The primary structure of the plasmid ColE1 DNA has been determined. The plasmid DNA consists of 6646 base pairs (molecular mass of 4.43 MDa) and is 48.46% in GC content. The φ80 trp insert of the composite plasmid of ColE1, pVH51, has also been determined. The determination of the nucleotide sequence of ColE1 DNA provides the basis for examining the relationships between the DNA sequence and the gene organization of the plasmid. The focus of this paper is to use this sequence data coupled with a review of the literature and our own work to examine the nine known functional regions of ColE1: imm (colicin E1 immunity), rep (replication function), Inc (plasmid incompatibility and copy number control), bom (basis of mobility), rom (modulator of inhibition of primer formation by RNA I), mob (plasmid mobilization), cec (determinant for conversion of plasmid multimers to monomers), exc (plasmid entry exclusion), cea (structural gene for colicin E1), and kil (structural gene for the Kil protein).

ColE1 is a small multiple copy plasmid of Escherichia coli which is nonconjugative and colicinogenic (1, 2). The process of DNA replication and the control of copy number for this plasmid have been studied extensively (for reviews, see Refs. 3–7). Considerable attention has also been directed on the primary structure of this plasmid 90% of the nucleotide sequence found between residues 1 and 3371 and part of the 480 trp region. Since pVH51 and its derivatives have been used in DNA physical studies (25, 26), we determined the sequence of the φ80 trp insert region (Fig. 2) so that the full sequence of this ColE1 derivative would be available to investigators.

EXPERIMENTAL PROCEDURES

Bacterial Strain and Growth—E. coli JC411 carrying either ColE1 or pVH51 (23) plasmid was used. Bacterial growth was in M9 minimal medium supplemented with casamino acids (Difco), glucose, and thiamine at 37 °C as described previously (27).

Plasmid Purification—ColE1 and pVH51 DNA were prepared as described previously (27, 28).

Chemicals and Enzymes—Dimethyl sulfate was purchased from Aldrich, hydrazine from Eastman Kodak Co., and pipidine from Fisher. MCB reagents were also used for the sequence analysis. Restriction endonucleases were purchased from New England Biolabs, T4 polynucleotide kinase from P-L Biochemicals, glyceraldehyde-3-phosphate dehydrogenase and 3-phosphoglycerate kinase from Calbiochem-Behring, and E. coli alkaline phosphatase from Worthington.

DNA Sequencing—Restriction maps of ColE1 DNA with various restriction endonucleases were obtained by standard methods of restriction analysis. DNA fragments were labeled at their 5'-ends with 32P and the nucleotide sequences were determined according to the method of Maxam and Gilbert (29) with a modification described previously (27).

Fig. 1 provides the strategy of sequencing ColE1 DNA and the φ80 trp insert of the mini-ColE1 derivative pVH51 (23) and the restriction sites used for the 5'-end labeling of restriction fragments for the sequence determination. Over 90% of ColE1 DNA was sequenced from both strands. More than 50% of the sequence data was duplicated by sequencing restriction fragments that overlap each other as shown in Fig. 1.

RESULTS

Complete Nucleotide Sequence of ColE1—The complete nucleotide sequence is presented in Fig. 2. The ColE1 DNA contains a unique EcoRI restriction site (30–32). This site will be used as a reference point in the presentation of the physical and genetic maps of ColE1. The imm proximal end is referred to as the left end, and the cea proximal end as the right end. The strand orientations from left to right are 5' to 3' for the upper strand and 3' to 5' for the lower strand. Separation of complementary strands in a cesium chloride buoyant density gradient by preferential binding of poly(UC) led to designation of the lower strand as the heavy strand and the upper strand as the light strand (33).

