Volume 15 Number 12 1987 Nucleic Acids Research

More than 150 nucleotides flanking the initiation codon contribute to the efficiency of the ribosomal binding site from bacteriophage T7 gene 1

Rolf M.Geisen, Hans-Peter Fatscher and Eckart Fuchs

Molekulare Genetik der Universität, Im Neuenheimer Feld 230, D-6900 Heidelberg, FRG

Received March 5, 1987; Revised and Accepted May 22, 1987

ABSTRACT

The ribosomal binding site (RBS) from gene 1 of bacteriophage T7 was isolated on fragments of differing length and cloned upstream of the mouse dihydrofolate reductase gene to control the translation of its sequence. A 29 base pair sequence containing all elements generally believed to be essential for the RBS's showed extremely low activity. Additional upstream and downstream sequences were required to obtain a several orders of magnitude higher efficiency. By contrast, areas further downstream than +112 nucleotides from the initiator proved to be inhibitory, whereas the presence of an upstream RNaseIII cleavage site showed a strong stimulatory effect. This suggests that tertiary structures are involved in the function of the RBS studied. The efficient RBS's were complexed by ribosomes at much lower concentrations of the mRNA than the weak ones.

INTRODUCTION

The initiation of translation at the ribosomal binding site (RBS) of *E. coli*, is accomplished by ribosomes according to Shine and Dalgarno (1) with the help of initiation factors and fMet-tRNA$_i$(2) . More careful investigations revealed that the sequences of the Shine-Dalgarno interaction and the initiation triplet are so variable that they alone cannot define an RBS (3,4). There must be additional information coded on the mRNA (5,6,7). Several features appear to be important: the spacing between the initiation triplet and the Shine-Dalgarno sequence, certain nucleotides flanking these sequences as well as the formation of secondary or tertiary structures (6).

The lytic bacteriophages can be expected to have some of the most efficient RBS's since they are not restricted by the equilibria of a cell. One of them, T7, belongs to the smallest phages with double stranded DNA as genome. Its DNA is fully sequenced (8,9). One of the most actively translated

© IRL Press Limited, Oxford, England.

genes of T7's early region is the gene 1 coding for the phage specific RNA polymerase. We have attempted to localize the sequences contributing to the activity of this gene's RBS by cloning fragments of different lengths around the start codon into the same position of the plasmid pDS1 (10), specifically in front of the mouse dihydrofolate reductase gene (dhfr) which does not have an RBS active in E.coli cells. The activity of the RBS increased by several orders of magnitude as the cloned fragment length increased from 29 to 205 nucleotides. Furthermore, sequences more than 112 bases downstream of the initiation triplet were inhibitory. A model of the initiation process will be discussed.

MATERIAL AND METHODS

Material

All chemicals were of the highest grade commercially available. Radiochemicals were purchased from Amersham-Buchler. (γ^{32}P) ATP had a specific activity of 3000 Ci/mmol. BamHI linkers were from PL-Biochemicals, nitrocellulose sheets from Schleicher and Schüll (Dassel,FRG).

Bacterial strains and plasmids

E.coli C 600 r^- Δ lac M15 (M. Zabin, UCLA) was the host strain for all plasmids. Cells were grown at 37°C in LB medium (11) supplemented with 0.1 mg/ml ampicillin and 0.025 mg/ml chloramphenicol. A derivative of the pDS1 plasmid (10) (Fig.1) was kindly provided by D.Stueber and H.Bujard, Heidelberg. The stock of T7 was obtained from F.W.Studier (Brookhaven, N.Y.).

DNA manipulation procedures

Cloning procedures were accomplished by standard techniques (11,13). Plasmid DNA was prepared as published (14) and further purified by centrifuging twice to equilibrium in CsCl-ethidium bromide gradients. For determination of plasmid copy number per cell, plasmid DNA was isolated from 5 x 10^9 cells and separated by gel electrophoresis. The stained bands were scanned on a Joyce-Loebl microdensitometer 3CS with electronic integrator MK2. Sequence determination of the cloned DNA fragments was done by the

dideoxy-method of Sanger et al. (11,15) similar to the procedure given in (16), using a 23mer priming oligonucleotide which was a gift from H. Bujard, Heidelberg.

