Volume 13 Number 11 1985 Nucleic Acids Research

Codon-defined ribosomal pausing in *Escherichia coli* detected by using the *pyrE* attenuator to probe the coupling between transcription and translation

Fons Bonekamp, Henrik Dalbøge Andersen[1], Thorkild Christensen[1] and Kaj Frank Jensen*

Institute of Biological Chemistry B, University of Copenhagen, Sølvgade 83, DK-1307 Copenhagen K, and [1]Nordisk Gentofte, Niels Steensens vej 1, DK-2820 Gentofte, Denmark

Received 18 April 1985; Accepted 9 May 1985

ABSTRACT

This communication describes an assay for the relative translation efficiency of individual codons which makes use of the pyrE attenuator to probe the coupling between transcription and translation at the end of an artificial leader peptide. By cloning of short synthetic DNA fragments the codons to be tested were placed in the middle of the leader peptide and the downstream transcription of a pyrE''lacZ gene was monitored by measuring β-galactosidase activity. The substitution, one by one, of three AGG codons for arginine with three CGT codons for the same amino acid residue was found to cause a two fold increase per codon of transcription over the pyrE attenuator, such that an eight fold higher frequency of pyrE expression was seen when all three AGG codons were replaced by CGT codons. No such effect of codon composition was observed, when the cells were grown with a low UTP pool which causes a reduction of the mRNA chain growth rate.

INTRODUCTION

During the past few years a steadily increasing amount of efforts has been made in order to understand the consequences of a biased use of codons. In general, the tendency towards the use of a more selective subset of codons increases with the expressivity of the genes (1). Moreover, as the preferred codons differ between organisms this may be an obstacle for expression of cloned genes in host cells, where the translation machinery is not optimized for the codon composition of the foreign gene.

Grosjean and Fiers (1) have treated this problem on a thermodynamic and statistical basis. They find that the rarely used codons are either decoded by minor tRNA species or they may be codons where the interaction with the anticodon is either too strong or too weak for a fast translation to proceed with a high fidelity. It can therefore be assumed that infrequently used

codons are poor or slow codons. This view has recently received support from the observations of Pedersen (2) who found that different protein chains are not elongated at the same rate in E.coli, at least when the genes are present on multicopy plasmids. In short, genes that contain more poor codons, as defined by Grosjean and Fiers, were found to be translated at a lower rate than genes containing few of these.

From studies of the polarity effects of non-sense mutations in operons it has long been known that a coupled translation is able to suppress mRNA chain termination at Rho-dependent transcription terminators inside genes (3). Recently, this was found also to be true for the Rho-independent transcription terminators at the pyrB attenuator and at the intercistronic attenuator in front of the pyrE gene (4-6). In the latter case, the translation of the preceeding gene stops eight nucleotide residues before the attenuator symmetry (5). The frequency of mRNA chain termination at these two attenuators appears to be controlled directly by the coupling between transcription and translation at the upstream side of the terminator structures. Changes in the UTP pool modulate the attenuation frequency through alterations in the saturation of RNA polymerase by UTP, which in turn alters the coupling because of a change in the mRNA chain elongation rate (4,5,8,9).

We have previously found indicative evidence suggesting that the codon composition in the translated region upstream to the pyrE attenuator has a profound effect on the frequency of mRNA chain termination and hence seems to affect the rate of polypeptide chain elongation (6). This observation is further explored in this communication which describes the use of the pyrE attenuator to probe for the translational efficiency of individual codons placed in an artificial leader peptide in front of the attenuator. The results show that the successive substitution, one by one, of three rare arginine codons (AGG) at position -17, -18, and -19 relative to the end of the leader peptide with frequently used arginine codons (CGT) increases transcriptional read-through over the pyrE attenuator by eight fold, when all three AGG codons are replaced.

MATERIALS AND METHODS

Materials:

Restriction endonucleases and T4-DNA ligase were obtained from New England Biolabs. Isopropyl-β-D-thiogalactoside (IPTG) and o-nitrophenyl-β-D-galactoside (ONPG), 5′-UMP, and antibiotics were purchased from Sigma.

Strains and plasmids:

The construction of strain SØ1256 Δ(*araD*139) Δ(*lac*)U169 *thi rpsL pyrB*) has been described previously (10). NF1830 is a *recA* derivative of MC1000 containing the episome F′*lacI*^{q1}*Z*::Tn*5*. The construction of plasmid pKCL107 was described by Bonekamp *et. al*. (6).

