Volume 12 Number 15 1984 Nucleic Acids Research

Expression of the *rpsO* and *pnp* genes: structural analysis of a DNA fragment carrying their control regions

Claude Portier and Philippe Regnier*

Institut de Biologie Physico-chimique, 13, rue Pierre et Marie Curie, 75005 Paris, France

Received 15 May 1984; Revised and Accepted 11 July 1984

ABSTRACT

Precise physical mapping of the genes rpsO and pnp coding respectively for ribosomal protein S15 and polynucleotide phosphorylase together with regions involved in the regulation of their expression has been obtained by the analysis of in vitro deletion mutants. The results suggest that each gene has its own promotor, but that there is coexpression of rpsO and pnp. The nucleotide sequence of rpsO and of the beginning of pnp is presented and includes the presumed regulatory regions of these genes. Several features of the sequence support the mapping experiments and are discussed in relation to the expression of the ribosomal and pnp genes.

INTRODUCTION

The genes rpsO and pnp have been shown to be adjacent and located near 69 min on the E. coli chromosome (1). The previously available evidence suggested that these genes are transcribed in the same direction and that they are adjacent on the chromosome (1,2). This juxtapositioning raises several questions : in particular are these two genes (one for a ribosomal protein and the other for an enzyme of unclarified functions but presumably involved in polynucleotide biosynthesis or metabolism) coexpressed ? Moreover what are the cell's relative requirements for their products and how is the expression of the two genes regulated to fulfill this need ? What is the type of regulation of these genes ? A translational control mechanism might be expected, at least in the regulation of the expression of the ribosomal protein S15 since a large number of ribosomal proteins are translationally autocontrolled (3-6). If this is the case for S15, is the expression of polynucleotide phosphorylase also affected by this type of control ?

As an approach to answer these questions, a set of in vitro constructed deletion mutants around the gene rpsO and the beginning of pnp has been constructed and examined for their effect on polynucleotide phosphorylase expression and the DNA sequence of the region determined.

© IRL Press Limited, Oxford, England.

MATERIAL AND METHODS

Enzymes and biochemicals

Restriction endonucleases were purchased from New England Biolabs, Boehringer Mannheim or Bethesda Research Laboratories. The T4 DNA ligase was obtained from Boehringer Mannheim. [^{32}P]ATP was purchased from New England Nuclear and acrylamide and bisacrylamide were products of Bio-Rad Company. All other materials were of reagent grade.

Bacterial strains and plasmids

Plasmids pBP 111 (or pBPΔ61) which carry the *pnp* and *rpso* genes (1) were grown in *E. coli* JC 357 (1) *arg, met, leu, his, rpsL, pnp :: Tn5, recA* or CSR 603 (*thr, leu, pro, phr, rec, argE, rpsL*) (1). pBPΔ61 carries a *Hpa I* deletion upstream of the *rpsO* and *pnp* genes and was occasionally used in cloning to eliminate certain restriction sites. Plasmid DNA was prepared and purified as described previously (1).

Construction of deletions in vitro around rpsO

Plasmid DNA was digested partially (or entirely) with the restriction endonucleases indicated and after deproteinisation ligated as described (1). After transformation and selection for antibiotic resistance (1) several colonies were isolated and their plasmid DNA characterised by digestion with appropriate restriction endonucleases followed by agarose gel electrophoresis (1).

Plasmid pBPΔ18 was isolated by ligating a *HindIII-PvuI* digest of pBR 322 DNA with pBPΔ61 DNA cleaved totally by *HindIII* and partially by *PvuI*. Transformants resistant to tetracycline (10 µg/ml) were isolated and one of the plasmids (pBPΔ18) characterized.

Determination of protein density from the autoradiogram

The method of Suissa (7) was used with a slight modification. The piece of the film corresponding to the radioactively labelled band was cut out and the gelatin dissolved in alkali directly in the test tube. The amount of silver grains was estimated by optical density at 500 nm.

Polynucleotide phosphorylase activity

Polynucleotide phosphorylase activity was determined using the exchange assay between UDP and [^{32}P]PO_4 as described (1). Proteins levels in the cell extracts were measured by the Biuret reaction.

