An efficient method for scoring base pair interactions

Frederick R. Blattner

Professor of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA

Received 31 October 1983; Accepted 20 November 1983

ABSTRACT

A simple method is described for speeding up the computation of base pairing interactions. It is especially effective on microcomputers.

In comparing nucleic acid sequences for homology it is often desirable to assign a separate score to the strength of interaction between each of the possible base pairs. Usually this is done with a two dimensional matrix in which the $i,j$th element contains the score assigned to the interaction of the $i$th with the $j$th base. If the 4 standard bases plus one symbol for an unknown base are allowed in the sequence, a 5 by 5 symmetric matrix is required. When searching large sequences for homology, a large number of comparisons must be done and the lookup of scores in the matrix comprises a major part of the task. I therefore investigated ways of improving the efficiency of the process. I found that if the 5 bases are represented by the numbers 0, 1, 3, 7, and 12, a 24 element linear array, could be constructed in which the score for the $ij$th interaction is stored in the $(i+j)$th element of the array. With this modification our sequence comparison program ran 6 times faster and required one less element of storage for the scoring array.

The key to the method is the selection of numbers to represent individual bases which have unique pairwise sums. The sets are constructed recursively. After obtaining such a set of $n$ elements, the $(n+1)$st set is found by adding, to the set of $n$ elements, the next largest integer which results in unique pairwise sums. The first 40 such numbers beginning with zero are listed in Table 1. There appears to be no simple pattern to the series, although if it is started with a number different from zero the entire series is simply offset. As shown in the table, memory required for a linear scoring matrix begins to exceed that required for a square matrix beyond 5 elements, but not drastically so.

A Z80 based microcomputer was used for the speed test. In
such computers, multiplication is accomplished by an algorithm based on addition so the increase in speed is primarily due to the elimination of the multiplication needed to address a two dimensional array. If a computer with hardware multiplication were to be used the speed increase would be less significant. Thus, this method is of greatest utility for microcomputers, since these rarely are equipped with hardware multiplication.

<table>
<thead>
<tr>
<th>SET SIZE</th>
<th>SET ELEMENT</th>
<th>MEMORY REQUIRED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>7</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>8</td>
<td>44</td>
<td>88</td>
</tr>
<tr>
<td>9</td>
<td>65</td>
<td>130</td>
</tr>
<tr>
<td>10</td>
<td>80</td>
<td>160</td>
</tr>
<tr>
<td>11</td>
<td>96</td>
<td>192</td>
</tr>
<tr>
<td>12</td>
<td>122</td>
<td>244</td>
</tr>
<tr>
<td>13</td>
<td>147</td>
<td>294</td>
</tr>
<tr>
<td>14</td>
<td>181</td>
<td>362</td>
</tr>
<tr>
<td>15</td>
<td>203</td>
<td>406</td>
</tr>
<tr>
<td>16</td>
<td>251</td>
<td>502</td>
</tr>
<tr>
<td>17</td>
<td>289</td>
<td>578</td>
</tr>
<tr>
<td>18</td>
<td>360</td>
<td>720</td>
</tr>
<tr>
<td>19</td>
<td>400</td>
<td>800</td>
</tr>
<tr>
<td>20</td>
<td>474</td>
<td>948</td>
</tr>
<tr>
<td>21</td>
<td>564</td>
<td>1128</td>
</tr>
<tr>
<td>22</td>
<td>592</td>
<td>1184</td>
</tr>
<tr>
<td>23</td>
<td>661</td>
<td>1322</td>
</tr>
<tr>
<td>24</td>
<td>774</td>
<td>1548</td>
</tr>
<tr>
<td>25</td>
<td>821</td>
<td>1642</td>
</tr>
<tr>
<td>26</td>
<td>915</td>
<td>1830</td>
</tr>
<tr>
<td>27</td>
<td>969</td>
<td>1938</td>
</tr>
<tr>
<td>28</td>
<td>1015</td>
<td>2030</td>
</tr>
<tr>
<td>29</td>
<td>1158</td>
<td>2316</td>
</tr>
<tr>
<td>30</td>
<td>1311</td>
<td>2622</td>
</tr>
<tr>
<td>31</td>
<td>1394</td>
<td>2788</td>
</tr>
<tr>
<td>32</td>
<td>1522</td>
<td>3044</td>
</tr>
<tr>
<td>33</td>
<td>1571</td>
<td>3142</td>
</tr>
<tr>
<td>34</td>
<td>1820</td>
<td>3640</td>
</tr>
<tr>
<td>35</td>
<td>1895</td>
<td>3790</td>
</tr>
<tr>
<td>36</td>
<td>2028</td>
<td>4056</td>
</tr>
<tr>
<td>37</td>
<td>2253</td>
<td>4506</td>
</tr>
<tr>
<td>38</td>
<td>2378</td>
<td>4756</td>
</tr>
<tr>
<td>39</td>
<td>2509</td>
<td>5018</td>
</tr>
<tr>
<td>40</td>
<td>2779</td>
<td>5558</td>
</tr>
</tbody>
</table>
Transcription analysis of the *lexA* gene of *Escherichia coli*: attenuation and cotranscription with the neighboring region

