Random screening of promoters from *Escherichia coli*
and classification based on the promoter strength*

Masayo Kubota, Yukiko Yamazaki and Akira Ishihama†

Department of Molecular Genetics, National Institute
of Genetics Mishima, Shizuoka 411, Japan

(Received 8 May 1991)

ABSTRACT

Five hundred fifty DNA fragments 100–500 base pairs in length were cloned from total chromosomal DNA of *Escherichia coli*, each capable of promoting the synthesis of β-lactamase when inserted upstream of the *ampC* structural gene without its own promoter in a promoter-probe plasmid. All clones in this library of putative promoters were classified based on the level of resistance to ampicillin, which ranged from 10 to more than 1,500 µg/ml. Most of the higher levels of drug resistance (more than 1,000 µg/ml) were due not only to an increase in gene expression but also to an increase in plasmid copy number. The DNA fragments which produced the highest level of drug resistance all mapped at 5.7 min on the *E. coli* chromosome and shared the same nucleotide sequence. In these fragments, a strong promoter was found, which carries an up stream AT-rich sequence in addition to −35 and −10 signals of the promoter consensus.

1. INTRODUCTION

Genetic and biochemical studies on a number of *Escherichia coli* promoters have generated the concept that, in promoters recognized by the major form (Eσ70) of RNA polymerase holoenzyme, two hexanucleotide consensus sequences exist at the −35 and −10 regions upstream of the transcription initiation site, separated by 16–18 bp (base pairs) whose sequence is usually less important (Robenberg and Court, 1979; Siebenlist et al., 1980; Hawley and McClure, 1983; Harley and Reynolds, 1987). In promoter complexes, the RNA polymerase associates with a region of about 60 bp of DNA including the −35 and −10 signals (for reviews see Chamberlin, 1974; von Hippel et al., 1984). Such promoter signals can be experimentally recognized by their ability to direct the expression of structural genes lacking their own promoter, when inserted upstream of these genes. For instance, genes encoding enzymes such as β-lactamase and β-galactosidase have been used for promoter screening in vivo.

* This is paper no. 10 in the series entitled “Promoter selectivity of *Escherichia coli* RNA polymerase”.

DDBJ accession number: The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and GenBank Nucleotide Sequence Database under the accession number D01031.

† Corresponding author.
Analyses of various *E. coli* promoters have indicated, however, that there are considerable differences in their sequences, which are considered to be related to variations in promoter strength. To test this idea, we developed the *in vitro* mixed transcription assay and determined the strength of a number of known promoters (for reviews see Ishihama, 1988; 1990). These observations, together with other measurements using the abortive initiation assay (McClure, 1985), indicated that each promoter indeed has specific and unique values for two major parameters determining its strength, *i.e.*, the affinity for RNA polymerase (parameter I) and the rate of open complex formation (parameter II). To reveal the detailed relationship between promoter strength and sequence, we have begun a systematic measurement of a collection of synthetic promoters with all possible sequences. Roles of individual bases within the promoter −35 signal have so far been clarified (Kobayashi et al., 1990). As a further extension of this line of studies, efforts are also being made in our laboratory to correlate these *in vitro* promoter strengths with their *in vivo* activities.

In parallel with the strength measurement of both known promoters and synthetic promoters, we have initiated the isolation of a collection of natural promoters from total chromosomal DNA of *E. coli*, the classification of these promoters based on their strength, and the identification of genes whose transcription is directed by these promoters. In this report, we describe an approach for constructing a promoter library using a promoter cloning plasmid carrying the ampC (*β*-lactamase) gene as a reporter gene for primary screening and the classification of these putative promoters based on the level of ampicillin resistance produced. Some representative promoters producing different levels of drug resistance have been mapped and sequenced.

2. MATERIALS AND METHODS

1. *Bacterial strains and plasmids*

   *E. coli* strain W3350 was used as a DNA donor strain for construction of a promoter collection. Strain DH1 was used as a recipient for transformation of promoter plasmids. Cells were grown at 37°C in LB-0.2% glucose medium. The transcriptional fusion plasmid pJAC4 (Jaurin, 1987) was used as a promoter-probe vector. The reporter gene present in this plasmid is ampC, the *β*-lactamase gene derived from the *E. coli* chromosome.