Plasmid constructions

Fragments containing RBS sequences were ligated to appropriate BamHI linkers and cloned into the BamHI site of the pDS1 vector (Fig.1). The KpnI-C fragment of T7 DNA was cut with HindII to create a 377 bp fragment for the construction of pl-377. pl-211 is derived from this fragment by a partial digest with HaeIII resulting in a 211 bp fragment. For the formation of pl-151 the 377 bp fragment was completely cut with HaeIII. pl-63 is a derivative of pl-151 by digestion of its 151 bp fragment with DdeI. For the construction of pl-29 a 29 bp sequence was chemically synthesized (17). It was a gift from Drs.W.Bannwarth and D.Stueber, Basel. All cloned RBS sequences were confirmed by DNA sequence analysis and were in frame with the dhfr gene.

Determination of protein and RNA synthesized

At various stages of growth, lysates of 2×10^8 cells were analysed by gel electrophoresis and staining with Coomassie brillant blue according to published methods (11,18,19). Further characterisation of the in vivo synthesized DHFR-fusion proteins was carried out by the immunological method of "Western blotting" (20). In vitro protein synthesis was according to the procedures described earlier (21,22) with purified plasmid DNA as template. Assay conditions were as given with the exception that 96 mM KCl and 12 mM NH_4Cl were used. For quantitative determination proteins were labelled by the addition of 10-25 µCi ^{3}H-leucine (65 Ci/mmol), separated by gel electrophoresis and visualized by fluorography (23). The protein bands were quantitated by microdensitometric tracing of the X ray films (Kodak XAR-5) and by cutting out bands from the PPO treated and dried gels and counting in a liquid scintillation counter.

RNA synthesis was determined in the coupled in vitro system by the addition of 0.2 µCi ^{14}C-UTP (53 Ci/mol) and with a DNA concentration between 0.015 and 0.016 mg/ml assay mixture. After incubation at 35 C, the assay mixture was precipitated with 0.3 M HClO , filtered on GF/A filters (Whatman)

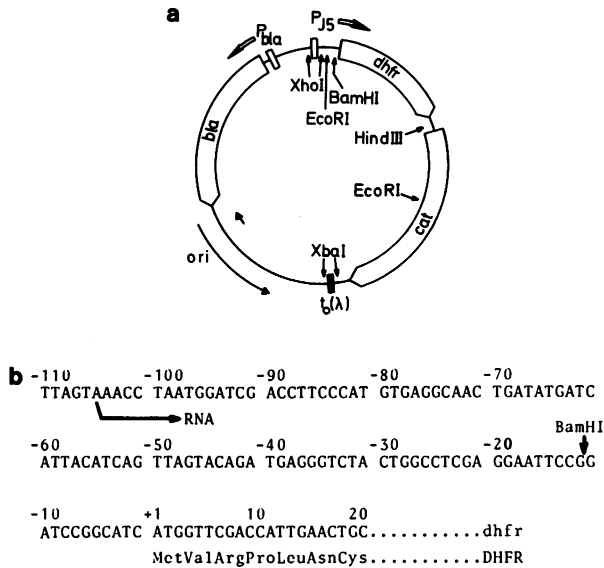

Figure 1. (a) Diagram of the pDS1 derivative used in this study. The plasmid carries the three genes, dihydrofolate reductase of the mouse (dhfr), chloramphenicol acetyl transferase (cat) and β-lactamase (bla), further the promoter J5 (PJ5) from coli phage T5 (12,26) and the promoter Pbla. The origin of replication (ori) is derived from pBR322. The bacteriophage lambda terminator t_0 is cloned in an XbaI site. Locations of relevant restriction sites are indicated. The RBS sequences were cloned into the unique BamHI site upstream of the dhfr initiation triplet. The sequence around this site from the 5'end of the mRNA to the beginning of the dhfr gene is shown in (b).

and counted. For inhibition of RNA synthesis, assay mixtures were divided in two after 10 min of incubation and one sample was supplemented with 2.5 µg lyophylized actinomycin C.

