JOURNAL OF BACTERIOLOGY, May 1983, p. 650–655
0021-9193/83/050650-06$02.00/0

Vol. 154, No. 2

Analysis of ColE1 Expression In Vitro After Chromosome Fragmentation

HUI-ZHU CHEN AND GEOFFREY ZUBAY*

Fairchild Center for Biological Sciences, Columbia University, New York, New York 10027

Received 15 November 1982/Accepted 22 February 1983

The RNA and protein products synthesized from ColE1 DNA were observed before and after cutting the DNA with different restriction enzymes. Synthesis was carried out in the DNA-directed coupled transcription translation system. The S-30 extracts used to catalyze synthesis were prepared from a *recB* mutant in which the linear DNA fragments resulting from restriction enzyme cleavage were spared from the usual degradation by exonucleolytic attack. By correlating the observed in vitro synthesized products with the location of the cleavage sites in the plasmid chromosome, it was possible to identify specific gene products. The *col* gene catalyzes the synthesis of numerous peptides in addition to the 56-kilodalton colicin protein encoded by this gene. Most of the subsidiary products appear to arise as the result of premature termination by a mechanism(s) which remains to be determined. A unique RNA and protein were characterized as products of the *imm* gene. The RNA has an estimated mass of 150 kilodaltons, and the protein has an estimated mass of 13 kilodaltons. From the DNA sequence of the chromosome, it was concluded that the transcripts from the *imm* and *col* genes must crisscross each other over a region of about 75 base pairs. Such a pattern of transcription might lead to interference of transcription of one gene by the other gene. Consistent with this hypothesis, it was found that *imm* gene transcription increased severalfold in vitro when the chromosome was cleaved in a way that eliminated transcription originating at the *col* gene promoter. Surprisingly, the increase in transcription by this mechanism did not result in a significant increase in the synthesis of the *imm* gene-encoded protein.

ColE1 is a small, closed, circular DNA duplex containing about 6,400 base pairs. This plasmid, or derivatives of it, has been the most popular vector used to promulgate foreign DNA in the form of hybrid plasmids. The main protein product of ColE1 has also excited the interest of many investigators seeking model systems for studying protein-membrane interactions. Relatively little attention has been focused on the intracacies of gene expression by the plasmid itself. It seems likely that a detailed investigation of the gene products of ColE1 under different conditions would reveal a complex interdependence of expression by the various genes and gene products and modes of gene expression which are presently poorly understood.

We chose to investigate the properties of ColE1 in a cell-free coupled transcription translation system containing an S-30 extract of *Escherichia coli*. In this system, synthesis of the DNA and protein products of the chromosome were observed under a wide variety of conditions and in the presence of different factors. In the past, quantitative studies in the coupled system were limited to the use of circular DNAs because of extensive degradation which resulted when linear DNAs were used. Recently, however, we have shown that when the necessary S-30 extracts are prepared from a *recB* mutant, linear DNAs are no longer degraded (12). This introduces a new dimension to the in vitro approach. Cleaved chromosomes can now be studied in detail to determine how their expression compares with that of intact chromosomes. Abortive synthesis of RNA and protein products should be associated with the cleaved gene, and decreased synthesis should be associated with other genes which are located at downstream sites within the same operon. The availability of hundreds of restriction enzymes with different sequence specificities provides the necessary set of tools for cleaving chromosomes at specific points.

In this paper, we present the results obtained by employing this approach to ColE1.

MATERIALS AND METHODS

The restriction enzymes *Eco*RI, *Hae*II, *Hha*I, *Pst*I, and *Sma*I were obtained from Bethesda Research Laboratories, Inc. All samples of restriction enzyme-