Identification of plasmid encoded gene products

Plasmid encoded proteins were labeled with [^{35}S] methionine (specific activity : 1143 Ci/mmole) in the maxicell system of Sancar *et* al. (8) and separated on a 13.5 % sodium dodecyl sulfate (SDS) polyacrylamide gel, as described (1).

DNA sequence analysis

After digestion with the appropriate restriction endonucleases, the DNA

fragments to be sequenced were separated on polyacrylamide gel and eluted by diffusion at 37°C. After labelling, the strands were always separated. The DNA sequence was determined by following the Maxam and Gilbert method (9) with only one modification : after elution of the separated strands from the polyacrylamide gel, DNA was purified from soluble acrylamide by precipitation with hexadecyltrimethylammonium bromide (SIGMA Chemical Company).

RESULTS

I. Construction of in vitro deletion mutants

Our previous results indicated that the genes pnp and rpsO are closely juxtaposed and are probably transcribed in the same direction from rpsO towards pnp (1). If each gene is an individual transcription unit, then the two control regions should be located upstream of pnp and of rpsO and they would be located either side of the rpsO structural gene (Fig. 1). On the other hand, if the two genes are expressed from the same promotor situated in front of rpsO, neither rpsO nor pnp should be expressed when this promotor is deleted.

A set of plasmids carrying in vitro deletions around these presumed control regions has been isolated and characterized (Fig. 1). The deletions start at sites near rpsO and remove all upstream regions of E. coli DNA stopping within pBR 322 vector DNA. pBPΔ100 and pBPΔ900 start at the DraI site before the rpsO structural

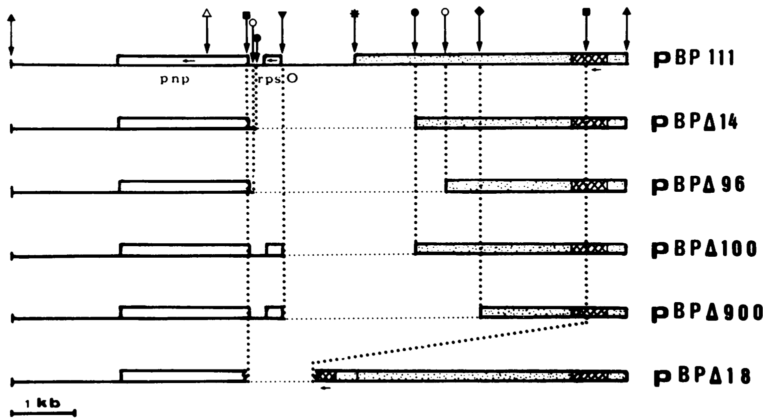

Fig. 1. Structures of the deleted plasmids constructed in this study. The bar with dotted points represents pBR322 and the light line cloned DNA. Genes correspond to empty bars except for β lactamase (cross-hatched bar). The dotted line represents the deleted region. Horizontal arrows indicate direction of transcription and broken bars a small deletions at the junction. Only relevant restriction sites are given.

gene and pBPΔ14 and pBPΔ96 at the Nrul and Mstl sites, respectively, in the intergenic region, between rpsO and pnp (Fig. 1). Attemps to replace the Pvul-HindIII fragment of pBP 111, carrying the rpsO gene, by the Pvul-HindIII fragment of pBR 322 plasmid encoding the beginning of β lactamase give plasmids pBPΔ18. This plasmid carries a deletion of about 40 bp (estimated by polyacrylamide gel electrophoresis) at the Pvul junction leading to the disappearance of the Pvul site and Rsal site (14 bp downstream of the Pvul site) (Fig. 4). Several attempts to isolate a plasmid without a deletion were unsuccessfull.

II. Deletion mapping of rpsO and pnp promotors

The proteins encoded by these mutant plasmids were analyzed on a 13.5 % polyacrylamide gel after labeling with [^{35}S] methionine in a maxicell system (Fig. 2). Lane b shows the presence of polynucleotide phosphorylase at the top of the autoradiogram and of S15 at the bottom, synthesized by the wild type plasmid pBP 111. The other major bands correspond to vector pBR 322 encoded proteins (lane a) : a product of the tet R gene, β lactamase and a small protein of about 5000 daltons.

