *et al.*, 1990). This ranking conforms well to the ranking of release factor-mediated termination efficiency at UAA and UGA codons (Poole *et al.*, 1995), although contradictory results are obtained for UAG codons (Pedersen and Curran, 1991). Release factor affinity for a stop codon is probably enhanced by specific interaction with a 3′ U base (Pedersen and Curran, 1991). Conversely, high suppression of stop codons followed by a 3′ A base is probably caused by preferential stacking between this base and the anticodon–codon complex of the ribosomal A-site decoding tRNA (Stormo *et al.*, 1986). Thus, the 3′ base has been shown to influence both release factor and suppressor tRNA decoding efficiencies.

The P-site codon and/or tRNA can, in some cases, influence the efficiency of suppressor tRNA decoding of a stop codon, indicating the importance of 5′ codon context (Miller and Albertini, 1983; Smith and Yarus, 1989; Buckingham *et al.*, 1990). We have shown recently that, in addition to such determinants, there is a significant contribution to termination efficiency by the penultimate amino acid at the C-terminal end of the nascent peptide (Mottagui-Tabar *et al.*, 1994). The acidic/basic property of this residue causes up to a 30-fold difference in UGA termination efficiency. Termination at UAG is less sensitive to the nature of the –2 amino acid.

Statistical analysis of *E.coli* genes shows that lysine codons are preferred as the –1 codon at the 5′ side of UAA (Brown *et al.*, 1990), whereas threonine and proline codons are underrepresented (Arkov *et al.*, 1993). One serine codon (UCC) and both phenylalanine codons are preferred as –1 codons relative to UGA in *E.coli* mRNAs (Arkov *et al.*, 1993). This difference in 5′-flanking codon preferences when UAA and UGA termination signals are compared could be connected with the higher frequency of UAA codons in highly expressed genes, hence reflecting a stricter requirement for efficient termination. The low frequency of UAG codons in *E.coli* genes makes it difficult to determine statistically any significant 5′ codon bias for this stop codon.

Escherichia coli mRNAs have many features which are important for translation efficiency. For the initiation step, several initiation codons with different efficiencies are used, besides AUG. Initiation is also dependent on an mRNA anchoring sequence at a proper distance upstream (Shine–Dalgarno sequence) (Shine and Dalgarno, 1974; Ringquist *et al.*, 1992) as well as downstream of the initiation codon (downstream box) (Sprengart *et al.*, 1990; Faxén *et al.*, 1991). For translation termination, UAA is the most frequently used codon in highly expressed *E.coli* genes. UAA is probably the most efficient termination signal, since both RF-1 and RF-2 are active at this stop

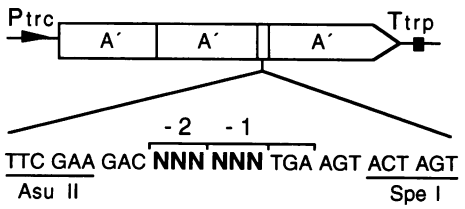

Fig. 1. The 3A′ reporter gene. The 3A′ gene is composed of three identical units coding for A′ domains. The A′ protein domain (previously referred to as Z or B′) is derived from the IgG binding B domain of protein A from *Staphylococcus aureus* (Nilsson *et al.*, 1987; Björnsson and Isaksson, 1993). Test codon sequences, designated NNN NNN TGA (NNN being any codon, and TGA the stop codon), were cloned into a linker region between the second and third A′ coding regions of the 3A′ gene, using the indicated *Asu*II and *Spe*I restriction sites. The influence on translation of test codons at positions –2 and –1 relative to UGA was evaluated by measurement of the molar amounts of 3A′ domain protein (translation readthrough) relative to 2A′ domain protein (translation termination) associated with each 3A′ allele (see Figure 2 and Table I).

codon (Scolnik *et al.*, 1968). Highly expressed genes also have a bias in sense codon utilization, with a preference for codons which are decoded by major tRNA species (Ikemura, 1992). Besides a bias in codon usage, a G–non-G–N sequence pattern has been found in gene coding sequences, which is independent of the gene expression level. This sequence pattern has been implicated in translational reading frame maintenance (Trifonov, 1987), but a recent experimental test failed to confirm that hypothesis (Curran and Gross, 1994). An alternative view is that the G–non-G–N pattern is a manifestation of codon usage corresponding to the most common amino acids in proteins (Yarus and Folley, 1985). Interestingly, this pattern is absent for the last five codons upstream of stop codons in *E.coli* (Brown *et al.*, 1990). This indicates that the last five amino acids in proteins are generally not important for protein function, or that some other selection constraint, such as translation termination, is a dominating factor.

We have tested here whether the last amino acid of the nascent peptide, like the penultimate amino acid (Mottagui-Tabar *et al.*, 1994), influences translation termination. Our data suggest a major influence on translation termination from the C-terminal amino acid. Furthermore, the effects of the –1 and –2 amino acid residues are cooperative. Our data also show that combinations of the last two amino acids, which are associated with inefficient translation termination, are avoided in highly expressed *E.coli* genes using the UGA A tetranucleotide termination signal, indicating a counter-selection during evolution.

Results

The in vivo translation termination assay

The 3A′ translation assay system was used for *in vivo* analysis of translation termination (Figure 1) (Mottagui-Tabar *et al.*, 1994). This system allows the quantification of both translation termination and readthrough protein products which originate from an allele with a particular stop codon context. Any change in mRNA level will affect the amount of both proteins equally. Variations in mRNA levels caused by transcription polarity and/or messenger stability therefore have a minimal influence on the quantitative results of the assay (Björnsson, 1994).

Translation effects associated with the last two codons upstream of UGA A

Codons representing the 20 common amino acids were placed at the ribosomal P-site (peptidyl-tRNA binding site; –1 codon) relative to UGA at the A-site (aminoacyl-tRNA acceptor site) and analysed for their effects on termination/readthrough at the stop codon. Table I shows that the effect of these codons is large when the –2 codon is GAC (Figure 2, lanes 2–6). This is true when readthrough of the UGA codon is caused both by the wild-type (Su^-) and the opal suppressor form (Su9) of $tRNA^{Trp}$(CCA). If GAC is replaced by UAC, readthrough is decreased or similar to the GAC-based context values, except being somewhat higher when the –1 codon belongs to the glycine codon family. When the –2 codon is CGC, the level of UGA readthrough is very low and mostly unaffected by –1 codons (Figure 2, lanes 7–11). The effects associated with the –2 and –1 codons are thus cooperative.

Table I shows a 2-fold effect on UGA readthrough by the codons within the proline family if they are at the –1 codon location. When the –2 codon is GAC (aspartic acid), a 3- to 4-fold difference is obtained within the glycine codon family at location –1, suggesting an influence of the P-site tRNA isoacceptor structure or codon–anticodon interaction. In contrast, readthrough is similar for the two threonine codons tested. This is also true for both glutamic acid codons. Interestingly, the serine codons AGU and UCC give similar readthrough values, even though the bases at all codon positions are different. All arginine codons, except CGA (Figure 2, lanes 6 and 11), give a similar low readthrough. In the case of the –1 codon CGA, it is probably the tRNA itself at the P-site which causes the extremely low UGA readthrough (see Discussion). Thus, even though different codon families give different effects on readthrough, the effects are quite similar for the codons within each family, with the exception of the few indicated codons. The data suggest that the major effect of varying the –1 codon on UGA termination/readthrough is due to the amino acid at position –1 in the nascent peptide.