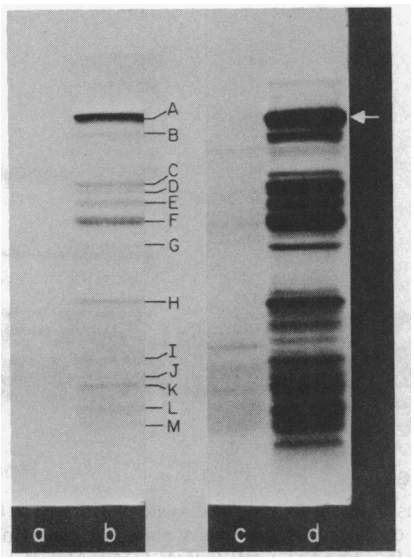

FIG. 1. Gel fluorogram of [^{35}S]methionine-labeled proteins made in the coupled system with intact ColE1 DNA. a, Complete system minus DNA; b, complete system with 50 μg of ColE1 DNA per ml added; c and d, the same as a and b, respectively, except for additional fluorography time. Only the main bands are labeled, and the M_r, values (kilodaltons) for each band are as follows: A, 56,000; B, 52,000; C, 42,000; D, 38,000; E, 36,000; F, 27,000; G, 24,000; H, 17,000; I, 15,000; J, 14,000; K, 13,000; L, 11,000; and M, 10,000. Molecular weights were estimated with the help of markers of known molecular weights. Such estimates are usually accurate to within 10%. Arrow indicates ColE1.

digested DNA were checked for complete digestion by gel electrophoresis.

The *E. coli recB21* strain used to make cell-free extracts for these studies has been described elsewhere (1, 12). RNA was electrophoresed on a 3 to 12% acrylamide gradient gel, and protein was electrophoresed on a 10 to 20% acrylamide gradient gel.

Preparation of extracts, conditions for cell-free synthesis, electrophoresis, and fluorography have been previously described in detail (12).

Isolation of individual segments of restriction enzyme-digested DNA was carried out by a modification of the procedure of Vogelstein and Gillespie (9). This rapid, quantitative technique involves binding of DNA to glass in the presence of NaI.

Colicin E1 protein was purified by the procedure of Schwartz and Helinski (7).

Preparation of antiserum to colicin E1 protein was carried out by vigorously mixing 1 ml of purified colicin protein suspension (1 mg/ml in 0.01 M potassium phosphate buffer [pH 7.0]) with an equal volume of Freund complete adjuvant (Difco Laboratories). Small portions of the homogenate were injected intracutaneously into a New Zealand white rabbit. After 6 weeks, 1 ml of the antigen suspension without adjuvant was injected intravenously as a booster, and 1 week later a small amount of blood (1 to 2 ml) was collected and the activity of antiserum was tested by the Ouchterlony method. Blood (50 to 100 ml) was collected from ear veins by cutting the veins or by using a large needle. The blood was left at room temperature for 30 min and then kept at 4°C overnight. The clot was spun down at 1,200 × *g* for 10 min, and the serum was stored at −20°C in small portions until use. The control serum was collected from the same animal before the immunization was started.

The antiserum was used in the following way. A 30-ml amount of IP wash buffer (20 mM Tris-hydrochloride [pH 7.5], 150 mM NaCl, 10 mM EDTA, 0.05% Nonidet P-40) and 10 μl of anticolicin serum were added to 10 μl of coupled protein synthesis mixture. After incubation for 1 h at room temperature, 25 μl of 10% staphylococcal cell suspension (Calbiochem) freshly washed with a buffer containing 25 mM HEPES (pH 7.5), 140 mM NaCl, 1 mM magnesium acetate, and 1 mM $CaCl_2$ was added. The mixture was then incubated for another 30 min. The precipitate was washed three times with 500 μl of IP wash buffer each time and suspended in 12 μl of the sample buffer for sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

RESULTS

Our results are shown in fluorograms or autoradiograms of gels containing electrophoresed products synthesized in the coupled transcription translation system. The DNA used to direct cell-free DNA and protein synthesis was either intact ColE1, restriction enzyme-digested ColE1, or fractionated segments of restriction enzyme-digested ColE1. When protein was being observed, detection was made possible by the addition of [^{35}S]methionine to the incubation mixture for cell-free synthesis; when RNA was being observed, [^{32}P]UTP was added.

Protein synthesis with ColE1 DNA or derivatives of ColE1 DNA. The pattern of proteins synthesized in the coupled system with intact ColE1 is shown in Fig. 1. The same synthetic mixture lacking the added ColE1 DNA gives a very low background, permitting unambiguous assignment of radioactive products. The number of proteins observed far exceeds the coding capacity of ColE1 (~6,400 base pairs). Only the main bands observed at the lower exposure (Fig. 1b) are labeled (A through M).