Toru Miki, Komei Shirabe, Yousuke Ebina and Atsushi Nakazawa

Department of Biochemistry, Yamaguchi University School of Medicine, Ube, Yamaguchi 755, Japan

Received 27 October 1983; Accepted 6 December 1983

**SUMMARY**

The *lexA* gene of *Escherichia coli* encodes a repressor of the genes whose expression is induced by the agents that result in DNA damage. *In vivo* transcripts of the *lexA* gene consisted of two species; mRNA-1 of 673 bases and mRNA-2 of approximately 3,000 bases. The transcription *in vivo* started at a site which was two-base pairs downstream from the *in vitro* initiation site reported previously. The majority of the transcription stopped at a series of T residues preceded by a dyad symmetry located immediately after the *lexA* gene. A small fraction of the transcription passed through the termination site to form the mRNA of downstream gene(s) which would be related to the "SOS functions".

**INTRODUCTION**

When *Escherichia coli* cells are treated with agents such as ultraviolet light (UV) or mitomycin C (MTC), a variety of cellular functions are induced. These "SOS functions" include exision repair, mutagenesis, inhibition of cell division, and prophage and colicin inductions (1). A set of genes responsible for "SOS functions" are repressed by the *lexA* gene product (2). Upon treatment of the cells by the agents which cause degradation of DNA, the *recA* gene product is "activated" to gain a protease activity and then cleaves *lexA* protein (3,4). The proteolytic inactivation of *lexA* protein, which is also the repressor of the *recA* gene, leads to increased synthesis of *recA* protein which stimulates the inactivation of *lexA* protein. Since the *lexA* gene itself is repressed by its own product, expression of the gene is also enhanced by the inducing treatments (5,6).

In a previous work (7), the *lexA* and its neighboring regions were organized. In the cell-free system for protein synthesis, two major protein products were identified to be encoded in this region: one is *lexA* protein (24 kilodaltons (kd)) and the other, a protein of 35 kd. The gene for the 35-kd protein, tentatively named as *lexD* here, was located in the downstream region of the *lexA* gene. Transcription R loop analysis of the *lexA-lexD*...
region showed that the transcription started from a site in front of the \textit{lexA} gene, and did not terminate at a fixed point. A small amount of longer \textit{in vitro} transcripts extended to the \textit{lexD} region. Thus a possibility arose that the \textit{lexD} region was cotranscribed with the \textit{lexA} gene.

This report describes the analysis of the \textit{lexA} transcription \textit{in vivo}. By using SI nuclease mapping and RNA blot hybridization techniques, we determined the initiation and termination sites for transcription of the \textit{lexA} gene. Evidence is also presented that the transcription of the \textit{lexA} gene is attenuated at a sequence located immediately after the \textit{lexA} coding frame.