The last two amino acids in the nascent peptide are UGA decoding determinants

We have found previously that the charge of the –2 amino acid in the nascent peptide affects termination at UGA (Mottagui-Tabar *et al.*, 1994). Results from further analysis (Table II, line 8) show that neutral –2 amino acids (acidic and basic amino acids excluded) give efficient termination at UGA if they have a hydrophilic side chain (e.g. glutamine), and inefficient termination if the side chain is hydrophobic (e.g. phenylalanine). For the –1 amino acids in the nascent peptide, no correlation is found between their acidic/basic or hydrophobic/hydrophilic character and decoding of UGA (Table II).

We next searched for correlations between UGA decoding and the involvement of the –2 and –1 amino acids in different natural secondary structures of proteins (Table II, lines 1–6). No such correlations are found for the –2 amino acids (NNN CCA UGA A contexts). However, for the –1 amino acids, a correlation is found between efficient UGA termination and high amino acid participation in α-helices (Figure 3A) and β-strands (Figure 3B) if aspartic

Table I. Influence of 5′ adjacent codons on UGA readthrough

Codon[a]	tRNA[a]	GAC NNN UGA A[b]		UAC NNN UGA		CGC NNN UGA A	
		trpT$^+$(Su$^-$)	*trpT*(Su9)	*trpT*$^+$(Su$^-$)	*trpT*(Su9)	*trpT*$^+$(Su$^-$)	*trpT*(Su9)
UUC	Phe	0.06	0.36	0.06	0.53		
UAC	$Tyr_{1,2}$	0.12	0.46				
UGC	Cys	0.05	0.43				
UGG	Trp	0.47	1.6	0.30	1.70		
CUC	Leu_2	0.05	0.18	0.03	0.15		
CCU	Pro_2	0.76	3.1	0.12	0.82		
CCC		0.46	2.6	0.13	0.61		
CCA	Pro_3	0.91[c]	4.6[c]	0.17	0.96	0.03	0.17
CCG	$Pro_{1,3}$	0.46	2.7	0.10	0.45		
CAC	His	0.10	0.79				
CAG	$Gln_{1,2}$	0.08	0.41				
CGU	Arg_2	0.06	0.40				
CGC		0.05	0.23	0.04	0.19	0.03	0.13
CGA		<0.01	0.06	<0.01	0.06	<0.01	<0.01
CGG	Arg_3	0.04	0.35				
AGA	Arg_4	0.08	0.27				
AGG	$Arg_{4,5}$	0.07	0.27				
AGU	Ser_3	0.06	0.59	0.04	0.35	0.02	0.11
AGC		0.10	0.79	0.07	0.50	0.02	0.10
UCC	Ser_5	0.11	0.57				
AUC	Ile_{1B}	0.06	0.30				
AUG	Met_m	0.08	0.27	0.04	0.26		
ACU	$Thr_{1,3}$	0.26	1.4				
ACC		0.23	1.3				
AAC	Asn	0.16	0.54				
AAG	Lys	0.09	0.36	0.05	0.22		
GUG	Val_1	0.04	0.26	0.05	0.34	0.02	0.15
GCC	Ala_2	0.10	0.46	0.11	0.38		
GAC	Asp_1	0.15	1.3	0.14	0.66		
GAA	Glu	0.13	0.66				
GAG		0.11	0.66	0.10	0.44		
GGU	Gly_3	0.07	0.48	0.19	0.92	0.01	0.15
GGC		0.09	0.65	0.20	1.4	0.03	0.13
GGA	Gly_2	0.21	2.0	0.36	2.4	0.03	0.17
GGG	$Gly_{1,2}$	0.16	1.5	0.23	1.7	0.05	0.09

UGA readthrough *in vivo* was measured using the 3A′ reporter system in *E.coli* strains XAc and CDJ64 (Mottagui-Tabar *et al.*, 1994). XAc has wild-type *trpT*$^+$(Su$^-$) tRNA-dependent readthrough of UGA. In CDJ64, readthrough of UGA is caused by the Hirsh suppressor form of the same tRNA [*trpT*(Su9)] (Hirsh, 1971; Coulondre and Miller, 1977). Transmission values (T = [3A′]/[2A′]) are molar ratios of the amounts of 3A′ domain protein (3A′) (translation readthrough product) relative to 2A′ domain protein (2A′) (translation termination product) associated with each stop codon allele of the 3A′ gene. Each transmission value is the average of at least four independent measurements and the standard error of the mean is 2–20%.

[a]Codons and tRNAs which correspond to location –1 relative to the UGA codon in the respective 3A′ mRNAs.

[b]Different –1 codons (NNN) tested in UGA A contexts with the indicated codon at location –2 (GAC, UAC, CGC).

[c]These values have been reported previously (Mottagui-Tabar *et al.*, 1994).

acid is at position –2 (GAC NNN UGA A). Low UGA termination efficiency is correlated with high amino acid participation in reverse turns (Figure 3C). If the –1 amino acid participation in α-helices is subdivided with respect to its location in α-helix structures (Williams *et al.*, 1987), a significant correlation is observed between efficient UGA termination and the –1 amino acid involvement in C-terminal and interior, but not N-terminal, parts of natural α-helices (Figure 3D–F and Table II). Most of these correlations are also observed when proline, glycine and leucine codons are excluded from our data set (not shown), especially the occurrence of amino acids in protein turn motifs (Williams *et al.*, 1987). Efficient termination at UGA is also correlated with high α-helical and β-strand propensities of the –1 amino acids in two model proteins (Table II, lines 10 and 11). Results using a strain with wild-type tRNA^{Trp}(Su$^-$) are similar to those presented in Figure 3 for the tRNA^{Trp}(Su9) strain (Table II), even though the readthrough levels in the former case are low and therefore more sensitive to experimental errors.

Transfer RNA selection and van der Waals volume of the last amino acid

The influence of the last amino acid on the relative readthrough efficiencies by two different forms of tRNA^{Trp} was analysed. For this purpose, ratios of the transmission values [T(Su9)/T(Su$^-$)] found for the different codon contexts in the *trpT*(Su9) and *trpT*$^+$(Su$^-$) *E.coli* strains were plotted against various physical properties of the –1 and –2 amino acids. For the –1 amino acids, a significant correlation is found between their van der Waals volumes and T(Su9)/T(Su$^-$) ratios in GAC NNN UGA A (Figure 4A), but not in UAC NNN UGA A (Figure 4B), CGC NNN UGA A or NNN CCA UGA A contexts (Table II). Thus, the efficiencies of two tRNA^{Trp} species are affected differently by the van der Waals volume of the last amino acid when the penultimate amino acid is aspartic acid. Data in Table II (line 9) indicate that it is the suppressor tRNA, and not the wild-type tRNA, which is sensitive to the amino acid volume. Small amino acids like glycine give more efficient suppressor tRNA

decoding of UGA than large amino acids like tryptophan (Figure 4A). A similar analysis of the other physical characters of the amino acids did not reveal any obvious influence on $T(Su9)/T(Su^-)$ ratios (not shown).