The plasmids pBPΔ100 and pBPΔ900 cause no detectable expression of S15 (Fig. 2 lanes f and g) suggesting that a promotor located upstream of the Dral site has been deleted. The DNA sequence (see below) shows that the Dral site is situated between the initiation codon of rpsO and its Shine and Dalgarno sequence. The complete lack of expression of rpsO in these deletions is thus probably due to the removal of both transcriptional and translational signals.

Despite the loss of the rpsO promotor, a strong band of protein corresponding to polynucleotide phosphorylase is present showing that pnp gene expression is not dependent on rpsO expression. This suggests the existence of a separate promotor for pnp.

Deletion of the entire rpsO gene, stopping at the Nrul site (pBPΔ14) or Mstl site (pBPΔ96) results in a much lower synthesis of polynucleotide phosphorylase (Fig. 2 lanes d and e). Plasmids pBPΔ14 and pBPΔ96 express this enzyme with no apparent decrease in its molecular weight, suggesting that the translational start of pnp is located downstream from the Mstl site. The difference in polynucleotide phosphorylase expression from the plasmids pBPΔ14 and pBPΔ100 cannot be due to different readthroughs from the pBR 322 part of the plasmids because both deletions are fused to the Nrul site of pBR 322 (Fig. 1). A simple explanation of these observations is that the pBPΔ14 deletion removes the pnp promotor and that the residual expression is due to readthroughs transcription from plasmid promotors.

The plasmid pBPΔ18, carrying a deletion covering the Pvul site causes the synthesis of a very weak protein band near the level of the wild type polynucleotide

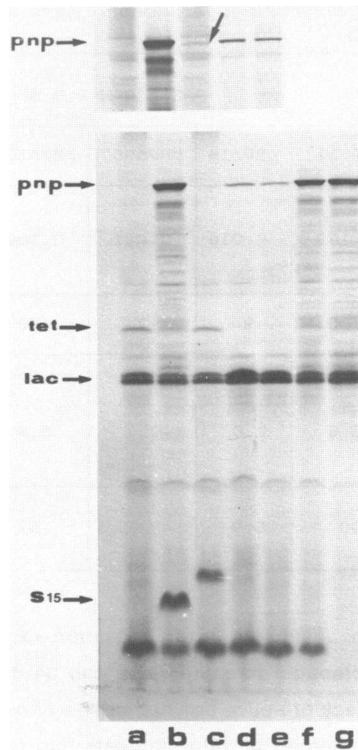

Fig. 2. Autoradiogram of a 13.5% SDS polyacrylamide gel electrophoresis of proteins labelled in maxicells and separated under 80 V during 20 h. The labelled bands, corresponding to the proteins encoded by the plasmid pBP 111 were previously identified as pnp : polynucleotide phosphorylase (1) and S15 : ribosomal protein S15 (2); tet : tetracycline resistance protein ; lac ; β lactamase. About 3 . 10^5 counts/min were loaded in each lane. Exposition of the dried gel was one day (four days for the inset). Lane a : pBR 322 (control) ; lane b : pBP 111 ; lane c : pBPΔ18 ; lane d : pBPΔ14 ; lane e : pBPΔ96 ; lane f : pBPΔ100 ; lane g : pBPΔ900. Inset : the autoradiogram was surexposed to show the synthesis of a shorter polynucleotide chain (arrow, lane c) at the level of polynucleotide phosphorylase.

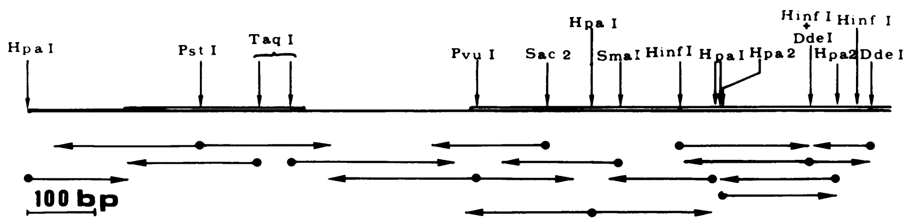

Fig. 3. Sequencing strategy for determining the sequence of the rpsO gene and beginning of the pnp gene. The arrows indicate the direction of sequencing and the length of sequence determined. 5'labelled ends are indicated by dots. Only relevant restriction sites are given.