\section*{MATERIALS AND METHODS}

\textbf{Bacterial Strains and Plasmids:} C600S (\textit{thr, leu, supE}) and W3110 (\textit{gal, lac}) are the laboratory stocks and were used as \textit{lexA}+ strains of \textit{Escherichia coli} K-12. DM1187 (\textit{spr-51, lexA3, tif-1, stfA11, sup-37, thr, leu, his, ilv (ts), pro, gal, rpsL}) (8) was used as a \textit{lexA}(Def) strain in which autogenous regulation of the \textit{lexA} gene is released. KT5371 (F’ \textit{gal3; trpE9851, Achl-bio, recA1}) carrying a defective \textit{\lambda} phage (a gift from K. Shigesada) was used as a \textit{recA}~strain. The genetic markers, except \textit{recA1}, of the strain are not relevant here. A \textit{Lac}~strain MC1000 ($\Delta$ \textit{lacX74, \Deltaara-leu, araD139, galU, galK, rpsL}) and a plasmid carrying the \textit{lacZ} structural gene, pMC1403, were gifts of M. Casadaban (9). The plasmid carries the most part of the \textit{lacZ} structural gene but the first eight codons for the amino-terminal end of \textit{\beta}-galactosidase are lacking (9). pLC44-14 and pMCR551 are the plasmid carrying the \textit{lexA} gene as described previously (7). pMCR780 is a plasmid carrying the tetracycline- and kanamycin-resistance genes (T. Miki, unpublished) and is a derivative of pKN410, a runaway-replication plasmid whose copy number increases at 40°C (10). Bacteria were grown in L broth (1\% tryptone, 0.5\% yeast extract, 0.5\% NaCl, and 0.1\% glucose, pH 7.2).

\textbf{Construction of Plasmids:} Molecular cloning was carried out as described previously (11). T4 DNA ligase and restriction enzymes were purchased from Takara Shuzo (Kyoto, Japan), except BclI, DdeI, NcoI (from Bethesda Research Laboratories), and Pvul (from New England BioLabs). Plasmids were constructed as follows. pMCR811 was constructed by cloning of the DNA fragment containing the \textit{lexA-lexD} region into pMC1403. pMCR551 was completely digested with EcoRI and then partially with AluI. The digested DNA was mixed with pMC1403 DNA which had been cleaved with EcoRI and Smal, and then ligated (Fig. 1A). The \textit{lacZ} deletion strain MC1000 was transformed to ampicillin resistance by the ligated DNA. From the transformants, pMCR811
was isolated as a plasmid carrying the DNA segment from the EcoRI to Alul-3 sites of pMCR551 (Fig. 1B). pMCR821 was constructed by subcloning of the major EcoRI-SalI fragment of pMCR811 (Fig. 1A) into a runaway replication vector pMCR780. pMCR819 was constructed by cloning of the 5.25-kb BamHI-BgIII fragment of pLC44-14 into the unique BamHI site on pMCR780.

**Inducing Treatments:** Induction was done by the addition of MTC (0.8 μg/ml) into the log-phase culture and shaking at 37°C for 1 hr, except in pMCR819-harboring cells which were induced by heat and MTC treatments as follows. W3110(pMCR819) was grown to the log-phase at 30°C, and the culture was shifted to 42°C. After 30 min, MTC (0.8 μg/ml) was added to the culture, and it was further shaken for 30 min. Then the culture was shifted down to 30°C and shaken for additional 30 min. β-Galactosidase activities before and after MTC treatment were determined according to Platt et al. (12).

