Title	ANALYSIS OF THE REGULATORY MECHANISM OF THE SOS-FUNCTIONS
Author(s)	Horii, Toshihiro
Citation	大阪大学, 1981, 博士論文
Version Type	VoR
URL	https://hdl.handle.net/11094/24599
rights	
Note	

Osaka University Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

Osaka University

ANALYSIS OF
THE REGULATORY MECHANISM OF
THE SOS-FUNCTIONS

BY

Toshihiro Horii

ACKNOWLEDGEMENTS

I would like to express my thanks to Professor Hideyuki Ogawa for constant guidance and encouragement in the course of the work. Thanks are also due to Dr. Tomoko Ogawa for valuable suggestions and discussions.

I am indebted to Professor Hiroshi Matsubara and Dr. Toshiharu Hase for collaboration in determination of the amino acid sequences and amino acid compositions. I also thanks Drs. Tateo Itoh and Jun-Ichi Tomizawa for collaboration in determination of the *recA*-mRNA sequence, Drs. Mitsuru Takanami and Kazunori Sugimoto for introducing me to the DNA nucleotide sequencing technique, Dr. Seiki Kuramitsu for valuable discussions, especially in Section 2, Professor Akira Nakazawa for communicating results prior to publication, and Misses Midori Nomura and Masako Sakaguchi for excellent assistances throughout the work described in Section 2.

Last, I expresses my thanks to my wife Hiroko for her encouragements throughout this work.

CONTENTS

Acknowledgements.

Abstract. -- 1.

Introduction. ------------------------------------- 2.

Materials and Methods. -------------------------------- 4.

Section 1. Organization of the recA gene. -------------- 14.

Summary. ---------- 15.

Introduction. ---------- 16.

Results. ---------- 17.

Considerations. ---------- 48.

Section 2. Organization of the lexA gene. -------------- 52.

Summary. ---------- 53.

Introduction. ---------- 54.

Results. ---------- 55.

Considerations. ---------- 81.

Discussion. ---------------------- 89.

References. -------------------------------- 93.

Publications. -- 99.

ABSTRACT

Escherichia coli exhibits a number of coordinately expressed responses, often called SOS-functions, when cellular DNA is damaged. I studied the regulatory mechanism of the expression of the "SOS-functions" through the structural analysis of the *recA* and *lexA* genes, whose products play key roles in the expression of "SOS-functions" : *recA* protein seems to work as a specific protease to the repressors of genes involved in the "SOS-functions", and *lexA* protein seems to work as a repressor of the genes involved in "SOS-functions".

The organizations of *recA* and *lexA* genes determined by the analyses of nucleotide sequences of the genes, of transcriptional products (mRNA) and of translational products (Proteins). From a comparison of the nucleotide sequence of the regulatory regions of the trascription initiation, the indication was obtained that both *recA* and *lexA* genes are repressed by the *lexA* protein. Both genes have homologous sequence comprised of 22 nucleotides at a promoter region and, this seems to be a common sequence for the binding of the *lexA* protein. A comparison of the amino acid sequences of *lexA* protein and λ repressor, these are susceptible to the proteolysis by the *recA* protein, revealed that a homologous sequence comprised of 11 amino acids was present in both proteins, this seems to work for the recognition sequence of *recA* protein

It is very plausible that the various genes involved in "SOS-functions" which are coordinately regulated by *lexA* protein carry a common 22 nucleotide sequence which I found in *recA* and *lexA* genes at their regulatory regions and that the repressor protein(s) of the genes involved in "SOS-functions" has also a common amino acid sequence found in both λ repressor and *lexA* protein.

INTRODUCTION

Escherichia coli exhibits various responses those are expressed coordinately, when cellular DNA is damaged. The typical functions in response to DNA damage are, enhancement of repair capacity (Clark and Margulies, 1965; Howard-Flanders and Boyce, 1966; Tomizawa and Ogawa, 1968; Jenkins and Bennett, 1976), inhibition of septum formation (Kirby *et al.*, 1967), induction of prophage replication (Brooks and Clark, 1967; Hertman and Luria, 1967), *recA* protein synthesis (Gudas and Pardee, 1976; Little and Kleid, 1977; Gudas and Mount, 1977; Emmerson and West, 1977), mutagenesis (Miura and Tomizawa, 1968) and colicin synthesis (Tessman *et al.*, 1978). These responses are induced by the treatments affecting in common to damage DNA or to interupt DNA replication. Such treatments include Ultraviolet (Clark and Margulies, 1965; Howard-Flanders and Boyce, 1966; Tomizawa and Ogawa, 1968) or X-ray (Young and Smith, 1973) irradiation; incubation with the chemical mutagens such as methyl methanesulfonate or antibiotics mitomycin C (Otsuji *et al.*, 1967); incubation with nalidixic acid (Cowlishaw and Ginoza, 1970); thymine starvation (Melechen and Skaar, 1962; Korn and Weissbach, 1962); or temperature elevation in certain mutants unable to synthesize DNA at high temperature (Monk and Gross, 1971; Noack and Klaus, 1972; Schuster and Beyers, 1973) Induction of all these functions in response to DNA damage (or arrested DNA replication) depends on the *recA* and *lexA* gene products.

This inducible system of the coordinately expressed responses was first recognized by M. Radman and E.M. Witkin. They refered to the group of inducible functions as "SOS-repair"(Radman, 1974) or "SOS-functions"(Witkin, 1976) and hypothesized that the *recA* and

lexA gene products jointly control "SOS-repair" or "SOS-functions". The designation "SOS" (international distress signal) implies that damage to DNA promotes the induction of expression of various functions, all of which are believed in their profits for the servival of the cell or of its prophage.

The regulatory mechanism for the expression of "SOS-functions" is sophisticated and unusual in contrast to other regulatory systems which control clusters of genes (operons) or unlinked genes (regulons), all of which synthesize products required of the same or closely related biochemical pathways.. The triggering events for the induction of "SOS-functions" seems to be metabolically diversed and they induce simultaneously the various genes whose products are mutually unrelated in their roles. Several models of SOS regulation were proposed (Gudas, 1976; Witkin, 1976; Gudas and Mount, 1977; Devoret, 1978), specifying the roles of recA and lexA gene products based on the genetical analyses of the mutations affecting some portion of the expression of "SOS-functions".

In this work, the entire nucleotide sequences of the recA and lexA genes and the primary structures of their products, which are directly involved in SOS regulation, were determined. In the light of new structural informations obtained here, the regulatory mechanism for dealing with the coordinate expression of the "SOS-functions" was discussed.

MATERIALS and METHODS

Bacterial Strains

The bacteria strains used are as follows: C600 (Appleyard, 1954) was used as wild-type strain; Q13 (Slater and Spiegelman, 1965) was used for the preparations of S-30, S-100 fractions and of the RNA polynerase; DM511 *tsl* (Mount et al., 1973), whose *lexA* product is thermosensitive; DM1187 *spr tif sfi*, whose *lexA* product leads a constitutive synthesis of *tif-1* protein (mutant *recA* protein) and of other SOS-functions without an inducing treatment (Mount, 1977); JM12 *tif-1* (Castellazzi et al., 1972), which leads to induction of own protien synthesis and other SOS-functions; JC1557 (Clark et al., 1966) carrying ColE2 plasmid, was constructed by ColI factor-mediated transfer from *Salmonella typhimurium*, LT2 *cycC-36* (ColE2)(Smith et al., 1963); N1790 *uvrA54 recA99* was constructed by co-transduction of a *recA* amber mutation, *recA99* with srl^+ gene to *srl* mutant of N179 *uvrA54* (Ogawa et al., 1968) using P1*kc* grown on DM455 strain (Mount, 1971); KS529 (Shimada et al., 1973) is a λ*c*I857 lysogen in which the phage DNA was inserted in the *malB* gene, this strain was used for the isolation of the transducing λ phage carrying the *lexA* gene; AN385 *ubiA*420 (Young et al., 1972) was used for the selection of the $ubiA^+$ transducing phages.

Preparation of DNAs

Closed circular duplex DNA was prepared as followed; cells were grown in M9 mediun (Anderson, 1946: 5 g NaCl, 1 g NH_4Cl, 15 g Na_2HPO_4, $12H_2O$, 3 g KH_2PO_4, 0.2 g $MgSO_4$,$7H_2O$ and 2 g glucose/l) fortified with 0.5 % Casamino acid (Difco) at 37°C. Chloramphenicol (180 μg/ml) was added at a cell density of 4 x 10^8/ml. After further 14 h incubation the cells were harvested by centrifugation, suspended and lysed as described in

Clewell and Helinski (1970). The resulted cleared lysate was layered onto 1 ml of saturated CsCl solution and spun at 20,000rpm for 16 h in SW27 rotor. Pelleted plasmid DNA was resuspended and concentrated in CsCl-EtBr density gradient. Covalently closed circular fraction was dialyzed and extracted with 80 % phenol containing 0.1 % sodium dodecylsulfate two times. After the removal of phenol with ether, the aqueous layer was dialyzed and concentrated by ethanol precipitation.

preparation of DNA fragment: The cleavage with the restriction enzyme was stopped by the addition of 1/10 volume of 3 M sodium acetate and 3 volume of ethanol. Resulted precipitates were resuspended with 5 % glycerol and 0.02 % Bromo-thimol blue and electrophoresed on 5 % polyacrylamide gel (0.5 x 12 cm), using 30 disc gel columns for 1mg of DNA, containing buffer (100 mM Tris-Borate, pH 8.3, and 2 mM EDTA). DNA bands were visualized with ethidium bromide (0.5 μg/ml) under a long-wave length UV light and DNA banding region was cut out. The gels cut out were homogenized with Dounce type homogenizer in a few milliliter of 10 mM Tris-HCl, pH 7.6 and 10 mM EDTA. After the incubation at 37°C for 14 h, gel-paste was removed by centrifugation, and then ethidium bromide was removed by phenol extraction. Finally DNA was purified by passage through Sephadex G-100 column (0.8 X 20 cm) after the concentration.

Construction of Plasmids

pMN1: λ*c*I857 phages were induced from KS529 and *ubiA*$^+$ transducing phages were selected on AN385 *ubiA*420. Among *ubiA*$^+$transducing phages a strain was found that suppresses the *tsl* phenotype of DM511 (λ) which is temperature-sensitive at 42°C. This λdl*exA*$^+$-8 phage was purified by CsCl equilibrium centrifugation and their DNA were circularized *in vitro*, ligated and then cleaved completely with Eco RI endonuclease. The cleaved fragments were ligated with the Eco RI-cleaved pBR322 (Bolivar et al., 1977) and a plasmid which suppressed the temperature sensitive character of DM511 strain at 42°C was selected. Plasmid thus obtained (pMN1) was 18 kb in length.

pTH227 construction:

Digestion with Exo III and S1: Plasmid pTH18 DNA linearized with Eco RI was partialy digested with Exo III and S1 at the ends according to a method described by Roberts et al. (1979). In a 100 µl of reaction mixture containing 100 mM Tris-HCl (pH 7.6) and 10 mM $MgCl_2$, 10 µg of pTH18 DNA linearized with Eco RI was digested with 60 units of Exo III for 5 min at 20°C (Exo III digests at a rate of about 18 bp. per min per end of DNA under this condition). Reaction was stopped by the addition of equal volume of x2 S1 buffer, which contains 100 mM NaOAc (pH 4.0), 300 mM NaCl, and 12 mM $ZnSO_4$. Then, 60 units of S1 nuclease was inactivated with a drop of phenol. After the extraction of phenol with ethyl ether, DNA was precipitated with ethanol.

Insertion of lac promoter: DNA fragment (95 bp.) bearing lac promoter was further purified from Alu I digests of 205 bp. Eco RI fragment that was isolated from pKB252. Thus obtained DNA fragment carrying blunt ends was inserted with T4 ligase into partialy resected pTH18 DNA with Exo III and S1 at its Eco RI ends. Since the formation of linear or circular products of ligation is a function of the concentration of DNA fragments (Dugaiczyk et al., 1975), ligation of blunt ended DNAs were carried out in two steps as follows; in a 100 µl of ligation buffer (50 mM Tris-HCl, pH7.6/ 10 mM $MgCl_2$/ 10 mM dithiothereitol/ 50 µM ATP), 2.5 µg of pTH18 DNA treated with Exo III and S1, and 0.12 µg of DNA fragment bearing lac promoter were ligated with 6 units of T4 ligase for 4.5 hr at 10°C. Under this condition, about an half ends of DNA were ligated. Then, 8 µl of this mixture was taken and circuralized with 6 units of ligase for 22 hr at 10 °C in a 100 µl of ligation buffer. About 200 colonies were obtained per each nano grammes of DNA ligated as described above.

Enzymes

RNA polymerase holoenzyme was prepared from E.coli Q13 by the method described in Burgess et al., (1969) and phospho-cellulose fraction was used in this experiment. The purity of this enzyme and the content of the sigma-subunit were analyzed by SDS polyacrylamide gel electrophoresis. The preparation used in this paper had 80 % purity and contained the sigma-facter at an enough amount. T4 polynucleotide ligase was prepared as described by Weiss (1971). The restriction enzymes Eco RI (Yoshimori, 1971) and Bam HI (Wilson and Young, 1975) were prepared as described. Hga I was prepared as described in Roberts et al.,(1975). The endonuclease HincII, AvaII and Hpa I were obtained from Bethesda Research Laboratories, Inc., Hinf I, Hae II Hae III and Hpa II were obtained from New England Biolab Inc., Alu I adn Pst I wer obtained from Boeringer Mannheim.

Digestion with Restriction Endonuclease and Determination of Electrophoretic Mobility of DNA Fragment

Enzyme digestions were carried out at 37 °C in reaction mixtures (50 to 200 ul) containing 10 mM Tris-HCl (pH7.6), 8 mM $MgCl_2$, 2 mM 2-mercaptoethanol and 100 ug/ml bovine serum albumin except that 100 mM Tris-HCl (pH7.6) was used for the digestion with Eco RI or Bam HI. Disc gels (0.5 x 12 cm) of 5 % or 10 % polyacrylamide (acrylamide/bis-acrylamide 19:1) in 36 mM Tris/32 mM KH_2PO_4/1 mM EDTA at pH7.8 (Oka, 1978) were used for determination of electrophoretic mobility of DNA fragments. The size of a DNA fragment was estimated from its mobility relative to the mobilities of the Hae III fragments of ColE1 DNA (Tomizawa,personal communication).

In vitro RNA Synthesis

For R-loop formation, in vitro transcription was carried out in 0.75 ml of reaction mixture containing 40 mM Tris-HCl (pH7.9), 8 mM $MgCl_2$, 100 mM KCl, 0.2 mM each of ATP,CTP and GTP, 0.2 mM 5-^{3}H-UTP (4 Ci/mole), 15 μg of Bam HI 3 kb-DNA fragment and 30 μg of RNA polymerase holoenzyme. After the incubation for 20 min at 37°C, 15 μg of DNase I (Sigma, RNase-free) was added and incubation was continued for 15 min further. The reaction products were collected by ethanol precipitation and fractionated on a sucrose gradient (5 to 20 % w/v in 30 mM Tris-HCl, pH 7.4, 1 mM $MgCl_2$) in SW 50.1 rotor. Centrifugation was carried out at 45,000 rpm for 3 h,10°C. A portion of each fraction was counted and peak fractions were collected and concentrated. For determination of RNA nucleotide sequence, the reaction of transcription was carried out as follows; the reaction mixture consisted of 40 mM Tris-HCl, pH7.9, 100 mM KCl, 10 mM $MgCl_2$, varing concentration of rNTPs including [α-^{32}P] UTP (35 Ci/mmol, Amersham), about 0.03 μg of DNA fragment and 0.2 units of RNA polymerase in 10 to 50 μl of the reaction mixture which was incubated for 30 min at 37°C. For labeling with [γ-^{32}P] ATP (16 Ci/mmol), 150 uM each of rNTPs, 0.3 μg of DNA, 2 unit of RNA polymerase in 300 μl of reaction mixture were incubated for 30 min at 37°C.

In vitro Protein Synthesis

S-30 and S-100 fractions were prepared from E. coliQ13 as described in Nierenberg (1963). RNA directed cell-free protein synthesis was carried out in a 50 μl reaction mixture containing 50 mM Tris-HCl (pH7.8), 42 mM NH_4Cl, 7.5 mM $MgCl_2$, 1 mM Dithiothleitol, 1 mM ATP, 0.2 mM GTP, 5 mM PEP (phosphoenolpyrubate), 30 μg/ml pyrubate kinase, 0.05 mM each of amino acids (except methionine), 0.03 mM ^{35}S-methionine (0.67 mCi/μmol), 0.2 volume of preincubated S-30 fraction and S-100 fraction and RNA which

was synthesized in vitro. Reaction mixtures were incubated at 34°C for 30 min and the reaction was stopped by the freezing the incubation mixture in liquid nitrogen.

Immuno Precipitation

The procedure of rapid isolation of antigen-antibody complex was followed as that described by Kessler (1975). The products of in vitro translation or the purified ^{35}S-labeled recA protein was incubated with an enough amount of the anti-recA serum in 0.5 ml of NET (150 mM NaCl, 5 mM EDTA, 50 mM Tris-HCl, 0.02 % sodium azide) buffer, pH7.4, containing 0.05 % Triton X100 at 37°C for 20 min, then treated with the cell corpuscles of Staphylococcuss aureus (Cowan I) prepared for an antibody adsorbent at 37°C for more 30 min. The cell corpuscles and antigen-antibody complexes were washed 3 times and added 100 µl of NET buffer containing 0.2 % SDS. After the incubation in boiling water bath, the released proteins from the complexes were electrophoresed on SDS polyacrylamide gel. By this procedure, 95 % of the total input radioactivity of the purified ^{35}S-labeled recA protein was recovered. Five centimeter (diameter 0.5 cm) 5 % polyacrylamide gel was subjected to electrophoresis according to the method of Weber and Osborn (1969). After the electrophoresis, gel was sliced 1.25 mm in thick and each sliced fraction was counted in Bray's solytion after hydrolysis in H_2O_2.

R-loop formation

The R-loop formation was carried out with 1.0 µg/ml of Bam HI 3 kb-DNA fragment and 10 µg/ml of RNA synthesized in vitro in 70 % formamide, 83 mM Pipes (pH7.8), 10 mM EDTA, at 48 °C for 12 h (Thomas et al., 1976). After the incubation, 20 µl of R-loop formation mixture was taken, diluted finally 50 % formamide and spread with cytochromeC on a hypophase of H_2O.

The DNA-RNA hybrids was picked up on a parlodion film, stained with uranyl-acetate and shadowed with Platinum. The electron micrographs were taken with JEM 100C at a magnification of 10,000x. Using ColE1 open circular DNA, 6.6 kb (Tomizawa, personal communication), as a marker each length of double stranded DNA, DNA-RNA region and single-stranded DNA was measured with.

Labeling of the Proteins Directed by the Plasmid and Autoradiography

The maxi-cell method (Sancar *et al.*, 1979) was generally followed in labeling the proteins directed by the plasmid. A bacterial strain, N1790 *uvrA54 recA99* was used as a host bacteria for each plasmid to be tested. Six ml of a log-phase culture (2 x 10^8/ml) grown in M9 medium supplemented with 0.5 % Casamino acid and 20 µg/ml of tryptophan, was irradiated with 180 J/M^2 fluence of UV light, and further incubated for 16h. After starvation for sulfate for 1 h, ^{35}S-methionine was added at a final concentration of 10 µCi/ml and incubation was continued further for 1 h. The labeled cells were washed three times and suspended in 0.2 ml of 50 mM Tris-HCl (pH 8.0) containing 10 mM EDTA. After addition of 40 µl of 5 mg/ml lysozyme, the sample was subjected to 3 cycles of freezing and thawing. After addition of 60 µl of 5 times concentrated lysing solution (Sancar *et al.*, 1979), 30 µl of sample was electrophoresed using 12.5 % of polyacryamide gel containing 0.2 % SDS. After electrophoresis the gel was immersed in the solution containing 10 % trichloroacetic acid, 10 % acetic acid and 30 % methanol for 60 min at room temperature. It was then treated with autoradiography enhancer, EN^{3}HANCE (New England Nuclear) for 1 h. After washing with cold water for 60 min, the gel was dried and contacted with an X-ray film (Kodak X-omat R-film) and exposed for 24 h at -80°C.