Table 1. Expression of polynucleotide phosphorylase with different plasmids (see the text).

	PLASMIDS					
	pBP 111	pBPΔ14	pBPΔ96	pBPΔ100	pBPΔ900	pBPΔ18
Amount estimated from autoradiogram	0.173	0.016	0.021	0.108	0.099	-
% of control	100	9	12	63	57	-
Specific activity in the crude extract (units/mg)	16.8	1.0	1.2	8.8	10.2	0
% of control	100	5.9	7.1	52	61	0

phosphorylase band but with a slightly faster migration rate (Fig. 2, lane c and inset). No activity of polynucleotide phosphorylase can be detected in the strain JC 357/pBPΔ18 (Table I). The lack of active polynucleotide phosphorylase together with the feeble expression of a truncated protein suggests that the beginning of the pnp gene has been affected by the deletion and that the translational start for polynucleotide phosphorylase is located in the vicinity of the PvuI site (see discussion). A 11 500 daltons protein clearly seen at the bottom of the lane c (Fig. 2) corresponds probably to the N-terminal part of β-lactamase cloned upstream of the pnp gene (Fig. 1). This truncated protein would have a theoretical molecular weight of 12 500 daltons which agrees well with the molecular weight estimated from the autoradiogram.

III. Quantitation of polynucleotide phosphorylase expression in the mutants

Polynucleotide phosphorylase activity produced by plasmids carrying the deletion mutants was measured in strain JC 357 (pnp^-). As predicted by the maxicell experiments, a much lower activity (about 6% of the activity from the wild type plasmid, pBP 111) was shown by plasmid pBPΔ14 and pBPΔ96 whose the deletion removes rpsO and stops before the pnp structural gene (Table I). The plasmids pBPΔ100 and pBPΔ900 with the shorter deletions stopping at the DraI site near the beginning of rpsO produces a higher specific activity of polynucleotide phosphorylase, but still only about 50% to 60% of the control.

The quantity of polynucleotide phosphorylase labelled in the "maxicell"

experiment was estimated from the intensity of the band on the autoradiogram. To correct for any variation in plasmid copy number, the polynucleotide phosphorylase levels were standardized by comparison to the level of β lactamase expression detected in the gel. The values observed (Table I) are very similar to the values derived from measurements of polynucleotide phosphorylase specific activity. The two methods are thus in agreement and show that removing DNA from in front of rpsO reduces pnp expression by nearly 50%. This suggests that pnp expression is the result of transcription from its own and the rpsO promotors.

Nucleotide sequence of a 1.25 Kb fragment carrying rpsO and the beginning of pnp

The sequencing strategy is given figure 3. The nucleotide sequence is shown figure 4. The DNA sequence was determined on both strands, except for two short regions : the first 32 nucleotides at the beginning of the sequence and 45 nucleotides between the TaqI restriction sites located inside the rpsO gene.

Only two long open reading frames, in the same direction, are present in this sequence. The other two phases contain multiple stop codons. The sequence corresponding to rpsO was previously identified as that around the PstI site (2). A second open reading frame starts at bp 514 and continues outside the presented sequence. This presumably corresponds to the pnp gene.