RESULTS

Experimental strategy

Different DNA fragments with the ribosomal binding site (RBS) of the T7 gene 1 were cloned into the plasmid pDS1 (Fig.1). The insertion in front of the mouse dhfr gene was chosen for the following reasons: i, this eukaryotic gene does not have an RBS active in E.coli cells and it did not show downstream sequences with prokaryotic RBS activity. ii, the DHFR protein is tolerated by E.coli cells in high amounts (10). iii, selection is possible

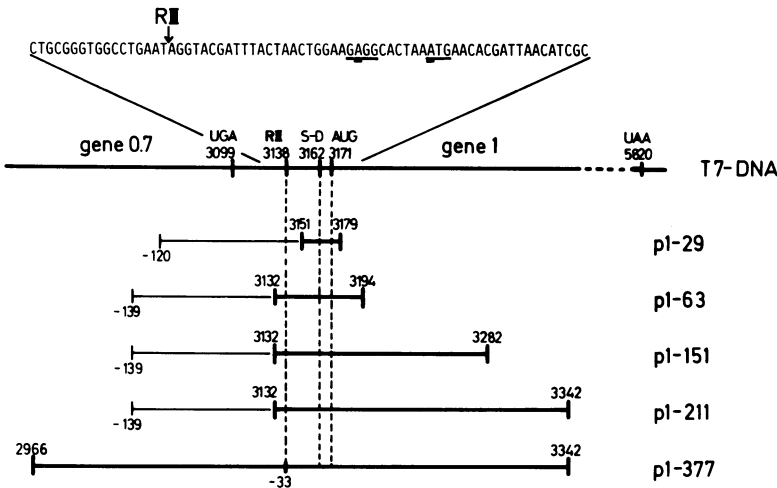

Figure 2. Fragments of the bacteriophage T7 genome cloned into the pDS1 vector. The first line represents part of the T7 DNA around the gene 1 RBS. Part of its sequence is shown in the insert. The Shine-Dalgarno sequence (S-D) and initiator triplet are underlined. Heavy lines represent the fragments cloned into the plasmids indicated on the right. The thin lines represent sequences derived from the vector plasmid. Numbers above the lines are the nucleotide numbers of the T7 genome (8,9), numbers below the lines indicate the 5'ends of the mRNA in relation to the translational start point. RIII represents the RNaseIII cleavage site. In pl-63, pl-151 and pl-211 this site is not active (8,25).

by the use of trimethoprim since the mouse enzyme is more resistant to the drug than the E.coli DHFR (24).

Five plasmids were constructed containing the RBS of gene 1 with different flanking regions (Fig.2). All mRNA's except that of pl-377 contained at their 5'ends a 99 nucleotides plasmid derived sequence. In pl-377 the mRNA was processed by RNaseIII at -33 nucleotides (8,25). Hybrid proteins were synthesized with 3 (in the case of pl-29), 8 (pl-63), 33 (pl-151) and 57 (pl-211 and pl-377) amino acids of the T7 gene 1 at their N-terminal ends.

In vivo synthesis of DHFR fusion proteins

The efficiencies of the different RBS's were determined after cell lysis and polyacrylamide gel electrophoresis in the stained gels. In Fig.3A lanes a and b the protein pattern of cells carrying plasmid pDS1 is shown. While the