Purifications of Proteins

W3623 cells freshly transformed with plasmid pTH227 were used because plasmid pTH227 is very unstable. This is probably due to abundant production of the lexA protein. Cells were grown in L-broth to a concentration 1×10^8 cells/ml then lactose was added to give a final concentration of 1% and continued to grow. At 4×10^8 cells/ml, cells were harvested by centrifugation and stored at -20°C. For the purification, all operations were carried out at 2°C and centrifugations were carried out at 12,000 xg for 20 min unless otherwise described. Frozen cells, 20 g, were suspended in 100 ml of buffer A (50 mM Tris-HCl pH 7.6/ 5 mM EDTA/ 100 mM NaCl/ 7 mM 2-mercaptoethanol/ 1 mM phenylmethylsulfonylfloride/ 5% glycerol), added 5 ml of lysozyme (5 mg/ml) and disruption was completed by sonication. After the addition of 5 ml of sodium deoxycholate (10 % wt/vol), lysate was centrifuged at 100,000 Xg for 45 min. The absorbance at 260 nm of resulting extract was adjusted to 160/ml with buffer A and added 2% of polymin P (pH 7.9, BASF) for an interval of 15 min to give a final concentration of 0.25% with stirring. After the stirring for an additional 15 min, precipitate was removed by centrifugation. To resulting suppernatant, crystalline ammonium sulfate was added slowly with stirring to 35% saturation, precipitate was discarded by centrifugation and then ammonium sulfate was added to give a 70% saturation. After centrifugation, the resulting pellet was dissolved in 50 ml of buffer B (10 mM Tris-HCl pH 7.6/ 2 mM EDTA/ 5 mM 2-mercaptoethanol/ 5% glycerol) containing 100 mM NaCl and dialyzed 16 hr against the same buffer. The dialyzed sample (Fraction I) was diluted with an equal volume of buffer B and applied to a 2.3 × 18 cm DEAE cellulose column previously equilibrated with buffer B containing 50 mM

NaCl. The column was washed with 150 ml of the same buffer and a linear gradient of 50-400 mM NaCl in buffer B (total volume, 700 ml). The fractions eluted between 100-150 mM NaCl were collected (Fraction II) and precipitated by a dialysis against 80% saturated ammonium sulfate solution containing 50 mM Tris-HCl pH 7.6. The resulting precipitate was collected by centrifugation and dissolved in 10 ml of buffer C (50 mM potassium phosphate pH 6.5/ 2 mM EDTA/ 100 mM KCl/ 5 mM 2-mercatoethanol/ 5% glycerol) and dialyzed 16 hr against the same buffer. The preparation was applied to a 1.6 × 5 cm phosphocellulose column previously equilibrated in buffer C, washed with 30 ml of buffer C and eluted with a linear gradint of 100-600 mM KCl in buffer C (total volume, 200 ml). The fractions eluted at 300 mM KCl were collected (Fraction III) and dialyzed against buffer B. Under this condition, _lexA_ protein was precipitated. The resulting precipitate was collected by centrifugation, dissolved in 4 ml of buffer C containg 500 mM NaCl and dialyzed against the same buffer. The preparation was applied to 1.2 × 50 cm Sephaacryl S-200 supper fine column previously equilibrated in buffer C containing 500 mM NaCl. The _lexA_ protein was eluted at 1.6 void volume (Fraction IV). The preparation was concentrated by dialysis against buffer B containing 100 mM NaCl and 50% (vol/vol) of glycerol, and storred at -20°C.

recA protein: The wild-type _recA_ protein was prepared from C600(pTM2) cells. The procedures used will be published elsewhere (Ogawa, Nakashita-Wabiko and Ogawa manuscript in preparation).

SDS-Polyacrylamide Gel Electrophoresis

Polyacrylamide slab gels with 3% stacking and 12.5% resolving gel were prepared according to the method of Laemmli (1970).

Where indicated, 10-25% gradient gel was used for resolving gel. After electrophoresis, the gel was stained with coomasie brilliant blue G-250.

Amino Acid Composition and Sequence Analysis

Proteins (20 μg) were hydrolyzed in 5.7 N HCl in evacuated, sealed tubes for 24 and 72 hr at 110°C. The amino acid composition of resulting hydrolysate was determined with an automated amino acid analyzer model A3300 (Irica Instruments Inc., Kyoto, Japan). The mannual Edman degradation (Blombäck et al., 1966) and identification of phenyl-thio-hydantoin (PTH) derivatives by thin layer chromatography (Niederwieser, 1972) were carried out as described. PTH derivatives were also identified by high performance liquid chromatography (Zimmerman, et al., 1977). The carboxyl-terminal sequence was determined by digestion with carboxypeptidase A (Ambler, 1972).

Determination of Nucleotide Sequence

Nucleotide sequence of DNA fragment was determined by the method of Maxam and Gilbert (1977). The 5' end of DNA fragment was labeled by using T4 polynucleotide kinase, for nucleotide specific modification, dimethyl-sulfate or hydrazine was used. After cleavage, the products were electrophoresed in a slab gel (0.1 x 30 x 40 cm) of 20% polyacrylamide (acryl/bisacrylamide, 19:1) containing 7 M urea, 100 mM Tris-borate (pH 8.3) and 2 mM EDTA. The nucleotide sequence near the 5'end of RNA was determined through analysis of the products of partial digestion of [γ-^{32}P]ATP-labeled RNA by RNase A (Seikagaku kogyo), RNase T1, RNase U2 (Sankyo) and RNase PhyI (P-L Biochemicals) described in Simoncsitz et al., (1977), and Itoh and Tomizawa (1980).

SECTION 1.

Organization of the *recA* Gene

Summary of Section 1.

The cloned Bam HI 3 kb-DNA fragment which contains the recA gene of E. coli was used to direct the in vitro RNA synthesis. By the hybridization of the synthesized RNA molecules with the Bam HI 3 kb-DNA fragment, the transcribed region containing the recA gene was mapped on the DNA fragment by electron microscopy. The results showed that the recA-mRNA is initiated at about 0.99 kb from the Bam HI site near the srl gene and streches about 1 kb in length and terminated.

The in vivo transcript from recA gene was also analyzed by the hybridization to the prove DNA immobilized on membrane filter The rate of recA-mRNA synthesis was increased more than 10-fold by the treatments those induce the SOS-functions and that the increment occured under the absence of protein synthesis. On the sucrose gradient, the in vivo transcript was co-sedimented with the in vitro product. These results indicate that the expression of recA gene is controled at a transcriptional level and its transcription is monocistronic.

For the structural analysis of the recA gene, nucleotide sequence of transcribed and surrounding regions of the gene was determined. The coding region of the recA gene comprises 1059 nucleotide residues and encodes a single protein of 353 amino acid residues. The amino acid sequence at amino-terminus of the recA protein agrees with the sequence predicted from the DNA sequence analysis except for the absence of formylmethionine in the purified protein. The site of initiation for in vitro synthesis of recA-mRNA has been determined by analysis of the 5'-nucleotide sequence of $[\gamma-^{32}P]$ATP-labeled transcripts. The promoter region shows a high degree of symmetry and contains sequences commonly found in recognition and binding sites for RNA polymerase.

Introduction

The *recA* gene of *E. coli* shows pleiotropic properties. In addition to playing an essential role in genetic recombination, the *recA* function is necessary for the induction of SOS-functions (Clark and Marglies, 1965; Witkin, 1976). Whenever the SOS-functions was induced, the increased production of *recA* protein is observed. Furthermore, *recA* protein is necessary itself for this increment (Gudas and Pardee, 1975; Ogawa, *et al.*, 1978).

The purified *recA* protein catalyses several reactions: the ATP hydrolysis depending on single-strand DNA (Ogawa, *et al.*, 1978); the uptake of single-strand DNA depending upon ATP by super helical DNA (Shibata *et al.*, 1979); the reannealing of single-strand of homologous DNA depending upon ATP (Weinstock *et al.*, 1979); and the proteolytic cleavage of λ repressor (Roberts et al., 1978).

To facilitate analysis of regulation of expression of the *recA* gene, a set of plasmids which carry the wild type *recA* gene, a *recA* deletion or a *recA* mutation have been constructed by *in vitro* gene manipulation techniques (Ogawa et al., 1978). One of these plasmids, pTM2, which contains a wild-type *recA* gene, is cleaved into two fragments by the resriction enzyme Bam HI. On the smaller fragment about 3 kb in length, the *recA* gene is located.

Using this cloned *recA* gene, the products of *in vitro* and *in vivo* transcription of *recA* gene were analyzed. The results showed that the *recA* gene is monocistronic and regulated at a transcriptional level. The region where transcribed on the cloned fragment was determined by R-loop method and the nucleotide sequence of the *recA* gene and its neighborhood was determined.

Results of Section 1

1. Transcription of *recA* gene in *in vitro*

The transcription of *recA* gene on the 3 kb-DNA fragment was carried out in *in vitro* by *E. coli* RNA polymerase. The detailed procedures are described in Materials and Methods. The synthesized RNA (about 15 μg, calculated form the incorporation of ^{3}H-UTP) was collected by the ethanol precipitation and centrifuged in a 5 to 20% sucrose density gradient. As a marker of sedimentation rate, R17 phage RNA (23S, about 3,000 nucleotides) was run in a separated tube simultaneously. Three distinct peaks of the synthesized RNA were observed in the gradient as denoted as No.1 No.2 and No.3 in the order of their magnitude in Fig. 1. The sedimentation value of the peak fractions No.3, corresponded to that of a reference, R17 RNA(23S), and the other two's were more smaller. Each pooled fraction, No.1, No.2 and No.3 contained 42%, 24% and 23% of the total synthesized RNA respectively. To confirm that the smallest RNA, the major fraction of the synthesized RNA, were a complete transcript of the *recA* gene, the translation experiments of these RNA were performed in *in vitro*.

The pooled RNA fraction, No.1, were dialyzed against 50 mM Tris-HCl buffer and added to the translation system (Nierenberg, 1963) at a final concentration of 30 μg/ml. After the incubation at 34°C for 30 min, the synthesized protein were precipitated with trichloroacetic acid (TCA) after hydrolysis aminoacyl-tRNA carrying labeled amino acid. About 19pM methionine was incorporated into an hot-TCA insoluble material. To determine what fraction is a complete *recA* protein, the reaction products were treated with the anti-*recA* protein serum and then the *Staphylococcus aureus* cells corpuscle were added to the mixture to

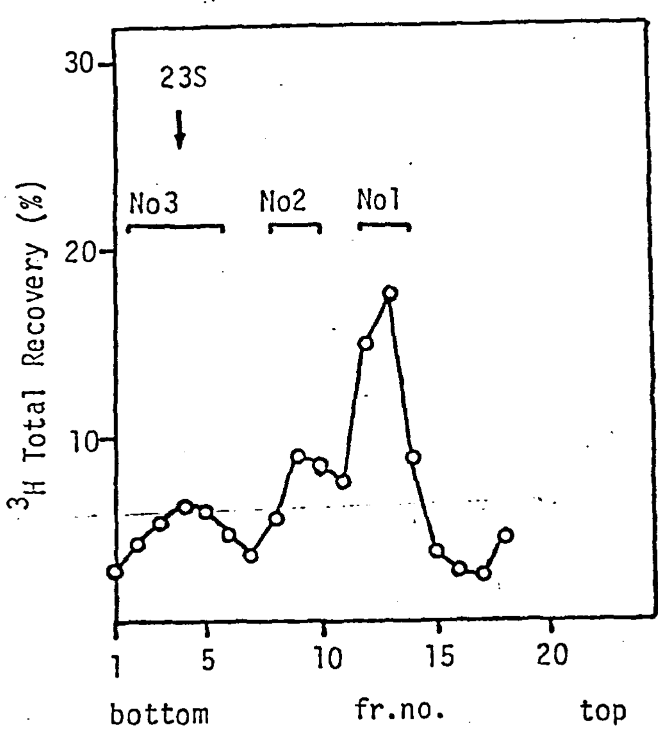


Fig. 1. Sedimentation analysis of the RNA synthesized in vitro.

In vitro transcription was carried out with E. coli RNA polymerase holoenzyme by the direction of BamHI 3 kb DNA fragment as described in Materials and Methods. The poduct was analyzed on 5-20% (W/V) sucrose gradient in SW50.1 rotor. Centrifugation was carried out at 45,000 rpm for 3 hr, 10°C. After fractionation, a portion of each fraction was taken and its acid-insoluble counts were scored. Fractions 12 to 14, 8 to 10 and 2 to 5 were pooled separately, and denoted as No.1, No.2 and No.3, respectively. Phage R17 RNA (23S) was run in parallel as a marker. The parcentage of total recovery was calculated from the total acid-insoluble count.

rapidly isolate antigen-antibody complex. The precipitated counts were found to be about 95% of the total hot-TCA insoluble material. The *recA* protein was released from the antigen-antibody complex by the addition of 0.2% SDS and then subjected to the SDS gel electrophoresis. A profile of the electrophoresis is shown in Fig. 2(a). Almost all counts (50%) were incorporated into a complete *recA* protein and the others were in smaller nacent *recA* protein. To compare with this result, about two times larger RNA fraction, No.2, were added in the translation system. The incorporated counts per RNA molecule were the same as that with the No.1 RNA fraction, and also a complete *recA* protein was synthesized as shown in Fig. 2(b). Therefore, the smaller RNA species, No.1, were found to be a complete RNA transcript of the *recA* gene.

2. Transcription of *recA* gene in *in vivo*

The size and the rate of in *in vivo* transcript from the *recA* gene was mesured. For specific detection of RNA from the *recA* region, two plasmid DNAs, pTM2 and ColE1::Tn3, were used as the proves. The plasmid DNA was cleaved with BamH1 and denatured in 0.3 M KOH at room temperature DNA was neutralized with HCl and immobilized on nitrocellulose filter by baking at 80°C. To the prove DNA filter, the RNA prepared from the cells pulse-labeled with ^{3}H-uridine for 1 min was hybridized. To investigate the specificity of the hybridization method used here, increasing amounts of the RNA were hybridized to excess amount of DNA, or excess amount of RNA was hybridized to the increasing amounts of DNA (Figure 3). The results showed that the fraction of the pulse-labeled RNA retained on ColE1::Tn3 DNA was negligibly small (0.01-0.03%) and the amount of DNA used here (2.5 μg/filter) was enough to bind all of the specific RNA. Using *in vitro* transcripts of the *recA* gene for

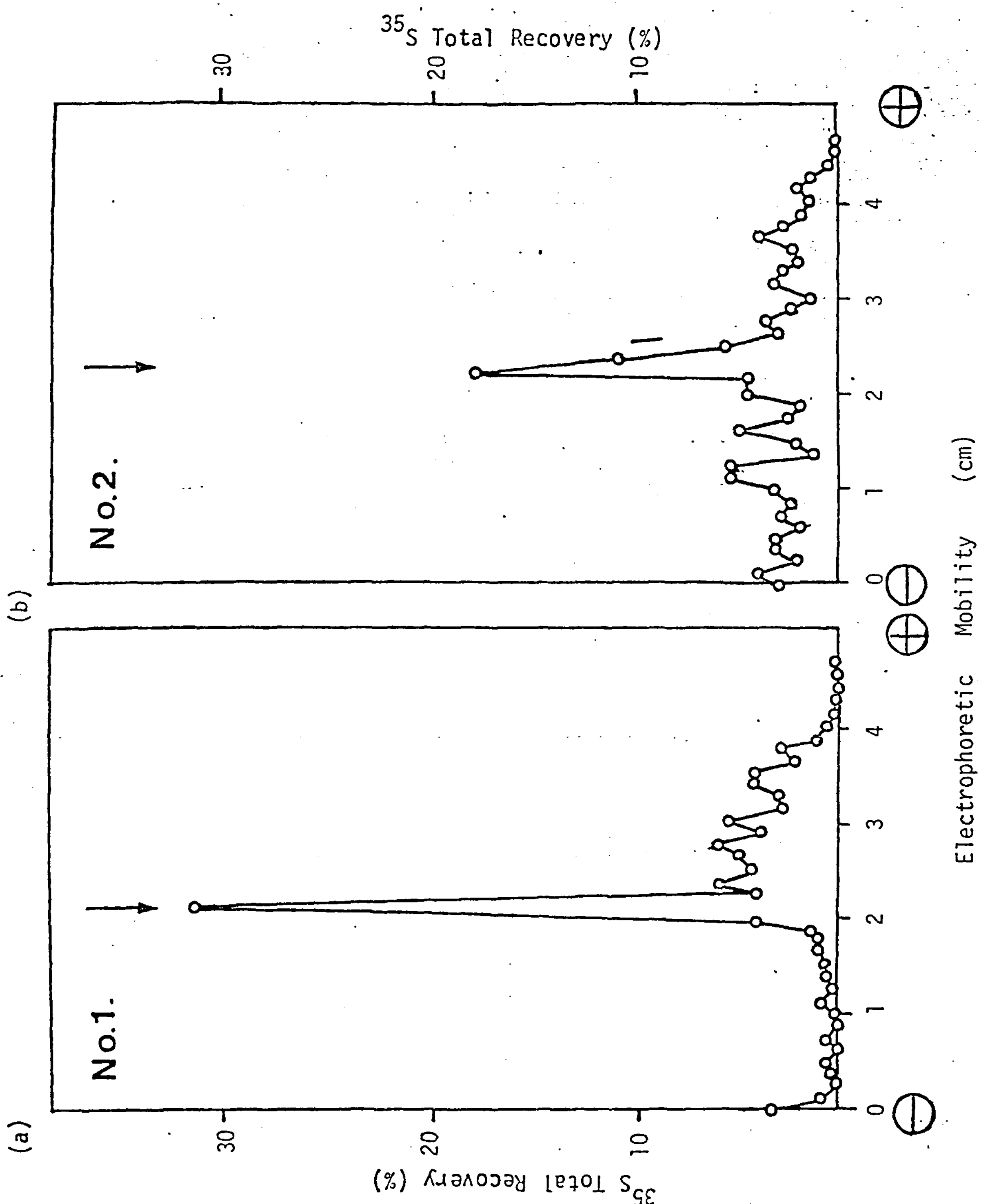


Fig. 2

Fig. 2. The SDS-polyacrylamide gel electrophoresis of the *in vitro* translation products precipitated with the anti-*recA* serum.

In vitro translation and the precipitation of the translated product with the anti-*recA* protein serum were carried out as described in Materials and Methods. *In vitro* translation was directed by the RNA pooled as No. 1 (panel (a)) or No. 2 (panel (b)). After the electrophoresis, gels were sliced in 1.25 mm thick and each fraction was counted. The allows indicate the position of the purified *recA* protein run in parallel. Total recovery (%) represents the percentage of the counts of each fraction to the total counts recovered after electrophoresis.