Plasmid constructions:

Oligonucleotides were synthesized by the procedure of Sanchez-Pescador and Ordea (11). Hybridization of a 20-mer and a 12 mer to form a double-stranded DNA fragment with *Eco*RI and *Pst*I overhangs (Figure 2) was performed by mixing equal amounts, heating to 90°C for 5 minutes and subsequently cooling on ice. Plasmid pKCL107 was digested with *Eco*RI and *Pst*I giving rize to three fragments: A large *Pst*I-*Pst*I fragment extending from the *Pst*I site in the cloning region to that in the β-lactamase gene, a large *Pst*I-*Eco*RI fragment containing the *pyrE′′lacZ* gene fusion, and a small *Eco*RI-*Pst*I fragment from the cloning region region of the artificial leader peptide. The two large fragments were purified by polyacrylamide gel electrophoresis, cut out from the gel and electroeluted into dialysis bags as described (12). The fragments were then ligated together with a large molar excess of the synthetic *Eco*RI-*Pst*I fragment. The ligated mixture was used to transform the *E.coli* strain SØ1256 (*pyrB*) selecting for ampicillin resistance. The plasmids (pCU201-204) were isolated from the transformants and that they had the desired and described structure (Figure 2) was controlled by nucleotide sequence determinations (13) from the *lacZ* start and across the *pyrE* attenuator. The episome F′*lacI*^{q1}*Z*::Tn*5* was now introduced in the plasmid-bearing strains by conjugation with NF1830.

Other procedures:

The media and growth conditions, induction of the *lac*-promo-

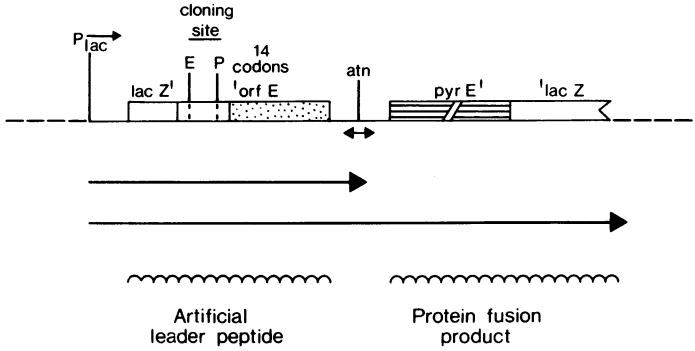

Figure 1: Structure and expression of the pyrE-lacZ gene fusions on plasmids. Plasmid pKCL107 carries resistance to ampicillin and its construction was described in ref.(6). The DNA containing the lacZ start and the cloning region is derived from pUC8 while the rest of the lacZ gene originates from pMLB1034. To construct plasmids pCU201-204 the EcoRI(R)-PstI(P) fragment in the linker was eliminated from pKCL107 and replaced by different synthetic DNA fragments. Atn denotes the UTP controlled pyrE attenuator, and the arrows below indicate the attenuated and the non-attenuated transcripts formed after induction by IPTG. Figure 2 gives the sequences of the leader region of plasmids pCU201-204.

ters with IPTG, sampling and assay procedures have all been described previously (6).

RESULTS

The pyrE gene is transcribed as the second gene of a bicistronic operon, orfE-pyrE, where the first gene (orfE) encodes a 238 amino acid residues long protein of unknown function (5). Expression of pyrE is controlled in the intercistronic space of the operon by a UTP modulated attenuation (5,6). To study this attenuation we have previously constructed a set of translational fusion vectors where expression of a pyrE''lacZ fusion is brought under control of the lac-promoter (6). Translation of a leader segment containing the first few codons of lacZ and a cloning region following it, 42 bp of DNA from the orfE end, and the pyrE attenuator is controlled by the lacZ translational start (Figure 1) which was varied in the different plasmids. It was found that pyrE expression and regulation required a translation in the leader segment (the end of orfE) close up to