It has been proposed, on stereochemical grounds, that the C-terminal end of the nascent peptide should form part of an α-helix at the peptidyltransferase centre in order to facilitate the transpeptidation reaction for all the 400 possible A- and P-site amino acid combinations (Lim and Spirin, 1986). Such stringency on the conformation of the nascent peptide might make UGA readthrough sensitive to physical properties of the amino acid carried by the A-site tRNA. For this reason, decoding of UGA by the *glyT*(Su^+UGA/G) suppressor tRNA (Murgola, 1985) was also measured for several representative –1 amino acids. The –1 amino acid effects on UGA readthrough of this suppressor tRNA are similar to those obtained for the *trpT*(Su9) suppressor (A.Björnsson, S.Mottagui-Tabar and L.A.Isaksson, unpublished results), despite the fact that glycine and tryptophan have quite different physical properties.

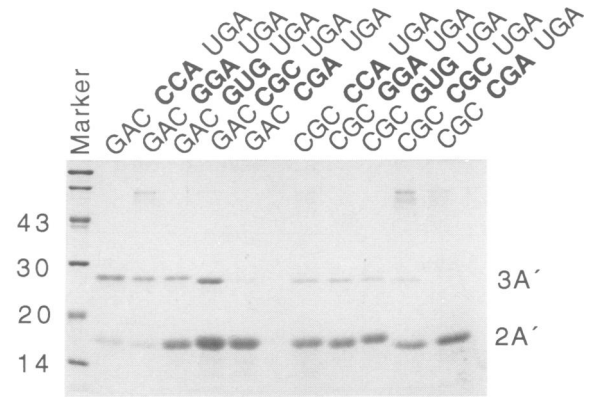

Fig. 2. SDS–PAGE analysis of protein products expressed from different NNN NNN UGA A codon context variants of the 3A′ gene. The –2 codon is either GAC (Asp) or CGC (Arg) and the –1 codon is CCA (Pro), GGA (Gly), GUG (Val), CGC (Arg) or CGA (Arg) when the ribosomal A-site codon is UGA. The 3A′ genes were assayed in a *trpT*(Su9) *E.coli* strain which gives measurable levels of UGA suppression for all of the tested UGA A-based codon contexts. Translation termination at the UGA codon gives 2A′ domain protein and translation readthrough at the stop codon gives 3A′ domain protein.

Inefficient translation termination at UGA correlates with a translation pause

The possibility that the amino acids at the C-terminal end of the nascent peptide cause ribosomal pausing at the UGA codon was considered. Such pausing could interfere with mRNA degradation. For this reason, cellular mRNA was investigated by Northern blot analysis. Two major RNA species are observed (Figure 5). The larger species is the 3A′ mRNA which is seen in all lanes. A 261 nt RNA fragment, denoted rpRNA (ribosomal pause RNA), is expressed in two cases (GAC CCA UGA and UAC CCA UGA context sequences, lanes 1 and 2). The 5′ end of the rpRNA is located 13 bases upstream of the UGA

Table II. Correlation analysis of amino acid influences on UGA readthrough[a]

Amino acid occurrence or property[b]	NNN CCA UGA A (n = 42)[c]		GAC NNN UGA A (n = 34)		UAC NNN UGA A (n = 20)		CGC NNN UGA A (n = 9)	
	Su^-	Su9	Su^-	Su9	Su^-	Su9	Su^-	Su9
α-Helix:	–0.13	–0.16	**–0.52**	**–0.66**	–0.51	**–0.61**	0.01	–0.04
N-terminal	0.10	0.05	–0.11	–0.19	–0.14	–0.32	0.12	0.24
Middle	–0.14	–0.18	**–0.56**	**–0.70**	–0.43	–0.49	0.03	0.04
C-terminal	–0.29	–0.30	**–0.63**	**–0.76**	**–0.67**	**–0.70**	–0.10	–0.20
β-Strand	–0.13	–0.11	**–0.51**	**–0.55**	–0.45	–0.32	–0.20	–0.07
Reverse turn	0.17	0.21	0.41	**0.55**	**0.65**	**0.65**	0.09	–0.08
Isoelectric point (pI)	**–0.72**	**–0.79**	–0.26	–0.36	–0.30	0.33	0.19	0.07
Hydrophobicity (ΔG°)	**–0.67**	**–0.63**	–0.09	–0.05	–0.09	–0.05	–0.23	–0.63
Van der Waals vol. (Å^3)	**–0.46**	**–0.47**	–0.24	**–0.45**	–0.48	–0.50	0.08	0.13
α-Helix (ΔΔG)	–0.35	–0.18	**–0.85**	**–0.89**	–0.18	–0.15	–0.18	–0.49
β-Strand (ΔΔG)	–0.47	–0.35	**–0.73**	**–0.81**	–0.37	–0.31	–0.32	–0.48
Van der Waals vol. (Å^3)[d]	0.09		**–0.52**		–0.16		–0.18	

[a]Correlation analysis of UGA readthrough and participation of amino acids in different secondary structures of natural proteins, or their respective physical properties. Individual codon values are used for the analysis (n = 34 for GAC NNN UGA A contexts). Values of the –1 codon CGA are excluded from the analysis, since the influence of this codon on UGA readthrough seems to be dominated by an unusual codon–anticodon interaction of the tRNA at the –1 location. Correlation coefficients (r) are judged as significant (shown in bold) if the derived probability (P) values as well as the probability values obtained from a non-parametric Kendall rank correlation test are in both cases ≤0.01.

[b]Correlation analysis of UGA readthrough and amino acid occurrence in natural α-helices (including N- and C-terminals and middle regions of α-helices), β-strands and reverse turns (Williams *et al.*, 1987; Creighton, 1993). Correlation analysis was also made on UGA readthrough and different physical properties of the respective amino acid: isoelectric point of the amino acid (Hardy, 1985); deduced hydrophobicity of the amino acid side chain [ΔG° (kcal/mol)] (Roseman, 1988) [charged amino acid residues D, E, H, K and R were excluded, since those residues are the major cause for the high correlation of –2 amino acid data to the isoelectric point (pI)]; van der Waals volume (Å^3) of the respective amino acid (Creighton, 1993); α-helix propensity [ΔΔG (kcal/mol)] of the 20 universal amino acids at position 44 in T4 lysozyme (Blaber *et al.*, 1994); β-strand propensity [ΔΔG (kcal/mol)] of the 20 universal amino acids at position 53 in the IgG binding domain of protein G (Minor and Kim, 1994).

[c]UGA A codon contexts are shown. The location of the varied codon (–1 or –2) and the number (n) of data points representing codons at the respective locations used for the analysis is indicated. Log(T) readthrough data from the *E.coli* strains XAc [*trpT*$^+$(Su^-)] and CDJ64 [*trpT*(Su9)] are used except when testing correlations to α-helical and β-strand propensities in model proteins (ΔΔG), where T values are used (lines 10 and 11). The –1 amino acid data used for the calculations are from Table I and the –2 amino acid data from a recent publication (Mottagui-Tabar *et al.*, 1994).