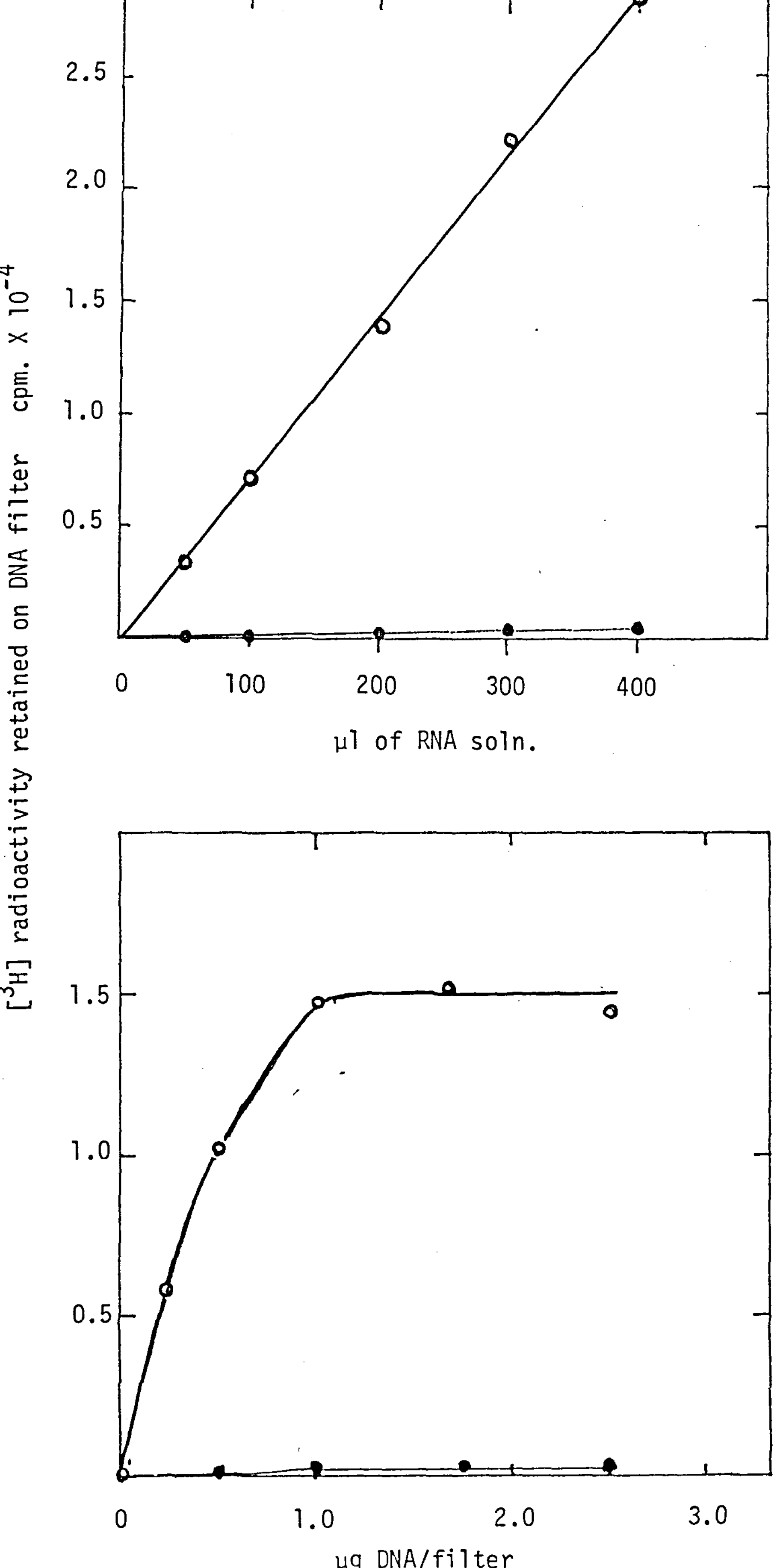


Fig. 3

Fig. 3. Hybridization of pulse-labeled RNA.

E. coli JM1 strain was grown in a 5 ml of minimal glucose medium to a cell density of 4 x 10^8/ml. Nalidixic acid was added at a final concentration of 40 μg/ml and incubation was continued. After 20 min, cells were pulse-labeled for 1 min with [^{3}H] uridine (30 μCi/mmol). Cells were rapidly chilled with ice flakes of the broth containing 10 mM NaN_3, washed and collected by centrifugation. The resulted cells were lysed in 2 ml of SSC containing, 10 mM EDTA, 10 mM NaN_3 and 0.25 % SDS by the incubation for 2 min in boiling water bath. The lysate was treated with phenol, precipitated with ethanol and resuspended in 2 ml of x 2 SSC (1 x 10^4 cpm/μl); (a), increasing amounts (50-400 μl) of labeled RNA preparation were hybridized to a excess amount of DNA (2.5 μg/filter) in 0.5 ml of x 2 SSC containing 1 mM EDTA and 0.1 % SDS for 15 h at 67.5 °C. (b), 200 μl of labeled RNA was hybridized to increasing amounts of DNA (0.25 - 2.5 μg/filter). All assays were duplicated and the average radioactivity hybridized to the DNA filter was shown. o—o ; pTM2 •—• ; ColE1::Tn3.

hybridization, the efficiency of this system was estimated as about 80% (data not shown). The amount of recA mRNA should, therefore, be 1.2-fold of the difference of the labeled RNA retained as hybrid with pTM2 and ColE1::Tn3.

The RNA prepared from the cells incubated with nalidixic acid (40 μg/ml) for 20 min was subjected to the 5-20% sucrose density gradient and aliquotes of each fractions were hybridized to pTM2 and ColE1::Tn3. As shown in Figure 4, the in vivo transcript from the recA region was co-sedimented with the major RNA about 1 kb in length of in vitro transcription products.(No.1 fraction in Fig. 1) This indicates that the major product of in vitro transcription is not an artifact and the transcription of recA gene in induced cell is confined just enough for coding the recA protein.

The inducibility of recA gene transcription was mesured by recA mRNA production before and after treatment of cells with nalidixic acid (40 μg/ml) or mitomycinC (2 μg/ml). As shown in Figure 5, the recA gene was transcribed at a rate below 0.05% during normal growth state, whereas the increase of recA mRNA synthesis rate began 3-5 min after the addition of nalidixic acid or mitonycicC and after the incubation for 40 min, the rate of recA mRNA synthesis reached more than 10-fold (0.5-0.6%)of total RNA synthesis . These results indicate that the expression of recA gene is regulated at a transcriptional level. To show wheather the translation is required for the induction of recA gene transcription, chloramphenicol (100 μg/ml) was added 5 min before the induction by nalidixic acid (Figure 5c,). The result showed that chloramphenicol has no effect on the induction of recA mRNA synthesis and translation is not required for this process.

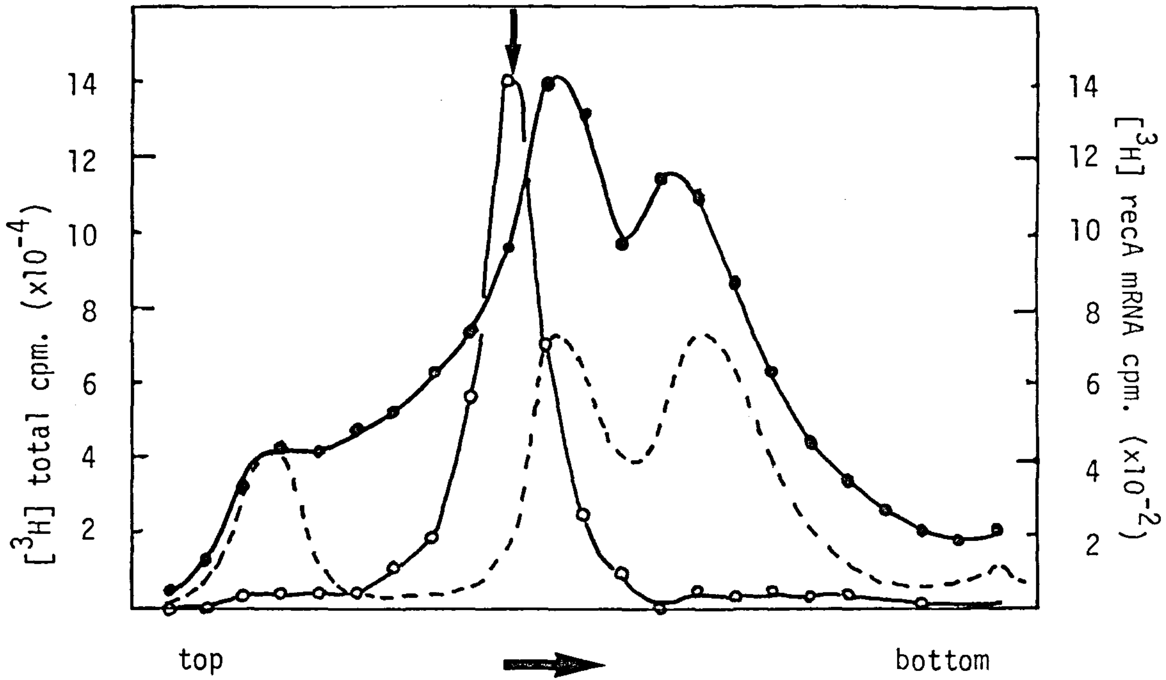


Figure 4. Sedimentaiton analysis of the RNA synthesized <u>in vivo</u>.

The pulse-labeled RNA was prepared as described in the legend to Fifure 3 . The product was analysed on 5 - 20 % (w/v) sucrose gradient in SW 50.1 rotor. Centrifugation was carried out at 45,000 rpm. for 3 hr at 10 °C. After fractionation, 50 μl portion of each fraction(200 μl) was hybridized to the DNA of pTM2 or ColE1::Tn3 and the difference of radioactivity retained on DNA filters was shown as recA mRNA (o—o). The allow indicates the position of major RNA product in <u>in vitro</u> transcription, which was run in parallel. Total radioactivity (•—•) and absorbance at 254 nm (-----) were shown.

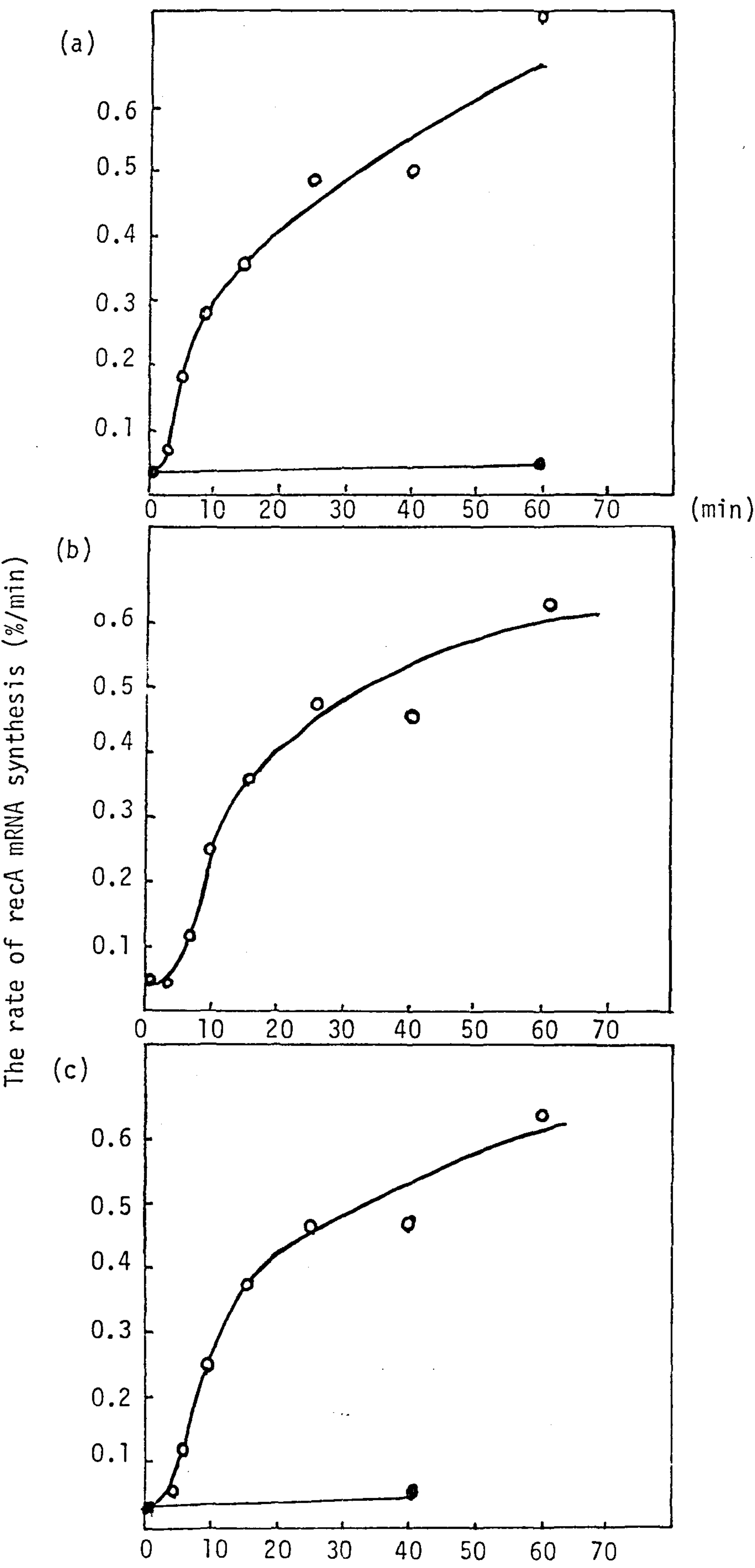


Fig. 5

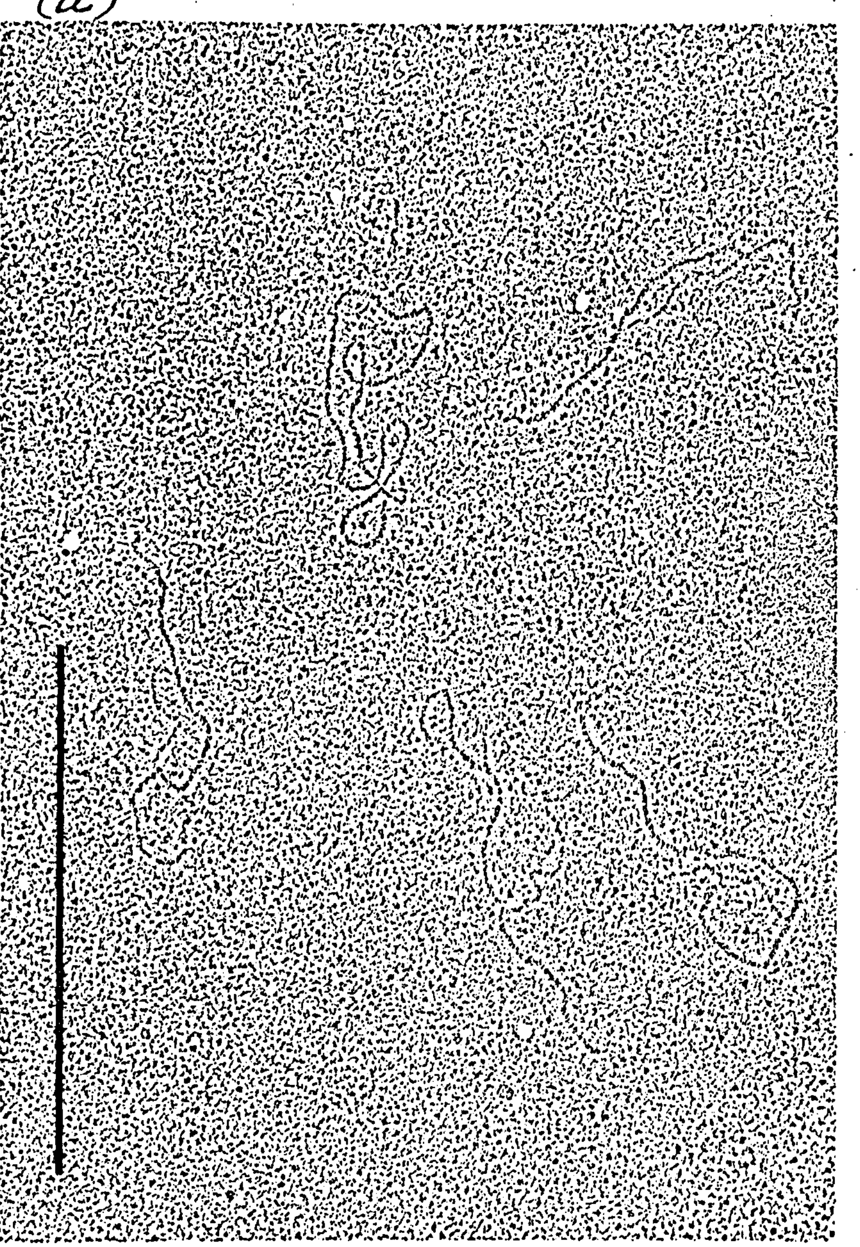
Figure 5. Kinetices of recA gene induction by nalidixic acid or mitomycinC.

The 25 ml cultures of strain JM-1 were grown to 4 x 10^8 cells/ml at 37°C. At 0 time, nalidixic acid or mitomycinC was added at a final concentration of 40 μg/ml or 2 μg/ml respectively. Aliquotes (2.5 ml) were subsequently removed and pulse-labeled with 25 μCi of [^{3}H]-uridine (30 μCi/mmol) for 1 min immediately. The RNA was prepared as described in legend to Figure 3. (a), o ; induced by nalidixic acid, ● ; not induced, (b), induced by mitomycinC, (c), o ; chloramphenicol was added at a final concentration of 100 μg/ml 5 min before the induction by nalidixic acid, ● ; chloramphenicol was added but not induced.