a) rpsO gene : the aminoacid sequence reported for the ribosomal protein S15 was compared with that predicted by the nucleotide sequence. Only one difference is observed : the DNA sequence predicts an extrahistidine residue between positions 44-45 of the published aminoacid sequence (10). The rpsO gene starts with an ATG at bp 148 coding for an N-terminal methionine not present in the protein S15. Thus, without its N-terminal methionine, S15 contains 88 aminoacids with a molecular weight of 10 137 daltons. The direction of transcription is counterclockwise relative to the E. coli genetic map. The DraI site is located 3 bp in front of the initiator codon, between the Shine-Dalgarno sequence for ribosomal binding, GGAG, and the initiation codon. Two possible transcription promotor sequences can be identified in front of the rpsO gene (Fig. 4). The TACACT sequence between bp 84-91 is identical to the Pribnow box sequences of lambda cin, $lacP_2$, RNA I (of pBR 322 and of clo DF). The corresponding -35 sequence (TTGTTT) has been found in the col E_1 P_2 promotor. Another potential Pribnow box promotor (TAGAAT) located at position 24 to 29 is the same as that found in the promotor $deoP_1$ and $deoP_2$ and the corresponding -35 sequence (TTGCGA) shares 5 out of 6 nucleotides in common with several other identified promotors (11).

The rpsO structural gene terminates with a TAA codon at bp 415. Thirty eight bp downstream from this termination codon, there is a stretch of thymidine residues preceded by a sequence of hyphenated dyad symmetry (Fig. 4). This is typical of a

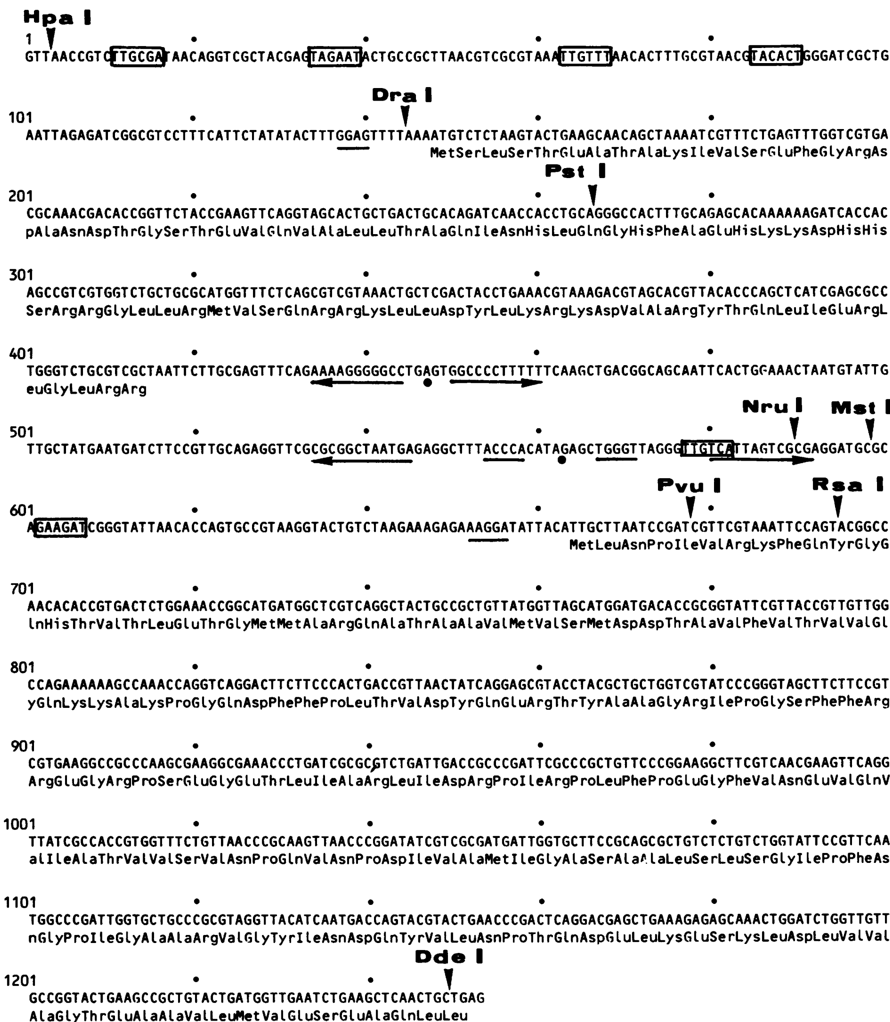

Fig. 4. Nucleotide sequence of rpsO and of the beginning of pnp gene. The nucleotides are numbered from the HpaI site (GCC ↓ AAC). The corresponding aminoacid sequences are written below the DNA sequence. Putative Shine-Dalgarno sequences are underlined. Regions of dyad symmetry or possible base-pairing are indicated by underlining with a dot at the center of symmetry. rpsO gene begins at position 148 and the pnp gene is presumed to start at position 664.