**Labeling of DNA Fragments:** The 3' ends of DNA fragments were labeled with [α-^32P]dCTP (3,000 Ci/mm; Radiochemical Centre, Amersham) and T4 DNA DNA polymerase (Takara Shuzo, Kyoto). The reaction mixture contained 63 mM MgCl₂, 6.3 mM β-mercaptoethanol, 0.1 mM each of dATP, dGTP, and dTTP, 0.5 μM [α-^32P]dCTP (30μCi), 0.5 pmol DNA fragment, and 3.2 units T4 DNA polymerase in 20 μl. After incubation at 11°C for 130 min, Na₂EDTA was added to give a concentration of 10 mM, and the mixture was heated at 65°C for 10 min. DNA was recovered by ethanol precipitation. The 5' ends of a DNA fragment were labeled with [γ-^32P]ATP (3,000 Ci/mm; Radiochemical Centre, Amersham) and T4 polynucleotide kinase (Takara Shuzo, Kyoto) as described (13). Each strand of the labeled fragments was separated essentially as described by Maxam and Gilbert (13) and extracted as described previously (14). Labeling of DNA fragments by nick translation was carried out with [α-^32P]dCTP and DNA polymerase I (Bethesda Research Laboratories) as described by Rigby et al. (15).

**Preparation of RNA:** RNA was prepared by the hot phenol procedure from induced or non-induced cells as described (16).  

**SI Nuclease Mapping:** The SI mapping assays were carried out as described by Aiba et al. (18). The ^32P-labeled probe DNA (5 x 10⁴ cpms) and 20 μg of RNA were denatured, hybridized, and digested with SI nuclease (Boehringer Mannheim). The products were analyzed by electrophoresis on 5% polyacrylamide gels containing 8 M urea at 40 mA for 14 hr. Radioactive bands were detected by autoradiography. Principles of the experiments are illustrated in Fig. 2. When the initiation or termination site was determined, the samples were co-electrophoresed with the labeled probe DNA.
that was chemically cleaved as described by Maxam and Gilbert (13).

RNA Blot Hybridization: The RNA blot hybridization was carried out as described by Thomas (18). Glyoxal-denatured RNAs were separated by agarose gel electrophoresis. RNA bands were blotted to nitrocellulose sheets and hybridized with DNA probes labeled by nick translation technique with \( \alpha^32P \) dCTP and DNA polymerase I (Bethesda Research Laboratories). Endogeneous 16S- and 23S-ribosomal RNAs in the RNA preparation and brome mosaic virus RNAs (a gift of I. Furusawa) were used as size markers.

RESULTS

**lexA mRNAs Consist of Two Species:** To know whether the *lexD* region is cotranscribed with the *lexA* gene, SI mapping analysis of the *lexA* transcript was performed. A plasmid pMCR8II carrying the *lexA* gene and *lexD-lacZ* fusion was constructed as shown in Fig. 1A, and the major EcoRI-Sall fragment was subcloned into a runaway-replication vector pMCR780, to produce pMCR821. The 2.8-kb MluI-C fragment carrying most