3. R-loop formation

These above recA-mRNA of the fractions No.1, No.2 and No.3, were hybridized with the 3 kb-DNA fragment according to the method, Thomas et al. (1976) respectively. The molar ratios of the RNA and DNA in each reaction mixture were 60 with the fraction No.1; 30 with the fraction No.2; and 20 with the fraction No.3(a). The observed molecules under the electron microscopy were scored and the results were summarized in Table 1. The most frequently observed structures were the eye-form (64%) when the RNA of fraction No.1 was hybridized with. A representative molecule of such structure is presented in Fig. 6. The lengths of each part were measured with 46 molecules as ColE1 DNA as a reference in length (6.6 kb) and the distributions of the lengths were shown in Fig. 7(a) by histograms. The linear part of the both sides were 0.99 ± 0.065 kb and 1.16 ± 0.080 kb respectively and the double-strand of the eye-form region was 1.04 ± 0.075 kb in length. The second major molecules were Y-like molecules (21% of scored molecules): one branch is a DNA and RNA hybrid and another is a displaced single-stranded DNA in the same length. The rest molecules (11%) were linear molecules, DNA and RNA hybrid or double stranded DNA (not hybridized with RNA). When the RNA in the No.2 fraction were hybridized with the 3 kb-DNA, the Y-like structure is most frequently formed (61%) and the eye-form and the linear form afre in the next place (17% and 15%, respectively). When the RNA in the fraction No.3, the largest molecules were hybridized, the linear structures were in the majority (62%) and Y-like form is in the second place (25%) and the eye-form is quite few (7%). Therefore, from the structures of hybrid molecules observed in the mojority, the length of RNA in each fraction could be estimated: the RNA in the fraction No.1 was about 1000 nucleotides; RNA in No.2, about 2,000

(a)



(b)

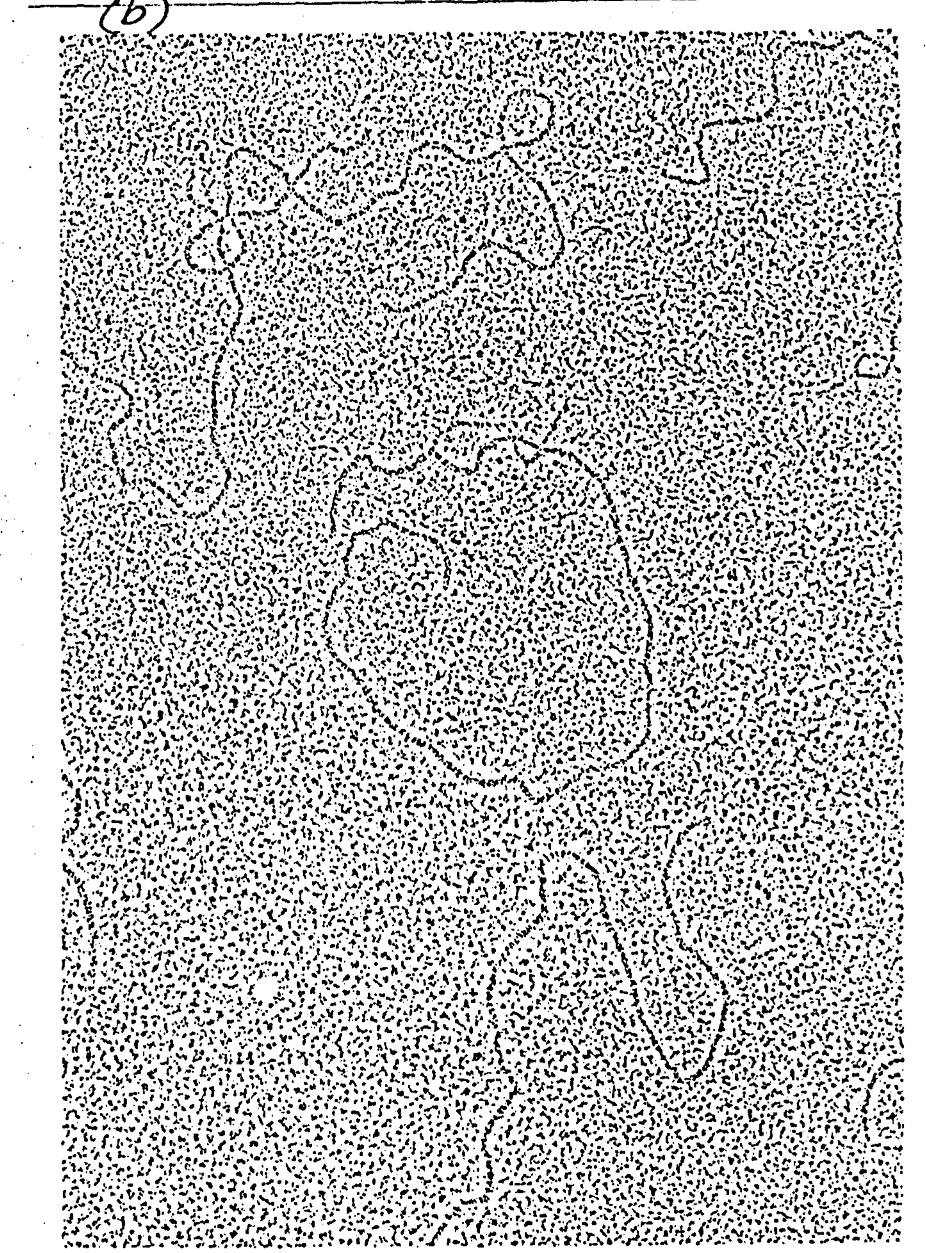


Fig. 6

Fig. 6 R-loop structure formed between the RNA synthesized *in vitro* and the Bam HI 3 kb-DNA.

(a) An eye-form structure: The most frequently observed R-loop structure formed between *in vitro* transcript of the Bam HI 3 kb-DNA and the Bam HI 3 kb-DNA. (b) A representative R-loop molecule formed between the pTM5 DNA cut with Bam HI and the *in vitro* transcripts of the Bam HI 3 kb-DNA. The R-loop formation was carried out as described in Materials and Methods. Bar represents 1 μm length.

structure	No1	No2	No3
	64% (136)	17% (36)	7% (8)
	21% (44)	61% (132)	25% (30)
	11% (23)	15% (32)	62% (75)
unknown	4% (8)	7% (17)	6% (8)
^{3}H-UMP incorporation	42%	24%	23%

Table 1. Observed frequency of the various molecules formed by hybridization of the recA-mRNA with the BamHI 3 kb-DNA.

The recA-mRNA synthesized *in vitro* was fractionated on the sucrose gradient (Fig. 1) and the pooled fractions No.1, No.2 and No.3 were hybridized with the BamHI 3 kb-DNA as described in Materials and Mathods. The structure of the observed molecules are depicted by the bold (the double-stranded region) and the dotted lines (single-stranded region). The number of molecules scored under the electron microscopy are described in parentheses. The molecules scored as unknown were highly branched or complexed with some molecules. ^{3}H-UMP incorporation denotes the fractional amount of the synthesized product contained in each pooled fraction.

nucleotides in No.3, 3,000 nucleotides.

The determination of the promoter side of the R-loop structure was made by hybridizing with pTM5 DNA which is a plasmid DNA carrying the 1.7 kb recA subfragment of 3 kb BamHI fragment containing the recA promoter site (Ogawa, T. et al., 1978). The pTM5 DNA were treated with BamHI and hybridized with No.1 fraction. A representative structure observed under the electron microscopy was shown in Fig. 6(b). An R-loop structure was seen but a small RNA tail was observed which was probably a sequence complementary to the 1.3 kb-subfragment lost in the pTM5 plasmid. The length from the end of the shorter arm to the begining point of R-loop structure was 0.98±0.06 and corresponded to the shorter arm of the R-loop formed with 3 kb-DNA. The summary of the above results is shown in Fig. 7(c). At about 0.99 kb in length from BamHI site in the side of srl gene, the promoter of the recA gene should be located and the length of the recA gene transcripts is 1.04 kb. Since this transcript can be translated to recA protein in in vitro, the size of the recA gene should be within 1.04 kb. The fact that the transcription of the recA gene from this promoter terminated effectively at 1.04 kb in length means that the induction of the recA gene transcript is limited to the recA gene itself and not extend to the down stream genes.

How is such RNA made that formes Y-like structure by hybridization with the 3 kb DNA fragment? Whether RNA correctly transcribed from the promoter of the recA gene but read through the termination site to the end of the DNA fragment or RNA erroneously transcribed from the end of the DNA fragment and stopped at the termination site? To distinguish between above two possibilities, the length of each part of the Y-like structure was measured. About 80% of the Y-like molecules (7 molecules

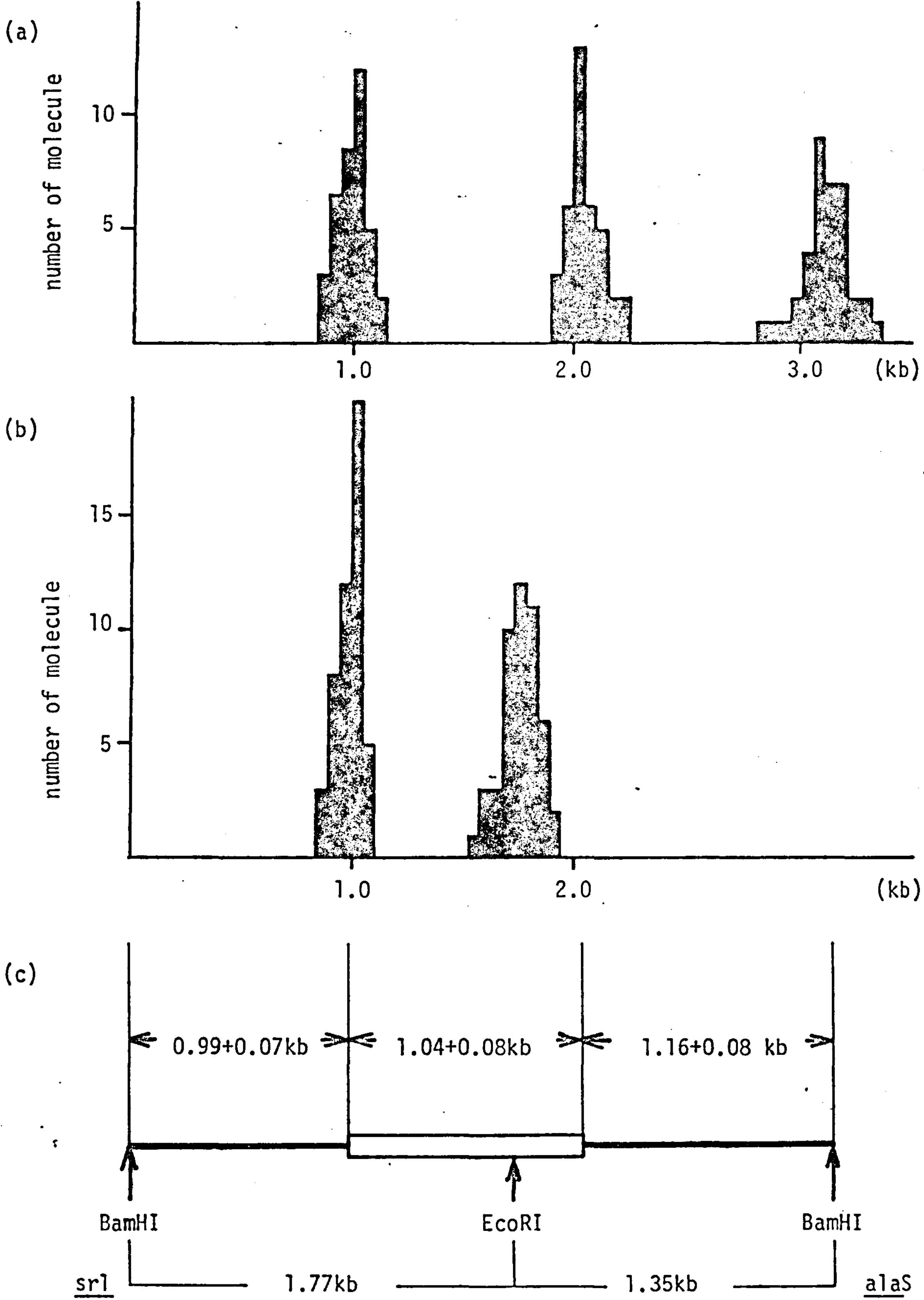


Fig. 7

Fig. 7. The location of the recA gene on the Bam HI 3 kb-DNA by R-loop mapping.

RNA transcribed *in vitro* from the Bam HI 3 kb- DNA was heteroduplexed with the Bam HI 3 kb-DNA.

(a) Length distribution of each part of the R-loop structure of the Bam HI 3 kb-DNA. The abscissa represents the length from the left end of the molecules, which are arranged so that the shorter arm (from the end to nearer branch point) is in the left. The location of each histogram, from the left, corresponds the length of the shorter arm, the length between two branch points and the length of the longer arm, respectively. Total molecules scored were 46.

(b) Length distribution of each part of the R-loop structure with the Bam HI cleaved pTM5 DNA. The representation is the same as the legend of (a). Total molecules scored were 48.

(c) A schematic representation of the location of the recA gene on the Bam HI 3 kb-DNA fragment. The location of the recA gene is indicated by the box. The orientation of the transciption is from left to right.

among 9 scored molecules) had 0.98±0.027 kb in length from the one end. Therefore, most of all RNA seems to start from the recA promoter site correctly but occasionally to fail to terminate at the usual termination signal immediately after the recA gene.

4. Cleavage map of the 3 kb BamHI fragments

To construct a cleavage map, the 3 kb BamHI DNA was digested with the restriction endonucleases EcoRI, HaeII, AvaII, HaeIII and HinfI. The sizes of the digestion products are presented in Table 2. Digestion with EcoRI yielded two fragments which were purified and then digested separately with one or more restriction enzymes. The sizes of fragments thus produced are shown in Table 3. The AvaII-B, HaeII-A, HaeIII-C and HinfI-C fragments found in digests of the intact 3 kb fragment, were not present in digests of the EcoR1-A or -B fragments (Table 3). This indicates that these fragments contain the EcoR1 sits (Fig. 8). Since the EcoR1-B fragment was not cleaved by the HaeII enzyme, the two HaeII cleavage sites on the 3 kb DNA fragment are located in the EcoR1-A fragment (Fig. 8, Table 2 and 3). Cleavage of the EcoR1-A fragment with a combination of the HaeII and AvaII enzymes did not give the AvaII-A fragment while the AvaII-D, HaeII-B and -C fragments were still generated (Table 3). Therefore all four AvaII DNA fragments can be arranged as shown in Fig. 8. From the results of double digestion of the EcoR1-A fragment by the HaeII and HinfI enzymes, the HinfI-D and -G fragments, and the HaeII-B and -C fragments can be mapped (Table 3 and Fig. 8). By a similar series of double digestion experiments, the locations of the HaeIII-B and -D and HinfI-F and -H fragments can be assigned using the AvaII and HaeII cleavage sites as the key map positions. Also, on the EcoR1-B fragment, the HaeIII and

Table 2

Enzyme	EcoRl	Haell	Avall	Haelll	Hinfl
Fragment A	1770*	2600*	1150	890	760
B	1350*	310	1070 (1089)	760	640
C		210	770	570 (599)	640 (652)
D			100 (101)	325 (318)	380
E				300 (302)	275 (280)
F				160 (166)	235 (228)
G				125 (132)	170
H					45 (51)

Table 2. Lengths of fragment produced from the 3 kb BamHI -DNA after treatment with restriction endonucleases.

The 3 kb BamHI -DNA was digested with each restriction enzyme. The sizes of the resultant fragments were estimated from their electrophoretic mobility. The numbers in parentheses are the sizes determined from the nucleotide sequence. For the calculation of the lengths shown here, 6.6 kg is assigned to the ColEl DNA (Tomizawa personal communication). *Those values are calculated as the sum of the lengths of sub-fragments formed by secondary digestion by other enzymes (see Table 3).

Table 3

	EcoRI A Fragment									EcoRI B Fragment					
Enzyme	AvaII	HaeII	HaeIII	HinfI	AvaII + HaeII	HinfI + HaeII	HaeIII + AvaII	HinfI + AvaII	HaeIII + HinfI	AvaII	HaeIII	HinfI	HaeIII + AvaII	HinfI + AvaII	HaeIII + HinfI
Fragments with heterologous ends*	470†	1180†	280†	290†	600 470†	290† 140 70	280† 180	290† 130 90	280† 240	600†	330†	360†	330† 125	360† 250	330† 135 115 35
Fragments with homologous ends**	AII-A AII-D	HII-B HII-C	HIII-B HIII-D HIII-E HIII-G	FI-B FI-D FI-F FI-G FI-H	HII-B HII-C AII-D	FI-B FI-F FI-H HII-B	HIII-B HIII-E HIII-G AII-D	FI-B FI-D FI-G FI-H	HIII-E HIII-G FI-D FI-F FI-G FI-H	AII-C	HIII-A HIII-F	FI-A FI-E	HIII-F AII-C	FI-A	FI-A

* Fragments with a BamHI end are not included.

** Fragments with a BamHI end are included.

† Fragments with an EcoRI end.

Table 3. Lengths of the fragments produced from the EcoRI-A and -B fragment after treatment with one or more restriction enzymes.

The EcoRl-A and -B fragments were digested with one or a various combination of the restriction enzymes and lengths of the fragments formed were determined. Each fragment is identified by a combination of the symbols under the number of base pairs. Symbols RI, AII, HII, HIII and FI represent the restriction enzyme EcoRl, AvaII, HaeII, HaeIII and HinfI, respectively, and, for example, AII-A means AvaII-A fragment. *Fragments with a BamHI are included with fragments with homologous ends. †Fragments with an EcoRl end.

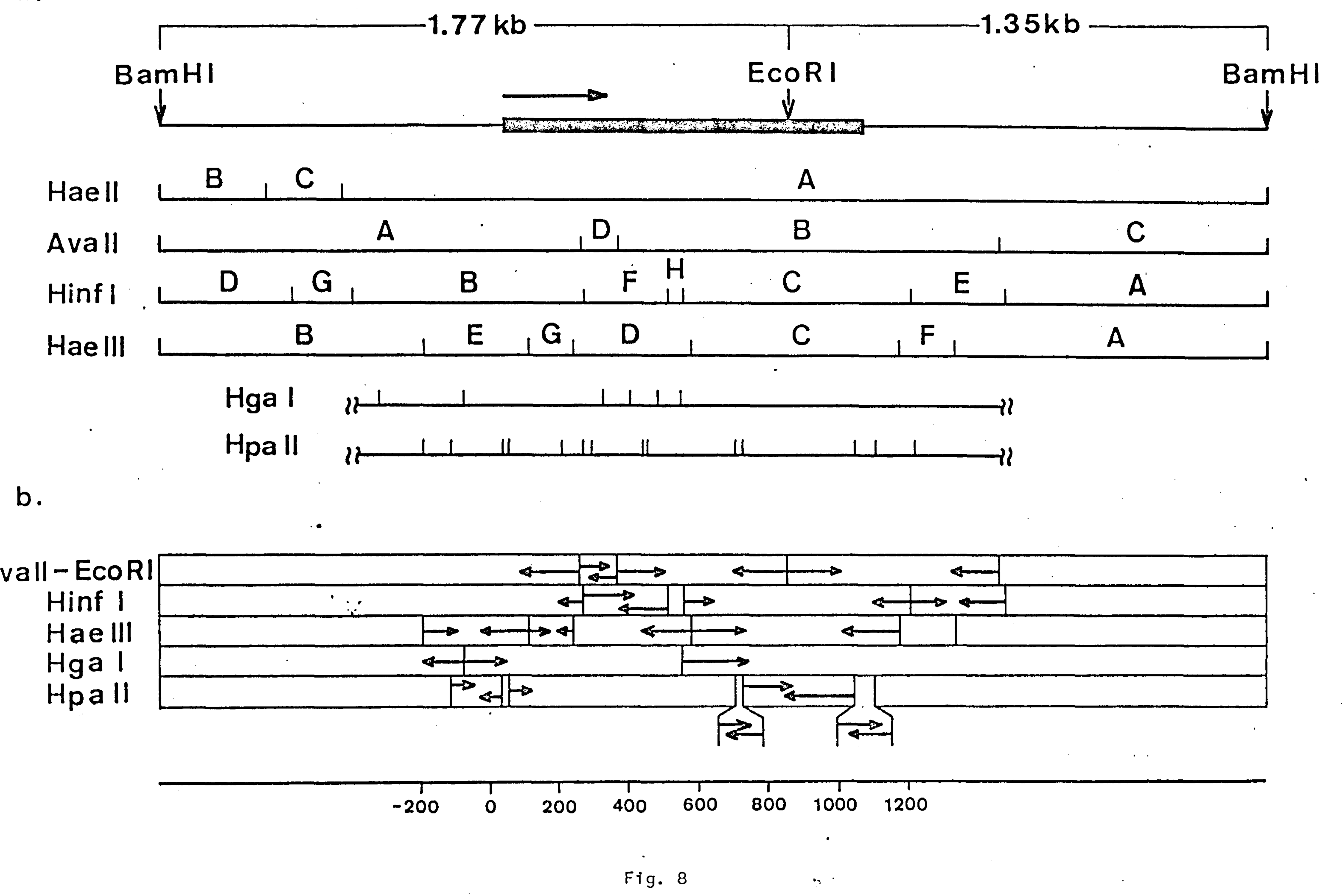


Fig. 8

Fig. 8. Cleavage maps and sequencing strategy

(a) Cleavage maps of the 3 kb BamHI -DNA segment. The map is constructed using the results presented in Tables 1 and 2, and the information obtained from the nucleotide sequence of Fig. 2. The thickened line indicates the region transcribed *in vitro* by RNA polymerase (15). The horizontal arrow indicates the direction of transcription. The positions of HgaI and HpaII sites are based on the nucleotide sequence.