The nucleotide sequence presented here and some partial sequences reported previously have the following differences. The sequence reported by Bastia (18) has a T/A base pair at position 1454 instead of C/G base pair. We also found three additional basepairs: C/G at positions 1415 and 1598 and A/T at position 1582. The latter addition eliminates the predicted PuvII restriction site which we did not find. An extra
FIG. 1. Sequencing strategy used to obtain the nucleotide sequence of ColE1 DNA and the ω80 trp insert of the plasmid pVH51. The left side shows the restriction sites and sequence overlaps that were used to determine the nucleotide sequence for the genetic regions rep, inc, rom, mob, and exc shown from top to bottom, respectively (see text for complete details of the genes coded by these regions). A dotted line specifies that the sequence was not read for that particular dotted segment of the 5′-end-labeled restriction fragment (see exc region). The right side shows the restriction sites and sequence overlaps that were used to obtain the nucleotide sequence for the cea region of ColE1 and the ω80 trp insert of pVH51 shown from top to bottom, respectively.

A/T base pair was reported by Machida et al. (34) between 3861 and 3862. The sequence we determined has a T/A base pair at 4785 as reported by van den Elzen et al. (21) which was not present in the sequence reported by Ebina et al. (35). We have observed CA/GT instead of the reported AC/TG as reported by Yamada et al. (22) at position 5519 and 5520 causing a single amino acid difference in the predicted colicin E1 protein.

Confirmation of part of the ColE1 nucleotide sequence data comes from protein sequence and composition data. Yamada et al. (22) compared the amino acid composition for colicin E1 from the nucleotide sequence with previously reported data on the analysis of amino acid composition and found the results to be in excellent agreement. The first 20 amino acids of the amino terminus of colicin E1 have been sequenced, and they are in exact agreement with the amino acid sequence.
Nucleotide Sequence of ColEl DNA

predicted from the nucleotide sequence data. An internal peptide of colicin E1 has been partially sequenced with exact agreement with the nucleotide sequence. In addition, the first 10 amino acids of the Rom protein (see below) are in agreement with the sequence predicted from the nucleotide sequence.

The high degree of verification described above coupled with the fact that 90% of the sequence was obtained from both strands and approximately 50% of the sequence was determined in duplicate provides for considerable confidence in the accuracy of the DNA sequence.

The determination of the complete nucleotide sequence of ColEl DNA shows that the DNA consists of 6646 base pairs (molecular mass of 4.43 MDa). The buoyant density of the DNA has been previously determined to be 1.7020 (37).

The nucleotide sequence of the φ80 trp insert of plasmid pVH51 was similarly determined (Fig. 2). The length of the φ80 trp insert in this composite plasmid is 473 base pairs, and the total number of base pairs of pVH51 is 3847 (molecular mass of 2.56 MDa).

2 W. A. Cramer, personal communication.

Coding Capacity of ColEl—To identify open reading frames for potential proteins, we employed the following criteria. 1) A Shine-Dalgarno sequence can be found within 3–10 base pairs upstream of the initiation codon (38, 39). 2) The number of amino acids in the coding region should be 40 or more. Within the limitations imposed above, a total of 23 open reading frames containing 27 translation initiation signals have been identified. The base sequence for these initiation sites together with location, orientation, size, and correlation with respective genes are presented in Table I. Four of the initiation signals were internal and in phase to longer reading frames. The arrangement of the potential 23 polypeptides with reference to the genetic organization of the plasmid is presented in Fig. 3.

DISCUSSION

Regions Involved in DNA Replication and Its Regulation—The region from position 600 to 1900 is involved in replication of ColEl DNA. No plasmid-encoded protein is required for replication (30, 40, 41). One of the transcripts (RNA II or primer transcript) from the region initiates at position 641 and proceeds rightward. The nascent transcript starts to form...
Fig. 2. Complete nucleotide sequence of ColEl DNA oriented from the unique EcoRI site with the placement of the imm gene at the left end and the cec gene at the right end. See Table 1 and Fig. 3 for the location of open reading frames, the gene or functional region assignment, and the specific reading frame for the coding region. After the ColEl sequence, the nucleotide sequence of the +E40 trp segment of the mini-ColEl plasmid pVH51 is shown. The plasmid pVH51 consists of a part of ColEl DNA and a segment of DNA from bacteria phage trp (see text for references). The 473-base pair fragment of +E40 trp is joined to the ColEl sequence at position 3371 (marked by an asterisk in the ColEl sequence) and position 6644 (EcoRI site) generating a 3847-base pair composite plasmid (pVH51).
Fig. 2—continued.
a persistent hybrid with the template DNA after synthesis of about 550 nucleotides. The hybridized RNA is cleaved by RNase H and used as the primer for DNA synthesis by DNA polymerase I (42-44). The position where the first deoxyribonucleotide is attached to the RNA primer is defined as the origin of DNA replication (17). The origin of DNA replication is located at one of the consecutive nucleotides AAC (the first A at position 1196) and is identical for in vivo (45) and for in vitro (17) replication. The process of primer formation can be affected by various mutations in the region between the primer promoter and the replication origin (44, 46, 47). Each mutation affects a specific step in the primer formation by altering the secondary structure of RNA I1 required for primer formation (44, 48).