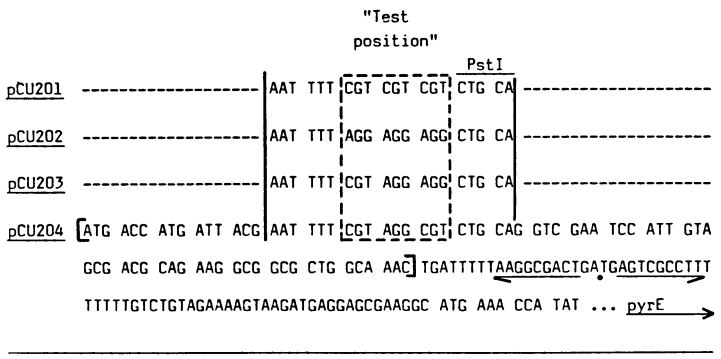

Figure 2: Nucleotide sequence of leader regions in pCU201-204. Open reading frames for the artificial leader peptides are shown in square brackets. The sequence is shown in full only for pCU204. The areas within the vertical solid lines indicate the synthetic oligonucleotides, while the area of three codons inside the broken lines shows the position where the codons are varied, i.e the "test position". Opposite arrows beneath the sequence of pCU204 denotes the symmetry region of the pyrE attenuator.

the attenuator structure. However at the same time we observed that the attenuation frequency differed as a result of variations in the middle of the translated leader segment, indicating that either the length or the codon usage in the artificial leader peptide can affect pyrE regulation (6). One of the constructed plasmids (pKCL107) contains a unique cleavage site for EcoRI in the leader and a PstI site, which apart from a site in the bla gene also is unique (Figure 1). This allows the substitution of the EcoRI-PstI fragment within the leader by synthetic DNA fragments of identical length, but with a varying codon content at a defined position (pCU201-204). As seen from Figure 2 the DNA fragments, which each consist of a 20-mer and a complementary 12-mer annealed to give an EcoRI and a PstI overhang, differ only in 1-3 codons inside what we have termed the "test position". The resulting artificial leader peptides all end like the orfE gene eight basepairs upstream to the pyrE attenuator symmetry, their 14 last codons being identical to the orfE end (Figure 2). We have chosen to test the most frequent and the most infrequent codons for arginine (CGT and AGG,

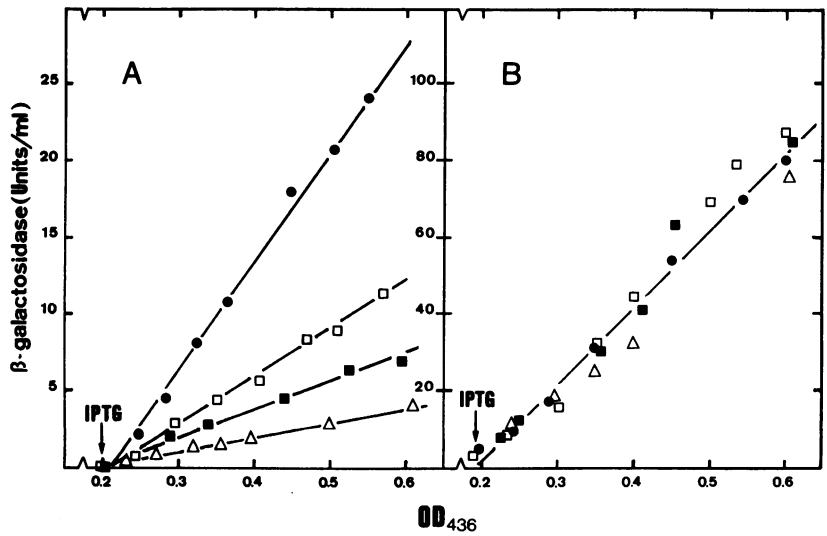

Figure 3: Expression and control of the hybrid pyrE-lacZ genes of plasmids pCU201-204 in strain S01256 (pyrB). Figure 3A shows the differential rate of appearence of β-galactosidase activity following induction by IPTG (0.1 mg/ml) of cultures grown with uracil (25 μg/ml). Figure 3B shows the result of induction by IPTG of cultures grown with 5'UMP (100 μg/ml). Closed circles pCU201; open triangles, pCU202; closed squares, pCU203; open squares, pCU204. The datapoints are means of duplicate assays.

respectively), since these have previously been shown to differ with regard to translational efficiency (15).

Plasmids pCU201-204 and the control plasmid pKCL11, which carries no pyrE DNA in its lacZ gene (6), were introduced into strain S01256 (pyrB) together with the episome F'lacI^{q1}Z::Tn5, which contains a promoter-up mutation and overproduces the lac repressor. Consequently, the lac promoter is repressed even when present on multicopy plasmids. The defective pyrB gene of strain S01256 enables us to study the attenuation frequency both at high and low UTP pools by using either uracil or 5'UMP as source of pyrimidine, respectively (6,14).

