A bacterial cell that synthesizes a protein containing the antigenic determinants of rat prolactin

Christopher R.Erwin*, Richard A.Maurer[†] and John E.Donelson*

Departments of Biochemistry* and Physiology and Biophysics[†], University of Iowa, Iowa City, IA 52242, USA

Received 13 March 1980

ABSTRACT

Bacterial minicells containing three different recombinant plasmids with rat prolactin cDNA sequences inserted at the Pst I site of pBR322 via the poly(dG):poly(dC) joining technique were examined for the expression of rat prolactin antigenic determinants. The three prolactin coding sequences were in the same orientation as the coding sequence of the ampicillin-resistance gene of pBR322. The presence of each of the three recombinant plasmids induced some prolactin synthesis by the bacteria as measured by immunoprecipitation with anti-prolactin antisera. About 10% of the protein synthesized from one of the plasmids, prl 3, precipitated with the antisera. These prolactin antigenic determinants were part of a larger fused protein.

INTRODUCTION

One of the goals of many recombinant DNA experiments is to achieve the functional expression of a foreign gene that has been cloned in a host cell. In particular much effort has been directed toward the expression of eukaryotic gene products in bacteria. At least some chromosomal genes of lower eukaryotes such as yeast (1-7) can be cloned and expressed directly in bacterial plasmids, presumably because the nucleotide sequences of the control regions surrounding the foreign gene can be recognized by the bacterial translation and/or transcription machinery. However, there have been no reports of the synthesis in bacteria of the protein product of a cloned chromosomal gene from a mammalian source. This is probably due to the presence of intervening sequences in many mammalian genes (8) that cannot be correctly removed by post-transcriptional processing by the bacterial cell.

Several different plasmids have been constructed in vitro which contain specific promoters located so that a particular coding sequence (for example a cDNA sequence) can be inserted at a restriction site down-stream from the transcription initiation point (9-11). However, the use of such plasmids is sometimes complicated by the additional steps required to insert the cDNA at the appropriate restriction site (12).

Recently it has become convenient and useful to clone cDNA sequences at the Pst I site of pBR322 (13) via the poly(dG):poly(dC) tailing technique. Since the Pst I site is within the ampicillin-resistance gene of pBR322, it is usually, but not always (14), inactivated by the foreign DNA insertion which provides a selection procedure for bacterial transformants that have foreign DNA inserts. And since the Pst I site is regenerated at the poly(dG):(dC) boundaries (15), the cloned DNA can be excised from the recombinant plasmid by Pst I digestion. Because of the ease in cloning DNA fragments at the Pst I site of pBR322 and in recovering the fragments from the recombinant plasmid, it would be useful if the expression of fragments cloned at this site could also be shown. Analysis of the gene expression after cloning at the Pst I site of pBR322 have produced two different kinds of results. Both the cloned rat proinsulin cDNA (14) and rat growth hormone cDNA (16) are expressed as part of a fusion product which also contains a segment of the β -lactamase polypeptide. The results of these studies suggest that proinsulin and growth hormone antigenic determinants were detected because of transcription and translation read through across the boundary between the β -lactamase and the cDNA coding sequences. In contrast Chang *et al.* (17) cloned mouse dihydrofolate reductase cDNA at the same pBR322 site by the same technique and demonstrated both enzymatic and immunologic activity of a protein that was the molecular weight expected for mouse dihydrofolate reductase. DNA sequence analysis of one recombinant revealed that the coding sequences for β -lactamase and dihydrofolate reductase were in different translation reading phases and that the methionine initiator codon for dihydrofolate reductase was immediately adjacent to the ten long dG:dC boundary. They hypothesize that the ten adjacent dG residues in the boundary hybridize to the Shine and Dalgarno sequence of CCUCC at the 3' end of the 16S ribosomal RNA (18) and serve to bind the ribosomes in position to initiate translation of the dihydrofolate reductase coding sequence directly.

In this report we provide another example of the expression of a mammalian cDNA that is cloned at the Pst I site of pBR322 via the poly(dG):poly(dC) technique. Three different recombinant plasmids containing the rat preprolactin coding sequence in the same orientation as the β -lactamase gene were examined. All were found to synthesize prolactin antigenic determinants but one synthesized substantially more than the other two. Experiments with bacterial minicells revealed that this particular amino acid sequence was part of a polypeptide that was larger than preprolactin. Hence our results are similar to those described for the

expression of the proinsulin and growth hormone coding sequences (14,16).