ρ-independent transcription terminator. The stem and loop structure predicted by this sequence is shown in the left of figure 5 and has a ΔG of −22.1 kcalories. mole^{-1}.

b) pnp gene : The results of the deletion mapping suggested that the beginning of the pnp structural gene is located near the PvuI site. However, no

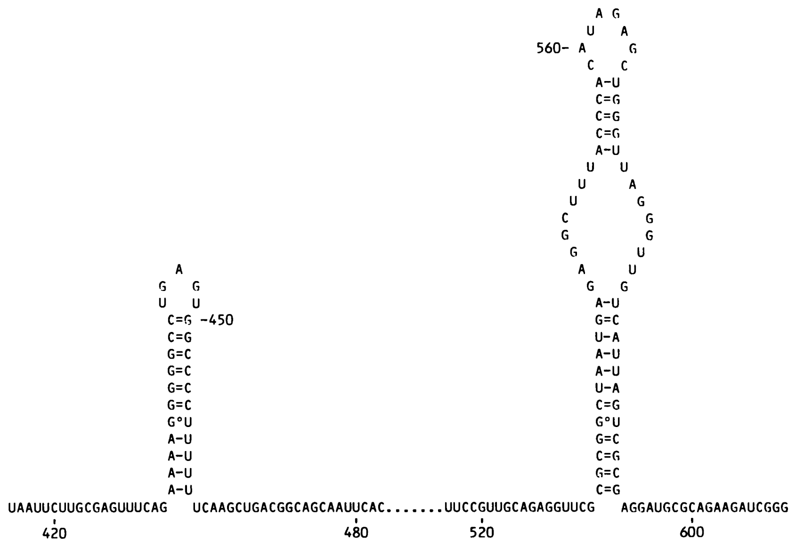

Fig. 5. Possible secondary structures for RNA transcripts within the rpsO-pnp intercistronic region predicted by the DNA sequencing data presented here. The structure on the left, is a possible terminator/attenuator site and the structure on right a possible RNase III processing site.

standard initiation codon is found in the near vicinity of this PvuI site; the nearest ATG's are 76 bp before or 47 bp after the PvuI site. The N-terminal protein sequence of polynucleotide phosphorylase was determined as : met-leu- ? -pro-phe (12). This does not fit with any of the ATG (or GTG) codons seen in this region but does seem to correspond to the use of a UUG located 10 bp upstream of this PvuI site. Initiation of this site would give an N-terminal sequence : met-leu-asn-pro-ile. A satisfactory Shine and Dalgarno sequence (AAGGA) precedes this initiation codon.

A putative promotor with a -10 and -35 sequences of GAAGAT and TTGTCA identical to these of the λ434 PRM promotor (11) is located at position 602 and 577. The NruI and MstI sites are found between these two sequences. As these were the two sites used in the construction of the pBPΔ14 and pBPΔ96 deletions, the considerable reduction in pnp expression from these plasmids is consistent with their location in the promotor region.

In addition, the sequences in this region (534-558) are highly complementary to those at position 567-592 and a stable ($\Delta G = -25$ Kcal. mole^{-1}) hairpin can be drawn in the mRNA (shown in the right of Fig. 5).

DISCUSSION

The results presented in this paper confirm the close proximity of the rpsO and pnp genes. The location of rpsO is unequivocally identified by comparison with the protein sequence. The comparison of the DNA sequence and N-terminal protein sequence of polynucleotide phosphorylase strongly suggests that pnp joins the small collection of proteins with the unusual initiation codon UUG (13). An in-phase ATG codon exists 76 bp before the PvuI site but is not preceeded by a recognizable Shine and Dalgarno sequence in the classic position upstream from this putative initiator codon. Moreover, the DNA sequence between this ATG and the PvuI site contains 26 codons, five of them (four AGA and one AGG) being very rarely or never used in strongly expressed proteins (14). The other in phase ATG, at position 727, seems to be too far from the PvuI site to be affected by the small in vivo deletion around this site isolated on pBPΔ18.