(b) The fragments used for determination of nucleotide sequence of the region that covers the transcribed region. The direction and the extent of the sequence determination are shown by the arrows. An arrow indicates a [5'-^{32}P]-labeled strand aligned in the 5' to 3' direction. Numbers on the bottom line indicate the distance (in nucleotides) from the predominant site for initiation of RNA synthesis *in vitro*.

HinfI sites are mapped from the three double digestions with AvaII and HaeIII enzymes, the AvaII and HinfI enzymes and HaeIII and HinfI enzymes. Although the HaeIII-G and -E fragments are located in the HinfI-B fragment, the relative position of the two can not be determined from the experiments described above. In addition, small fragments (less than 30 base pairs) could have been missed. Assignment of the precise locations of the cleavage sites is based on the nucleotide sequence, when possible. A restirction enzyme map of the region covering the recA gene has recently been reported elsewhere (Sancar and Rupp, 1979).

5. The nucleotide sequence of the recA gene

Making use of the restriction enzyme map of the 3 kb BamHI fragment (Fig. 8a), we next determined the nucleotide sequence of the middle region of the fragment which was already though to be the location of the recA gene. The DNA fragments used in the sequence determination are indicated in Fig. 8b. The sequence is shown in Fig. 9. Inspection of the sequence shows that the region could encode a protein containing 353 amino acids.

6. Amino acid sequence of the N-terminal region and the amino acid content of the recA protein

The size of the protein (37,800 dalton) predicted from the nucleotide sequence coinsides with that of the recA protein (39,000 daltons measured by gel electrophoresis) isolated from bacteria. To show that this region in fact encodes the recA protein, the amino acid sequence of five residues from the N-terminus of the recA protein synthesized in vivo was determined (data not shown). The sequence is NH_2-Ala-Ile-Asp-Glu-Asn-. This agrees with the sequence predicted from

```
           -100                                                    -50                                                      +1
TGCTTCAGCGGCGACCGTGATGCGGTGCGTCGTCAGGCTACTGCGTATGCTTGCAGACCTTGTGGCAACAATTTCTACAAAACACTTGATACTGTATGAGCATACAGTATAATTGCTTCAACAGAACATATT

                    +50                                                    +100
GACTATCCGGTATTACCCGGCATGACAGGAGTAAAAATGGCTATCGACGAAAACAAACAGAAAGCGTTGGCGGCAGCACTGGGCCAGATTGAGAAACAATTTGGTAAAGGCTCCATCATGCGCCTGGGTGAA
                                     fMetAlaIleAspGluAsnLysGlnLysAlaLeuAlaAlaAlaLeuGlyGlnIleGluLysGlnPheGlyLysGlySerIleMetArgLeuGlyGlu
  +150                                                     +200                                                     +250
GACCGTTCCATGGATGTGGAAACCATCTCTACCGGTTCGCTTTCACTGGATATCGCGCTTGGGGCAGGTGGTCTGCCGATGGGCCGTATCGTCGAAATCTACGGACCGGAATCTTCCGGTAAAACCACGCTG
AspArgSerMetAspValGluThrIleSerThrGlySerLeuSerLeuAspIleAlaLeuGlyAlaGlyGlyLeuProMetGlyArgIleValGluIleTyrGlyProGluSerSerGlyLysThrThrLeu
                    +300                                                     +350                                                     +400
ACGCTGCAGGTGATCGCCGCAGCGCAGCGTGAAGGTAAAACCTGTGCGTTTATCGATGCTGAACACGCGCTGGACCCAATCTACGCACGTAAACTGGGCGTCGATATCGACAACCTGCTGTGCTCCCAGCCG
ThrLeuGlnValIleAlaAlaAlaGlnArgGluGlyLysThrCysAlaPheIleAspAlaGluHisAlaLeuAspProIleTyrAlaArgLysLeuGlyValAspIleAspAsnLeuLeuCysSerGlnPro
                                     +450                                                     +500
GACACCGGCGAGCAGGCACTGGAAATCTGTGACGCCCTGGCGCGTTCTGGCGCAGTAGACGTTATCGTCGTTGACTCCGTGGCGGCACTGACGCCGAAAGCGGAAATCGAAGGCGAAATCGGCGACTCTCAC
AspThrGlyGluGlnAlaLeuGluIleCysAspAlaLeuAlaArgSerGlyAlaValAspValIleValValAspSerValAlaAlaLeuThrProLysAlaGluIleGluGlyGluIleGlyAspSerHis
        +550                                                     +600                                                     +650
ATGGGCCTTGCGGCACGTATGATGAGCCAGGCGATGCGTAAGCTGGCGGGTAACCTGAAGCAGTCCAACACGCTGCTGATCTTCATCAACCAGATCCGTATGAAAATTGGTGTGATGTTCGGTAACCCGGAA
MetGlyLeuAlaAlaArgMetMetSerGlnAlaMetArgLysLeuAlaGlyAsnLeuLysGlnSerAsnThrLeuLeuIlePheIleAsnGlnIleArgMetLysIleGlyValMetPheGlyAsnProGlu
                     +700                                                     +750                                                     +800
ACCACTACCGGTGGTAACGCGCTGAAATTCTACGCCTCTGTTCGTCTCGACATCCGTCGTATCGGCGCGGTGAAAGAGGGCGAAAACGTGGTGGGTAGCGAAACCCGCGTGAAAGTGGTGAAGAACAAAATC
ThrThrThrGlyGlyAsnAlaLeuLysPheTyrAlaSerValArgLeuAspIleArgArgIleGlyAlaValLysGluGlyGluAsnValValGlySerGluThrArgValLysValValLysAsnLysIle.
                                          +850                                                     +900
GCTGCGCCGTTTAAACAGGCTGAATTCCAGATCCTCTACGGCGAAGGTATCAACTTCTACGGCGAACTGGTTGACCTGGGCGTAAAAGAGAAGCTGATCGAGAAAGCAGGCGCGTGGTACAGCTACAAAGGT
AlaAlaProPheLysGlnAlaGluPheGlnIleLeuTyrGlyGluGlyIleAsnPheTyrGlyGluLeuValAspLeuGlyValLysGluLysLeuIleGluLysAlaGlyAlaTrpTyrSerTyrLysGly
           +950                                                     +1000                                                    +1050
GAGAAGATCGGTCAGGGTAAAGCGAATGCGACTGCCTGGCTGAAAGATAACCCGGAAACCGCGAAAGAGATCGAGAAGAAAGTACGTGAGTTGCTGCTGAGCAACCCGAACTCAACGCCGGATTTCTCTGTA
GluLysIleGlyGlnGlyLysAlaAsnAlaThrAlaTrpLeuLysAspAsnProGluThrAlaLysGluIleGluLysLysValArgGluLeuLeuLeuSerAsnProAsnSerThrProAspPheSerVal
                               +1100                                                    +1150                                                    +1200
GATGATAGCGAAGGCGTAGCAGAAACTAACGAAGATTTTTAATCGTCTTGTTTGATACACAAGGGCGCATCTGCGGCCCTTTTGCTTTTTTAAGTTGTAAGGATATGCCATGACAGAATCAACATCCCGTCG
AspAspSerGluGlyValAlaGluThrAsnGluAspPhe
```

Fig. 9. The nucleotide sequence of the <u>recA</u> gene and its neighborhood, and the amino acid sequence of the encoded protein. The DNA sequence is numbered beginning at the predominant site of initiation of <u>in vitro</u> transcription.

the nucleotide sequence, except for the absence of a formylmethionine residue at the N-terminal end. The amino acid composition of the purified recA protein also agreed with the composition of the protein predicted from the nucleotide sequence, again with the exception of a formylmethionine residue (Table 4).

7. The site of initiation of transcription

It has been shown that in vitro synthesis of the recA messenger RNA starts approximately 1 kb to the right of the BamHI site within the EcoR1-A segment. Inspection of the nucleotide sequence of this region reveals a sequence which has many similarities to sites known to act as promoters for E. coli RNA polymerase (discussion below). Considering this information, we determined a site of initiation of transcription in the HaeIII-E fragment that contains the suspected region.

RNA was synthesized on the fragment with RNA polymerase. In a reaction mixture the concentration of each rNTP was 10 μM including [α-^{32}P]UTP. Three other reaction mixtures contained 200 μM each of ATP, GTP or CTP and other rNTPs at 10 μM each. Electrophoretic analysis of the products from these reactions showed that a single species of RNA of approximately 95 nucleotides was always the predominant product (Fig. 10). An increase in the ATP concentration stimulated RNA synthesis about 5-fold while increases in the GTP or CTP concentration had small stimulatory or inhibitory effects respectively. These results suggest that the transcription started mostly with ATP, and set the site of transcription initiation at around the position No.1 shown in Fig. 9.

The exact position of initiation was determined by the analysis of the nucleotide sequence of the RNA of approximately 95 nucleotides at

Table 4

Amino acid composition of *recA* protein

Amino acid	Residues per molecule	
	Predicted value[a.]	Measured value[b.]
Lysine	27	26.4
Histidine	2	2.6
Arginine	14	15.2
Asparagine	15	34.7
Aspartic acid	20	
Threonine	17	17.0[c.]
Serine	20	18.8
Glutamine	13	42.9[c.]
Glutamic acid	30	
Proline	10	10.6
Glycine	35	34.2
Alanine	38	38.0
Cysteine	3	N.D.[d.]
Valine	22	24.0
Methionine	9	8.6
Isoleucine	27	25.1
Leucine	31	30.0
Tyrosine	7	6.8
Phenylalanine	10	9.8
Tryptophan	2	N.D.[d.]
	352	

a. Predicted value was from the DNA sequence analysis described in Fig. 9

b. Acid hydrolysis was performed for 24 and 72 h. The value of threonine and serine were obtained by extrapolation to zero time of hydrolysis. Values of valine and isoleucine were of 72 h-hydrolysate.

c. Sum of acid and amide forms.

d. Not determined.

1 2 3 4

—a

—b

Fig. 10

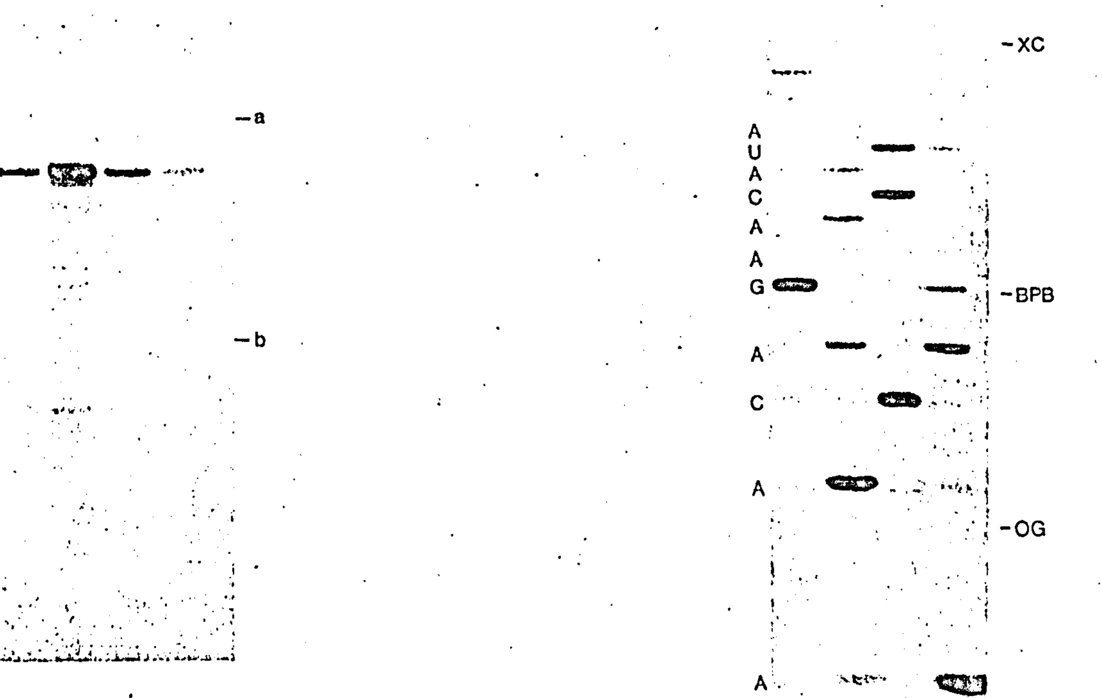


Fig. 11

Fig. 10. Autoradiograph of transcripts synthesized from the Hae III-E fragments by RNA polymerase. Transcripts were labeled with [α-^{32}P]UTP in the presence of 10 μM of each rNTP except that the concentrations of ATP, GTP and CTP were 200 μM for transcripts shown in lane 2, 3 and 4, respectively. Electrophoresis in an 8 % acrylamide-urea gel was carried out at 1000 volts for 3 h. The positions indicated by a and b are location of bands formed by RNA-1 of ColE1 (108 nucleotides, Morita and Oka, 1979) and 4S RNA of phage λ (77 nucleotides, Rosenberg et al.,1976).

Fig.11. Autoradiograph of partial digests by various RNases of [α-^{32}P]ATP labeled nucletodies synthesized from the HaeIII-E fragments. [α-^{32}P]ATP labeled transcripts from the HaeIII-E fragments were fractionated in an 8% gel as described in Fig.10 and RNA of about 95 nucleotide long was extracted electrophoretically and concentrated by ethanol precipitation in the presence of 20μg of E. coli tRNA. The precipitate was dissolved in water, divided in 4 equal portions and dried. The sample was dissolved in 4μl of the buffer to which 1μl of an RNase solution was added. The conditions used (personal communication of Drs. T. Itoh and J. Tomizawa) were as follows: lane G, 0.03μg of RNase T_1 in 0.1M Tris-HCl, 10mM EDTA, pH 7.5; lane A, 0.1 unit of RNase U_2 in 20mM sodium citrate, pH 3.5; lane C,U, 3 μg of RNase A in the same buffer for RNase T_1 digestion and lane -C, 0.5 units in 10mM sodium acetate, 1mM EDTA, pH 5.0. Samples were incubated for 30 minutes at 0°C except that the reaction with RNase Phy I was carried out at 25°. Gel electrophoresis was carried out in 25% acrylamide-urea gel at 2000 volts for 5 hours. XC, BPB and OG indicate the positions of xylencyanol FF, bromphenol blue and orang G, respectively. It has been shown that orang G moves slightly faster than dinucleotide pentaphosphate and the mobility of bromphenol blue is similar to that of hexanucleotide nanophosphate (personal communication of Dr. J. Tomizawa).

its 5' end. The sequence was deduced from the products of partial digestion (Simoctizt, et al., 1977) of the [α-^{32}P]ATP-labeled RNA by various RNases (T. Itoh and J. Tomizawa, personal communication). The result in Fig. 11 shows that the 5' sequence of the major labeled RNA is pppAAGCAGAACAUA--. The presence of faint bands such as di-, tri- and tetranucleotides formed by digestion with RNase A, RNase U_2 and RNase T_1, respectively, suggests the existence of a minor transcript that begins at a position corresponding to the next A residue to the major initiation site. Therefore, transcription initiates most frequently at position 1 and much less frequently at position 2 in the nucleotide sequence (Fig. 9).

Considerations

In vitro transcription of the cloned 3 kb-DNA fragment

The major RNA molecules synthesized by E. coli RNA polymerase holoenzyme had about 1 kilo base long and found to be transcribed from the middle part of the 3 kb-DNA by the R-loop formation with the DNA. The transcription starts from a fixed point, 0.99 kb from the BamHI site near the srl gene, and terminates at another fixed point. Since the length of the transcription is just enough for coding only recA protein, the starting point should be the promoter of the recA gene and another the terminal point of the transcription. In this transcription system, the termination was occured without rho-factor. This fact indicates that the termination of recA mRNA synthesis is rho- independent.

Transcriptional regulation of the recA gene

Some of the treatments to induce "SOS-functions"; addition of nalidixic acid of mitomycin C, increased the rate of recA mRNA synthesis immediately. This fact showed that the expression of recA gene is mainly controled at a stage of transcription. The size of recA mRNA synthesized in in vivo co-sedimented with the major product, No.1 fraction, of in vitro transcription on a sucrose gradient. Therefore, it was concluded that the in vitro transcript is not an artifact and the recA gene is transcribed monocistronic in cells. In high induction state, the monocistronic transcription will ensure that the induction is confined to the recA gene, and will provide the effective and economical induced synthesis in a cell. Moreover, this monocistronic transcription of the recA gene excludes a cotranscription model for the induction of "SOS-functions": some genes responsible for "SOS-functions"

locate at the down stream of the recA gene and simultaneously transcribed by a polycistronic manner from the recA promoter whenever the induced synthesis of the recA protein occurs (Gudas and Pardee, 1975).

Structural informations of recA gene

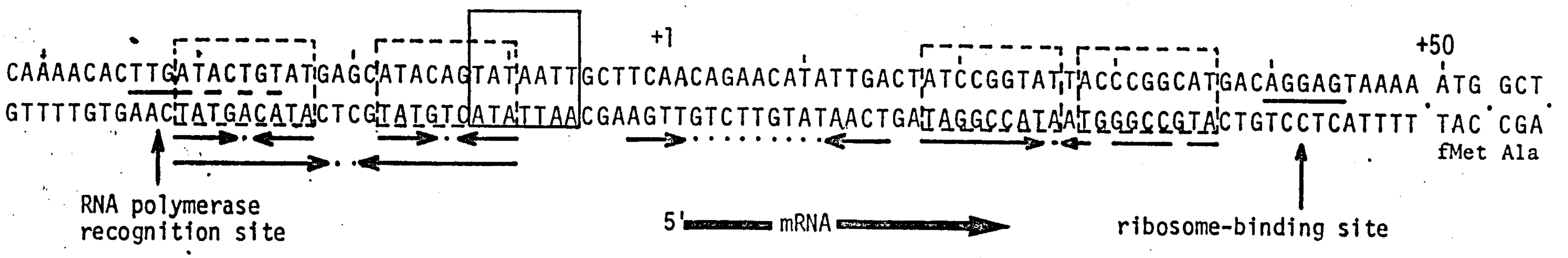
(1) Regulatory region:

The observation that the recA gene is transcriptionaly regulated indicates the presence of specific sequence(s) that is involved in a regulatory mechanism. It has been suggested that the lexA protein inhibits the expression of the recA gene (Gudas, 1976). The complex structure with dyad symmetries in the promoter region (Figure 12) could contain the site where the lexA protein interacts with the DNA.

(2) Initiation of transcription:

The known promoter regions for E. coli RNA polymerase have certain structural homologies, particularly in two regions about 10 and 35 nucleotides upstream of the site where transcription begins. The latter probably provides a site recognized by the RNA polymerase, while the former is where the polymerase binds (Gilbert, 1976; Rosenberg and Court, 1980). Approximately 10 nucleotides upstream of the position where transcription of the recA gene begins, there is the sequence TATAATT which matches the general structure of the RNA polymerase binding site (Pribnow, 1975; Schaller et al., 1974). The same sequence is present in the C_{17} promoter of the phage λ (Rosenberg et al., 1978). About 20 nucleotides further upstream, a sequence that includes highly common TTGA and surrounding exists in the recognition region. These arrangements of nucleotides frequently found in promoter regions are indicated in Fig. 12.