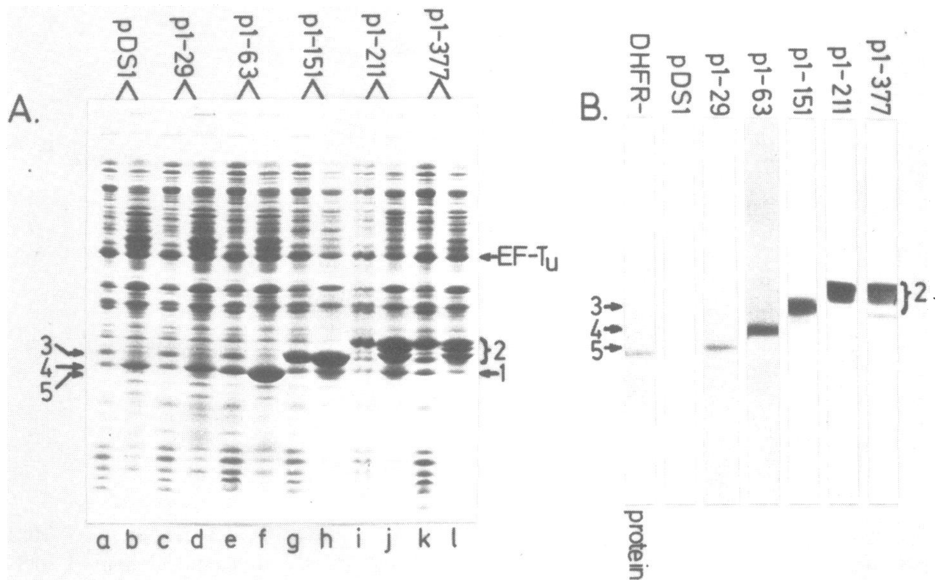

Figure 3. Analysis of cell lysates by polyacrylamide gel electrophoresis. (A) 2 x 10^8 E.coli M15 cells, transformed with the plasmids indicated, were harvested at a concentration of 8 x 10^7/ml in (a, c, e, g, i, and k), and at 5 x 10^9/ml in (b, d, f, h, j and l). After resuspending in sample buffer they were boiled for 3 min and separated on 12.5 % polyacrylamide gels (18).(B) Analysis by "Western blotting" of the plasmid derived proteins as given in "Methods". The lane designated "DHFR-protein" presents as a reference band the separation of purified DHFR protein. The other lanes give the proteins synthesized in cells transformed with the vector (pDS1) or the constructs as indicated. The plasmid derived proteins are indicated by arrows: chloramphenicol acetyl transferase (1), DHFR hybrid proteins from p1-377 and p1-211 (2), from p1-151 (3), p1-63 (4) and p1-29 (5).

chloramphenicol acetyl transferase gene product can be seen as a band, no DHFR hybrid protein is visible. The efficiency of the cloned RBS in cells carrying p1-29 was so low that the pattern of the early log phase cells (lane c) does not show a DHFR hybrid band. Only in the overnight culture (lane d) a faint band can be seen. Plasmid p1-63 produced a hybrid protein which migrated with the chloramphenicol acetyl transferase protein. Little DHFR hybrid protein was found in the early log phase (lane e), whereas quite large amounts of the protein accumulated in the overnight culture (lane f).

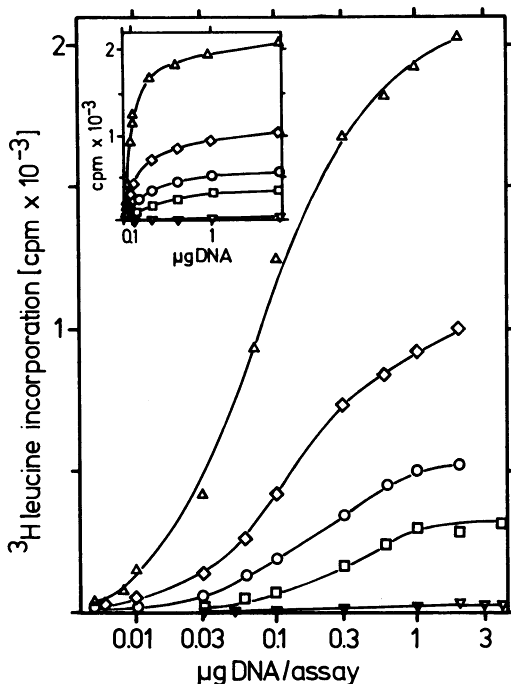

Figure 4. In vitro synthesis of DHFR hybrid proteins in a coupled system with purified plasmid DNA as template as given in "Methods". The dependence of protein synthesis on DNA concentration is shown in a semilogarithmic plot. The insert represents a linear plot of the same results. The templates were from plasmids: pl-377 (△), pl-211 (○), pl-151 (◇), pl-63 (□) and pl-29 (▽).

In contrast to pl-29 and pl-63 , pl-151 already produced large amounts of DHFR hybrid protein in the early log phase and much more in the overnight culture (lanes g and h). This was also true for pl-211 (lanes i and j) and pl-377 (lanes k and l). A number of bands with slightly lower molecular weights were formed by these two plasmids. These proteins were partials of the hybrid protein since they crossreacted with antibodies specific for DHFR protein as could be shown by "Western Blot" analysis (Fig.3B). The copy number of the 5 plasmids and pDS1 in the transformed cells was determined to be about 40. It was about the same at all growth phases.