The PvuI site is located 7 bp downstream of the pnp initiator codon. Comparison of the sequences of the β-lactamase and pnp genes, at their PvuI sites, shows that the attempt to join the two genes at this site should result in an in- phase fusion giving a hybrid protein. Other examples of similar large hybrid protein carrying a signal peptide are known to be lethal for the cell (15, 16). This probably explains why the only plasmids isolated were those which carry mutations (deletions) eliminating this hybrid protein like pBPΔ18. The feeble synthesis of a truncated protein from this plasmid could be due to a restart from one of the other in-phase ATG codons after the PvuI site.

The presence of promotor sequences just upstream of each of the two genes is suggested both by deletion analysis and the DNA sequence. The deletions carried by the plasmids pBPΔ100 and pBPΔ900 stop just before the initiation codon of rpsO and eliminate entirely rpsO expression. As mentionned before, this complete lack of expression presumably results from the deletion of the rpsO promotor. If readthrough from any pBR 322 promotor exists, it is not sufficient to permit the synthesis of a protein missing a Shine and Dalgarno sequence.

These plasmids however still synthesize appreciable amounts (60% of wild type levels) of polynucleotide phosphorylase. That this level of expression is not due to readthrough from pBR 322 promotors is shown by the fact that deletions of plasmids pBPΔ14 and pBPΔ96 which delete from the NruI (or MstI) sites in front of pnp to the NruI (or MstI) sites in pBR 322, synthesize much less polynucleotide phosphorylase (6% of the wild type level). These deletions leave the initiation codon and the Shine and Dalgarno sequence of the pnp gene intact and fuse, in the case of pBPΔ14, to the same NruI site in the vector as pBPΔ100. Thus it is reasonable to think that the 6% expression from pBPΔ14 could be due to readthrough from pBR 322 promotors but that

the 60% expression from pBPΔ100 is due to the presence of a pnp promotor present in pBPΔ100 but removed in pBPΔ14. Alternatively the weak expression from pBPΔ14 and pBPΔ96 could be explained by the fact that the sites of fusion are located in the region of the putative pnp promotor identified on the DNA sequence (Fig. 4). Fusions in this region might be expected to give a "down" mutation phenotype. In either case, these observations strongly imply the existence of a functional pnp promotor.

The differences in pnp expression between the wild type plasmid (pBP111) and pBPΔ100 and pBPΔ900 could be interpreted as evidence that the rpsO promotor (deleted in pBPΔ100 and pBPΔ900) contributes to pnp expression in plasmid pBP111. Another explanation, that the extraexpression seen with pBP111 is due to transcription from pBR322 promotors in the HindIII-NruI fragment seems unlikely : no promotors have been identified so far in this region transcribing from NruI towards HindIII (i.e. antisense to the tetracycline resistance cistron) (17, 18).

If this coexpression is verified, it will be interesting to see what effect rpsO has on the regulation of the expression of pnp. Other ribosomal proteins studied have been shown to be translationally autocontrolled (3-6). The observations of Takata (19, 20), showing that S15 is not subject to gene dosage, would suggest some autoregulation. The model for translational autocontrol suggests there exist some homology between the rRNA binding site for the ribosomal protein and its mRNA. However no obvious sequences corresponding to these of the rRNA protected by S15 from kethoxal (21) can be observed in the DNA sequence near rpsO. This possibility of translational control is currently under investigation.

In addition two possible transcriptional regulatory mechanisms are suggested by the sequence. There is a potential terminator structure at the end of rpsO which would support the idea that rpsO and pnp are independently expressed. However there are several examples where terminators are only partially effective and act as transcriptional attenuators, e.g. the operons coding for the sigma and beta subunits of RNA polymerase (22, 23).