Similar fluorograms are shown for proteins synthesized in the presence of intact ColE1 and ColE1 digested with different restriction enzymes (Fig. 2). Cleavage of the chromosome with each enzyme resulted in the loss of some of the bands seen when the intact chromosome was used to direct synthesis. With very few exceptions, no new bands were created as a result of the cleavages. The exceptions include a new band around M_r 54,000 when *Eco*RI-cleaved DNA was used (Fig. 2A). The interpretations of the changes in band pattern are discussed below.

Effect of addition of Colicin E1-specific antiserum on the banding pattern. Purified colicin E1

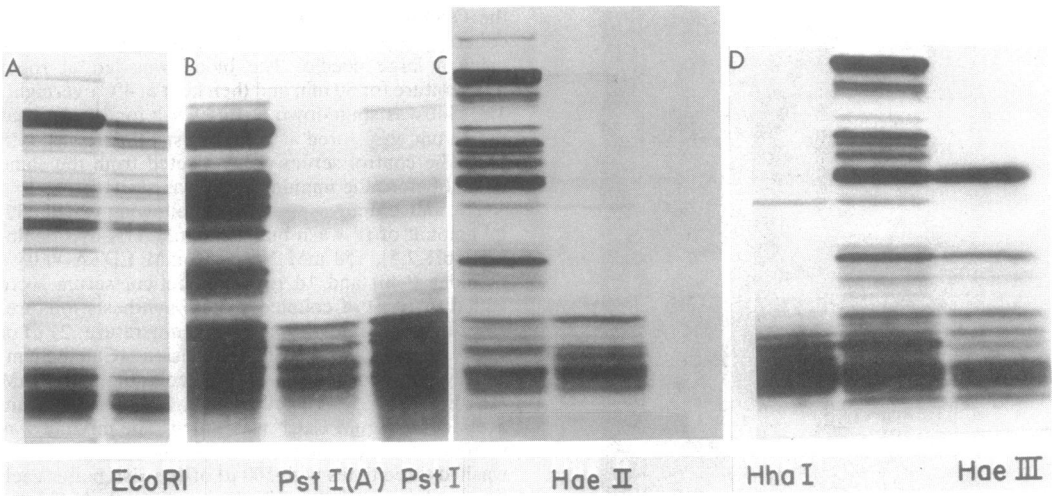

FIG. 2. Gel fluorograms of [^{35}S]methionine-labeled proteins made in the coupled system with intact ColE1 DNA or restriction enzyme-digested ColE1 are comigrated to eliminate any ambiguity about the assignment of bands. Five different restriction enzymes were used: (A) *Eco*RI (5′-GAATTC), (B) *Pst*I (5′-CTGCAG), (C) *Hae*II (5′-PuGCGCPy), (D) *Hha*I (5′-GCGC), and *Hae*III (5′-GGCC).

protein was used to make a specific rabbit antiserum as described above. The ^{35}S-labeled proteins made in the ColE1 DNA-directed coupled system were mixed with a neutral rabbit serum and a colicin E1 antiserum as described above. The resulting precipitates were suspended and electrophoresed alongside untreated portions from the same incubation (Fig. 3). It can be seen that several bands near the top of the gel pattern were selectively precipitated by the specific antiserum, indicating an immunological relatedness.

RNA synthesis with intact ColE1 and restriction enzyme-digested ColE1. Observations of RNA products made in the coupled system were facilitated by the addition of [^{32}P]UTP to the incubation mixtures. After electrophoresis, the gels were autoradiographed. The results for RNA synthesized in the presence of ColE1 and restriction enzyme-cleaved ColE1 are shown in Fig. 4. RNA products synthesized in the presence of intact plasmid pBR322 are also shown.