Table I
IN VITRO SYNTHESIS OF DHFR HYBRID PROTEINS

Plasmid DNA as template	^{3}H leucine incorporated			
	cpm (plateau values)	%	cpm/µg DNA	%
pl-377	2000	100	12000	100
pl-211	520	26	1800	15
pl-151	1010	50	4300	36
pl-63	320	16	410	3
pl-29	<20	<1	<20	<0.2

In vitro synthesis of the DHFR fusion proteins

The RBS efficiencies of the different constructs were quantified in an in vitro protein synthesizing system from E. coli Q13 with purified plasmid DNA as template. The incorporation of ^{3}H leucine into proteins was measured (21,22). The results are summarized in Fig.4 and Table I. A semilogarithmic plot is used to spread the points more evenly over the graph which represents averages of 13 independent experiments. In all 5 cases the synthesis of DHFR fusion proteins approached plateau values at high DNA concentrations which is more obvious in a linear plot (Fig.4, insert). These plateau values can be taken as a measure for the efficiency of the cloned RBS's. They are summarized in Table I in the first three columns. The plasmid pl-29 DNA dependent synthesis of DHFR fusion protein was below the limits of sensitivity. The efficiency of the RBS's increased with increasing fragment length with one exception: pl-211 showed a lower activity than pl-151, i.e. the downstream sequence of +112 to +172 T7 gene 1 nucleotides proved to be inhibitory. The efficiency increased by roughly a factor of 2 at each step from plasmid pl-63 to pl-211 to pl-151 and finally to pl-377.

A further, more sensitive method for evaluation of the RBS efficiencies can be taken from the slopes of the curves given in Fig.4. These values are listed in Table I, right columns. Relative to the most active plasmid DNA of pl-377, pl-29 showed an almost three orders of magnitude lower activity. pl-63 was about 30 times, pl-211 by a factor of 6-7 and pl-151 by almost a factor of 3 less efficient than pl-377.

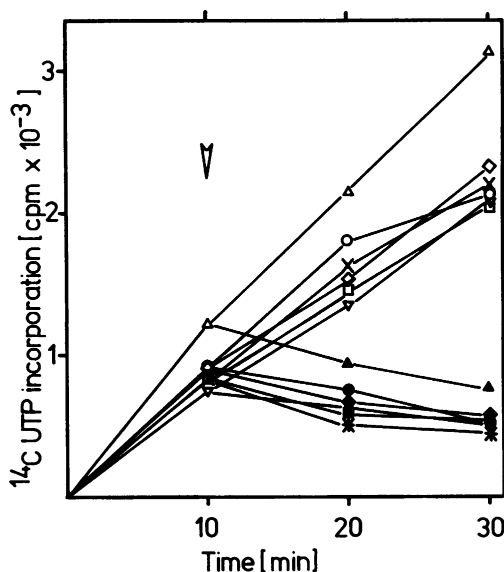

Figure 5. Kinetics of in vitro RNA synthesis in the coupled protein synthesizing system with purified plasmid DNA as template as given in "Methods". After 10 min of incubation, assay mixtures were divided and in one half of the samples RNA synthesis was stopped by the addition of actinomycin C (arrow). The templates were from the following plasmids: pl-377 (△, ▲), pl-211 (○, ●), pl-151 (◇, ◆), pl-63 (□, ■), pl-29 (▽, ▼) and pDS1 (✕, ✱). The filled symbols give the incorporated radioactivity after inhibition of RNA synthesis.

Transcriptional activity of the plasmids

All plasmids contained the same promoters: J5 located 99 base pairs upstream of the cloned fragments with an about 10 times higher activity than the bla promoter (26). To test for a possible transcriptional - translational interdependence as shown for the lac system (27) we measured the plasmid-DNA dependent total RNA synthesis in the in vitro system (Fig.5). The stability of the RNA was determined by the addition of actinomycin C to a fraction of all assay mixtures after 10 min of synthesis and further incubation. All the plasmids showed the same transcriptional activities . The only exception was pl-377 which was about 10-20 % more active than the other plasmids due to the weak E.coli C promoter just upstream of the

RNaseIII site (8,28). The degradation of the RNA and the dependence of RNA synthesis on DNA concentration was the same within error limits for all 5 cases.

DISCUSSION

Three classes of prokaryotic RBS's can be defined: those close to the 5'end of an mRNA, those present intercistronically and the RBS's of RNA phages which have a very sophisticated tertiary RNA structure (29) and almost no turnover rate. The T7 gene 1 RBS belongs to the first class because of an RNase III cleavage site just upstream of the cistron (8,25). The RNaseIII is active in the *in vitro* system as was demonstrated earlier (22). The concentration of this enzyme in *E.coli* cells is high enough to cut RNA produced by high expression multicopy plasmids (30). Constructs with an active RNaseIII site showed the highest activity of the tested RBS's (compare pl-377 with pl-211).