(3) Termination of transcription:



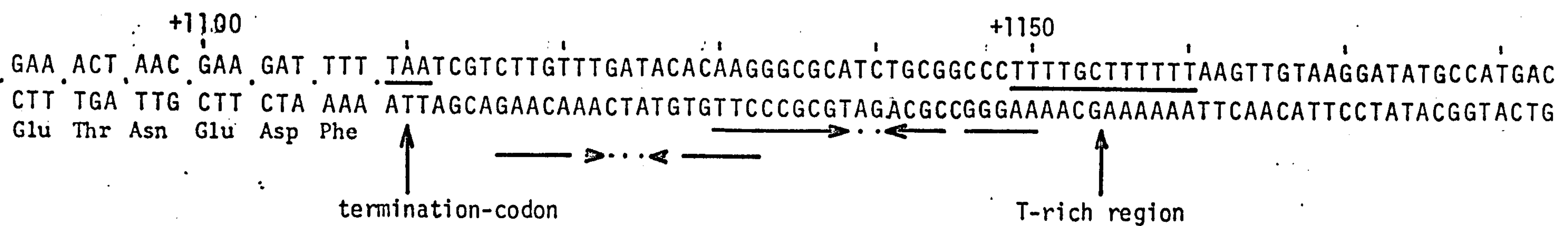


Fig. 12. Nucleotide sequence of the regions of initiation and termination of transcription. The heptanucleotide sequence that probably serves as the binding for the RNA polymerase is boxed with solid lines. The presumed RNA polymerase recognition site, the ribosome-binding site as well as the termination codon and the T-rich sequence in the region of transcription termination are underlined. The regions of dyad symmetries are indicated by arrows and homologous dyad symmetries are boxed with broken lines. Some amino acid residues encoded in the regions are also presented.

It was shown that *in vitro* transcription of the 3 kb *Bam*HI fragment frequently terminates after synthesis of RNA of about 1 kb (without a termination factor *rho*). In the nucleotide sequence of the suspected region of transcription termination, there is a T-rich stretch preceeded by two sequences with dyad symmetry. This region probably signals termination of transcription (Gilbert, 1976; Adhya and Gottesman, 1978). Of the two symmetrical sequences, the one closer to the T-rich stretch is larger and is richer in G-C pairs. If transcription terminates about 20 nucleotides downstream of the center of the symmetry (Rosenberg and Court, 1980), it probably terminates in the T-rich stretch.

(4) Translation:

About 50 nucleotides downstream from the initiation site of transcription, there is an AUG codon which begins the structural portion of the *recA* gene. About 10 nucleotides upstream from the initiation codon, there is the sequence AGGAG which could serve as a ribosome binding site (Shine and Dalgarno, 1975). Starting at this AUG, a protein of 353 amino acids would be made. No protein containing more than 46 amino acid residues could be made by reading the transcript in a different frame.

SECTION 2.

Organization of the *lexA* Gene

Summary of Section 2.

Using a cloned fragment containing the *lexA* gene of *E. coli*, the entire nucleotide sequence of the *lexA* gene has been determined. The probable coding region of the *lexA* gene contains 606 nucleotide residues and encodes a single protein of 202 amino acids. The initiation site of *in vitro* transcription of the *lexA* mRNA has been determined by analysis of the 5'-end nucleotide sequence. Comparison of the DNA sequence of the promoter region of the *lexA* gene with that of the *recA* gene revealed the presence of sequences that are common to both.

For the analysis of *lexA* protein, the coding region of *lexA* gene was fused with the promoter of *lacZ* gene of *E. coli* to increase the *lexA* protein production. In the cells harboring a plasmid pTH227, which bears the *lexA*-*lac* fusion gene, the *lexA* protein was induced by the addition of lactose up to 1.5 % of the total cellular proteins. Utilizing the cells harboring this plasmid, *lexA* protein was purified nore than 96 % in purity. The amino-terminal sequence and amino acid composition of the purified *lexA* protein were analyzed. The results were in agreement with the prediction from the nucleotide sequence analysis.

The purified *lexA* protein was cleaved into two peptides with the wild-type *recA* protein in the presence of ATP and single-strand DNA. The cleaved site of the *lexA* protein was determined by the analysis of the amino acid sequence at the ends of cleaved products. The results indicated that the cleavage occurs at one site between the residues of no.84 Ala and no.85 Gly. From a comparison of amino acid sequences around the cleavage site of λ repressor and *lexA* protein, a homologous amino acid sequence composed of 11 amino acids was found at adjacent amino-terminal side of the cleaved site.

Introduction

The coordinate expression of SOS-functions is regulated by at least two gene products, those are recA and lexA proteins. Genetical and biochemical studies suggested that the lexA protein is a negative regulator for recA gene (Mount, 1977; Gudas and Mount, 1977; McEntee, 1977), lexA gene itself (Little and Harper, 1979; Brent and Ptashne, 1980), and colicin E2 gene (Tessman et al., 1978). In addition, recent studies using Mu(lac::Tn3) suggested that the expression of both uvrA gene (Kenyon and Walker, 1980) and uvrB gene (Fogliano and Schendel, 1981) are also controled by the recA-lexA regulation system. On the other side, the recA protein seems to work for the induction of SOS-functions: the purified recA protein cleaves the λ repressor (Craig and Roberts, 1980), the P22 phage repressor (Phizicky and Roberts, 1980) and the lexA protein (Little and Mount, 1980) in the presence of ATP and single-strand DNA.

In this section, to obtain an insight into molecular mechanism of the coordinate regulation of SOS-functions, the structure analysis of the lexA gene and purification and characterization of the lexA protein were carried out. Comparison of the nucleotide sequence of the promoter region of the lexA gene with that of recA gene revealed the presence of sequences that are common to both. The purified lexA protein is found to be cleaved by recA protein in the presence of the single-strand DNA and ATP. The amino acid sequence at the cleavage site was determined and some similarity between the amino acid sequence of the lexA protein and the λ repressor was found.

Results of Section 2

1. Construction of a plasmid pGC3 carrying the lexA gene

An Eco RI DNA fragment derived from the lexA transducing λ phage, λdlexA^{+}-8 (see Materials and Methods) was inserted into the Eco RI site of pBR322 plasmid (Bolivar et al., 1977) and a plasmid which suppressed the tsl mutation (temperature-sensitive mutation of lexA gene) of DM511 strain was selected . The resulting plasmid, pMN1, carrys a 15 kb DNA fragment of the lexA transducing λ phage DNA in addition to the lexA gene. To remove the portion of λ phage DNA of the plasmid, pMN1 DNA was partially digested with Pst I, self ligated and then transformed to DM511 tsl bacteria. Tetracycline resistant and tsl^{+} (survival at 42°C) transformants were selected. A plasmid, pMN102, thus obtained contains a bacterial DNA fragment of about 3 kb in place of the fragment between Eco RI and Pst I sites of pBR322. The cleavage map of the pMN102 with Hinc II and Hpa I is presented in Fig. 13. The 3 kb-bacterial fragment contains two Hpa I sites and two additional Hinc II sites (Hinc II enzyme can recognize Hpa I sensitive sites).

The plasmid pMN102 DNA was, then, partially digested with Hinc II and the resulting fragments were ligated and used to transform DM511 cells From the independent tetracycline-resistant transformants, four smaller plasmids were obtained. The cleavage maps of these four plasmids pGA1, pGA2, pGA3 and pGC3 are shown in Fig. 13. Only pGC3 could confer the ability to complement the tsl mutation of the DM511 strain.

The phenotypes conferred by these plasmids in various host cells were examined and the results obtained were summarized in Table 5. Plasmids pMN102 and pGC3 are able to complement the tsl mutation: substantially reduced tif filamentous death of JM12 cells and the turbid plaque formation of λ phage on DM1187 spr tif sfi cell (STS-cell). pGA1, pGA2, pGA3 or the control plasmid pBR322 did not affect on these properties. A wild-type

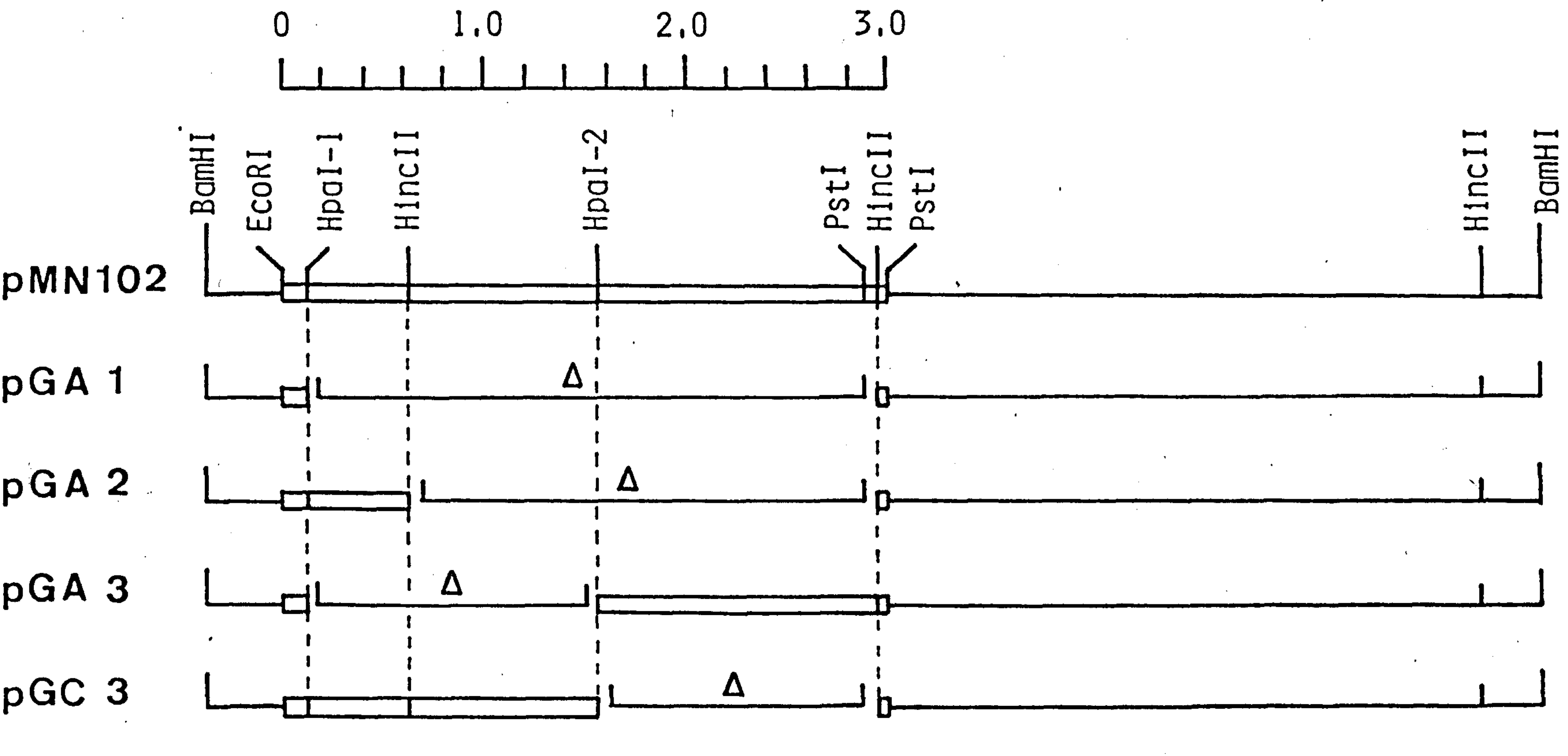


Fig. 13. Pysical maps of the plasmid pMN102 and its derivatives. Open and solid bars represent the DNA from λdlexA+-8 and pBR322 DNA regions respectively. Plasmids are presented in linear form opened at the Bam HI site in the pBR322 DNA. The brackets with triangles indicate regions deleted from pMN102 DNA by the partial digestion with Hinc II and subsequent self-ligation. Hinc II can cleave Hpa I-sensitive site as well. Upper line shows the distance in kilobases (kb) from the Eco RI site.

Table 5 Phenotypes Conferred by the Recombinant Plasmids

Plasmids	Complementation for tsl[a]	Suppression of Colicin E2 Production[b]	Formation of λ Turbid Plaque on STS strain[c]	UV Resistance[d]	Suppression of tif-killing[e]
pMN102	+	+	+	s	+
pGC3	+	+	+	s	+
pGA1	-	-	-	r	-
pGA2	-	+	-	r	-
pGA3	-	-	-	r	-
pMN223	+	+	+	s	+
pMN206	+	+	(+)	(s)	-
pMN300	-	-	-	r	-
pMN400	+	-	-	r	-
pBR322	-	-	-	r	-

Legend to Table 5.

a. Plus means that the plasmid could suppress the host killing of DM511 at 42°C.

b. Colicinogenic strain JC1557 (ColE2) was transformed with the plasmids indicated, and spontaneous colicin E2 production was examined by over-layered with an indicator strain, CL142 (Ozeki et al., 1962). Plus means that the plasmid could suppress the spontaneous colicin E2 production.

c. λ wild-type phage was plated with DM1187 harboring each of plasmids and incubated at 32°C. Plus means that the λ phage could make turbid plaque. Parenthesis means not distinctive turbid plaque nor clear plaque, "intermediate".

d. "s" means sensitive and "r" resistant as wild-type. Parenthesis means intermediate between "s" and "r".

e. Plasmids were introduced into the _tif-1_ cells (JM12) and resulting transformed cells were plated on minimal plate containing 0.1 % Casamino acid and 75 μg/ml adenine at 42.5°C. Plus means that the plasmid could suppress the host killing of JM12 at 42.5°C.

cell (C600) carrying plasmid pMN102 or pGC3 becomes sensitive to UV-irradiation (about 12 times). One of the reasons is that the increase of lexA gene product overrepresses the recA gene expression and thus sppresses the host SOS-functions. Other plasmids could not confer this property to their host. One of the SOS-functions is the induction of production of various colicines. The plasmids pMN102, pGC3 and pGA2 conferred the ability to suppress completely spontaneous production of colicin E2, while pGA1, pGA3 or pBR322 did not. It is interesting that the plasmid pGA2 that could not confer the lexA^{+} phenotypes in other respects suppressed activity of the colicin E2 production. This raised a possibility that the lexA protein directly involves in the expression of the colicin E2 gene (see Consideration). The results described above indicate that the Eco RI-Pst I, 1.6 kb fragment of plasmid pGC3 contains the complete lexA gene.

2. Fine mapping and cloning of the sub-fragment of Eco RI-Pst I 1.6 kb DNA

A more precise cleavage map of the 1.6 kb chromosomal region of pGC3 was constructed using Alu I, Hinf I and Taq I restriction enzymes. The results are shown in Fig.14. To minimize the flanking region of the lexA gene, sub-fragments derived from the 1.6 kb Eco RI-Pst I fragment by cleavage with these enzymes were re-cloned. Purified Eco RI-Pst I 1.6 kb fragment was partially digested with Alu I and then ligated with Hinc II-partial digests of pBR322 DNA which had been cleaved previously with Eco RI or Pst I. Plasmids that carry the exogenous fragment in the β-lactamase gene of pBR322 were selected by isolating transformants resistant to tetracycline and sensitive to ampicillin simultaneously. The DNAs of plasmids thus obtained were analyzed by cleaving with various restriction enzymes. The results show that three types of plasmids each represented by pMN206, pMN223 and pMN300 were obtained. The cleavage maps of these plasmids are shown in Fig. 14. By similar procedures, the plasmid which has a 1.5 kb sub-fragment generated by the Hpa I partial digestion of Eco RI-Pst I fragment in the

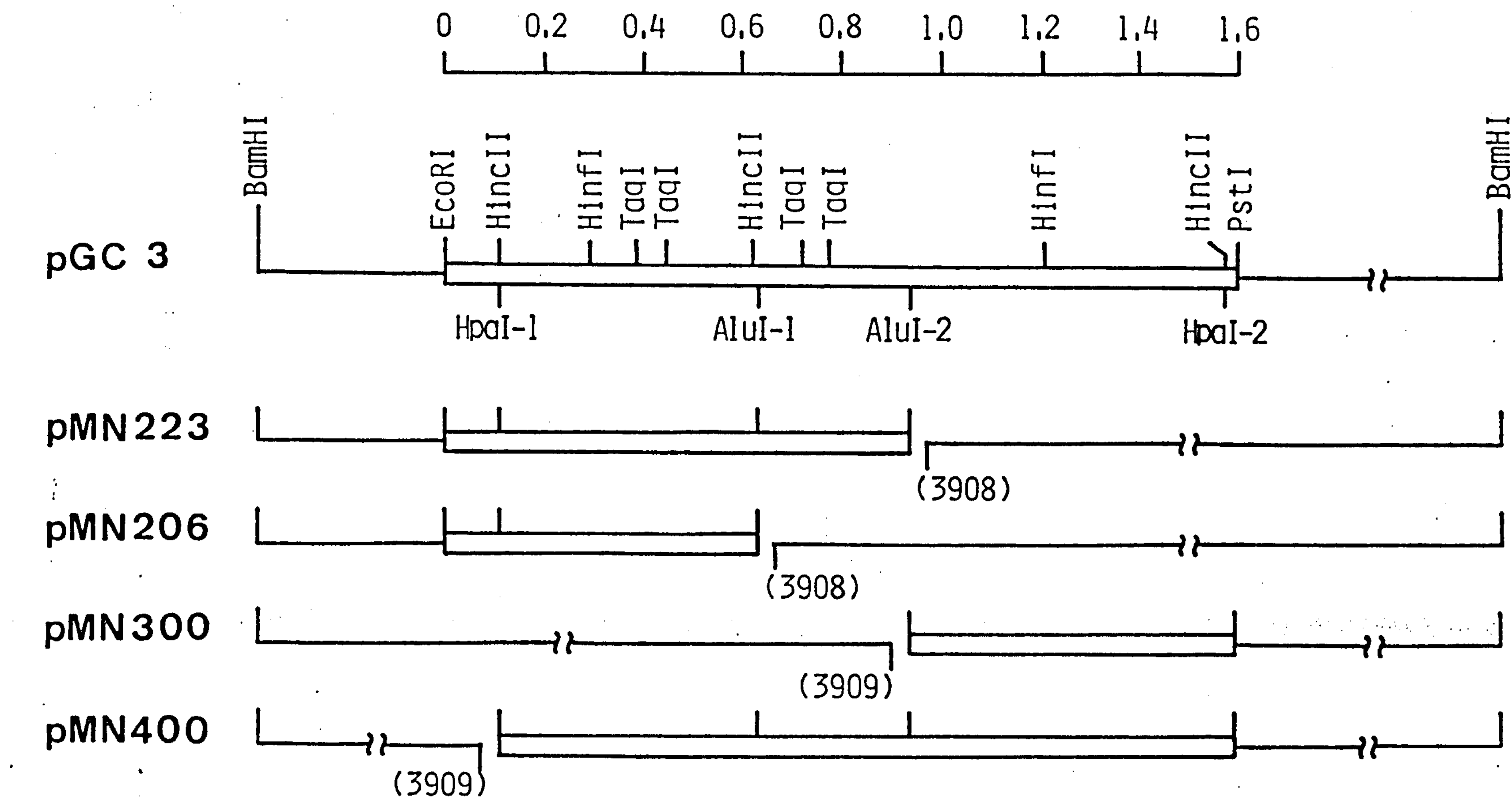


Fig. 14. Fine cleavage map of the chromosomal region of pGC3 and of other plasmids carrying a sub-fragment of the lexA region. Upper line shows the length of DNA in kb. Chromosomal DNA regions (open bars) of the plasmids are aligned in accordance with the derivation from the chromosomal region of pGC3. Solid bars represent pBR322 DNA regions. Numbers in the parentheses are the nucleotide number of cutting site of enzymes in the pBR322 nucleotide sequence (Sutcliffe, 1978).