2. *Cloning of promoter fragments*

   Total chromosomal DNA of *E. coli* W3350 was digested with *Sau3AI* and fractionated by electrophoresis on a 1.5% agarose gel. Fragments 100–500 bp in length were eluted from the gel, purified by passing through a QIAGEN-tip (QIAGEN, USA), and ligated with pJAC4, which had previously been treated with *BamHI* and *E. coli* alkaline phosphatase. After transformation into *E. coli*
DH1, colonies were isolated on LB plates containing 50 μg/ml kanamycin and 10 μg/ml ampicillin. Plasmid DNA was isolated from each transformant and the size of DNA insert was determined by gel electrophoresis.

3. Classification of promoter plasmids

Recombinant plasmids carrying *E. coli* DNA fragments were used to re-transform *E. coli* DH1 and the level of ampicillin resistance of each transformant was determined by measuring the ability to form colonies on LB plates containing 50 μg/ml kanamycin and various concentrations of ampicillin ranging from 10 to 5,000 μg/ml.

4. Mapping of promoter fragments

Phage lambda DNA was prepared from phage lysates of Kohara's mini-set clone library (Kohara et al., 1987). Approximately 5 to 50 ng of lambda DNA was spotted onto nylon membranes (Hybond™, Amersham, UK). For mapping of promoter fragments, the inserts were isolated from each recombinant plasmid by cleaving at *Sph*I and *Eco*RI sites, which are located near the insert-vector junctions in pJAC4 (see Fig. 1), and labeled by the random primed incorporation of digoxigenin-labeled deoxyuridine triphosphate using a commercially available kit (Boehringer, Mannheim, Germany). Labeled fragments were hybridized with the DNA membranes. Prehybridization was carried out overnight at 68°C for 3–12 hr in 5×SSC, 0.5% blocking reagent (Amersham, UK), 0.1% N-laurylsarconine (Na salt) and 0.2% SDS. Hybridization was carried out overnight at 68°C in the same solution containing labeled probe. Membranes were washed twice with 2×SSC, 0.1% SDS for 5 min at room temperature, followed by a wash with 0.1×SSC and 0.1% SDS at 68°C for 15 min.

5. DNA sequencing

DNA sequence was determined by Sanger's dideoxy chain termination method (Sanger et al., 1977) with two oligonucleotide primers, *i.e.*, (5')AGCGTCGT-TTTGAAACAT(3') and (5')CAGGATGCGTCCGGCGTA(3'), which anneal to the anti-sense and sense strands, respectively, of pJAC4 just upstream of the vector-insert junctions. Plasmid DNA was directly used as the template for sequencing. Sequences were determined in both directions.

6. Primer extension analysis

For determination of transcription initiation sites, the primer extension analysis was performed according to Fujita and Ishihama (Fujita and Ishihama, 1987). The oligonucleotides used for DNA sequencing were labeled with [γ-32P]ATP and T4 polynucleotide kinase and used as primers for reverse transcription. Total cellular RNA was prepared from *E. coli* W350 by the hot phenol method (Nevins, 1987) and used as the template. Mixtures of 7.5 pmol labeled oligonuc-
leotides and 150 µg total RNA were incubated overnight at 30°C and then adjusted to 50 mM Tris-HCl (pH 8.2), 100 mM KCl, 10 mM MgCl₂ and 0.28 mM each of dATP, dGTP, dCTP and dTTP in a final volume of 50 µl. After addition of 20 units of highly purified AMV reverse transcriptase, DNA synthesis was carried out for 90 min at 42°C. Products were analyzed by electrophoresis on an 8% polyacrylamide sequencing gel containing 7 M urea, and autoradiography.

7. Enzymes and chemicals

Restriction enzymes, E. coli alkaline phosphatase, T4 polynucleotide kinase and T4 DNA ligase were purchased from Takara, Japan. Reverse transcriptase was purified from avian myeloblastosis virus (AMV) according to the method by Kato et al. (1984). Labeled and unlabeled nucleotides were purchased from Amer- sham, UK, and PL-Phamacia, USA, respectively.

The gene mapping membrane carrying all λ clones of the Kohara mini-set library was prepared by T. Nagata-Suzuki of the Gene Library Laboratory of this institute.