[d]Readthrough ratios [$T(Su9)/T(Su^-)$] as a function of the van der Waals volume of the amino acids (see also Figure 4).

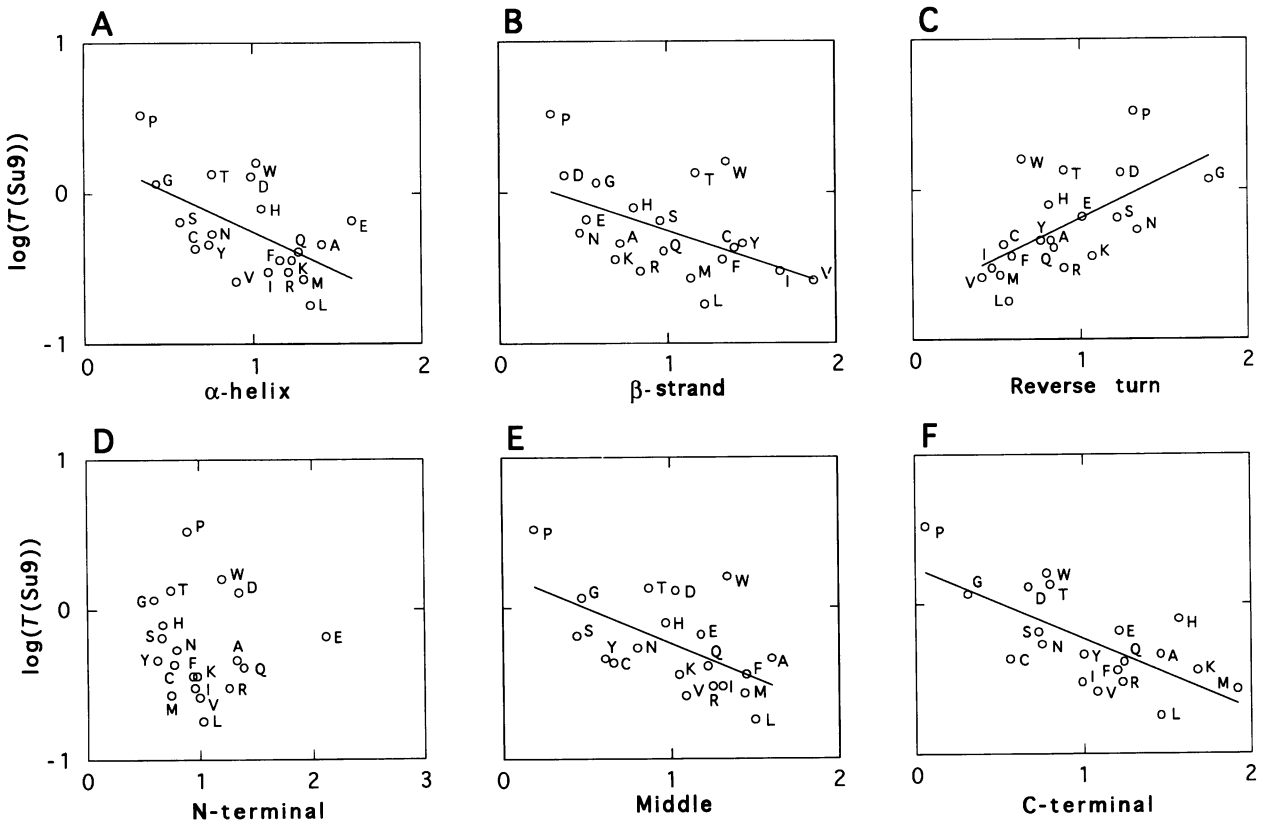

Fig. 3. UGA transmission values [log(T(Su9))] obtained for different –1 codons (GAC NNN UGA A contexts; Table I) as a function of the respective amino acid frequency, as found in different secondary structures of proteins: (**A**) α-helix, (**B**) β-strand, (**C**) reverse turn, and the (**D**) N-terminal, (**E**) middle and (**F**) C-terminal parts of α-helices (Williams *et al.*, 1987; Creighton, 1993) (Table II). Averages of transmission values are plotted for those amino acids which are coded by more than one –1 codon (Table I), except for the arginine codon CGA which is excluded (Table II). Significant correlations are indicated by a regression line. Amino acids are indicated by the standard single letter code.

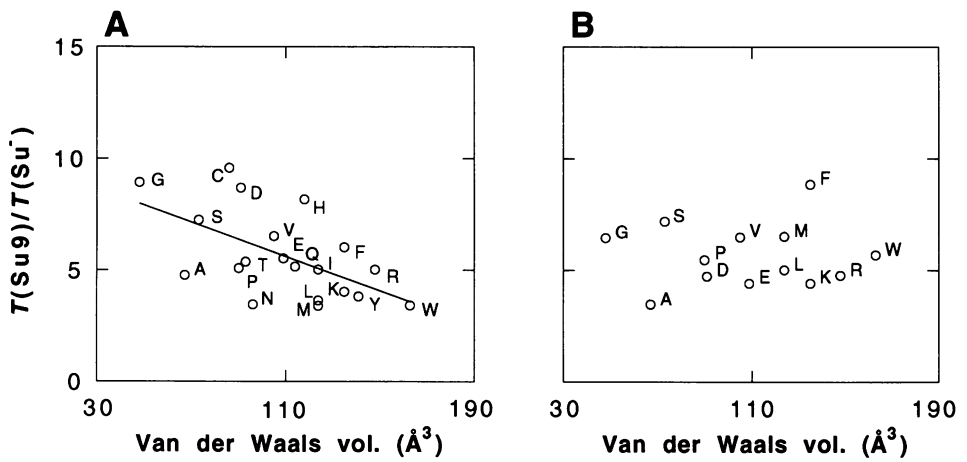

Fig. 4. Transmission ratios T(Su9)/T(Su$^-$) as a function of van der Waals volume (Å^3) of individual –1 amino acid residues (Creighton, 1993). The transmission ratios T(Su9)/T(Su$^-$) are calculated using transmission values presented in Table I. Averages of transmission values are plotted for those amino acids which are coded by more than one –1 codon (Table I), except for the arginine codon CGA which is excluded (see Table II). A significant correlation is indicated by a regression line. Amino acids are indicated by the standard single letter code. UGA readthrough ratios plotted for (**A**) GAC NNN UGA A codon contexts and (**B**) UAC NNN UGA A codon contexts.

codon. Its steady-state level is increased by inefficient termination and decreased in strains with the tRNA^{Trp}(Su9) suppressor. These data and results from pulse–chase experiments indicate that the rpRNA is a decay intermediate of 3A′ mRNA that is stabilized by a ribosomal pause at the stop codon (A.Björnsson and L.A.Isaksson, in preparation). The two codon contexts which give detectable amounts of rpRNA are the ones with lowest efficiency in translation termination. Furthermore, when these two contexts are compared, the context with lower termination efficiency (GAC CCA UGA A) also gives a higher level of rpRNA.