DISCUSSION

Different functional regions (genes) in the ColE1 chromosome have been identified by isolating variants which contain transposons inserted at discrete locations or by comparing the behavior of deletion mutants with the wild-type plasmid (2). In this way, four genes have been identified, *col*, *imm*, *rel*, and *inc* (Fig. 5). The *col* gene encodes a 56-kilodalton protein, colicin, that can be secreted; it kills cells which do not contain the plasmid. The *imm* gene encodes an immunity factor which protects the plasmid-containing cells against the toxic effects of colicin. The *inc* gene encodes a function that modulates the number of plasmids in any given cell and inhibits the replication of other homologous plasmids present in the same cell. The *rel* gene encodes a factor which permits transfer of the plasmid during conjugation. In addition to these four genes, the origin of replication (*ori*) has been identified. Two transcripts which are involved in replication and control of replication are made from overlapping sites located on one side of the *ori* locus.

The cleavage sites for a number of restriction enzymes have been determined in relation to the different functional regions of the chromosome (Fig. 5). Our strategy was to identify the specific gene products by determining the effects of restriction enzyme cleavage on the RNA and protein products synthesized in the coupled transcription translation system. This work has been greatly aided by the availability of DNA sequence data for about 60% of the chromosome (5, 6, 10).

Observations relating to the *col* gene. Three restriction enzymes, *Eco*RI, *Pst*I, and *Sma*I, make cleavages in the *col* gene exclusively (Fig. 5). The precise location of these cleavage points is derived from sequence data (10). These restriction enzymes have been used to determine the gene products associated with this gene. The *col* gene contains a single, long reading frame for a 523-residue polypeptide which reads in the counterclockwise direction. This size is consistent with the 56-kilodalton protein which is the main product of this chromosome (Fig. 1, arrow). The single restriction site for the *Eco*RI enzyme interrupts this reading frame 21 amino

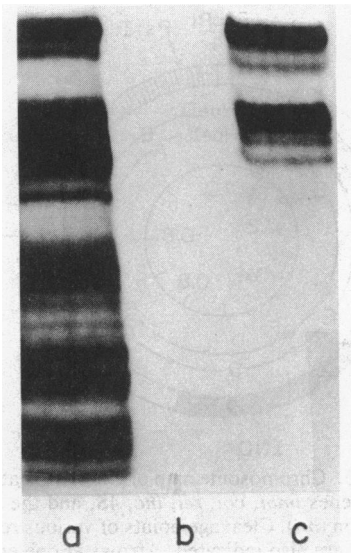

FIG. 3. Gel fluorograms of [^{35}S]methionine-labeled proteins made in (a) untreated control, (b) protein precipitated by control rabbit serum, and (c) protein precipitated by colicin protein antiserum. The procedures for preparation and use of antiserum are given in the text.

acids before the terminus. The protein pattern observed when the chromosome is cleaved by this enzyme differs from the normal pattern by the disappearance of the 56-kilodalton band and the emergence of a new band about 2 kilodaltons smaller, with a greatly lowered intensity (Fig. 2A). The remainder of the pattern is qualitatively similar in appearance to the normal pattern except for the absence of a faint band (labeled G in Fig. 1). Loss of the 56-kilodalton band is clearly consistent with the synthesis of a somewhat smaller colicin-related protein lacking 21 C-terminal amino acids. The greatly reduced intensity of the 54-kilodalton band compared with the 56-kilodalton band can be attributed to two factors. First, the colicin gene is expressed at a higher level when the DNA is supercoiled, which is the case only when the intact chromosome is used (11). Second, termination and release of the polypeptide fragment made when the cleaved chromosome is used are probably slowly blocking the retranslation of the abortive mRNA.