3. RESULTS AND DISCUSSION

1. Construction of a promoter collection

In order to construct a collection of E. coli promoters, total chromosomal DNA

![Diagram](image)

Fig. 1. Strategy of cloning and analysis of promoter plasmids. Sau3AI fragments of total E. coli DNA of 100–500 bp in length were cloned into a promoter-probe vector pJAC4 at BamHI site within a multi-cloning site. For gene mapping and DNA sequencing, inserts were isolated following digestion of plasmid DNA with SphI and EcoRI. SD, ribosome binding site. XhoI, PstI and BglII are other major restriction enzyme sites within the cloning vector.
was digested with Sau3AI and fragments 100–500 bp in length were inserted into a promoter-cloning vector pJAC4, at a BamHI site located just upstream of the β-lactamase (ampC) gene (Fig. 1). Since the ampC gene on pJAC4 retains its own ribosome-binding signal, recombinant plasmids carrying suitably located promoters on the inserts could be selected by screening transformants resistant to both kanamycin and ampicillin (Jaurin, 1987). First, five hundred fifty independent transformants were screened, which were resistant to a low concentration (10 μg/ml) of ampicillin. The proportion of AmpR recombinants was approximately a few percent of the total recombinants carrying E. coli DNA fragments. This suggests that, in the E. coli chromosome, a promoter exists in every 0.5–1.0×10⁴ bp segment (total 500–1000 promoters), although a certain class of promoters might have not been isolated by the method employed (see DISCUSSION below). Since the ampicillin resistance thus observed might have arisen not only from promoter insertion but also from chromosomal mutations in the recipient cells leading to ampicillin resistance, plasmids were recovered from these transformants and checked for the presence of DNA inserts. Moreover,

Fig. 2. Classification of transformants on the basis of ampicillin resistance. Five hundred and fifty plasmid clones with putative promoter activity were isolated from transformants resistant to 10 μg/ml ampicillin. Each transformant was then grown in LB plates containing various concentrations of ampicillin ranging from 10 to 1,500 μg/ml, and the maximum concentration that allowed cell growth was determined. Plasmids isolated from all these transformants were re-transformed into E. coli DH1, and the level of ampicillin resistance of each transformant was checked again by measuring the ability to form colonies on LB plated containing 50 μg/ml kanamycin and various concentrations of ampicillin. Four hundred and seventy one clones which exhibited the same level of ampicillin resistance after plasmid isolation and retrasformation are classified based on the drug resistance level.
detailed analyses of the recombinant plasmids carrying *E. coli* DNA fragments with putative promoter activity (hereafter designated “promoter plasmids”) were performed only after re-transformation of isolated promoter plasmids into *E. coli* DH1.

Because a linear correlation exists between the levels of β-lactamase production and ampicillin resistance (Jaurin et al., 1982), promoter activity can be estimated roughly by measuring the maximum concentration of ampicillin where the corresponding transformant can grow. Thus, each transformant was exposed to a series of concentrations of ampicillin. The results, summarized in Fig. 2, indicated that the level of ampicillin to which promoter plasmid-containing cells were resistant ranged from 10 μg/ml to more than 1,500 μg/ml. Approximately 60% of total recombinant promoter plasmids gave resistance to 10–200 μg/ml ampicillin, and more than 77% produced resistance to concentrations between 10 and 500 μg/ml; but only about 5% conferred resistance to ampicillin at concentrations higher than 1,000 μg/ml. It should be important to note that, at the ampicillin concentration (50 μg) which is widely used for gene cloning, about one fifth of the *E. coli* putative promoters with weak promoter activity escaped screening.

![Graphs A and B](image)