The downstream sequences were very important for the efficiency of the RBS's: with the extension of this region up to +112 nucleotides as in pl-151 the activity increased by a factor of at least 10. There are other reports on the influence of downstream sequences on the initiation of translation: The C to U transition in position +11 of a cloned interferon gene (31), mutations in the region from +35 to +44 in the galK gene (32) and a suppressor mutation at position +64 in the 0.3 gene of T7 which partially restored wild type expression of the mutated initiation codon (33).But all these published effects do not operate as far downstream as that described here. Sequences further downstream between + 112 (pl-151) and +172 nucleotides (pl-211) show inhibitory effects. They possibly influence the translational elongation in the way as postulated by Yamamoto et al. (34). This inhibitory sequence is also present in pl-377 since a fragment with the upstream sequence of pl-377 and the downstream sequence of pl-151 has a higher efficiency than pl-377 (data not shown).

The downstream effects studied here, which are very distant from the sequence on the mRNA initially bound by the ribosomes, must be due to tertiary structures of the mRNA. An analysis of the secondary structure of

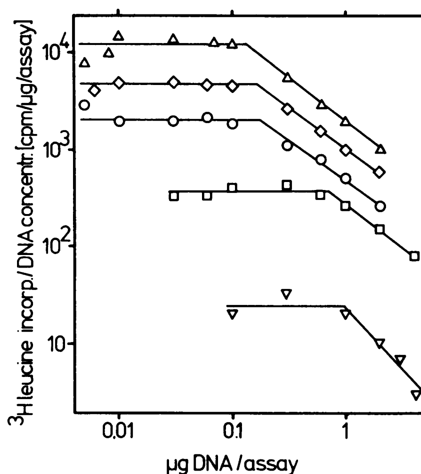

Figure 6. Double logarithmic plot of the in vitro synthesized DHFR hybrid proteins per concentration of DNA template versus µg of template DNA per assay mixture. The templates were from the following plasmids: pl-377 (△), pl-211 (○), pl-151 (◇), pl-63 (□) and pl-29 (▽).

the tested sequences with a computer program (35), however, did not reveal an obvious structural feature which might be responsible for the changes in efficiency. The Shine-Dalgarno sequence of pl-29 possibly located in a 5 base pair double helical structure did not cause the low activity of this plasmid (data not shown).

The in vitro experiments offer a further way to analyse the initiation process. RNA synthesis in the system was linearly dependent on the DNA concentration in the non saturating range and it was about the same in all plasmids tested. Therefore the DNA concentration can be taken as a measure of the amount of mRNA synthesized in the system. We did not find an interdependence of transcription and the efficiency of translation as published for the lacZ gene (27). The efficiencies of the different RBS fragments are presented by the plateau values as shown in Fig.4. These efficiencies become even more pronounced when the initial slopes of the curves are taken, i.e. when the amount of synthesized DHFR hybrid protein per µg of template DNA is calculated (Table I). The effects are more obvious when these

slopes are drawn in a double logarithmic plot as shown in Fig.6. The saturating range for the DNA concentration was reached much earlier by the highly active plasmids as compared to the weaker ones at the same ribosome concentration. Additionally, the onset of significant protein synthesis lies at much lower DNA concentrations for the strong versus the weak RBS's. Since the DNA concentration parallels the mRNA amount in the system, the same conclusions can be drawn for the mRNA dependent protein synthesis. These results suggest that the efficient RBS's were found and complexed by the ribosomes at much lower concentrations of the mRNA than the weak ones. Therefore, the increased strength of the most efficient RBS's is at least partly due to a high rate of complex formation between these parts of the mRNA and the ribosomes with the help of initiation factors. This means that the so called "on rate" may play an important role in the translational initiation process. Sequences far downstream (up to + 112 nucleotides) stimulate the activity of the RBS's by the formation of mRNA tertiary structures as discussed above. For the process of translation these structures have to be unfolded so that the codons can be read. This has to occur during the second part of the initiation process ("off rate"). The ribosomes therefore have to introduce conformational changes in the mRNA. In this way the nucleic acid - nucleic acid interaction of the first stage of recognition would be disrupted. Perhaps these changes are performed by the most efficient RBS's more easily than by the weak ones.

ACKNOWLEDGEMENTS

We thank Dr.H.Bujard for valuable discussions and criticisms and for the gift of the vectors, Dr.V.Bosch for critical reading of the manuscript and Drs.D.Stueber and W.Bannwarth for the gift of synthetic oligonucleotides. G.Kaelberer is acknowledged for her excellent technical assistance and M.Lanzer for his help with the method of "Western Blotting".