A region of ColEl DNA which specifies the 5'-portion of RNA I1 is involved in regulation of plasmid DNA replication. The region also specifies a small RNA called RNA I which is about 108 nucleotides long (49). Transcription of RNA I starts at position 751 and proceeds in the direction opposite to the RNA I1 transcription and terminates near the site where transcription of RNA I1 starts. Thus, RNA I is complementary to the 5"portion of RNA I1. RNA I binds to RNA I1 (50, 51), and the binding prevents formation of a secondary structure of RNA I1 that is required for the transcript to form the primer for DNA synthesis by cleavage with RNase H.\(^3\) The inhibitory activity of RNA I is incompatibility group-specific and can be altered by certain mutations in the region that specifies both RNA I and RNA I1 (51-55).

Presence of another regulatory factor is suggested by the fact that deletion of the HaeII C segment from position 1371 to 2687 increases the plasmid copy number (56). This region specifies a small protein of 63 amino acids which is encoded between nucleotide positions 1803 and 1614 (36, 57, 58). The protein increases the rate of binding of RNA I to RNA I1 and thereby enhances the inhibitory activity of RNA I on primer formation (59). The gene for the protein is named rom, (RNA protein has no effect on in vitro transcription of RNA I or RNA I1 (59, 60), eliminating the contention that it acts as a repressor of the primer transcription (Rop) (57). The observation that led to the proposal of the repressor function can be explained by the increased rate of binding of RNA I to RNA I1 in the presence of the Rom protein (59). Thus, replication of ColEl DNA is controlled by binding RNA I to RNA I1, the rate of binding being affected by the Rom protein. Replication and its control are also affected by the supercoiling of the DNA because the promoter activities of RNA I (61, 62) and RNA I1 (63) are strongly affected by supercoiling.

The phenomenon of "amplification," the continuous replication of ColEl in bacteria treated with chloramphenicol (64), can be explained by inhibition of the synthesis of Rom protein, metabolic instability of RNA I, and increased availability of replication proteins due to cessation of synthesis of the host chromosome (36).

\(^3\) H. Masukata and J. Tomizawa, unpublished data.
The first nucleotide of the initiation codon specifies the position of the open reading frame. R and L refer to the rightward (top strand) and leftward (bottom strand) orientation of the reading frame. Open reading frames without a serial number are internal, in-frame, coding sequences of the preceding frame. The Shine-Dalgarno (38) sequences are underlined. The last three columns show the number of amino acids coded by the respective open reading frame, the polypeptide molecular mass (daltons), and the gene or functional region assignment, respectively. See text for the criteria used to determine open reading frames.