**Fig. 3.** Size and copy number of the promoter plasmids. Cells of *E. coli* DH1 carrying one of the promoter plasmids were mixed with pJAC4-carrying DH1 cells, and two kinds of plasmids were isolated simultaneously. [A] The sizes of DNA inserts in the promoter plasmids were determined by measuring electrophoretic mobility on polyacrylamide gels. [B] The copy numbers of promoter plasmids in transformants were determined by measuring the amounts of DNA, relative to pJAC4 isolated simultaneously from pJAC4-transformed DH1.
In order to test the possibility that the ampicillin resistance was due not only to the insertion of promoters, but also to an increase in gene dosage, the copy number of the promoter plasmid relative to that of pJAC4 was determined for some representatives from each resistance group. Among about 110 transformants examined, most of those showing resistance to ampicillin concentrations higher than 800 $\mu$g/ml were found to contain increased numbers of the promoter plasmid (Fig. 3). Within the transformants having the same plasmid copy number as pJAC4, the range of promoter activity as measured by the level of ampicillin resistance was roughly 100-fold. This might suggest that E. coli promoters vary by at least 100-fold in terms of their intrinsic activities. Although we used a low concentration (10 $\mu$g/ml) of ampicillin, the weakest promoters such as those expressed once per generation might not be detected. Likewise, the strongest promoters might have escaped the present cloning into a multi-copy plasmid due to interference with cell growth by depleting RNA polymerase. Some promoters may require limiting transcription factors, and therefore be less efficient when present in high copy. Use of a low copy number plasmid may improve the promoter cloning. In addition, the cloning strategy employed here does not exclude influence of DNA insertion upstream of the ampC ribosome-binding signal on the translation efficiency of ampC mRNA, for example, due to insertion of additional ribosome-binding sequence, initiation and/or termination codons, or to formation of mRNA secondary structures.

2. **Characterization of some promoters**

Examination of the size of DNA inserts indicated that the majority of promoter plasmids carried inserts of the expected size (100–500 bp in length); but many plasmids recovered from transformants showing resistance only to low levels of ampicillin (less than 100 $\mu$g/ml) contained inserts larger than 1,000 bp (Fig. 3). Following analyses of restriction enzyme sites and, in some cases, nucleotide sequences, these large inserts were found to be attributable either to insertion of multiple DNA fragments or to contamination with large fragments in the original DNA samples. The low levels of drug resistance observed for transformants carrying these large inserts might be due to: inefficient elongation of transcription within long 5'-flanking sequences upstream of the ampC gene (when promoters are located near the upstream end of the inserts); interference with translation of the ampC mRNA carrying long leader sequences (although the reporter ampC gene carries its own SD signal); or decreased stability of the fused ampC mRNA.

In order to determine the genes from which these promoters were derived, we first tried to locate the relevant fragments on the E. coli chromosome by hybridization of labeled inserts with a mini-set library of lambda phage clones, which covers the entire chromosome of E. coli (Kohara et al., 1987). So far, we have examined representative promoter plasmids from each of the groups showing different levels of ampicillin resistance. The results are summarized in Table 1.
Table 1. Mapping of cloned promoters

<table>
<thead>
<tr>
<th>Ampicillin resistance (µg/ml)</th>
<th>Clone number</th>
<th>Size of insert (bp)</th>
<th>Relative copy number</th>
<th>Lambda clone number</th>
<th>Map position (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50–75</td>
<td>519</td>
<td>840</td>
<td>1.5</td>
<td>416/417</td>
<td>52</td>
</tr>
<tr>
<td>150–175</td>
<td>335</td>
<td>160</td>
<td>0.8</td>
<td>451/452</td>
<td>59</td>
</tr>
<tr>
<td>250–300</td>
<td>472</td>
<td>190</td>
<td>1.7</td>
<td>416/417/420</td>
<td>52</td>
</tr>
<tr>
<td>400–450</td>
<td>220</td>
<td>280</td>
<td>1.5</td>
<td>573/574</td>
<td>81</td>
</tr>
<tr>
<td>450–500</td>
<td>290</td>
<td>270</td>
<td>0.7</td>
<td>470</td>
<td>63</td>
</tr>
<tr>
<td>600–650</td>
<td>308</td>
<td>315</td>
<td>1.0</td>
<td>316</td>
<td>36</td>
</tr>
<tr>
<td>700–750</td>
<td>430</td>
<td>250</td>
<td>1.0</td>
<td>132/133</td>
<td>7</td>
</tr>
<tr>
<td>750–800</td>
<td>3</td>
<td>410</td>
<td>1.0</td>
<td>157</td>
<td>321/322</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>675/676</td>
<td>0</td>
</tr>
<tr>
<td>800–850</td>
<td>125*</td>
<td>210</td>
<td>1.5</td>
<td>126/127</td>
<td>5.7</td>
</tr>
<tr>
<td>850–900</td>
<td>318</td>
<td>810</td>
<td>1.5</td>
<td>254</td>
<td>28</td>
</tr>
<tr>
<td>900–950</td>
<td>492</td>
<td>365</td>
<td>5.0</td>
<td>176</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>622</td>
<td>74</td>
</tr>
<tr>
<td>950–1000</td>
<td>413</td>
<td>255</td>
<td>6.0</td>
<td>132/133</td>
<td>7</td>
</tr>
<tr>
<td>1100–1200</td>
<td>267*</td>
<td>210</td>
<td>5.5</td>
<td>126/127</td>
<td>5.7</td>
</tr>
<tr>
<td>1300–1400</td>
<td>270*</td>
<td>210</td>
<td>7.7</td>
<td>126/127</td>
<td>5.7</td>
</tr>
<tr>
<td>&gt;1500</td>
<td>456*</td>
<td>210</td>
<td>8.6</td>
<td>126/127</td>
<td>5.7</td>
</tr>
</tbody>
</table>