β-lactamase gene of pBR322 was also obtained (pMN400 in Fig. 14). These plasmids were tested to determine whether or not they carry the whole lexA gene. The pMN223 which contains Eco RI-Alu I-2, about 950 bp fragment (Fig. 14) conferred the same characteristics to the host bacteria as pGC3 did, whereas pMN300 that contained the other part, Alu I-2-Pst I fragment, did not (Table 5). Therefore, the lexA gene seemed to be located between the Eco RI site and the Alu I-2 site. In addition pMN206 showed the ability to complement the tsl mutation and suppress the colicin E2 production, but pMN400 showed the ability to complement the tsl mutation but not suppress the colicin E2 production (Table 5). These results suggest that the Alu I-1 and Hpa I-1 restriction sites (Fig.14) are located in the lexA gene and presumably near both ends of the gene. This conclusion was supplemented by the fact that pGA2 which deletes 17 bp inside Eco RI-Alu I-1 fragment at the Alu I-1 site lost a complementation ability for the tsl mutation but still confers the ability to suppress colicin E2 production.

3. Analysis of proteins synthesized from the constructed plasmids in cells

To examine the proteins synthesized from the lexA gene cloned in the recombinant plasmids, we used the maxi-cell technique (Sancar et al., 1979; see Materials and Methods). The results of autoradiography are shown in Fig. 15. In the cells bearing the plasmids pMN102, pGC3 and pMN223 which express the lexA function, the protein with the molecular weight of 24,000 dalton was synthesized instead of β-lactamase (Fig. 15, lane 2, 3 and 6). While in the cells bearing pGA3 which deletes the region assigned to the lexA gene, that protein band was missing (Fig. 15, lane 5). These results suggest that the protein with a molecular weight of 24,000 corresponds to the lexA gene product. The same conclusion was obtained recently by Little and Harper, (1979); and Brent and Ptashne (1980). The slightly larger proteins with molecular weights about 25,000 and 26,000 were

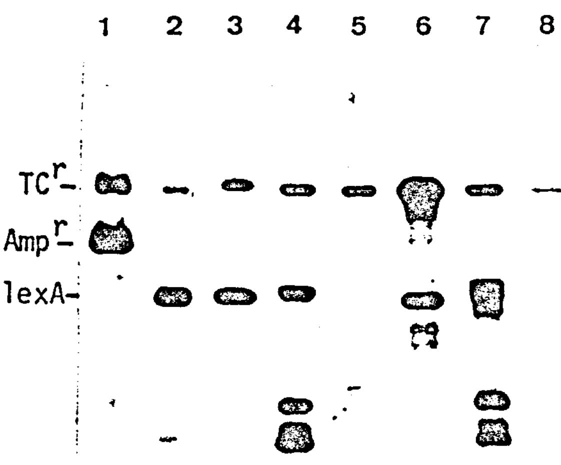


Fig. 15. Autoradiography from SDS-gel electrophoresis of ^{35}S-labeled polypeptides synthesized from the recombinant plasmids.

Methods of labeling for proteins, electrophoresis and autoradiography were described in Materials and Methods. Molecular weights of labeled proteins were estimated from comparison of their mobility with that of marker proteins of known molecular weight. TC^r and Amp^r indicate the products from the genes conferring tetracycline and ampicillin resistance; and lexA indicates the wild-type lexA product with a molecular weight of 24,000. The lanes contained the proteins from the cells harboring the following plasmids: (1) pBR322; (2) pMN102; (3) pGC3; (4) pGA2; (5)pGA3; (6) pMN223; (7) pMN206; (8) pMN400.

synthesized in the cells containing pGA2 and pMN206 respectively (Fig. 15, lane 4 and 7), and the plasmid pMN400 conferred none of the distinctive bands except protein from the tetracycline resistant gene (Fig. 15 lane 8). Since both plasmids, pMN206 and pGA2 probably lost the translation termination site of the lexA gene, the translation of the nascent polypeptide of the lexA gene was continued into an adjacent region and a protein larger than the lexA protein was formed. On the other hand, the pMN400 plasmid probably lost the promoter of the lexA gene and could not synthesize the lexA protein.

4. Initiation site of transcription of the lexA gene

Absence of the synthesis of the lexA protein from pMN400 suggested that the promoter of the lexA gene exists in the Eco RI-Hpa I-1 segment of the cloned fragment. Therefore, to determine the initiation site of the transcription, mRNA was synthesized from the Eco RI-Hpa I fragment (111 bp) with RNA polymerase and the products were analyzed by electrophoresis. The major transcripts synthesized in the standard reaction (see Materials and Methods) were about 25-35 nucleotides long (not shown). In the reaction mixture containing each of rNTP at restricted concentration (10 μM), small amounts of RNA of about 25 nucleotides was synthesized (Fig. 16 A, lane 1, R3). An increase in the concentration of GTP or CTP had only small stimulatory effect on the synthesis of R3 RNA (Fig. 16 A, lane 3 and 4), whereas, with increased ATP concentration longer RNA products about 30-35 nucleotides long were synthesized (Fig. 16 A,lane 2, R1 and R2) as in the normal reaction and the synthesis of R3 RNA was also stimulated. These findings suggest that the transcription started mostly with ATP.

To analyze the nucleotide sequence of the 5' end of the major products R1, R2 and R3, RNAs synthesized in the presence of [γ-^{32}P]ATP were isolated from the gel and partially digested with various RNases as described in Scimocszit et al., (1977);and Itoh and Tomizawa (1980). No difference

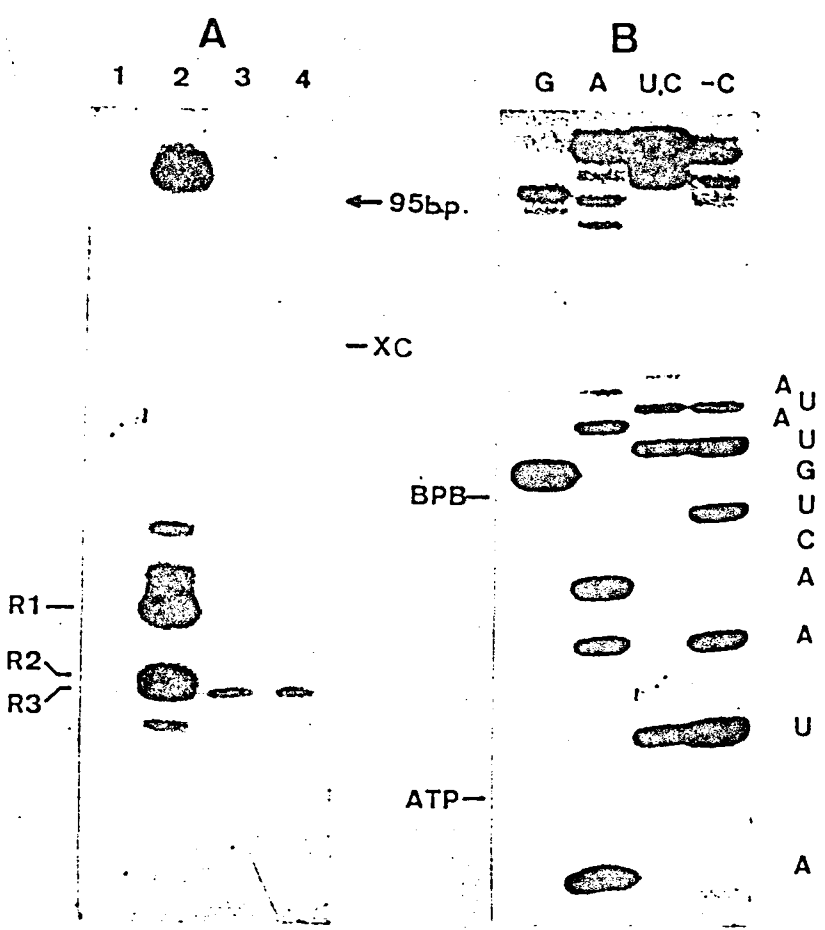


Fig. 16

Fig. 16. Gel electrophoretic patterns of RNA transcribed *in vitro* from the *Eco* RI-*Hpa* I fragment.

A. Autoradiograph of transcripts. Transcripts were labeled with [α-^{32}P]UTP in the presence of 10 μM of each rNTP (lane 1), except that the concentration of ATP, GTP and CTP were 200 μM for transcripts shown in lane 2, 3 and 4 respectively. Electrophoresis in a 8 % acrylamide/urea gel was carried out at 1000 volts for 3 h. An arrow indicates the position of about 95 nucleotides RNA synthesized from the Hae III fragment of *recA* gene of E. coli (Section 1). R1, R2 and R3 indicate the major products whose nucleotide sequence at their 5' ends were analyzed. XC indicates the position of xylene cyanol FF.

B. Autoradiography of partial digests of the transcripts. Transcripts labeled with [γ-^{32}P]ATP were electrophoresed in 8 % gel as shown in panel A. The regions corresponding to R1, R2 and R3 were cut out from the gel. RNAs were extracted and concentrated by ethanol precipitations respectively. RNases used for partial digestions were RNase T1, (G); RNase U2, (A); RNase A, (U,C); and RNase PhyI, (-C). The procedures of RNase digestions were carried out as described in Simoncsizt et al., 1977; and Itoh and Tomizawa, 1980. Gel electrophoresis was carried out in 25 % acrylamide/urea gel at 1000 volts for 10 h. The representative autoradiograph is shown (see the text). BPB and ATP indicate the positions of bromphenol blue and ATP respectively.

was observed in the 5' sequence of three major RNA products. The result in Fig. 16 B shows that the 5' sequence of the major products (Fig. 16 A) is the same and pppAUAACUGCACA... . The heterogeneity of lengths of the RNA products may be due to the premature termination near the end of the template DNA fragment.

5. Nucleotide sequence of the *lexA* gene

Using the DNA fragments obtained from the cleavage with variety of combination of restriction enzyme *Eco* RI, *Hinf* I, *Hinc* II, *Alu* I and *Taq* I, the entire nucleotide sequence of the *Eco* RI-*Alu* I-2 about 950 bp region was determined. To verify the determination of the sequence, *Hpa* II fragments were also sequenced, The whole sequence determined is shown in Fig. 17. The DNA sequence is numbered beginning at the predominant site of initiation of transcription *in vitro*. One possible frame in which a sequence of amino acids covered *Hpa* I-1 site and *Alu* I-1 site was described under the DNA sequence (Fig. 17). This *lexA* polypeptide comprised 202 amino acid residues and had a calculated molecular weight of 22,358 .

6. Construction of *lac*-*lexA* fusion gene

To increase the production of *lexA* protein, the regulatory region of *lexA* gene was removed and the *lexA* coding region was fused with the *lac* promoter bearing UV5 mutation. The procedure to construct the *lac*-*lexA* fusion gene is diagrammed in Fig. 18.

Eco RI-*Bam* HI DNA fragment of pHK10, which carries the *bla* gene at its *Bam* HI side, was inserted into a plasmid pGC3 (Fig.14) that carries *lexA* gene. DNA of the obtained plasmid pTH18 was linearized with *Eco* RI. Then about 90 bp of DNA from *Eco* RI site was resected using Exo III and S1 nuclease to remove the control region of the transcription of the *lexA*

```
         -50                                                   +1
GAATTCGATAAATCTCTGGTTTATTGTGCAGTTTATGGTTCCAAAATCGCCTTTTGCTGTATATACTCACAGCATAACTGTATATACACCCAGGGGGCGGAATGAAA
                                                                                                     fMetLys
      +50                                                   +100
GCGTTAACGGCCAGGCAACAAGAGGTGTTTGATCTCATCCGTGATCACATCAGCCAGACAGGTATGCCGCCGACGCGTGCGGAAATCGCGCAGCGTTTGGGGTTCCGT
AlaLeuThrAlaArgGlnGlnGluValPheAspLeuIleArgAspHisIleSerGlnThrGlyMetProProThrArgAlaGluIleAlaGlnArgLeuGlyPheArg
    +150                                                 +200                                              +250
TCCCCAAACGCGGCTGAAGAACATCTGAAGGCGCTGGCACGCAAAGGCGTTATTGAAATTGTTTCCGGCGCATCACGCGGGATTCGTCTGTTGCAGGAAGAGGAAGAA
SerProAsnAlaAlaGluGluHisLeuLysAlaLeuAlaArgLysGlyValIleGluIleValSerGlyAlaSerArgGlyIleArgLeuLeuGlnGluGluGlu
                                                  +300                                             +350
GGGTTGCCGCTGGTAGGTCGTGTGGCTGCCGGTGAACCACTTCTGGCGCAACAGCATATTGAAGGTCATTATCAGGTCGATCCTTCCTTATTCAAGCCGAATGCTGAT
GlyLeuProLeuValGlyArgValAlaAlaGlyGluProLeuLeuAlaGlnGlnHisIleGluGlyHisTyrGlnValAspProSerLeuPheLysProAsnAlaAsp
                                           +400                                             +450
TTCCTGCTGCGCGTCAGCGGGATGTCGATGAAAGATATCGGCATTATGGATGGTGACTTGCTGGCAGTGCATAAAACTCAGGATGTACGTAACGGTCAGGTCGTTGTC
PheLeuLeuArgValSerGlyMetSerMetLysAspIleGlyIleMetAspGlyAspLeuLeuAlaValHisLysThrGlnAspValArgAsnGlyGlnValValVal
                                    +500                                             +550
GCACGTATTGATGACGAAGTTACCGTTAAGCGCCTGAAAAAACAGGGCAATAAAGTCGAACTGTTGCCAGAAAATAGCGAGTTTAAACCAATTGTCGTTGACCTTCGT
AlaArgIleAspAspGluValThrValLysArgLeuLysLysGlnGlyAsnLysValGluLeuLeuProGluAsnSerGluPheLysProIleValValAspLeuArg
              +600                                                 +650
CAGCAGAGCTTCACCATTGAAGGGCTGGCGGTTGGGGTTATTCGCAACGGCGACTGGCTGTAACATATCTCTGAGACCGCGATGCCGCCTGGCGTCGCGGTTCGTTTT
GlnGlnSerPheThrIleGluGlyLeuAlaValGlyValIleArgAsnGlyAspTrpLeu
       +700                                                 +750
TCATCTCTCTTCATCAGGCTGTCTCGATGGCATTCCTCACTTCATCTGATAAAGCACTCTGGCATCTCGCCTTACCCATGATTTTCTCCAATATCACCGTTCCGTTGC

    +800                                             +850
TGGGACTGGTCGATACGGCGGTAATTGGTCATCTTGATAGCCGGGTTTATTTGGGCGGCGTGGCGGTTGGTGCAACCGGCGACCAGCT
```

Fig. 17. Nucleotide sequence of lexA gene and a possible amino acid sequence of lexA protein.

The base sequence of DNA fragment from the Eco RI cleavage site to AluI-2 site described in Fig. 14 is presented. Nucleotides are numbered from the site corresponding to 5' end of the in vitro transcript.

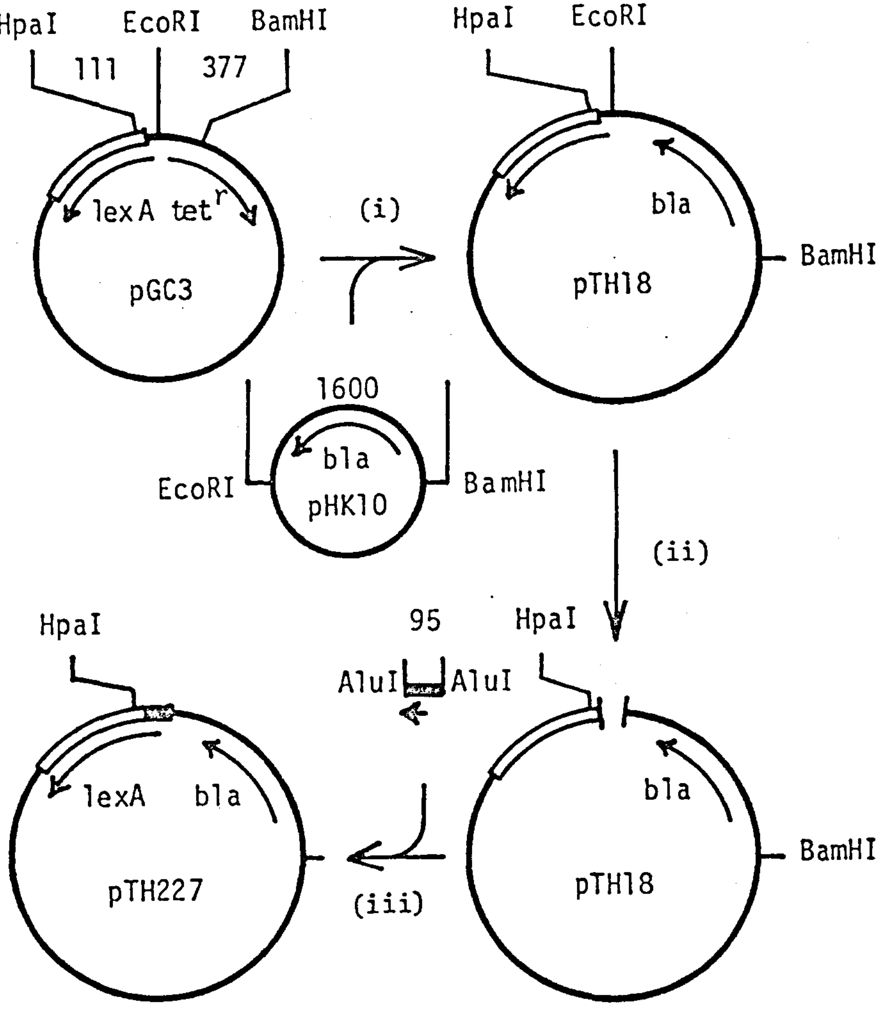


Figure 18. Procedure for construction of a plasmid carrying the lac-lexA fusion gene.

The approximate locations of several restriction endonuclease cleavage sites are shown for the plasmid pGC3, for the plasmid pHK10 which is used for a source of DNA fragment bearing the bla gene, and for a DNA fragment bearing the promoter of the lacZ gene which is isolated from the plasmid pKB252. The location and orientation of the tet^r, lexA, bla and lac-promoter are indicated by allows. Box represents the coding region of lexA gene. Distances are indicated in base pairs. i), The fragment bearing the bla gene was ligated into Eco RI and Bam HI sites of pGC3. ii), The plasmid pTH18 thus obtained was opened at the Eco RI site and resected about 90 bp. from both ends by Exo III and S1 nucleases. iii), DNA fragment bearing the lac promoter was inserted into the partially resected plasmid DNA.

gene. The transcriptional control region of the lexA gene is located in the region between Eco RI and Hpa I sites (111 bp) and the initiation codon, ATG, is located at 100 bp apart from Eco RI site (Fig. 17). After ligation of Alu I fragment bearing lac promoter to the resected end, the plasmid DNA was recircularized with T4 ligase. Transformation of DM511 tsl (temperature-sensitive mutant of lexA gene) cells was carried out with this plasmid DNA with selection of transformants for ampicilin resistance and viability at 42°C. About a quarter of the transformant colonies thus obtained appeared blue on an agar-plate containing 5-chloro-4-bromo-3 indolyl-β-D-galactoside, on which the cells carrying the lac UV5 promoter form blue colonies (Bachmann et al., 1976). Since the increased level of the production of lexA protein seems to make a cell sensitive to UV irradiation as described above, the colonies which were sensitive to UV irradiation were selected. Four strains among 76 showed higher sensitivity to UV irradiation than the strain carrying the original plasmid pTH18. Finally, the efficiency of lexA protein synthesis of each candidates was examined by maxi-cell system described by Sancar et al., (1979) (data not shown) and one strain carrying plasmid pTH227 was selected.

With SDS-polyacrylamide gel electophoresis of total cellular protein from a strain harboring a plasmid pTH227, the lexA gene product was found to be easily identifiable (Fig. 19 A). When the synthesis of the lexA protein was induced by lactose with this strain the amount of the lexA protein reached to about 1.5 % of total cellular protein (Fig. 19 A lane b). This indicates that the expression of the lexA gene on the plasmid pTH227 is under the regulation of lactose repressor. DNA sequence analysis of the joining region of the lexA gene with lac promoter of the plasmid, pTH227, showed that the lac promoter wasinserted just in front of the Shine-Dalgarno sequence of the lexA gene and the coding region of lexA gene was remained intact (Fig. 20 and 21).