MATERIALS AND METHODS

Plasmid DNA Isolations. One liter overnight cultures of *E. coli* strain χ 1776 or HB101 containing the different plasmids were grown in LB medium supplemented with 20 μ g tetracycline/ml (and with 100 μ g diaminopimelic acid/ml and 50 μ g nalidixic acid/ml in the case of χ 1776). The plasmid DNAs were isolated by Brij 58 detergent lysis of the cells and CsCl-ethidium bromide gradient centrifugation as described previously (19).

Minicell Isolation and Incubation. The *E. coli* minicell producing strain χ 1411 was obtained from Roy Curtiss. Transformation of this strain with the plasmids pBR322, prl 1, prl 2 and prl 3 was performed as described previously (20). After transformation with selection for tetracycline resistance, small scale plasmid DNA isolations were performed (21) and the plasmids digested with *Eco* RI and/or *Hinc* II followed by agarose gel electrophoresis to verify that the plasmids had not undergone deletions or sequence rearrangements during the transformation. No evidence for such deletions or rearrangements was detected. The minicells were purified by a slight modification of the procedure of Kool *et al.* (22) as described previously (20). They were incubated with [35 S]methionine (Amersham, 500 Ci/mmol) for 90 minutes at 37°C with aeration and crude minicell lysates prepared as described previously (20,21).

Immunoprecipitations and Sodium Dodecyl Sulfate Gel Electrophoresis. Aliquots of minicell extracts were combined with 2 μ g of purified rat prolactin and an excess of anti-prolactin. The purification of rat prolactin and the preparation of anti-prolactin have been described previously (23). After incubation overnight at 4°C the immunoprecipitates were washed as described previously (24). Immunoprecipitates were dissolved in NCS for scintillation counting or in 1% sodium dodecyl sulfate, 4 M urea, 1% β -mercaptoethanol, 0.01 M Tris, pH 7.4 for analysis on sodium dodecyl sulfate, 10% polyacrylamide gels using the buffers of Laemmli (25). The polyacrylamide gel was prepared for fluorography as described by Bonner and Laskey (26).

Containment. Permission was obtained from the University of Iowa Bio-hazards Committee to conduct the experiments described here under P2 + EK1 containment (NIH Guidelines for Recombinant DNA Research, December 1978). Permission was granted to lower the containment from P2 + EK2 to P2 + EK1 on the basis that DNA sequence determinations and restriction enzyme analy-

sis showed unambiguously that the plasmids prl 1, prl 2 and prl 3 contained rat prolactin (cDNA) coding sequences.

RESULTS AND DISCUSSION

We have previously described the construction and identification of about 150 rat prolactin cDNA clones (27). The cDNA molecules were inserted at the Pst I site of pBR322 (13) by the poly(dG):poly(dC) joining technique (15). The sequence of the cDNA insert in one of the plasmids, prl 1, was determined and found to contain all of the preprolactin coding sequence except for the codons for the first 13 amino acids of the signal peptide at the N-terminus (27). Sequence determinations of the cDNA insert of another plasmid, prl 2, revealed that it contains the complete coding sequence for preprolactin including 17 nucleotides preceding the methionine initiator codon (Figure 1). In addition, restriction enzyme analyses were conducted on seven other prolactin cDNA inserts to determine which portion of the preprolactin coding sequences had been cloned. The five largest recombinant plasmids were then examined for the orientation of the prolactin coding sequence relative to the β -lactamase coding sequence. This was accomplished by digestion of the recombinant plasmid with Hinc II which cleaves twice in the pBR322 sequence (29) and once in the prolactin coding sequence (see

Figure 1. Restriction enzyme maps of the prolactin cDNA inserts of prl 1, prl 2, and prl 3 showing the sizes and relative amounts of prolactin coding sequence that each contains. For simplicity only the Hinc II and the Alu I sites are shown. The numbers refer to sizes in base pairs of the restriction fragments. When the number is not preceded by the symbol ~, the number is known from DNA sequence determination; when the number is preceded by ~, the fragment size is estimated from polyacrylamide gel analysis. The double lines at the ends of the three cDNAs indicate the poly(dG):poly(dC) boundaries. Each of the six boundaries is known to be between 20 and 32 base pairs from acrylamide gel analysis but the exact number is not known for any of the boundaries.