Another interesting potential secondary structure is predicted by the complementary sequences located after the terminator. This structure seems characteristic of sites recognized by RNase III. Whether this enzyme has a role in polynucleotide phosphorylase expression has yet to be determined.

In conclusion, we would like to point out that the features of the rpsO-pnp region described here show remarkable similarity to that of the RNA polymerase β subunit operon (22). Both start with gene(s) for ribosomal protein(s) successively followed by a terminator structure which probably functions as an attenuator, a possible RNase III processing site and a promotor permitting potential independent expression of a gene involved in RNA metabolism.

ACKNOWLEDGEMENTS

We wish to thank Dr J. Plumbridge for careful reading and useful criticisms of the manuscript and M. Duponchelle for editorial assistance.

This work was supported by grants to Professor M. Grunberg-Manago from the "Centre National de la Recherche Scientifique" (Groupe de Recherche n°18, ATP "Microbiologie 1982"), from the "Ministère de la Recherche et de l'Industrie" (Convention "Action Biologie Moléculaire" de la Mission des Biotechnologies, n° 82 V 1289), from the "Institut National de la Santé et de la Recherche Médicale" (contrat libre n°823.008, et contrat de Recherche Externe n° 831.013), from the "Fondation pour la Recherche Médicale", from the Society E.I. Du Pont de Nemours and Company, and from the University Paris VII.

*Philippe Regnier is a fellow of University Paris VII.

REFERENCES

1. Portier, C., Migot, C. and Grunberg-Manago, M. (1981) Mol. Gen. Genet. 183, 298-305.
2. Portier, C. (1982) Gene 18, 261-266
3. Nomura, M., Yates, J.L., Dean, D. and Post, L. (1980) Proc. Natl. Acad. Sci. USA 77, 7084-7088.
4. Fuduka, R. (1980) Mol. Gen. Genet. 178, 483-486
5. Brot, N., Caldwell, P. and Weissbach, H. (1980) Proc. Natl. Acad. Sci. USA 77, 2592-2595
6. Yates, J.L., Arfsten, A.E. and Nomura, M. (1980) Proc. Natl. Acad. Sci. USA 77, 1837-1841.
7. Suissa, M. (1983) Anal. Biochem. 133, 511-514.
8. Sancar, A., Hack, A.M. and Rupp, W.D. (1979) J. Bacteriol. 137, 692-693.
9. Maxam, A.M. and Gilbert, W. (1980) Methods in Enzymology, Moldane, K. and Grossman, L. Eds., Vol. 65, pp. 499-560, Academic Press, New-York.
10. Morinaga, T., Funatsu, G., Funatsu, M. and Wittmann, H.G. (1976) FEBS Lett. 64, 307-309.
11. Hawley, D.K. and Mc Clure, W.R. (1983) Nucleic Acids Res. 11, 2237-2255.
12. Portier, C. (1975) Biochimie 57, 545-550
13. Kozak, M. (1983) Microbiol.Rev. 47, 1-45
14. Grossjean, H. and Fiers, W. (1982) Gene 18, 199-209.
15. Ito, K., Bassford, P.J.J. and Beckwith, J. (1981) Cell. 24, 707-717.
16. Bassford, P.J., Silhavy, T.J., Becwith, J.R. (1979) J. Bacteriol. 139, 19-31.
17. Stüber, P. and Bujard, H. (1981) Proc. Natl. Acad. Sci. USA 78, 167-171.
18. Backman, K. and Boyer, H.W. (1983) Gene 26, 197-203.
19. Takata, R., Aoyagi, M. and Mukai, T. (1982) Mol. Gen. Genet. 188, 334-337.
20. Takata, R. (1978) Mol. Gen. Genet. 160, 151-155.
21. Müller, R., Garrett, R.A. and Noller, H.F. (1979) J. Biol. Chem. 254, 3873-3878.
22. Barry, G., Squires, C. and Squires, C. (1980) Proc. Natl. Acad. Sci. USA 77, 3331-3335.
23. Burton, Z.F., Gross, C.A., Watanabe, K.K. and Burgess, R.R. (1983) Cell. 32, 335-349.