<table>
<thead>
<tr>
<th>Serial No. of reading frames</th>
<th>Position of initiation codon sequence</th>
<th>Sequence preceding start codon</th>
<th>No. of amino acids</th>
<th>Polypeptide molecular mass Da</th>
<th>Gene or functional region</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0400L</td>
<td>TATATGCGCAGGGTTTTATTTATG</td>
<td>113</td>
<td>13,434</td>
<td>imm</td>
</tr>
<tr>
<td>2</td>
<td>0448R</td>
<td>CATTTCCTTAAAGGTGGTTATG</td>
<td>45</td>
<td>4,869</td>
<td>kil</td>
</tr>
<tr>
<td>3</td>
<td>0605R</td>
<td>GAAAAGATATAGGGTCCTTCTTG</td>
<td>96</td>
<td>10,208</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>1189R</td>
<td>TGTCAGGGGAGCAGGGGCTATG</td>
<td>53</td>
<td>6,002</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>1241L</td>
<td>TATAAGGCGGAGAAGAAGTATG</td>
<td>123</td>
<td>13,763</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1347R</td>
<td>CGACCGAGCTAGGGGCTGATG</td>
<td>59</td>
<td>6,930</td>
<td>?</td>
</tr>
<tr>
<td>7</td>
<td>1408L</td>
<td>TACGCCATGCGGGTGTAGAAATG</td>
<td>56</td>
<td>6,448</td>
<td>?</td>
</tr>
<tr>
<td>8</td>
<td>1609R</td>
<td>ACCGCTCCCGGAGCTGATG</td>
<td>45</td>
<td>5,146</td>
<td>mob 1</td>
</tr>
<tr>
<td>9</td>
<td>1709L</td>
<td>GTCACCGAGAGGAGCTGATG</td>
<td>69</td>
<td>7,624</td>
<td>?</td>
</tr>
<tr>
<td>10</td>
<td>1803L</td>
<td>CCCCTTACCGGAGAGCTGATG</td>
<td>63</td>
<td>7,283</td>
<td>rom</td>
</tr>
<tr>
<td>11</td>
<td>1843R</td>
<td>GATTTTCTGTCTTATGGGTAATG</td>
<td>115</td>
<td>12,997</td>
<td>mob 2</td>
</tr>
<tr>
<td>12</td>
<td>1945R</td>
<td>GTCACACGGAGCTGATG</td>
<td>81</td>
<td>8,997</td>
<td>mob 2 int.</td>
</tr>
<tr>
<td>13</td>
<td>2180R</td>
<td>CAGGGACAGGGAGGCGGGATG</td>
<td>517</td>
<td>57,895</td>
<td>mob 3</td>
</tr>
<tr>
<td>14</td>
<td>2214R</td>
<td>TATGCCAGGGAGAAGGTTGTG</td>
<td>61</td>
<td>6,930</td>
<td>mob 4</td>
</tr>
<tr>
<td>15</td>
<td>2396R</td>
<td>CGCCGGAGAGCAGTGAAGGATG</td>
<td>445</td>
<td>50,171</td>
<td>mob 3 int.</td>
</tr>
<tr>
<td>16</td>
<td>2868R</td>
<td>CAGGCTGGAAGAGCGAGATTTATG</td>
<td>172</td>
<td>19,548</td>
<td>mob 6</td>
</tr>
<tr>
<td>17</td>
<td>3090R</td>
<td>AAGGGCTGAAAGAGCGAGATG</td>
<td>98</td>
<td>11,189</td>
<td>mob 6 int.</td>
</tr>
<tr>
<td>18</td>
<td>3393R</td>
<td>ATCTGAGGGATGCAAGTGT</td>
<td>77</td>
<td>9,350</td>
<td>mob 7</td>
</tr>
<tr>
<td>19</td>
<td>3657R</td>
<td>GCTGATACGGACGCAGAATG</td>
<td>53</td>
<td>5,937</td>
<td>mob 8</td>
</tr>
<tr>
<td>20</td>
<td>3786R</td>
<td>TGGCGAGCTCTGAGATTAAATG</td>
<td>56</td>
<td>6,111</td>
<td>mob 9</td>
</tr>
<tr>
<td>21</td>
<td>4258R</td>
<td>TCATCGAGAGCAATAGCTATG</td>
<td>129</td>
<td>14,464</td>
<td>?</td>
</tr>
<tr>
<td>22</td>
<td>4362L</td>
<td>TTTAAGGTAGGGAGCAGCTATG</td>
<td>139</td>
<td>15,870</td>
<td>exc 2</td>
</tr>
<tr>
<td>23</td>
<td>4862L</td>
<td>TCTAGGACAGCTGTCAGCTTTATG</td>
<td>139</td>
<td>16,705</td>
<td>exc 1</td>
</tr>
<tr>
<td>24</td>
<td>4892L</td>
<td>CACTCTTGGGAGCAGCTAAATAATG</td>
<td>128</td>
<td>15,381</td>
<td>exc 1 int.</td>
</tr>
<tr>
<td>25</td>
<td>5140R</td>
<td>AAGTCGAAGAGGATTTTATATG</td>
<td>522</td>
<td>57,349</td>
<td>cca</td>
</tr>
<tr>
<td>26</td>
<td>5435R</td>
<td>GCCTCAGGAGATACCGTGAATG</td>
<td>72</td>
<td>8,786</td>
<td>?</td>
</tr>
<tr>
<td>27</td>
<td>5852R</td>
<td>AAGGCTGCTGGAAGAGAAGATG</td>
<td>53</td>
<td>6,692</td>
<td>?</td>
</tr>
</tbody>
</table>