![Fig. 1. Restriction Maps of the Plasmids and DNA Fragments Used: Open, hatched, and closed portions represent vector (pBR322 or pACYC177), lac operon, and *lexA-lexD* region DNAs, respectively. (A) Construction of pMCR8II. DNA segment between EcoRI and AluI-3 sites of pMCR551 was cloned into pMC1403 as described in MATERIALS AND METHODS. Specific activity of \( \beta \)-galactosidase in MC1000(pMCR8II) was low (2.1 units per absorbance at 660 nm of the culture) and increased two fold by MTC treatment. Direction of transcription of the *lexA* gene is indicated by arrows. (B) Plasmids and DNA fragments used. Locations of the *lexA* and *lexD* structural genes are shown under the restriction map of the *lexA-lexD* region. Locations of DNA fragments used in this study are shown by ---.](image-url)
part of the \( \text{lexA} \) gene and the \( \text{lexD-lacZ} \) fusion (Fig. 1B) was isolated from pMCR821. The 3' ends of the fragment were labeled with \( ^{32}\text{P} \). Both the strands were isolated and used as probes. Principles of the experiments are illustrated in Fig. 2. Each probe was hybridized with RNA prepared from MTC-induced or non-induced cells without \( \text{lexA} \)-plasmids, and the DNA-RNA hybrids were digested with nuclease SI to eliminate single-stranded portions. If the transcription stopped at a site immediately after the \( \text{lexA} \) gene, the product of SI digestion would be approximately 550-base pairs (bp) long (Fig. 2; product a). On the other hand, if the transcription further proceeded up to the \( \text{lexD} \) region, the product of SI digestion would be much longer. The RNA, however, should be hybridized only with the \( \text{lex} \)
part (corresponding to the major MluI-BamHI subfragment) of the sense strand of the MluI-C fragment, because the remaining portion of the probe was derived from the lacZ gene. Thus the length of the product of SI digestion should be approximately 1.6 kb (Fig. 2, product b). The product could be separated by electrophoresis from the probe, which had been protected from SI digestion by a small amount of the complementary strand contaminating the probe preparation. Results are shown in Fig. 3. By using the slow strand of $^{32}$P-labeled MluI-C fragment as a probe, an expected product of 1.6-kb was detected (lane r, band 2). These results clearly showed that the lexA and lexD genes were cotranscribed. A shorter product of approximately 550 bases was also found (lane r, band 3). Detection of the two transcripts of the lexA gene suggested that the shorter transcript was produced by attenuation of the transcription or by processing of the longer transcript. The band 1 was considered to be the probe protected from SI digestion by the contaminating complementary strand, because the band was also detected in the assay without RNA (lane p) and had the same mobility as the intact probe (compare lanes d, q, and r). The SI assays by using the fast strand showed no distinct band (lanes m-o), indicating that this strand was the anti-sense strand. The assay by using the major MluI-Pvul subfragment of 1.7 kb as a probe also showed the longer and shorter transcripts (lanes k and l), although the longer one was mostly hidden by the band of the single-stranded MluI-Pvul subfragment (compare lanes e, k, and l). The lowest band in the lanes k and l was identified to be the double-stranded subfragment by comparing lanes e (the denatured subfragment) and f (the double stranded subfragment) with lanes k and l. Synthesis of both the transcripts was stimulated by MTC treatment of the cells (compare lanes q and r).

**lexA Transcription Extends to 3-kb Downstream from the Promoter:** The SI mapping data showed that the longer transcript extended, at least, to the Alul-3 site (see Fig. 1B) in the lexD region. To estimate sizes of the mRNAs, lexA mRNAs were analyzed by blot hybridization. The HincII-C and HaeII-A fragments shown in Fig. 1B were labeled with [$\alpha$-$^{32}$P]dCTP by nick-translation, and used as probes for lexA mRNAs. Total cellular RNAs were isolated from MTC-treated or non-treated cells of W3110 and the strain harboring lexA-plasmids. The RNAs were separated by agarose gel electrophoresis under denatured conditions, transferred to nitrocellulose sheets, and then hybridized with the probes. Results are shown in Fig. 4. Two RNA species from W3110(pMCR551) were hybridized with the HincII-C
Fig. 3. SI Mapping Analysis of *lexAm*RNAs: (A) Reference DNAs: a) denatured MluI-C fragment, b) native (double-stranded) MluI-C fragment, c) fast strand (the strand which migrated faster than the other strand in the strand separation gel) of the MluI-C fragment, d) slow strand of the MluI-C fragment, e) denatured major MluI-PvuI subfragment, f) native major MluI-PvuI subfragment, g) denatured minor MluI-PvuI subfragment, h) native minor MluI-PvuI subfragment (The fragment ran off from the gel.), and i) denatured preparation of the BamHI-cleaved MluI-C fragment. The lengths of single stranded markers are shown on the left side of the figure in kilobases.

(B-D) SI mapping analysis of the *lexA* mRNA: The MluI-PvuI major fragment and the fast and slow strands of the MluI-C fragment were used as probes in B, C, and D, respectively. Lanes j, m, and p; without RNA, and lanes k, n, and q; with RNA from non-induced cells, and lanes l, o, and r; with RNA from MTC-induced cells of C600S.