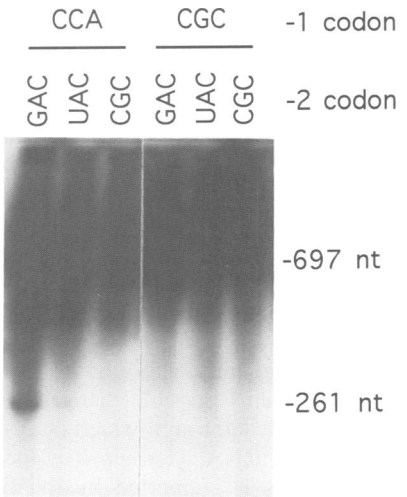

Fig. 5. Northern blot analysis of mRNA expressed from alleles of the 3A′ gene in the *trpT*$^{+}$(Su^{-}) *E.coli* strain XAc. The codons present at the –1 and –2 locations relative to UGA are indicated. The 3A′ mRNA (697 nt) is seen in all lanes, whereas a smaller RNA fragment (261 nt) is found at a detectable level only in connection with the 3A′ contexts GAC CCA UGA A and UAC CCA UGA A.

Translation termination at natural UGA codon contexts

UAA followed by U gives efficient termination and is the most frequent translation termination signal in highly expressed genes (Brown *et al.*, 1990). In contrast, UGA followed by A is a poor termination signal in *E.coli* and is rarely found in highly expressed genes. We wanted to examine if the effect on termination of C-terminal amino acids is correlated with the expression levels of genes using the UGA A termination signal. Therefore, readthrough was determined for different C-terminal dipeptides encoded by natural *E.coli* genes using this termination signal. The corresponding readthrough values were plotted against the CAI (codon adaptation index) value(s) assigned to these genes (Brown *et al.*, 1993) (Figure 6). The CAI value serves as an index for the level of gene expression (Sharp and Burgess, 1992). As can be seen, about half of the dipeptides encoded by genes with CAI values <0.4 are associated with inefficient termination at UGA A. All the 10 genes with higher CAI values (expressed at a higher level) encode C-terminal dipeptides which are associated with efficient termination at UGA A. This correlation is found irrespective of whether readthrough is caused by the wild-type (Su^{-}) or UGA suppressing form of tRNA^{Trp} (Su9) (Figure 6A and B). These data suggest that highly expressed genes ending with the weak termination signal UGA A have evolved to use C-terminal amino acids which promote efficient termination.

Discussion

The last two amino acids in the nascent peptide as codon context determinants

A strong effect on termination efficiency at UGA A, in both UGA suppressor-negative and suppressor-positive *E.coli* strains, is obtained by varying the codons at the –1 and –2 positions at the 5′ side of the stop codon. For the –2 codon, we have suggested earlier that it is the corresponding amino acid, and not the codon itself, which

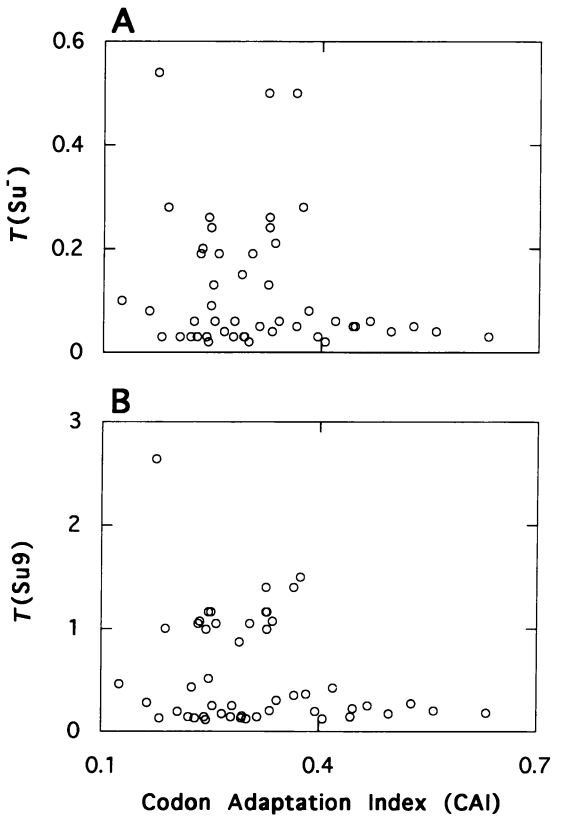

Fig. 6. UGA readthrough associated with 32 different C-terminal dipeptides as a function of the CAI for UGA A ending *E.coli* genes, encoding the same terminal dipeptides (Brown *et al.*, 1993). The UGA readthrough values corresponding to the C-terminal dipeptides were determined by a 3A′ assay, as described in Table I. The points with the 10 highest CAI values in the *trpT*$^{+}$(Su^{-}) *E.coli* strain are: ECOHU49§10, NC, 0.63, 0.03; ECNUSB§3, KA, 0.56, 0.04; ECOGROELA§1, SY, 0.53, 0.05; ECHTRA§1, NL, 0.50, 0.04; ECOLUXH§5, AS, 0.47, 0.06; ECNU00§5, YK, 0.45, 0.05; ECO110K§45, RF, 0.44, 0.05; ECMTHM§1, SG, 0.42, 0.04; ECORPSDOA2§1, RA, 0.41, 0.02; ECOPHN§9, NQ, 0.40, 0.03, and the points with the 10 highest transmission values are: ECOPAPG§2, FP, 0.17, 0.54; ECOTYSPDH§2, AP, 0.37, 0.50; ECTSRX§1, AP, 0.33, 0.50; ECOGLNB§1, VP, 0.37, 0.28; ECOGLPDA§2, IP, 0.19, 0.28; ECOGICA§1, GG, 0.24, 0.26; ECD1CABC§6, WP, 0.33, 0.24; ECO110K§16, WP, 0.25, 0.24; ECORFABC§4, SP, 0.34, 0.21; ECOPHNAQ§20, PW, 0.24, 0.20, where locus, dipeptide (amino acid single letter code), CAI value and transmission value are indicated in that order. Most of the C-terminal dipeptides have a similar ranking in the *trpT*(Su9) strain (not shown). (**A**) UGA readthrough in a *trpT*$^{+}$(Su^{-}) strain. (**B**) UGA readthrough in a *trpT*(Su9) strain.

is a determinant for context effects (Mottagui-Tabar *et al.*, 1994). The –1 codon context effect described here is of similar magnitude to the previously described –2 context effect. The codons within a –1 codon family (i.e. code for the same amino acid) give similar UGA decoding effects, although exceptions are found, suggesting that the identity of isoaccepting tRNA at the –1 codon location is also important (Table I, see below). Furthermore, a significant correlation between readthrough influence of the –1 codon and secondary structure propensities of the corresponding amino acid is observed (Figure 3 and Table II). These findings indicate that also in the case of the –1 codon, it is the amino acid, and not the codon, which is the major determinant for context effects. Thus, both the penultimate and the last amino acids at the C-terminal end of the nascent peptide affect efficiency of termination at UGA A.