The *Pst*I and *Sma*I enzymes both produce cleavages very early in the long reading frame (Fig. 5); the *Pst*I enzyme produces two additional cleavages. The results of cleavage by either of these enzymes on the observed protein pattern are quite similar; only the results obtained with the *Pst*I-cleaved chromosome are shown (Fig. 2B). All bands above the I band (M_r, 15,000; Fig. 1) are eliminated by this treatment. We concluded that all proteins larger than 15 kilodaltons are encoded by the *col* gene. The presence of all of these proteins, except for the 56-kilodalton protein in the pattern observed when the chromosome is cleaved with *Eco*RI, indicates that none of them terminates at the end of the main reading frame. Therefore, they must arise either from premature termination or degradation of a larger protein. Degradation cannot be ruled out; however, it has been shown that all of these proteins are completely stable in the S-30 incubation system by postincubation for 2 h when further synthesis of labeled protein is prevented by the absence of [^{35}S]methionine. Furthermore, introduction of the serine protease inhibitor *p*-toluene sufonyl fluoride (59 μM) during synthesis does not result in any alteration of the observed banded profile by the gel techniques used here. However, run-off transcripts made from restriction enzyme-cleaved templates might be observable. Inspection of the ColE1 sequence predicts that the run-off transcripts in a *Sma*I-digested DNA or the A fragment obtained from a *Pst*I digest should have 290 and 223 bases, respectively. Weak bands of RNA in this size range have been observed on autoradiograms of the ^{32}P-labeled RNA (data not shown). Further

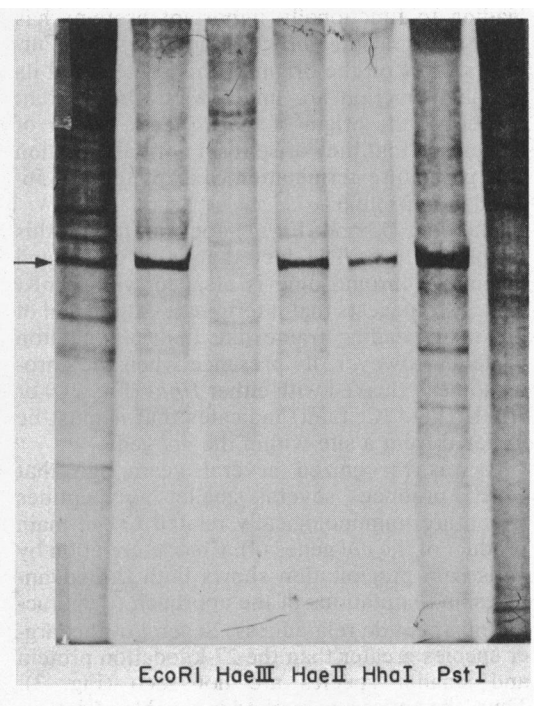

FIG. 4. Gel autoradiogram of [^{32}]PUTP-labeled RNA made in the coupled system. Templates used were either intact ColE1 (far left), restriction enzyme-cleaved ColE1 (as indicated), or intact pBR322 (far right). Arrow points to the presumptive *imm* mRNA.

characterization of these bands has not been made; thus, this interpretation is only tentative.

The *Hae*III enzyme makes at least nine cleavages in the ColE1 chromosome but only one in the *col* gene. The single cleavage in the *col* gene is located near the middle of the main reading frame. The largest protein that could be made from the abortive fragment containing the promoter-proximal part of the *col* gene would have an M_r of about 28 kilodaltons. In fact, the 27-kilodalton protein (band F) and other proteins with a lower molecular weight known to be associated with the *col* gene are made, but no *col* gene-related proteins with a higher molecular weight are made (Fig. 2D). It has also been shown that these colicin-related proteins are made when the *Hae*IIIA restriction fragment is used. Due to the size of the F protein and the size of the *Hae*III-cleaved *col* gene, it seems highly likely that the initiation point for F protein synthesis is the same as that for the 56-kilodalton protein. Clearly, the same rather than out-of-phase reading frames must be used because inspection of the sequence shows that there is no alternative. The F protein is made in appreciable amounts, and we suspect it results from premature termination. In this regard, it should be noted that premature termination leading to functionally important proteins has been seen in other procaryotic genes (3). Further studies on the origin of this protein and its possible function are under way. Our current model for the origin of the *col* gene family of proteins is that they arise from normal initiation and premature termination, except for the 56-kilodalton protein.

The faint G band (Fig. 1) does not fit into this simple picture. The fact that it is eliminated when the chromosome is cleaved with *Eco*RI (Fig. 2A) suggests that it terminates at the end of the main reading frame like the 56-kilodalton protein. However, its presence when the chromosome is cleaved with either *Hha*I (Fig. 2D) or *Hae*II (Fig. 2C, faint) indicates that it must be initiated from a site within the *col* gene.