Figure 1). In circular pBR322 the distances between the two Hinc II and the one Pst I site (at which the cDNA is cloned) are 296 base pairs and 2959 base pairs respectively. In the individual cDNA inserts the distance from the single Hinc II site to the Pst I ends was known from the sequence determinations and/or restriction enzyme analyses (see Figure 1). Therefore it was possible to tell the orientation of the cDNA sequence from a Hinc II digest of the recombinant plasmid. For example, Hinc II digestion of prl 1 yielded fragments of 920 base pairs (296 of pBR322 and 624 of cDNA) and 3215 base pairs (2959 of pBR322 and 256 of cDNA) as well as the 1106 base pair fragment that is between the two Hinc II sites in circular pBR322 and does not contain the Pst I site. This established the prolactin and β -lactamase coding sequence to be in the same orientation. By identical analyses the prolactin coding sequences in prl 2 and prl 3 were also found to be in the same orientation. The other two cDNA inserts investigated were in the opposite orientation. Figure 1 shows a comparison of the cDNA inserts in prl 1, prl 2 and prl 3.

The original plasmid pBR322 and the three recombinant plasmids prl 1, prl 2 and prl 3 were each introduced into the *E. coli* minicell strain, χ 1411, and the plasmids re-isolated from individual transformants to verify by restriction enzyme digestion that the plasmid had not undergone sequence rearrangements in χ 1411 (See Methods and Materials). Minicells containing each of the four plasmids were prepared by three successive sucrose gradient centrifugations as described previously (20) and incubated with [35 S]methionine in minimal salts medium containing all of the unlabeled amino acids except methionine. Crude lysates were prepared and two aliquots removed from each of the four lysates to determine (i) the total amount of radioactivity incorporated into protein precipitable with trichloroacetic acid and (ii) the radioactivity precipitable with rabbit antisera rat prolactin. Table I shows an example of one such experiment.

In minicells containing prl 3, 9.7% of the protein whose synthesis is directed by the plasmid also immunoprecipitates with the prolactin antisera. This represents a 247 fold increase above the amount of immunoprecipitable protein found in minicells containing pBR322. Minicells containing prl 1 and prl 2 also appear to synthesize a small amount of prolactin (14 fold and 9 fold above the pBR322 background respectively). Duplicate experiments to the one shown in Table I yielded similar results, i.e., plasmid prl 1 directs the synthesis of slightly more prolactin than does prl 2 while substantially more prolactin is synthesized from prl 3 than either prl 1 or prl 2.

TABLE I

A comparison of the relative amounts of radioactive protein synthesized in different plasmid-containing *E. coli* minicells that immunoprecipitate with rabbit anti-rat prolactin antisera. The precipitations with trichloroacetic acid (TCA) to determine the total precipitable radioactivity and the immunoprecipitation were each performed on 5% of the minicell lysate. For the calculations the amount of TCA-precipitable radioactivity was normalized to 35,550 cpm, the amount present in 5% of the lysate of prl 1-containing minicells.

Plasmid in the minicells	Number of cpm in the immunoprecipitate	Percent plasmid-directed protein that immunoprecipitates	Fold increase above pBR322 background
pBR322	14	0.04%	1
prl 1	196	0.6 %	14
prl 2	132	0.4 %	9
prl 3	3462	9.7 %	247

Figure 2 shows an autoradiogram of the [^{35}S]polypeptides synthesized from the plasmids listed in Table I. The migration of the most prominent polypeptide of pBR322 (lane B) corresponds to the migration of a protein with a size of about 300 amino acids which is consistent with the size of the ampicillin-resistance gene product of pBR322, a 286 amino acid β -lactamase (28). This polypeptide is missing, or at least greatly reduced, in minicells containing the recombinant plasmids, as expected since the cDNA inserts are cloned at the Pst I site located in this gene. In minicells containing these prl plasmids the β -lactamase is replaced with a large number of polypeptides that migrate as if they are about the same size as prolactin (198 amino acids) but do not precipitate with prolactin antisera (lanes C through e). Since the Pst I cleavage site occurs in the codon for amino acid number 182 of the β -lactamase gene, these polypeptides are probably incomplete β -lactamase chains whose synthesis was terminated shortly after the ribosome entered the poly(dG) boundary or passed through the boundary into the cDNA. In the acrylamide gel shown in Figure 2 polypeptides smaller than about 150 amino acids were expected to migrate with the ion front. This was confirmed by the fact that the hemoglobin chains used as standard molecular weight markers in the gel did run with the ion front (not shown). Therefore, polypeptide products of pBR322 or of the prl plasmids smaller than about 150 amino acids could not be detected in this gel. A final