probe which was located within the *lexA* structural gene (Fig. 4A, lanes a and b; indicated by 1 and 2), indicating that these RNAs were the transcripts of the *lexA* gene. Synthesis of both the mRNAs was
Fig. 4. Detection of Intact \textit{lexA} mRNA\textsubscript{s} by Blot Hybridization: RNA was extracted from the cells of W3110 harboring pMCR551 (lanes a and b), pMCR819 (lanes e and f), or no plasmid (lanes c, d, g, and h). Before extraction, the cells were induced by MTC (lanes b, d, and h) or MTC and heat (lane f). After electrophoresis under the denatured conditions, the bands were blotted to a sheet of nitrocellulose paper and hybridized with the nick-translated HincII-C (A) and HaeIII-A (B) fragments. Locations of the bands for size marker RNAs are indicated by arrows: 16S- (1541 bases), and 23S- (2904 bases) ribosomal RNAs (19,20) were used in A and B. Four species of brome mosaic virus RNAs B1 (3270 bases), B2 (2970 bases), B3 (2250 bases), and B4 (840 bases) (21) were used in A. The film for B was overexposed to show the band for mRNA-2. Locations of the bands for mRNA-1 and mRNA-2 are indicated as 1 and 2, respectively.

MTC-inducible (compare lanes a and b). The size of the shorter transcript, mRNA-1, was estimated to be approximately 650 bases and agreed well with calculation from the data of SI mapping analysis. The longer transcript, mRNA-2, was estimated to be approximately 3,000 bases long, indicating that the transcription terminated near the PstI site which corresponded to an end of the insert of pMCR551. The major band on the top seemed to be formed by hybridization of pMCR551 DNA with the probe, because these bands were not observed when RNAs from the cells without the plasmid were used (see, lanes c and d). In our hands, multicopy plasmids tend to contaminate RNA preparations during the extraction procedure. With RNAs from the plasmid-free cells, mRNA-2 was failed to be detected (lanes c and d). To determine whether transcription for mRNA-2 terminated at a site in the insert of pMCR551 or not, RNA prepared from W3110 harboring pMCR819 which carried the longer insert, the BamHI-BgIII fragment (see Fig. 1B),
was used. The Haelll-A fragment (Fig. 1B) was used as a probe for the hybridization experiment. The sensitivity of detection would increase by using the longer probe. The lexA mRNAs could be efficiently amplified by heat treatment of the cells harboring pMCR819, because the vector portion of the plasmid was derived from a runaway replication plasmid. Indeed, higher amounts of both the mRNA species were detected as shown in Fig. 4B (lanes e and f). The size of mRNA-2 was nearly the same as that in the former experiments with RNA from pMCR551-harboring cells. In the results shown in Fig. 4B, both the mRNA species were detected even in RNA from the plasmid-free cells (lanes g and h), indicating that mRNA-2 was also synthesized from the lex operon on the chromosome.

Transcription Initiation Site In Vivo Is 2-bp Downstream from That In Vitro: The initiation site of in vitro transcription of the lexA gene was previously determined (7,22). To know whether the initiation site in vivo is different from that in vitro, the in vivo site was determined by SI mapping. The 180-bp EcoRI-MluI fragment containing the lexA promoter (Fig. 1B) was labeled at the 5' ends, and the sense strand (identified by sequence analysis) was used as a probe. The probe was hybridized with RNA extracted from DM1187 which was expected to synthesize lexA mRNAs at a high level. After SI digestion, the sample was electrophoresed together with the chemically-cleaved probe DNA. As shown in Fig. 5, the probe protected by mRNA gave an intense band, when 1,000 units of SI was used under standard conditions (lanes 1 and 3). Several faint bands above the intense band were observed after digestion with 500 units of SI (lane 2), indicating that this amount of the enzyme was insufficient. No bands shorter than the main band were observed, suggesting that degradation of the RNA did not occur during the experiments. When transcription started at a site further upstream from the EcoRI site, a full length of the probe would be protected by the mRNA. The intensities of the bands for the intact probe in the SI assays with and without RNA were almost the same (data not shown), indicating that the intact probe had been protected by the contaminating complementary strand from SI digestion. Thus only slight level of transcription, if any, started at a site further upstream from the lexA promoter and extended to the lexA gene. The length of the major product of SI assay was determined by comparison with the fragments obtained by chemical cleavage of the original probe (see lanes T and 3). The 1.5-bp correction, however, is usually applied in making this comparison (23), because the mode of reactions of SI digestion and chemical
Fig. 5. SI Mapping Analysis of the Initiation Site for lexA Transcription:
The sense strand of the 181-bp EcoRI-MluI fragment (Fig. 1) labeled at its
5' ends was hybridized with 20 μg of RNA extracted from DM1187 and digested
with 1000 units (lanes 1 and 3) or 500 units (lane 2) of SI nuclease at 37°C
for 45 min. A portion of each sample was analyzed by electrophoresis. One
tenth of the sample in lane 1 was electrophoresed in lane 3. The probes
chemically cleaved by the G, A>G, C, and C+T reactions (13) were electro-
phoresed in lanes G, A, C, and T, respectively. A part of the sequence
determined from the cleavage patterns is shown on the right side with that
of the complementary strand. The transcription initiation site is shown by
an arrow head.