Figure 3 shows the differential rate of appearence of β-galactosidase activity from the hybrid pyrE''lacZ genes of the plasmids upon induction of the lac-promoter by isopropyl-β-D-thiogalactoside (IPTG). The slopes of the curves are given in Table 1. The following conclusions can be drawn:

1. Since virtually no β-galactosidase activity is present

Table 1: Effect of different arginine codons on the regulation of pyrE attenuation.

Plasmid	No. of CGT codons	Differential rate of β-galactosidase activity[a] (units/ml x OD_{436}) with: uracil		5'UMP
pCU201	3	71		ca.180
pCU204	2	31	(2.3)[b]	ca.180
pCU203	1	19	(1.7)	ca.180
pCU202	0	9	(2.1)	ca.180
pKCL11[c]	-	190	-	190

a) Taken from the slopes of Figure 3.
b) The numbers in parentheses give the positive effect (fold) of substituting one additional AGG codon with a CGT codon in the leader.
c) Control plasmid without any pyrE DNA.

with any of the plasmids at the time of inducer (IPTG) addition, the pyrE''lacZ gene is expressed only from the lac-promoter in all of the plasmids.

2. Relative to uracil, the use of 5'UMP as pyrimidine source leads to an increased synthesis of β-galactosidase activity in all plasmids containing pyrE DNA. The control plasmid pKCL11, does not respond to these manipulations (Table 1).

3. The substitution of the three CGT codons in the "test position" (Figure 2) by AGG codons results in an increase of the frequency of mRNA chain terminations at the pyrE attenuator by roughly a factor two per codon, but only when uracil serves as source of pyrimidine, and the UTP pool hence is high. There is no effect of the codon composition in the leader when RNA polymerase transcribes slowly due to a reduced UTP pool established by the use of 5'UMP as pyrimidine source.

These observations strongly indicate that individual codons can influence the coupling between transcription and translation and hence the rate of polypeptide chain elongation.

DISCUSSION

It has previously been shown that translation of the orfE message preceeding pyrE and of the leader-peptide preceeding pyrB is an absolute prerequisite for the regulation and expression of these two genes (5,16). The regulation of these genes is established by attenuation sensitive to variations in the degree of coupling between the leading ribosome and the transcribing RNA polymerase (4-6). The results of the present paper show that apart from the cellular concentration of UTP, which was shown to be a dynamic modulator of attenuation by affecting the elongation rate of RNA polymerase, also the static codon composition of the leader-peptide is important for the coupling between translation and transcription. Efficient mRNA chain termination did not occur at the pyrE attenuator during growth on uracil if the artificial leader-peptide utilized only CGT codons for the arginines in the "test position", but could be established if these codons were replaced by AGG codons. It appears therefore that rare codons like (AGG), which we propose function as ribosomal pause sites, have to be present in the translated leader region in order to allow RNA polymerase to escape efficiently from the leading ribosome at conditions of high cellular UTP pools. No increase in termination (i.e. decoupling) was observed by the poor arginine codons (AGG) when the cells were grown with UMP. This condition of growth results in a reduced UTP pool (10) and RNA polymerase slows down its transcription rate or may even elicit a pause (4,8,9). This finding suggests that the pausing made by the ribosomes at the 1-3 AGG codons is short compared to the reduction in the mRNA chain growth rate caused by the low UTP pool.

Other factors have to be considered as possible explanations for the different behaviour of the pyrE attenuator in plasmids pCU201-204. These include possibilities for secondary structure formation, mRNA stability, and protein stability. However, for the reasons below we consider these factors unlikely as explanations for our observations:

First, since the cloned fragments were assymmetric in their design there are no obvious possibilities for different secondary structure formation in plasmids pCU201-204. Likewise, we

consider it unlikely that secondary structures might explain the simple relation between the number of AGG codons and the frequency of mRNA termination at the pyrE attenuator. Second, if the cloning of the various fragments in the leader region resulted in a change in mRNA stability this should only be manifested when cells are grown with uracil as pyrimidine source, and not when 5'UMP serves this purpose. Third, the different fragments are cloned in a leader region and do thus not change the sequence of the protein product of the pyrE''lacZ hybrid gene which is used to monitor pyrE expression.