It was recognized several years ago that ColE1 produces several smaller polypeptides which are immunologically related to the main product of the *col* genes (4). Characterization by antiserum precipitation shows both the advantages and limitations of the approach to characterizing peptide relatedness. Several of the larger species greater than the 27-kilodalton protein and smaller species are not seen (Fig. 3). Thus, the antiserum method is capable of detecting relatedness for some but not all *col* gene-encoded proteins.

Observations relating to the *imm* gene. The sequence of the *imm* gene is known from the work of Oka et al. (5). It contains a 113-codon

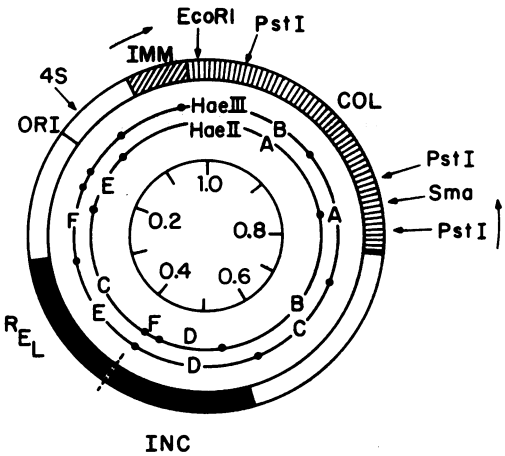

FIG. 5. Chromosome map of ColEl indicating position of genes *imm*, *col*, *rel*, *inc*, 4S, and the origin of replication (*ori*). Cleavage points of various restriction enzymes are also indicated. Arrows at one end of the *imm* and *col* genes indicate the directions of transcription.

uninterrupted reading frame believed to encode the immunity protein. The size of the *imm* gene mRNA estimated from the presumptive initiation and termination sites for transcription is about 425 bases. Of the restriction enzymes tested (*Pst*I, *Sma*I, *Eco*RI, *Hae*II, *Hae*III, and *Hpa*I), only *Hae*III has a recognition sequence within the gene (Fig. 5). Correlated with this is the finding that the K protein is missing when ColE1 is cleaved with *Hae*III but not when it is cleaved with the other restriction enzymes (cf. Fig. 2A through D). The estimated size of this protein from gels (M_r, ~13 kilodaltons) is in reasonable agreement with that expected for a protein with 113 amino acid residues. Further evidence that the K band encodes the *imm* protein comes from the observation that the purified *Hae*IIA restriction fragment directs synthesis of the K protein (data not shown). It can be seen that the *Hae*IIA restriction fragment includes the *imm* gene (Fig. 5) but not other genes of ColE1. These observations on the *imm* gene protein confirm the suggestion of Inselburg and Applebaum (2) based on products of mutant ColE1 that a 14-kilodalton polypeptide is associated with the immunity function.

Further observations relating to the messenger for the *imm* gene have been made. The *Hae*IIA restriction fragment directs the transcription of an RNA molecule with a molecular weight of about 150 (data not shown). The *imm* mRNA is the only known complete transcript contained in the *Hae*IIA restriction fragment; therefore, the observed band is clearly a strong candidate for the *imm* mRNA molecule. It can be calculated that a 425-base mRNA should

produce a molecule of about this size. Further support that this RNA is the main product of the *imm* gene comes from the observation that all the restriction digests used, except the *Hae*III digest, direct the synthesis of such an RNA (Fig. 4). The amount of presumptive *imm* mRNA produced by the cleaved chromosomes is always considerably greater than that synthesized from intact ColE1 (Fig. 4). In this connection, it should be pointed out that the transcripts for the *imm* and the *col* genes crisscross each other, resulting in an overlap of about 75 base pairs. The synthesis of larger amounts of *imm* mRNA on cleaved DNAs where the *col* promoter has been removed may reflect the absence of interfering transcription originating at the *col* gene promoter. Further studies are necessary to confirm this interpretation.