REFERENCES

1. Shine,J. and Dalgarno,L. (1975) Nature 254, 34-38.
2. Ganoza,M.C., Marliere,P., Kofoid,E.C. and Louis,B.G. (1985) Proc. Natl. Acad.Sci.USA, 82,4587-4591.

3. Gold,L., Pribnow,D., Schneider,T., Shinedling,S., Singer,B.S. and Stormo,G. (1981) Ann.Rev.Microbiol. 35, 365-403.
4. Stormo,G.D., Schneider,T.D., Gold,L.M. (1982), Nucleic Acids Res. 10, 2971-2996.
5. Stormo,G.D., Schneider,T.D., Gold,L. and Ehrenfeucht,A. (1982) Nucleic Acids Res. 10, 2997-3011.
6. Kozak,M. (1983) Microbiol. Rev. 47, 1-45.
7. Kolaskar,A.S. and Reddy,B.V.B. (1985) Nucleic Acids Res.13 185-194.
8. Dunn,J.J. and Studier,F.W. (1983) J.Mol.Biol. 166, 477-535.
9. Moffatt,B.A., Dunn,J.J. and Studier,F.W. (1984) J.Mol.Biol. 173, 265-269.
10. Stueber,D. and Bujard,H. (1982) EMBO J. 1, 1399-1404.
11. Maniatis,T., Fritsch,E.F., and Sambrook,J. (1982) Molecular Cloning, Cold Spring Harbor Laboratory Press, N.Y.
12. Gentz,R. and Bujard,H. (1985) J.Bact. 164, 70-77.
13. Wu,R. (ed.) (1979) Meth.in Enzymol. Academic Press, Vol. 68.
14. Birnboim,H.C. and Doly,J. (1979) Nucleic Acids Res. 7, 1513-1523.
15. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) Proc.Natl.Acad.Sci.USA 74, 5463-5467.
16. Chen,E.Y. and Seeburg,P.H. (1985) DNA 4, 165-170.
17. Bannwarth, W. and Iaiza, P. (1986) DNA 5, 413-419.
18. Laemmli,U.K. (1970) Nature 227, 680-685.
19. Studier,F.W. (1973) J.Mol.Biol. 79, 237-248.
20. Burnette,W.N. (1981) Analyt.Biochem. 112, 195-203.
21. Fuchs,E. (1976) Eur.J.Biochem. 63, 15-22.
22. Fuchs,E., Hirth,K.P., Henrich,B. and Kälberer,G. (1980) Eur.J.Biochem. 113, 61-66.
23. Laskey,R.A. and Mills,A.D. (1975) Eur.J.Biochem. 56, 335-341.
24. Chang,A.C.Y., Nunberg,J.H., Kaufman,R.J., Erlich,H.A., Schimke, R.T. and Cohen, S.N. (1978) Nature 275, 617-624.
25. Dunn,J.J. and Studier,F.W. (1973) Proc.Natl.Acad.Sci. USA 70,1559-1563.
26. Deuschle, U., Kammerer, W., Gentz, R. and Bujard, H. (1986) EMBO J. 5, 2987-2994.
27. Stanssens, P., Remaut, E. and Fiers, W. (1986) Cell 44, 711-718.
28. Stahl,S.J. and Chamberlin,M.J. (1977) J.Mol.Biol. 112, 577-601.
29. Kastelein,R.A., Berkhout,B., Overbeek,G.P. and van Duin,J. (1983) Gene 23,245-254.
30. Panayotatos, N. and Truong, K. (1985) Nucl.Acids Res 13, 2227-2240.
31 Tessier,L.H., Sondermeyer,P., Faure,T., Dreyer,D., Benavente,A., Villeval,D., Courtney,M. and Lecocq,J.P. (1984) Nucleic Acids Res. 12, 7663-7675.
32. Dreyfus,M. and Kotlarz,D. (1985) J.Mol.Biol. 182,411-417.
33. Dunn,J.J., Buzash-Pollert,E., and Studier,F.W. (1978) Proc.Natl.Acad.-Sci.USA 75, 2741-2745.
34. Yamamoto,T., Suyama,A., Mori,N., Yokota,T. and Wada,A. (1985) FEBS Lett. 181, 377-380.
35. Zuker,M. and Stiegler,P. (1981) Nucleic Acids Res. 9, 133-148.