Statistical analysis of the nucleotide sequences of E.coli genes has revealed a strong relation between codon composition and mRNA expressivity (1,17,18) and that genes encoding abundant proteins selectively use codons corresponding to major tRNAs (19,20). Furthermore, it has been suggested that translation modulation by the use of isoaccepting tRNAs might be part of a mechanism to keep expression of regulatory genes low (21).

The first direct evidence that codon usage can effect protein synthesis was obtained from experiments where insertion of four AGG codons in the reading-frame for chloramphenicol transacetylase decreased the yield of protein product at very high transcription rates (15). Furthermore, the observation that the cloning of lacI and bla on multicopy plasmids increased the time needed for their expression was interpreted as being a result of an overall drain of the rare tRNAs cognate to the rare codons in the genes (2).

With the increasing interest in the expression of eukaryotic genes in prokaryotes an understanding of the consequences of codon usage for the expressivity of a cloned gene in the host is clearly needed. The assay described in this communication, where the codons to be studied are placed in the leader area in front of the pyrE attenuator offers an opportunity for a systematic study of codon induced effects on the coupling between transcription and translation, and thereby a measure for the relative translation kinetics of the individual code-words of the genetic code. It shall now be interesting to learn whether the most frequently used codons for different aminoacids are translated at different rates. The system described here

probably carries no constraints on the choice of codons to be tested, since these are incorporated in a leader-region of no other function than to probe the coupling between transcription and translation. Therefore, as opposed to other systems (15), the changes of aminoacid residues do not influence the stability and activity of protein product to be monitored. One disadvantage of the system is that the pyrE attenuation frequencies cannot be transformed into step-times for the ribosomes at the tested codons, but hopefully future experimentation will make this calculation possible.

ACKNOWLEDGEMENTS

This work was supported by grants from the Danish Natural Science Foundation to K.F.J. and from the Carlsberg Foundation by a Scholar stipendium to F.B. We want to thank Lise Schack for excellent technical assistance.

*To whom correspondence should be sent

REFERENCES

1. Grosjean, H. and Fiers, W. (1982) Gene 18, 199-209.
2. Petersen, S. (1984) EMBO J. 3, 2895-2898.
3. Adhya, S. and Gottesman, M. (1978) Ann.Rev.Biochem. 47, 967-996.
4. Turnbough, C.L.Jr., Hicks, H.L. and Donahue, J.P. (1983) Proc.Natl.Acad.Sci.USA 80, 368-372.
5. Poulsen, P., Bonekamp, F. and Jensen, K.F. (1984) EMBO J. 3, 1783-1790.
6. Bonekamp, F., Clemmesen, K., Karlstrom, O. and Jensen, K.F. (1984) EMBO J. 3, 2857-2861.
8. Jensen, K.F., Neuhard, J. and Schack, L. (1982) EMBO J. 1, 69-74.
9. Jensen, K.F., Fast, R. and Larsen, J.N. (1985) J.Bacteriol. (submitted).
10. Jensen, K.F., Larsen, J.N., Schack, L. and Sivertsen, A. (1984) Eur.J.Biochem. 140, 343-352.
11. Sanchez-Pescador, R. and Ordea, M.S. (1984) DNA 3, 339-343.
12. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual. Published by Cold Spring Harbor Laboratory Press, NY.
13. Sanger, F., Nicklen, S. and Coulson, A.R. (1977). Proc.Natl.Acad.Sci.USA 74, 5463-5467.
14. Poulsen, P., Jensen, K.F., Valentin-Hansen, P., Carlsson, P. and Lundberg, L.G. (1983) Eur.J.Biochem. 135, 223-229.
15. Robinson, M., Lilley, R., Little, S. Emtage, J.S. Yarranton, G., Stephens, P., Millican, A., Eaton, M. and Humphreys, G. (1984) Nucl.Acids Res. 12, 6663-6671.

16. Clemmesen, K., Bonekamp, F., Karlstrom, O. and Jensen, K.F. (1985) J.Bacteriol. (submitted)
17. Gouy, M., and Gautier, C. (1982) Nucl.Acids Res. 10, 7055-7074.
18. Grantham, R., Gautier, C., Jacobzone, M. and Mercier, R. (1981) Nucl.Acids Res. 9, r43-r74.
19. Ikemura, T. (1981) J.Mol.Biol. 151, 389-409.
20. Ikemura, T. (1981) J.Mol.Biol. 146, 1-21.
21. Koningsberg, W. and Nigel Godson, C. (1983) Proc.Natl.Acad.Sci.USA, 80, 687-697.