The production of increased amounts of *imm* mRNA is not paralleled by an increased synthesis of the *imm* protein. Approximately the same amount of *imm* protein is made with intact ColE1 as with ColE1 cleaved by *Pst*I or *Hae*II. This puzzling observation may be explained by some sort of translational control of *imm* protein synthesis. One possibility is that a product(s) of the *col* gene serves to activate translation of the *imm* gene messenger. This would be a convenient way of regulating the ratio of *imm* and *col* gene proteins.

In vivo observations in which levels of the *imm* protein are measured when the *col* gene has been interrupted should relate to this question. However, the observations which have been made thus far appear to support opposing views. Thus, Dougan and Sherratt (1a) have observed a large reduction in the levels of immunity in several instances in which the Tn*I* transposon has been inserted into the *col* gene. In a seemingly contradictory observation, Shafferman et al. (8) found that deletion of a 120-base-pair segment near the end of the *col* gene results in a 30-fold increase in immunity. It seems likely that this complex question relating to the regulation of *imm* gene expression is most likely to be resolved by further investigations into the cell-free coupled system, where it is easiest to measure the levels of both the RNA and protein products in the presence of various factors.

Measurements and interpretations on the other gene products made from ColE1 in the coupled system will be reported elsewhere.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant 5 R01 GM-29217 from the National Institutes of Health and grant PCM-80-21625 from the National Science Foundation.

We thank R. Spangler and T. Segawa for preparing colicin protein antiserum.

LITERATURE CITED

1. **Chen, W. Z., and G. Zubay.** 1982. Prokaryotic coupled transcription-translation. Methods Enzymol **101**:674–690.

1a. **Dougan, G., and D. Sherratt.** 1977. The transposon TnI is a probe for studying Col E1 structure and function. Mol Gen. Genet. **151**:151–160.

2. **Inselburg, J, and B. Applebaum.** 1978. Proteins synthesized in minicells containing plasmid ColE1 and its mutants. J. Bacteriol. **133**:1444–1451.

3. **Kastelein, R. A., E. Remant, W. Fiers, and J. van Duin.** 1982. Lysis gene expression of RNA phage MS2 depends on a frameshift during translation of the overlapping coat protein gene. Nature (London) **295**:35–41.

4. **Meagher, R. B., R. C. Tait, M. Betlack, and H. W. Boyer.** 1977. Protein expression in minicells by recombinant plasmids. Cell **10**:521–536.

5. **Oka, A., N. Nomura, M. Morita, H. Sugisaki, K. Sugimoto, and M. Takanami.** 1979. Nucleotide sequence of small Col E1 derivatives: structure of the regions essential for autonomous replication and colicin E1 immunity. Mol. Gen. Genet. **122**:154–159.

6. **Patient, R. K.** 1979. Characterization of in vitro transcription initiation and termination sites in Col E1 DNA. Nucleic Acids Res. **6**:2647–2665.

7. **Schwartz, S. A., and D. H. Helinski.** 1971. Purification and characterization of colicin E1. J. Biol. Chem. **246**:6318–6327.

8. **Shafferman, A., S. Cohen, and Y. Flashner.** 1978. A DNA segment within the colicin E1 structural gene in Col E1 affecting immunity to colicin. Mol. Gen. Genet. **164**:259–264.

9. **Vogelstein, B, and D. Gillespie.** 1979. Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. U.S.A. **76**:615–619.

10. **Yamada, M., Y. Ebina, T. Miyata, T. Nakazawa, and A. Nadazawa.** 1982. Nucleotide sequence of the structural gene for colicin E1 and predicted structure of the protein. Proc. Natl. Acad. Sci. U.S.A. **79**:2827–2831.

11. **Yang H.-L., K. Heller, M. Gellert, and G. Zubay.** 1979. Differential sensitivity of gene expression in vitro to inhibitors of DNA gyrase. Proc. Natl. Acad. Sci. U.S.A. **76**:3304–3308.

12. **Yang, H.-L., L. Ivashkiv, H.-Z. Chen, G. Zubay, and M. Cashel.** 1980. Cell-free coupled transcription-translation system for investigation of linear DNA segments. Proc. Natl. Acad. Sci. U.S.A. **77**:7029–7033.