Figure 2. A fluorogram of a sodium dodecyl sulfate-acrylamide gel of polypeptides labeled during incubation of plasmid-containing minicells with [^{35}S]methionine. Minicell lysates were either applied directly to the gel (capital letters) or precipitated with prolactin antisera and the immunoprecipitate applied to the gel (small letters). The minicells contained the plasmids pBR322 (lanes B and b), prl 1 (lanes C and c), prl 3 (lanes D and d) and prl 2 (lanes E and e). Lanes A and F are [^{14}C]prolactin standards. The arrow indicates a polypeptide synthesized from prl 3 (lane D) which precipitates with prolactin antisera (lane d) and is not present among the polypeptides synthesized from pBR322 (lanes B and b). The numbers on the left indicate the approximate size of the polypeptides in amino acid residues as determined by the migration of standard length polypeptides.

general point about the gel shown in Figure 2 is that some of the larger and the fainter polypeptides observed in lanes C and D are probably due to bacterial cell contamination of the minicell preparations. The minicell preparation shown in lane D appeared to have considerably greater bacterial cell contamination. Although the minicell isolation involved three successive sucrose gradient centrifugations, there was still a variable amount of bacterial contamination in the minicells used for the radiolabeling.

Figure 2 shows that minicells with prl 3 (lane D) contain a polypeptide of about 350 amino acids that is not observed in minicells with the other plasmids. This polypeptide precipitates with prolactin antisera (lane d).

The most likely explanation for the occurrence of this polypeptide is that it is a fusion product of the first 182 amino acids of β -lactamase and the prolactin sequence. Unlike prl 1 and prl 2 DNA, no DNA sequence determinations have been conducted on prl 3 so the exact codon at which the cDNA sequence begins is unknown. The restriction enzyme analyses of prl 3 indicate that its prolactin coding sequence begins within the codons for the last few amino acids of the signal peptide and extends through the termination codon (see Figure 1) for a total codon number of about 200. This suggests that the fused product should be about 382 amino acids ($182 + 200$) plus the amino acids coded for by the poly(dG) boundary. Since these boundaries are between 20 and 30 nucleotides (again based on restriction enzyme analysis), the fused product should be about 390 amino acids. The results shown in Figure 2 indicate the fused product is about 350 amino acids. The reason for the difference in these two values is not clear. However the conclusion of Table I, i.e., that prolactin antigenic determinants are synthesized from prl 3, is supported by the acrylamide gel analysis. It should be noted that the amount of the fused protein from prl 3 (lane D) is not as great as the amount of β -lactamase from pBR322 (lane B). This could indicate either (i) that not as much of the fused protein is synthesized or (ii) that the fused protein is degraded more quickly in the minicells than β -lactamase. The amount of additional radioactivity in the area of migration for the 200 amino acid polypeptides is the same for prl 3 as for prl 1 and prl 2 even though much more prolactin antigen is made from prl 3 than from prl 1 and prl 2. This radioactivity could be incomplete β -lactamase chains or the partially degraded fused protein. Since it is in minicells containing prl 3 as well as in those containing prl 1 and prl 2, the most likely explanation is that it represents incomplete β -lactamase chains and that in minicells containing prl 3, translation frequently stops near the codon for amino acid 182 of β -lactamase even though the coding sequence for prolactin is in the same translation reading frame on the other side of the poly(dG) boundary.

The immunoprecipitations shown in Table I indicate that prl 1 and prl 2 also direct the synthesis of small amounts of prolactin antigen. This was not detected in the fluorogram of Figure 2 and probably reflects the fact that at least seventeen times less immunoprecipitated radioactivity was applied to the gel for prl 1 and prl 2 than for prl 3.