For the –2 amino acid residue, its acidic/basic as well as hydrophobic/hydrophilic properties are important for termination efficiency (Table II). Thus, acidic residues, such as aspartic acid and glutamic acid, give inefficient termination, and basic residues, such as arginine and lysine, give efficient termination at UGA A. Of the neutral amino acids, phenylalanine (hydrophobic) promotes inefficient, and glutamine (hydrophilic) promotes efficient termination (Mottagui-Tabar *et al.*, 1994). For the –1 amino acid residue, a high propensity for an ordered structure (α-helix, β-strand) favours termination. Thus, the physical properties of the –1 and –2 amino acids which influence termination at the UGA A termination signal are different. However, the effect of changing the –1 codon is dependent on the nature of the –2 codon (Table I). This cooperative effect indicates that the two amino acids act together as a context effector in translation.

How do the two C-terminal amino acids in the nascent peptide influence translation termination?

Cellular RNA was analysed for the appearance of a small RNA species, denoted rpRNA, which is a degradation product arising from 3A′ mRNA. This RNA is indicative of a ribosomal pause at the internal stop codon in the 3A′ mRNA (A.Björnsson and L.A.Isaksson, in preparation). Figure 5 shows that rpRNA is found only for contexts which give inefficient UGA termination. Inefficient termination can be explained by high tRNA decoding activity, low RF-2 activity, or both. Less efficient RF-2 binding, in contrast to increased tRNA decoding activity, will make UGA recognition slower. Decreased RF-2 binding could cause ribosome pausing at the stop codon, thus protecting it from degradation and leading to the accumulation of rpRNA. The data therefore indicate that contexts which are associated with inefficient termination are also associated with low RF-2 efficiency. This could result from a disturbed interaction between the penultimate amino acid residue (negative charge, hydrophobicity) and some component in the RF-2-mediated termination reaction. The cooperative nature of the nascent peptide effect indicates that the last amino acid is also involved, but maybe only indirectly (secondary structure propensity), by positioning of the interacting penultimate residue. As a result of low RF-2 efficiency, the competing UGA readthrough reaction by $tRNA^{Trp}$ (CCA) would be favoured.

We note that the van der Waals volume of the last amino acid residue can influence the selection of wild-type and suppressor forms of $tRNA^{Trp}$ (CAA) differently when these tRNAs compete with RF-2 (Figure 4 and Table II). A decrease in van der Waals volume of the amino acid seems to increase the efficiency of the suppressor tRNA. This indicates that the tRNA in the EF-Tu ternary complex or some other part of the ternary complex could interact directly with the C-terminal end of the nascent peptide during tRNA selection at the ribosomal A-site.

The exceptions: influences of P-site tRNA on UGA decoding

Translational readthrough is low for all six arginine codons at the –1 codon position, especially for CGA (Table I and Figure 2), suggesting that arginine itself is associated with an efficient termination. The three arginine codons, CGU, CGC and CGA, are exclusively decoded by $tRNA^{Arg2}$ (ICG), which is the only tRNA in eubacteria with inosine at the first base position in the anticodon (Osawa *et al.*, 1990). The I–A base pair seems to be less stable than I–U and I–C during anticodon–codon interaction (Crick, 1966; Munz *et al.*, 1981). This wobbling deficiency of $tRNA^{Arg2}$ at the –1 location (P-site) has been implicated in decreased readthrough of UAG at the ribosomal A-site (Curran, 1994). Most probably, such inefficient wobbling is also the explanation for the extremely low UGA readthrough found to be associated with the CGA codon.

The effects of the –1 glycine codons are too variable to only result from the amino acid glycine in the nascent peptide. Instead, it seems possible that $tRNA^{Gly3}$ has an effect on termination/readthrough which is different from that of $tRNA^{Gly1}$ and $tRNA^{Gly2}$ when it is located at the ribosomal P-site. This is supported by the finding that termination at UAG by RF-1 is affected differently by a GGA decoding mutant form of $tRNA^{Gly3}$ at the ribosomal P-site, compared with the normal GGA translator $tRNA^{Gly2}$ (S.Zhang, M.Rydén-Aulin and L.A.Isaksson, in preparation). These data are in line with the observation that the structure of the P-site tRNA can influence ribosomal A-site decoding (Smith and Yarus, 1989).

Highly expressed E.coli genes encode C-terminal amino acids which give efficient termination

We have analysed here how different C-terminal amino acids in the nascent peptide, corresponding to the –2 and –1 codons, influence translation termination at UGA A. Our data set includes 32 different dipeptides, which are found at the C-terminal ends of proteins encoded by UGA A ending *E.coli* genes. The UGA readthrough values associated with the C-terminal dipeptides were analysed by plotting against the CAI values for genes encoding the same dipeptide. This index reflects the gene expression level (Sharp and Burgess, 1992). The results suggest that highly expressed genes have a bias against C-terminal dipeptides, which give inefficient termination (Figure 6). This indicates that the last two amino acids in the nascent peptide are under selection constraints, so that high gene expression is not disturbed by inefficient translation termination. Furthermore, the same constraint could explain the absence of the G–non-G–N sequence pattern five codons upstream of stop codons (Brown *et al.*, 1990). However, only a 3-fold effect on termination is observed when the –3 codon is varied, indicating that the nascent peptide effect on UGA termination levels off for amino acids further upstream from the C-terminal end of the nascent peptide (Mottagui-Tabar *et al.*, 1994).

Statistical analysis of codons corresponding to the last amino acids of proteins shows that UGA codons in *E.coli* have a preference for a 5′ adjacent serine codon (UCC) and both phenylalanine codons in *E.coli* mRNAs (Arkov *et al.*, 1993). Our data show that the –1 serine codon UCC has a similar effect to that of the serine codon AGU, both giving efficient UGA termination (Table I). The phenylalanine codon tested at the –1 position also gives efficient translation termination. Analysis of the 5′ adjacent codon to natural UAA stop codons reveals a preference for lysine codons. Threonine and proline codons are underrepresented (Arkov *et al.*, 1993). Our data indicate that lysine gives efficient, whereas threonine and proline

give inefficient, termination at UGA. UGA is rarely used as a stop codon in highly expressed genes. Instead, UAA is used in 92% of such genes (Brown *et al.*, 1990). Since RF-2 is also used for termination at UAA, the preference for lysine and the avoidance of threonine and proline at the 5′ side of this codon could contribute to higher RF-2 efficiency at UAA in mRNAs of highly expressed genes.

The nascent peptide and regulation of gene expression

The leader sequence of the *cat* gene in Gram-positive bacteria plays a role in ribosome stalling which is necessary for gene induction (Alexieva *et al.*, 1988). This is caused by an N-terminal pentapeptide sequence in the nascent peptide which inhibits the ribosomal peptidyltransferase centre (Gu *et al.*, 1993). The ribosome can also 'hop' over a 50 base long mRNA region containing several stop codons in the bacteriophage T4 topoisomerase mRNA. This event is dependent on a stretch of 15 amino acids in the nascent peptide coded by the mRNA ~19 codons upstream of the bypass region (Weiss *et al.*, 1990). Our results suggesting that the last two amino acids affect termination at UGA A provide an additional example of how the nascent peptide can affect gene expression.