cleavage are different. With this correction, the initiation site in vivo
was determined to be 2-bp downstream from that in vitro (Fig. 7B). By
using RNA from a recA~ strain KT5371, the same result as in Fig. 5 was
obtained (data not shown).

lexA Transcription Is Attenuated at a Series of T Residues Immediately
After a Dyad Symmetry: The results obtained so far indicated that the
majority of the transcription from the lexA promoter terminated at a site,
t1, located about 650 bp from the initiation site. A 187-bp Ddel-NcII
fragment (Fig. 1B) containing t1 was labeled with 32P at the 3' ends and
the sense strand (identified by the sequence analysis) was used as a probe
to determine the termination site. The probe was hybridized with RNAs
prepared from MTC-induced or non-induced cells of C600S and RNA from cells
of KT5371 (a recA~ strain). After SI digestion, samples were
electrophoresed with the chemically-cleaved probe DNA. As shown in Fig. 6,
lane 3, some bands were seen as the probe protected by mRNAs from SI
Fig. 6. SI Mapping Analysis of the Termination Site 1 for the lexA Transcription: The sense strand of the 188-bp Ddel-NcoI fragment (Fig. 1B) labeled at its 3' ends was hybridized with 20 μg of RNA prepared from non-induced (lane 1) or MTC-induced (lane 2) cells of C600S, or from the cells of KT5371 (lane 3). After SI digestion, a portion of each sample was electrophoresed with the probe chemically cleaved by the G, A>G, C, and C+T reactions (13), in lanes G, A, C, and T, respectively. A part of the sequence determined is shown on the left side with that of the complementary strand. The transcription termination sites are indicated by arrow heads. The larger arrow head represents the predominant termination site.

cleavage. To determine a correction factor for SI mapping analysis, the labeled Ddel-NcoI fragment was cleaved by BstNI and electrophoresed with the chemically-cleaved probe DNA. The resulting Ddel-BstNI subfragment migrated slower than the band corresponding to the BstNI cleavage site in the sequence ladder. The difference in the mobility was calculated to be approximately 1.2 nucleotides equivalent. Since BstNI and SI digestions leave the same ends (5'-P and 3'-OH), a correction factor of 1.2-bp was applied to estimate the termination site. Although the multiple termination sites were seen, the transcription of the lexA gene appeared to stop predominantly at the T residue located 675-bp downstream from the in vitro initiation site. In MTC-treated wild-type cells, predominant termination site was also the same (lane 2). Some faint bands under the main band appeared to be formed by degradation of mRNA-1. This is because the intensities of the faint bands were lower when RNA prepared immediately before hybridization was used (lane 3), but higher when RNA was used after
Fig. 7. Initiation and Termination Sites for the lexA Transcription: (A) Transcription map of the lexA operon. Two species of lexA mRNAs are shown under the restriction map of the lexA-lexD region. The locations of the promoter (P) and terminators (t1 and t2) are indicated by vertical arrow heads. (B) Transcription initiation site and termination site 1 of the lexA operon. Nucleotides are numbered from the site corresponding to 5' end of the transcript synthesized in vitro (II, 22). The initiation and termination sites are shown by arrow heads. The minor termination sites are indicated by small arrow heads. Dyad symmetries are shown by arrows. Pribnow box (PB), ribosome-binding site (SD) (II, 22), lexA-protein-binding site (3, 4), and relevant restriction sites are shown. The first and stop codons of the lexA gene are shown by fMet and End, respectively. The nucleotide sequence of the t1 region determined here agreed with that determined by others except for two positions. Markham et al. (24) and Horii et al. (22) reported lack of A at position 656 and C instead of T at position 677, respectively.