In conclusion, the prolactin antigenic determinants detected in prl 3-containing minicells are part of a polypeptide that is larger than preprolac-

tin and most likely are part of a fused translation product containing a β -lactamase sequence in the N-terminal portion and prolactin sequence in the C-terminal portion. These findings demonstrate that minicell-producing bacteria are very useful for analysis of such fused proteins. The fused β -lactamase-rat prolactin product is similar to the β -lactamase-proinsulin (14) and β -lactamase-rat growth hormone (16) hybrid polypeptides. Unfortunately the assays available to detect the biological activity of prolactin (30) require relatively large amounts of material and were not feasible to try using the minicell system. Nevertheless, the experiments described here and those reported from other laboratories (14,16) suggest that any cDNA coding sequence, when inserted at the Pst I site of pBR322, can be expressed as part of a large fused translation product if it is in the same translation reading frame as the β -lactamase gene.

ACKNOWLEDGMENTS

This research was supported by NSF grant PCM 76-13451 to J.E.D. and NIH grant AM 21803 to R.A.M. C.R.E. was supported by NIH predoctoral genetics training grant GM 07091.

REFERENCES

1. Struhl, K., Cameron, J.R. and Davis, R.W. (1976) *Proc. Natl. Acad. Sci. USA* **73**, 1471-1475.
2. Carbon, J., Ratzkin, B., Clarke, L. and Richardson, D. (1977) in Miami Winter Symposium (eds. W.A. Scott and R. Werner) **13**, 59-72. Academic Press, New York.
3. Ratzkin, B. and Carbon, J. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 487-491.
4. Struhl, K. and Davis, R.W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5255-5259.
5. Clarke, L. and Carbon, J. (1978) *J. Mol. Biol.* **120**, 517-532.
6. Dickson, R.C. and Markin, J.S. (1978) *Cell* **15**, 123-130.
7. Citron, B.A., Feiss, M. and Donelson, J.E. (1979) *Gene* **6**, 251-264.
8. Abelson, J. (1979) *Ann. Rev. Biochem.* **48**, 1035-1069.
9. Fraser, T.H. and Bruce, B.J. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5936-5940.
10. Roberts, T.M., Kacich, R. and Ptashne, M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 760-764.
11. Bernard, H.-U., Remaut, E., Hershfield, M.V., Das, H.K., Helinski, D.R., Yanofsky, C. and Franklin, N. (1979) *Gene* **5**, 59-76.
12. Bahl, C.P., Marians, K.J. and Wu, R. (1976) *Gene* **1**, 81-92.
13. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., and Boyer, H.W. (1977) *Gene* **2**, 95-113.
14. Villa-Komaroff, L., Efstratiadis, A., Broome, S., Lomedico, P., Tizard, R., Naber, S.P., Chick, W.L. and Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3727-3731.
15. Chang, A.C.Y. and Cohen, S.N. (1978) *J. Bact.* **134**, 1141-1156.

16. Seeburg, P.H., Shine, J., Martial, J.A., Ivarie, R.D., Morris, J.A., Ullrich, A., Baxter, J.D. and Goodman, H.M. (1978) *Nature* 276, 795-798.
17. 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.
18. Shine, J. and Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* 71, 1342-1346.
19. Wensink, P.C., Finnegan, D.J., Donelson, J.E. and Hogness, D.S. (1974) *Cell* 3, 315-325.
20. Miller, D.L., Gubbins, E.J., Pegg III, E.W. and Donelson, J.E. (1977) *Biochemistry* 16, 1031-1038.
21. Gubbins, E.J., Newlon, C.S., Kann, M.D. and Donelson, J.E. (1977) *Gene* 1, 185-207.
22. Kool, A.J., Van Zeven, M.S. and Nijkamp, H.J.J. (1974) *J. Bact.* 118, 213-224.
23. Maurer, R.A., Stone, R. and Gorski, J. (1976) *J. Biol. Chem.* 251, 2801-2807.
24. Maurer, R.A. (1979) *Mol. Cell. Endocrinol.* 13, 291-300.
25. Laemmli, U.K. (1970) *Nature* 227, 680-685.
26. Bonner, W.M. and Laskey, R.A. (1974) *Eur. J. Biochem.* 46, 83-88.
27. Gubbins, E.J., Maurer, R.A., Hartley, J.L. and Donelson, J.E. (1979) *Nucleic Acids Res.* 6, 915-930.
28. Sutcliffe, J.G. (1978) *Proc. Natl. Acad. Sci. USA* 75, 3737-3741.
29. Sutcliffe, J.G. (1979) *Cold Spring Harbor Symposium on Quantitative Biology* 44, 77-90.
30. Nicoll, C.S. (1967) *Endocrinology* 80, 641-655.