The van der Waals volume of the C-terminal amino acid of the nascent peptide influences differently the efficiency of UGA decoding by the wild-type or the suppressor form of $tRNA^{Trp}$ (CAA) (Figure 4 and Table II). This observation indicates that the choice of isoacceptor tRNAs during translation of sense codons might also be affected by the nascent peptide. Such an effect on tRNA selection rates could mean that the nascent peptide influences the rate of translation elongation, thereby affecting the level of gene expression.

Materials and methods

Materials

Restriction enzymes, T4 DNA ligase and the Wizard DNA miniprep kit were purchased from Promega, oligodeoxynucleotides from KEBO, the Sequenase kit from Amersham and the Geneclean kit from Bio 101.

Growth media and bacterial strains

LB and M9 growth media were made as described previously (Miller, 1972). All 20 L-amino acids were added to the M9-based medium at recommended concentrations (Neidhardt *et al.*, 1974), together with ampicillin (200 μg/ml) and IPTG (0.5 mM). The *E.coli* strain MC1061 was used for gene cloning work (Sambrook *et al.*, 1989), and XAc and CDJ64 were used as suppressor-negative [$trpT^{+}$(Su^{-})] and suppressor-positive [*trpT*(Su9)] *E.coli* strains, respectively (Coulondre and Miller, 1977).

Cloning of UGA A codon context variants

Standard recombination techniques were used (Sambrook *et al.*, 1989) and all synthetic sequences cloned were sequenced using a Sequenase kit (ABPO2 primer) (Björnsson and Isaksson, 1993). All variants of the NNN NNN UGA A (NNN denoting a variant codon) sequence were cloned from complementary oligodeoxynucleotides with *Asu*II and *Spe*I overhangs into pAB93 (Björnsson and Isaksson, 1993; Mottagui-Tabar *et al.*, 1994).

3A′ gene protein assay

Up to forty 3A′ plasmid strains were assayed each time (Björnsson and Isaksson, 1993). Inoculation was from overnight cultures into M9 assay media (10 ml) and cells were grown with good aeration at 37°C. Cells were harvested in the mid-log growth phase at an A_{540} of 0.5, placed on ice and pelleted by centrifugation. Bacterial pellets were resuspended in 1 ml of 10× TST [0.5 M Tris (pH 7.4), 1.5 M NaCl, 0.5% (w/v) Tween 20] and the cells were lysed by incubation at 90°C for 8 min. Cell debris was eliminated by centrifugation at 13 000 r.p.m. for 10 min at 4°C. A′ protein was purified from supernatants using IgG-containing mini-columns [filters from a Wizard DNA miniprep system (Promega) with 100 μl of IgG Sepharose® 6 Fast Flow (Pharmacia) per column] and a vacuum manifold system (Promega). Each column was first washed with 2 ml of 0.5 M HAc (pH 3.2) and 2 ml of 1× TST. The A′ protein-containing supernatant was then added to each column, followed by several washing steps; first 2 ml of 1× TST, then 2 ml of 1× TS (same composition as 1× TST but without Tween 20), and finally 2 ml of 5 mM NH_4Ac (pH 5.5). A′ protein was eluted from each column with 0.1 ml of 0.5 M HAc (pH 3.2). IgG columns were washed with 2 ml of 1× TST after each use and stored in 10× TST at 4°C. The eluate was dried in a Speed Vac prior to SDS–PAGE analysis. The A′ proteins were then processed and quantified as described previously (Björnsson and Isaksson, 1993).

RNA preparations and Northern blots

Escherichia coli cultures (20 ml) were inoculated using the M9-based medium and growth conditions described above. Total RNA was prepared using the hot phenol method (von Gabain *et al.*, 1983). Samples containing 5 μg of total RNA were fractionated on 6% Sanger gels in a minigel system at 15 mA/gel until the xylene cyanol marker dye was running out of the gel (Midget system from Pharmacia), followed by blotting in a Midget minigel blotting unit at 20 V overnight. The ^{32}P-labelled oligo-probe ABS01 (5′-CGTTGTTCTTCGTTTAAGTTAGG-3′) which is complementary to a sequence within each A′ domain of the 3A′ mRNA was used for hybridization (20 pmol/filter) (Nilsson *et al.*, 1987; Björnsson and Isaksson, 1993).

Acknowledgements

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

References

Alexieva,Z., Duvall,E.J., Ambulos,N.P.,Jr, Kim,U.J. and Lovett,P.S. (1988) Chloramphemicol induction of cat-86 requires ribosome stalling at a specific site in the leader. *Proc. Natl Acad. Sci. USA*, **85**, 3057–3061.

Arkov,A.L., Korolev,S.V. and Kisselev,L.L. (1993) Termination of translation in bacteria may be modulated via specific interaction between peptide chain release factor-2 and the last peptidyl-transfer RNA(Ser/Phe). *Nucleic Acids Res.*, **21**, 2891–2897.

Björnsson,A. (1994) *The Influence of Codon Context on Translation Termination and Messenger RNA Levels*. Thesis, Stockholm University, Akademitryck, Edsbruk, Sweden.

Björnsson,A. and Isaksson,L.A. (1993) UGA codon context which spans three codons; reversal by ms^2i^6A37 in tRNA, mutation in *rpsD*(S4) or streptomycin. *J. Mol. Biol.*, **232**, 1017–1029.

Blaber,M., Zang,X.-J. and Matthews,B.W. (1994) Structural basis of amino acid alpha helix propensity. *Science*, **260**, 1637–1640.

Bossi,L. (1983) Context effects: translation of UAG codon by suppressor tRNA is affected by the sequence following UAG in the message. *J. Mol. Biol.*, **165**, 73–87.

Bossi,L. and Roth,J.R. (1980) The influence of codon context on genetic code translation. *Nature*, **286**, 123–127.

Brown,C.M., Stockwell,P.A., Trotman,C.N.A. and Tate,W.P. (1990) The signal for the termination of protein synthesis in procaryotes. *Nucleic Acids Res.*, **18**, 2079–2086.

Brown,C.M., Dalphin,M.E., Stockwell,P.A. and Tate,W.P. (1993) The translational termination signal database. *Nucleic Acids Res.*, **21**, 3119–3123.

Buckingham,R.H., Sörensen,P., Pagel,F.T., Hijazi,K.A., Mims,B.H., Brechemier-Baey,D. and Murgola,E.J. (1990) Third position base changes in codons 5′ and 3′ adjacent UGA codons affect UGA suppression *in vivo*. *Biochim. Biophys. Acta*, **1050**, 259–260.

Coulondre,C. and Miller,J.H. (1977) Genetic studies of the lac repressor. III. Additional correlation of mutational sites with specific amino acid residues. *J. Mol. Biol.*, **117**, 525–576.

Creighton,T.E. (1993) *Proteins: Structures*. W.H.Freeman and Co., New York.

Crick,F.H.C. (1966) Codon–anticodon pairing: the wobble hypothesis. *J. Mol. Biol.*, **19**, 548–555.

Curran,J.F. (1994) Decoding with the A:I wobble pair is inefficient. *Nucleic Acids Res.*, **23**, 683–688.