storage (lanes 1 and 2). The predominant termination site was located at the first base on a series of T residues preceded by a dyad symmetry (Fig. 7B). The structure is characteristic of factor-independent terminators (25).

DISCUSSION

We previously analyzed the chromosomal region containing the lexA gene and suggested the existence of a gene, which we named as lexD here, coding for a protein of 35 kd (7). The lexD gene was located downstream
from the *lexA* gene by the analysis of the truncated gene product directed by a recombinant plasmid (7). By analysis of *in vivo* mRNAs, we have shown here that the region containing the *lexD* gene is cotranscribed with the *lexA* gene. The initiation and termination sites of the transcription of the *lexA* gene have also been determined (Fig. 7).

The two species of *lexA* mRNAs were detected by RNA blot hybridization (Fig. 4). Because the initiation sites of the mRNAs were the same (Fig. 5), only the termination sites should be different, suggesting the existence of attenuation or processing mechanism. There is a dyad symmetry followed by a series of T residues on which the majority of transcription stopped. The sequence is characteristic of factor-independent terminators or attenuators (25). It has been postulated that RnaselII can cleave a large stem-and-loop structure in *rpsJL-rpoBC* mRNA, and this processing event regulates the gene expression (26). The dyad symmetry preceeded by the *lexA* gene, however, is considered to be too small to be cleaved by RnaselII. Thus the *lexA* mRNA is thought to be formed by attenuation of transcription rather than by processing of mRNA-2.

The DNA region corresponding to mRNA-2 is long enough (approximately 3 kb) to code for some proteins including the *lexA* and *lexD* gene products. Since transcription of the *lexD* region is very weak in non-induced cells (Fig. 3, lane q), the attenuation mechanism should serve a tight repression of the genes including the *lexD* gene. After MTC- or UV-induction, transcription of the *lexA* gene is stimulated and a fraction of the transcript passes through the "attenuator" to allow the *lexD* protein to be synthesized. Kenyon and Walker (27) identified a damage-inducible locus *dinF* close to the *lexA* gene. According to the results described here, *dinF* insertions could be located at any sites in the *lexD* region. Walker and his coworkers (28) have recently determined the position of one of the *dinF* insertions to be proximal to the KpnI site which is very close to the junction between the *lexD* and *lacZ* portions on pMCR811 (see Fig. 1B). They reported that insertion of Tn5 into the *lexA* gene in the *dinF* insertion strain reduced the β-galactosidase activity, suggesting that the inserted Tn5 stopped transcription of the *dinF-lacZ* fused operon. Their results support our conclusion on cotranscription of the *lexA* and *lexD* genes in genetic criteria. After the position of the *lexA* and its neighboring genes were determined on the physical map (7), we noticed that Lightner et al. (29) cloned the other neighboring genes, *plsB* and *dgk* from pLC44-14 carrying the *lexA* gene. From their and our restriction mapping data, it is clear now that the *plsB* and
dgk genes are located counter-clockwise to the lexA gene, in contrast to the standard genetic map (30). Thus the gene order is \textit{malB-plsB-dgk-lexA-lexD (dinF)-dnaB-uvrA}.

ACKNOWLEDGMENTS

We are grateful to Dr. M. Casadaban for MC1000 and pMC1403, to Dr. K. Shigesada for KT5371, and to Dr. I. Furusawa for brome mosaic virus RNAs. This work was supported by grants-in-aid for Scientific Research from Ministry of Education, Science and Culture, Japan.

REFERENCES