The T2 terminator of cea-kil transcription as described below. Point mutations in the palindrome allows the transcription to read through the region (see Fig. 5 below). The transcript can be processed to primer and the primer formation is not inhibited by RNA I. Consequently, the mutant plasmid is compatible with other CoIE1 derivatives. These results suggest that the presence of the palindrome ensures the functional independence of the origin sequence. Although some open reading frames are present in the origin region, a comparative study on the nucleotide sequences of the regions of CoIE1 and related plasmids pBR322, RSF1030, p15A, and CloDF13 reveals that none of the possible frames specify identical proteins or those with similar size and amino acid sequence (65). Probably this region does not specify a functional protein.

**Region for Conjugative Mobilization of the Plasmid**—As indicated, CoIE1 is a nonconjugative plasmid. However, it can be transmitted efficiently from cells carrying a conjugative plasmid such as F or R factor (for reviews, see Refs. 66 and 67). Models for the mobilization of small, nonconjugative plasmids for transfer have postulated the following steps (11, 13, 18, 66–69). 1) Mobilization proteins bind at the origin of transfer and introduce a single-stranded nick. 2) The nick generates a relaxed complex, coupled with the covalent attachment of a protein to the 5’ terminus of the cleavage site of the strand to be transferred. 3) The protein-bound strand is displaced and moves across the cell wall to the recipient cell. 4) After strand transfer, the complementary strand is synthesized.

The bom (basis of mobility) region is a cis-requiring region for plasmid mobility (11). Within this region is the relaxation or nic site located between the G and C residues at 1466 with cleavage occurring on the lower strand (18). Three polypeptides with molecular weights of 11,000, 16,000, and 60,000 can associate to CoIE1 DNA at this site (68). Using complementation analysis, the mobilization region involved in plasmid transfer was divided into at least three complementation groups (12). RNA polymerase binding assays show that there are at least four segments of DNA within the mobilization region which are potential promoters.4 Transcription that

---

4 P. T. Chan and J. Lebowitz, manuscript in preparation.
The exc Region—Inseldburg (15) designated the property of ColEl DNA that confers resistance to conjugal transfer of DNA between mating pairs carrying homologous ColEl plasmids as incompatibility. Obviously this designation generates confusion with incompatibility that results from the inhibition of formation of primer RNA by RNA I described above. For F plasmids, the loss of conjugal transfer between cells carrying homologous plasmids has been named entry exclusion previously by Naumova et al. (72) to avoid possible confusion.

The region between 3800 and 4500 base pairs of ColEl was identified as coding for the exc function by deletion analysis (15). For this region we find two open reading frames that we designate exc 1 and exc 2 reading leftward (Table 1 and Fig. 3). The open reading frame for exc 2 extends from positions 4362 to 3945, allowing for the synthesis of a 139-amino acid polypeptide. Electron microscopy identified an RNA polymerase-binding site that maps close to the exc 2 region (28), and ternary complex formation shows two closely spaced diverging promoters in this region. The promoter for RNA rightward synthesis is very weak, and transcription would proceed into a region that has no apparent coding capacity. The leftward promoter would lead to a coding sequence for a 139-amino acid polypeptide. A membrane location is suggested for the presumptive Exc 2 protein based upon a potential lipoprotein signal sequence at its amino terminus. This sequence contains the Leu--...Cys sequence which may be a recognition sequence for diglyceride modification, a prerequisite for processing of a prolipoprotein (73).