* Clones 125, 267, 270 and 456 shared the same DNA insert but exhibited different levels of ampicillin resistance.

Several of the promoter plasmids yielding high levels of ampicillin resistance, e.g., p125, p267, p270 and p456, hybridized with the same lambda clone, indicating that there is a strong promoter(s) at 5.7 min region on the E. coli chromosome. In addition, there might be some source of increased plasmid copy number in this region, because the plasmid copy numbers were higher for at least three of these (p267, p270 and p456) than for the parent plasmid pJAC4.

Some plasmids such as p430, p3 and p492 hybridized to two sets of lambda phages, each carrying a different region of the chromosome. Sequence analysis showed that, in those cases, two different fragments had been inserted together into a single recombinant plasmid. For instance, clone 3 contained two E. coli DNA fragments: one from the cet gene at 0 min and another yet unpublished sequence from 11 min. In these cases, promoter strength must be determined after recloning each fragment.

Finally, DNA sequences were determined for some of these promoter fragments, to further identify their origin. In addition, the start site of transcription was established for some promoters, using the primer extension method. Four plasmids, p125, p267, p270 and p456, which showed the highest level of ampicillin resistance were found to share the same sequence (Fig. 4). Three of these (p267,
Fig. 4. DNA sequence of the promoter plasmids with high levels promoter activity, and distribution of A+T sequences. DNA sequences of inserts within p267, p270 and p456 were determined by the dideoxy chain termination method as described in Materials and Methods. All these promoter plasmids carried the same 
Stu3A fragment shown. The start site of transcription determined by the primer extension method according to Fujita and Ishihama (1987) is shown by arrows, while the promoter −35 and −10 signals are shown by bars. The distribution of A+T (or G+C) in this DNA sequence region was analyzed using a 10 bp window.

p270 and p456) share exactly the same sequence in the same orientation. The different levels of drug resistance might be due, at least in part, to the difference in plasmid copy number. When the plasmid p456 was reintroduced into the original recipient strain, the drug resistance was found to be within the range of 1200–1300 μg/ml. This indicates that there was an additional mutation(s) in the chromosomal DNA of this particular transformant. Within the DNA insert in these plasmids, just upstream of the transcription start site determined by reverse transcriptase mapping, are −10 and −35 regions with 5/6 and 4/6 identities to the consensus promoter sequence, and an 18 bp spacing. In addition, an AT-rich sequence is present upstream of the consensus promoter as has previously been found for genes expressed at high levels in exponentially growing cells (Lamond and Travers, 1983; Buck et al., 1986; Jo et al., 1986).

In contrast, p125, showing weaker drug resistance, was again found to include the same sequence which had, however, been inserted in the opposite direction. A promoter-like sequence can be found to exist even in this backward direction although the homology is less than that in the forward direction present in p267, p270 and p456. The sequence present in these four plasmids could not be found in DNA data bases (GENBANK, EMBL and DDBJ). Quantitative comparison of the promoter strength between the strong promoters isolated in this study and those of known genes is in progress using the in vitro mixed transcription assay.

We thank B. Jaurin for providing us with promoter-probe plasmid pJAC4, T. Nagata-Suzuki for
preparation of DNA filters, and R. S. Hayward for advices and critical reading of the manuscript. This work was supported by Grants-in-Aid from the Ministry of Education, Science and Culture, a NIG Cooperative Research Program and a Joint Studies Program of The Graduate University for Advanced Studies.

REFERENCES