Curran,J.F. and Gross,B.L. (1994) Evidence that GHN phase bias does not constitute a framing code. *J. Mol. Biol.*, **235**, 389–395.

Faxén,M., Plumbridge,J. and Isaksson,L.A. (1991) Codon choice and potential complementarity between messenger RNA downstream of the initiation codon and bases 1471–1480 in 16S-ribosomal RNA affects expression of *glnS*. *Nucleic Acids Res.*, **19**, 5247–5251.

Goldstein,J.L. and Caskey,C.T. (1970) Peptide chain termination: effect of protein S on ribosome binding of release factors. *Proc. Natl Acad. Sci. USA*, **67**, 537–543.

Grentzmann,G., Brechemier-Baey,D., Heurgué,V., Mora,L. and Buckingham,R.H. (1994) Localization and characterization of the gene encoding release factor RF3 in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **91**, 5848–5852.

Grentzmann,G., Brechemier-Baey,D., Heurgué-Hamard,V. and Buckingham,R.H. (1995) Function of polypeptide chain release factor RF-3 in *Escherichia coli*. *J. Biol. Chem.*, **270**, 10595–10600.

Gu,Z., Rogers,E.J. and Lovett,P.S. (1993) Peptidyl transferase inhibition by the nascent leader peptide of an inducible *cat* gene. *J. Bacteriol.*, **175**, 5301–5313.

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

Hirsh,D. (1971) Tryptophan transfer RNA as the UGA suppressor. *J. Mol. Biol.*, **58**, 439–458.

Ikemura,T. (1992) Correlation between codon usage and tRNA content in microoganisms. In Hatfield,D.L., Lee,B.J. and Pirtle,R.M. (eds), *Transfer RNA in Protein Synthesis*. CRC Press Ltd, Boca Raton, FL, pp. 87–111.

Kopelowitz,J., Hampe,C., Goldman,R., Reches,M. and Engelberg-Kulka,H. (1992) Influence of codon context on UGA suppression and readthrough. *J. Mol. Biol.*, **225**, 261–269.

Lim,V.L. and Spirin,A.S. (1986) Stereochemical analysis of ribosomal transpeptidation; conformation of nascent peptide. *J. Mol. Biol.*, **188**, 565–577.

Mikuni,O., Ito,K., Moffat,J., Matsumura,K., McCaughan,K., Nobukuni,T., Tate,W. and Nakamura,Y. (1994) Identification of the *prfC* gene, which encodes peptide-chain-release factor 3 of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **91**, 5798–5802.

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

Miller,J.H. and Albertini,A.M. (1983) Effects of surrounding sequence on the suppression of nonsense codons. *J. Mol. Biol.*, **164**, 59–71.

Minor,D.L.,Jr and Kim,P.S. (1994) Measurement of the beta-sheet-forming propensities of amino acids. *Nature*, **367**, 660–663.

Mottagui-Tabar,S., Björnsson,A. and Isaksson,L.A. (1994) The second to last amino acid in the nascent peptide as a codon context determinant. *EMBO J.*, **13**, 249–257.

Munz,P., Leupold,U., Agris,P. and Kohli,J. (1981) *In vivo* decoding rules in *Schizosaccharomyces pombe* are at variance with *in vitro* data. *Nature*, **294**, 187–188.

Murgola,E.J. (1985) tRNA, suppression, and the code. *Annu. Rev. Genet.*, **19**, 57–80.

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

Nilsson,B., Moks,T., Jansson,B., Abrahamsén,L., Elmblad,A., Holmgren,E., Henrichson,C., Jones,T.A. and Uhlén,M. (1987) A synthetic IgG binding domain based on staphylococcal protein A. *Protein Engng*, **1**, 107–113.

Osawa,S., Muto,A., Ohama,T., Andachi,Y., Tanaka,R. and Yamao,F. (1990) Prokaryotic genetic code. *Experientia*, **46**, 1097–1106.

Pedersen,W.T. and Curran,J.F. (1991) Effects of the nucleotide 3′ to an amber codon on ribosomal selection rates of suppressor transfer RNA and release factor-1. *J. Mol. Biol.*, **219**, 231–241.

Poole,E.S., Brown,C.M. and Tate,W.P. (1995) The identity of the base following the stop codon determines the efficiency of *in vivo* translational termination in *Escherichia coli*. *EMBO J.*, **14**, 151–158.

Ringquist,S., Shinedling,S., Barrick,D., Green,L., Binkley,J., Stormo,G.D. and Gold,L. (1992) Translation initiation in *Escherichia coli*—sequences within the ribosome-binding site. *Mol. Microbiol.*, **6**, 1219–1229.

Roseman,M.A. (1988) Hydrophilicity of polar amino acid side-chains is markedly reduced by flanking peptide bonds. *J. Mol. Biol.*, **200**, 513–522.

Salser,W. (1969) The influence of the reading context upon the suppression of nonsense codons. *Mol. Gen. Genet.*, **105**, 125–130.

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

Scolnik,E., Tompkins,R., Caskey,T. and Nirenberg,M. (1968) Release factors differing in specificity for terminator codons. *Proc. Natl Acad. Sci. USA*, **61**, 768–774.

Sharp,P.M. and Burgess,C.J. (1992). Selective use of termination codons and variations in codon choice. In Hatfield,D.L., Lee,B.J. and Pirtle,R.M. (eds), *Transfer RNA in Protein Synthesis*. CRC Press Ltd, Boco Raton, FL, pp. 397–425.

Shine,J. and Dalgarno,L. (1974) The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl Acad. Sci. USA*, **71**, 1342–1346.

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

Sprengart,M.L., Flatscher,P.H. and Fuchs,E. (1990) The initiation of translation in *E.coli*: apparent base pairing between the 16S rRNA and downstream sequences of the mRNA. *Nucleic Acids Res.*, **18**, 1719–1723.

Stormo,G.D., Schneider,T.D. and Gold,L. (1986) Quantitative analysis of the functional relationship between nucleotide sequence and functional activity. *Nucleic Acids Res.*, **14**, 6661–6679.

Trifonov,E.N. (1987) Translation framing code and frame-monitoring mechanism as suggested by the analysis of mRNA and 16S rRNA nucleotide sequences. *J. Mol. Biol.*, **194**, 643–652.

von Gabain,A., Belasco,J.G., Schottel,J.L., Chang,A.C.Y. and Cohen,S.N. (1983) Decay of mRNA in *Escherichia coli*: investigation of the fate of specific segments of transcripts. *Proc. Natl Acad. Sci. USA*, **80**, 653–657.

Weiss,R.B., Huang,W.M. and Dunn,D.M. (1990) A nascent peptide is required for ribosomal bypass of the coding gap in bacteriophage T4 gene *60*. *Cell*, **62**, 117–126.

Williams,R.S., Chang,A., Juretic,D. and Loughran,S. (1987) Secondary structure predictions and medium range interactions. *Biochim. Biophys. Acta*, **916**, 200–204.

Yarus,M. and Folley,L.S. (1985) Sense codons are found in specific contexts. *J. Mol. Biol.*, **182**, 529–540.

Received on September 13, 1995; revised on November 28, 1995