The possible reading frames for exc 1 start at nucleotide positions 4862 and 4829 and stop at 4445. Polypeptides of 139 and 128 amino acids would be synthesized respectively. An in vivo study revealed a leftward mRNA that initiates at position 4887 (74), which is close to the potential ribosome-binding site at position 4878 (Table 1). The spacing between the potential GCA ribosome site and the 4862 start codon is 12 nucleotides, suggesting some caution in this potential polypeptide assignment versus the start site at 4829 which has a GGA spaced 10 nucleotides from the initiation codon that would fit the protein synthesis criteria better. There is no signal sequence for either the 128- or 139-amino acid polypeptide of exc 1.

It would be worthwhile to note that the F plasmid produces two gene products, trh S and trh T (M, 18,000 and 25,000, respectively), that are responsible for reduced formation of stable mating pairs between cells carrying homologous plasmids (16). The latter protein is located in the outer membrane (16). It therefore appears reasonable, by analogy with the F...
plasmid, to propose that exc 1 and exc 2 function to reduce formation of stable mating pairs.

The cer Determinant—Although ColE1 is maintained stably under normal growth conditions, some deletion and substitution derivatives are relatively unstable. This instability appears to be correlated with plasmid multimerization (14). ColE1 is stable because it encodes a cis acting determinant, cer (ColE1 resolution sequence) that is necessary for recA, recE, and recF independent recombination events which convert multimers to monomers. The determinant is located in the HpaII E fragment from position 3687 to 4063. When this region is cloned into plasmid vectors, it greatly increases their stability (14).

The cea-kil and Immunity Regions—The region between positions 5000 and 622 contains the genes coding for colicin E1, the immunity and Kil proteins (Fig. 4). The cea structural gene 5140-60 codes for colicin E1, a polypeptide of 52 amino acids (22). The kil gene (445-583) codes for a polypeptide of 45 amino acids (75, 76). These genes, cea and kil, form an operon oriented from left to right (see below). The imm gene (400-61) is oriented from right to left (15, 19, 70, 77) and does not belong to the cea-kil operon.

Differentiation between cea and kil lethality functions upon SOS induction (78) was revealed by a physiological analysis of mitomycin C-induced cells carrying ColE1 plasmids with differing combinations of the three genes cea, kil, and imm. Most of the cells containing ColE1 cea- kil+ imm+ plasmids are killed within 1 h after induction; macromolecular synthesis stops, active transport is lost, and the membrane becomes abnormally permeable (75). These physiological events are also observed for induced cells containing cea+ kil- imm+ plasmids, whereas cells carrying cea- kil+ imm+ plasmids remained unaffected although the mitomycin C induction can cause a 100-fold increase in colicin E1 synthesis (75). These results strongly suggest that killing can be attributed solely to the kil function and not to the presence of colicin E1. The imm gene cannot prevent killing of those cells carrying cea- kil+ imm+ plasmids.

The observation that colicin E1 is not released to the medium from kil- bacteria (75) suggests that the kil gene acts to produce membrane damage which serves to release colicin E1 from the cells. This concept is reinforced by the similarity between the Kil protein and an inducible protein, designated protein H, produced by the bacteriocinogenic plasmid Clo-DF13, which has been shown to be involved in the lysis of the bacterial host (79, 86). Between the carboxy-terminal portion of the Kil protein and the corresponding portion of protein H, 29 of 32 amino acid residues are identical (80). The Kil polypeptide has a signal sequence containing the Leu---...Cys site (487-501) which may be used for diglyceride modification, suggesting the possibility that it may be processed into a lipoprotein (73).

Colicin E1 acts by forming membrane channels leading to a depolarization of the cell membrane potential and dissipation of the proton electrochemical gradient of the cell (10). Cells carrying cea+ kil- imm+ plasmids are killed when the SOS response is induced. This effect of colicin E1 is not prevented by adding trypsin to the medium, indicating that cell death cannot be due to the exogenous action of released colicin E1 (75). Furthermore, the induced damage occurs in tolC cells, carrying cea+ kil- imm+ plasmids (75). Because tolC is a strain specifically tolerant to exogenous colicin E1, the result indicates that the attack is from inside the cell. The role of immunity appears to be the prevention of lethality resulting from attack by endogenous and exogenous colicin E1. How immunity blocks the channel-forming action of colicin E1 is unknown.

In summary, these results support the following physiological model for mitomycin C-induced cell killing involving wild-type ColE1. The SOS response leads to the production of colicin E1 and Kil proteins. The lysis of cells, mediated by the kil function, leads to a release of colicin E1, which attacks cells not carrying the immunity function, thereby offering a selective advantage for surviving immune cells.

The key features of the control of transcription for the cea-kil operon have been elucidated from a number of studies (Fig. 4). Two transcripts of 1700 and 2200 nucleotides were identified both initiating at position 5065, 75 base pairs upstream from the amino-terminal codon (5140-5142) of colicin E1 (81). The promoter for these transcripts was designated the P1 promoter (35, 82). Studies on transcription from restriction fragments show an additional promoter designated P2 that promotes transcription from position 5149 (35, 82). This cannot be an alternative cea promoter because the sequence overlaps the colicin E1 structural gene sequence. On supercoiled DNA, transcription starts exclusively from the P1 promoter (81). A physiological role for the P2 promoter remains unclear.

The P1 promoter region contains a lex operator sequence (35, 82), and thus synthesis of colicin E1 and Kil proteins is induced upon treatment of E. coli cells with DNA-damaging agents or inhibitors of DNA replication. The sequence of the LexA protein-binding site is composed of two overlapped SOS boxes, CTTGATATAAAACATGGTTATATGACAG-(nucleotide positions 5063-5093). Purified LexA protein binds to the cea-kil control region and inhibits mRNA synthesis from the P1 promoter (83).

Dependence of induced synthesis of colicin E1 on cAMP in adenylate cyclase-deficient mutant (cya-) cells has been known (84), and stimulation of the synthesis of cea-kil in vivo mRNAs occurs upon cAMP addition (81). Sequences homologous to the binding site for the complex of cAMP and cAMP receptor protein in the operator region of the lac and gal operons (85) can be found at 47 and 70 base pairs (4978-4982 and 4950-4954) upstream from the −35 site of the P1 promoter (81).

The two in vivo transcripts of 1700 and 2200 nucleotides terminate at two sites, T1 and T2 (Fig. 5). The former corresponds to a location of a possible ρ-dependent terminator which extends from nucleotide positions 85 to 124. This region contains a region of dyad symmetry (Fig. 5A) followed by the sequence CAAAACAA which is similar to a common sequence CAACTCA found in other ρ-dependent terminators (86).
will affect the ratio of the 1700-2200 nucleotide mRNAs. In the kil function, leads to the release of colicin El. The imm of the cell membrane by colicin El acting from outside as well as from inside the cell.

expression of the cea-kil operon is repressed by the LexA protein. Some transcripts must circumvent the termination of the plasmid. It is evident from this review that many corresponding names of mutant plasmids allowing transcriptional read-through are blocked by p-dependent transcription termination at the T1 terminator. The T2 terminator is independent of p-factor in in vitro experiments (42, 81). Point mutations in the dyad symmetry (Fig. 5B) allow read-through of the transcription. The T2 terminator is independent of p-factor in in vitro experiments (42, 81). Point mutations in the dyad symmetry (Fig. 5B) allow read-through of the transcription. The T2 terminator is independent of p-factor in in vitro experiments (42, 81). Point mutations in the dyad symmetry (Fig. 5B) allow read-through of the transcription.

Concluding Remarks—The determination of the nucleotide sequence of ColEl DNA serves as the foundation for understanding control regions involved in the expression of proteins and cis acting determinants for this important plasmid.

REFERENCES