A

lexA

a. b. c.

B

lexA

a. b. c. d.

Fig. 19

Fig. 19. SDS-polyacrylamide gel electrophoresis of the lexA protein.

A: Production of the lexA protein in the cells harboring plasmid pTH227.

Cells were cultured in 5 ml of L-broth as described below and collected by centrifugation at a concentration of 4 x 10^8 cells/ml. The cells were resuspended in 0.5 ml of 50 mM Tris-HCl (pH 8.0) containing 10 mM EDTA and added 0.1 ml of 5 mg/ml lysozyme. After 3 cycles of the freezing and thawing, 120 μl of lysing solution (300 mM Tris-HCl pH 6.8/ 10 % SDS/ 25 % 2-mercaptoethanol/ 25% glycerol/ 0.002 % bromophenol blue) was added and heated for 2 min in boiling water. From each samples, 30 μl was taken and electrophoresed on a 12.5 % SDS-polyacrylamide gel. Contents of the lexA protein were assessed by a densitometry of a gel. (a), W3623 (pTH227) cells grown in L-broth containing 1 % of glucose; (b), W3623 (pTH227) cells grown in L-broth and added lactose at 1 x 10^8 cells/ml and continued to grow at 4 x 10^8 cells/ml; (c), W3623 (without a plasmid) cells grown as b. ;

B. Purification of the lexA protein.

Aliquots of fraction I-IV (see Materials and Methods) were electrophoresed as described above. Proteins in each fractions was quantified by a method of Lowry et al., (1951). (a), Fraction IV, 6 ug (b), Fraction III, 7 ug (c) Fraction II, 14 ug (d) Fraction I, 25 ug.

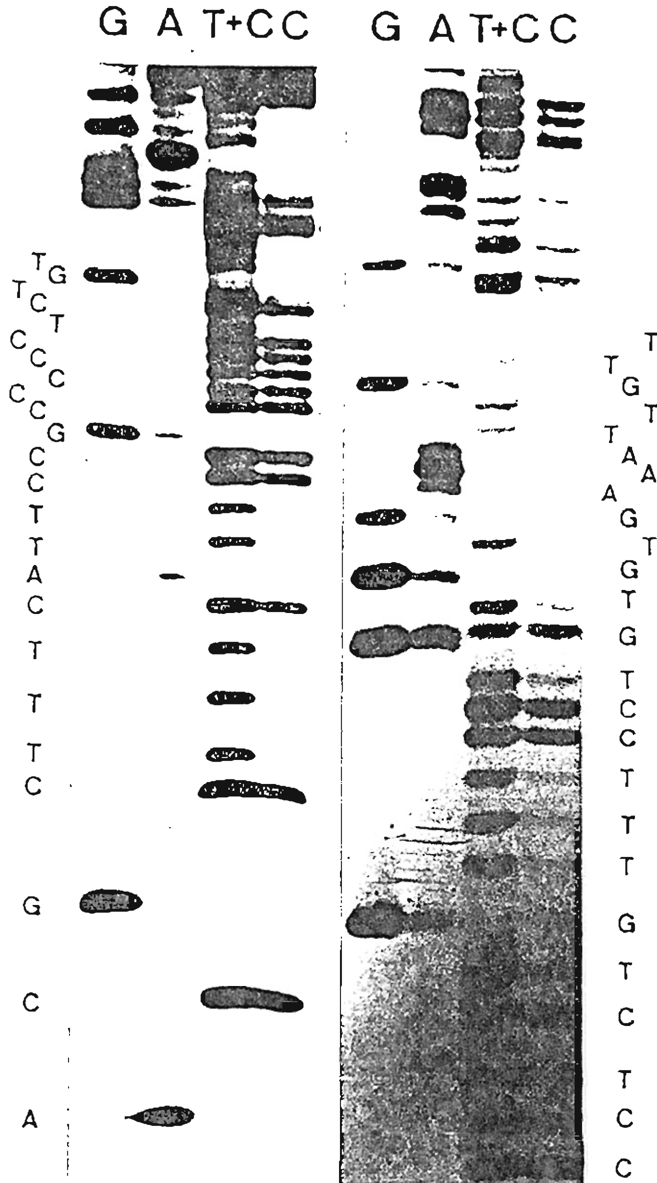


Fig. 20 Autoradiography for DNA sequence analysis of the fused region of lac promoter and the lexA gene.

DNA was sequenced from Hpa I site labeled with [γ-^{32}P] ATP and T4 kinase to the upstram of the lac-lexA fused gene by a method of Maxam and Gilbert (1977). Electrophoreses were carried out for 6 h. (left), and for 20 h. (right) at 1000 volts.

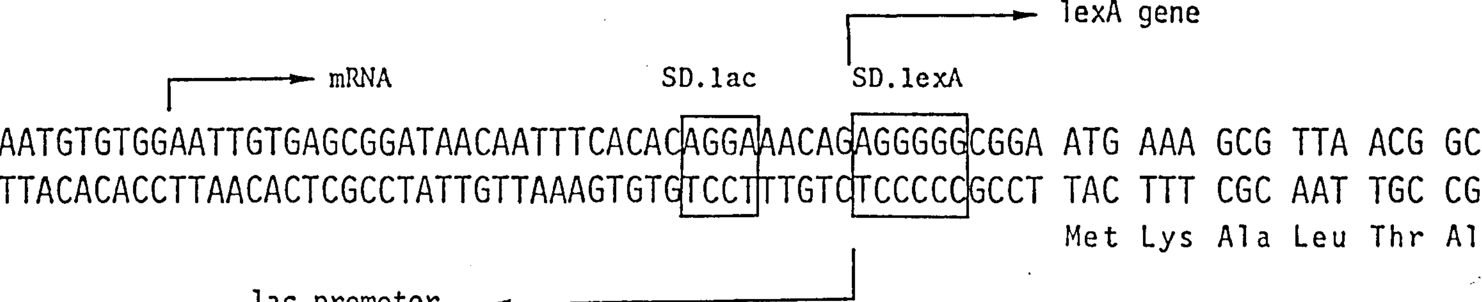


Fig. 21 Nucleotide sequence of the fused region of lac promoter and the lexA gene.

The initiation site of transcription of lac promoter and Shine-Dalgarno sequences are indicated and amino acid sequence of amino-terminus of the lexA protein is described under the DNA sequence. Original sequence of lac promoter bearing UV5 mutation was described in Reznikoff and Abelson (1978).

7. Purification, amino-terminal sequence and amino acid composition of the purified lexA protein

When the cells harboring the constructed plasmid pTH227 was cultured by adding lactose, the amount of the lexA protein in the cells increases as mentioned above and could be estimated by SDS-polyacrylamide gel electrophoresis even in a cell crude lysate. Therefore, using this strategy, the lexA protein was purified by examining its amount by SDS-polyacrylamide gel electrophoresis and obtained 1 mg sample from 20 g of the induced cells in more than 96 % purity (Fig. 19 B). To confirm that the purified protein is really the product of the lexA gene, the amino acid sequence of five residues from the amino-terminus for the purified protein (Fraction IV) was determined by the manual Edman degradation (Fig. 23 A). The obtained sequence was NH_2-Met-Lys-Ala-Leu-Thr-. This agreed with the sequence predicted from the nucleotide sequence obtained previously (Fig. 17) except that the formylmethionine is deformylated. The amino acid composition of the protein also agreed with the predicted values from the nucleotide sequence analysis (Table 6). Therefore, it was concluded that the purified protein was the product of the lexA gene itself.

8. Proteolytic cleavage of lexA protein with recA protein

The purified lexA protein was subjected to a proteolytic cleavage with the wild-type recA protein. The products of the reaction were analysed on a SDS-polyacrylamide gel electrophoresis. As shown in Fig.22. the lexA protein was cleaved into two peptides with the recA protein in the presence of ATP and single-strand DNA. The molecular weights of the cleaved products were estimated as 13,500 and 9,000 on a SDS-polyacrylamide gel electrophoresis described by Weber and Osborn (1969) (data not shown). Under the condition as in Fig. 22, 70 ng of lexA protein was cleaved with 1 µg of wild-type recA protein per min at 37°C. This rate continued for

Table 6

Amino acid composition of lexA protein

	Residues per molecule	
	Analytical value[a]	Predicted value[b]
Lys	10.5	11
His	5.0	5
Arg	15.6	15
Asn	17.8[c]	6
Asp		12
Thr	6.0	6
Ser	9.2	9
Gln	29.1[c]	13
Glu		16
Pro	9.7	9
Gly	17.5	17
Ala	15.8	16
Cys	N.D.[d]	0
Val	19.5	19
Met	3.9	5
Ile	11.9	13
Leu	21.3	22
Tyr	1.1	1
Phe	5.3	6
Trp	N.D.[d]	1
Total residues		202

a. Acid hydrolysis was performed for 24 and 72 hr. The values of threonine and serine were obtained by extrapolation to zero time of hydrolysis. Values of valine and isoleucine were of 72 hr-hydrolysate.

b. Predicted value was from the DNA sequence analysis reported previously (Horii, Ogawa and Ogawa, 1981).

c. Sum of acid and amide forms.

d. Not determined.

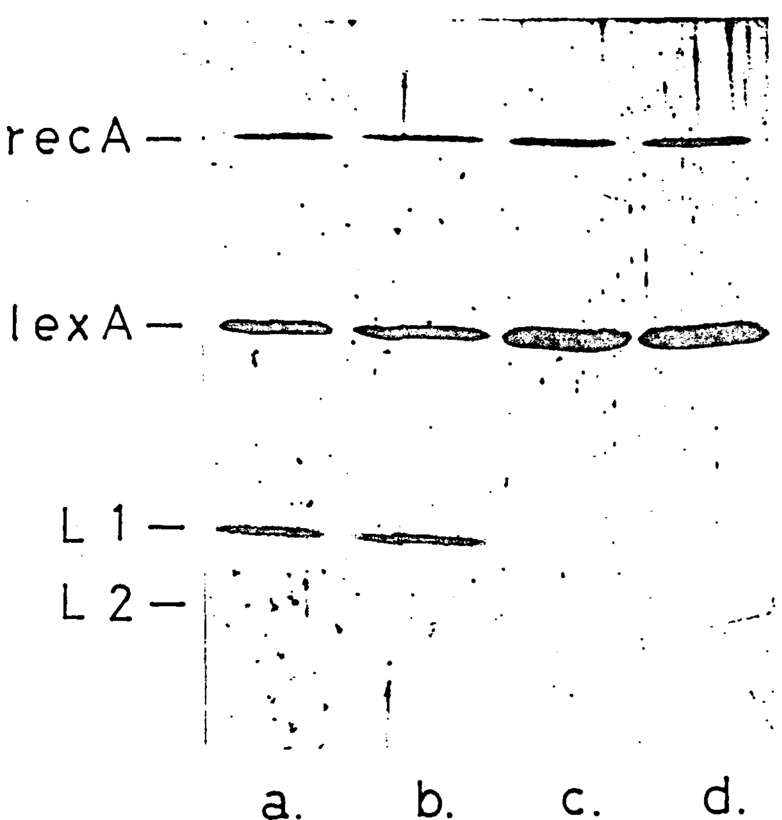


Figure 22. Proteolytic cleavage of the purified lexA protein.

Complete reaction mixtures of 100 μl contained 25 mM Tris-HCl pH 7.4/ 10 mM $MgCl_2$/ 50 mM NaCl/ 1 mM dithiothreitol/ 1 mM ATP/ 0.05 μg heat denatured calf thymus DNA/ 1 μg purified recA protein/ 10 μg purified lexA protein (Fraction IV). After incubation for 60 min at 37°C, reaction mixtures were added 20 μl of lysing solution (see legend to figure 2), heated for 2 min in boiling water and analysed on SDS-polyacrylamide 10-25% gradient gel. a, complete; b, 1 mM +ATP-γ-S instead of ATP; c, -ATP; d, -calf thymus DNA;

more than 180 min.

To determine the cleavage site in tne lexA protein, amino-terminal sequence of the cleaved products were analyzed by Edman degradation. The lexA protein, 10 nmoles was cleaved with recA protein, 1 nmoles in the presence of both ATP-γ-S and heat denatured calf thymus DNA for 2 h at 37°C. The reaction mixture was dialyzed against 5 % of formic acid, lyophilized and subjected to 5 steps of Edman degradation. The amino acids were released in pairs by each step of the degradation as follows: Met-Gly, Lys-Glu, Ala-Pro, Leu-Leu and Thr-Leu in sequence. Since the native lexA protein have a sequence of NH_2-Met-Lys-Ala-Leu-Thr- at amino-terminus (Fig. 23 A), another sequence was deduced as Gly-Glu-Pro-Leu-Leu- (Fig. 23 B). This sequence was found at the region from the residue Gly^{85} to Leu^{89} for once in the lexA protein. This indicates that the cleavage was occured between the residues of Ala^{84} and Gly^{85} For the determination of carboxyl terminal sequence of the cleaved products, carboxypeptidase A was used. The cleaved or native lexA protein, 50 µg, was incubated with 10 µg of carboxypeptidase A in a 40 µl reaction mixture containing 100 mM Tris-HCl (pH 8.0) for 1 h at 37°C. In this reaction, two alanines and one valine were released, those were corresponding with the residues Ala^{84} , Ala^{83} and Val^{82} ., whereas the expected amino acids,leucine, tryptophan and aspartic acid, for the carboxyl-terminus of native lexA protein, were not released from the the cleaved products nor the native lexA protein. This is probably due to such conformational characterictics of the carboxyl-terminus region of lexA protein that it prevents the attack of peptidase. These facts strongly suggest that the recA protein cleaved the lexA protein at a single site and released two peptides composed of 84 and 118 amino acid residues. The molecular weights calculated for the peptides of 84 and 118 amino acids from nucleotide sequence were 9,201

and 13,175 respectively, and these were in accord well with the estimated values for the cleaved products as described above.

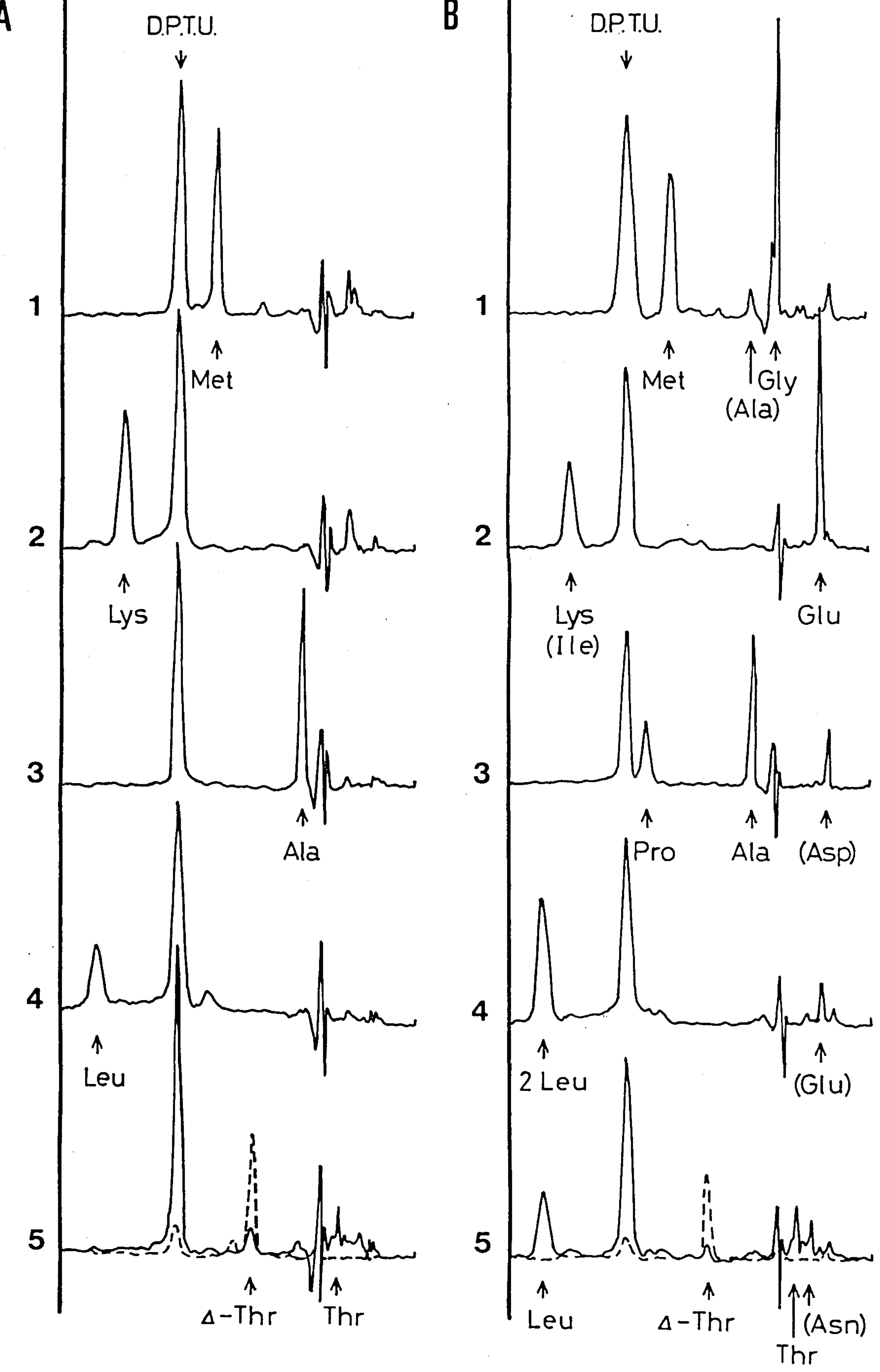


Fig. 23

Fig. 23. High performance liquid chromatography for determination of the amino terminal sequences of native (A) and cleaved (B) lexA protein.

The lexA protein (Fraction IV), 10 n moles, was cleaved with the recA protein, 1 n moles, in a 1.5 ml of reaction mixture containing 25 mM Tris-HCl pH7.4/ 10 mM $MgCl_2$/ 50 mM NaCl/ 1 mM dithiothreitol/ 1 mM ATP-γ-S/ 1 μg heat denatured calf thymus DNA. After incubation for 2 h. at 37°C, reaction mixture was dialyzed against 5 % (vol/vol) of formic acid in membrane tubing of molecular weight cut off 3,500 (Spectrapor), lyophilized and subjected to 5 steps of Edman degradation. Phenylthiohidantoin (PTH) derivatives released by each step of Edman degradations were determined by high performance liquid chromatography. PTH derivatives were detected by the absorbance 269 nm (solid line). Specifically, dehydro (Δ)-PTH Thr was detected by the absorbance at 320 nm (broken line). D.P.T.U. means diphenylthiourea, which is a in the reaction of Edman degradation. The identification procedure used for amino acids is not applicable to the following pairs of amino acids both can not be distinguished; Lys/Ile, Val/Met, Asp/Cmc. When needed, identification was corroborated by the thin layer chromatography. Amino acids denoted in parentheses in panel B were corresponding to the amino terminal sequence of recA protein, NH_2-Ala-Ile-Asp-Glu-Asn-, which was described in Fig. 9.

Considerations

Phenotypes conferred by the sub-cloning plasmids

The recombinant plasmids which contain lexA gene confer the ability to suppress the killing caused by tif-1 and tsl mutations. These plasmids also render the wild-type cell substantially sensitive to UV irradiation and STS-cell susceptible to the turbid plaque formation of λ phage. These effects caused by the presence of plasmid containing lexA gene are in agreement with the results reported previously (Brent and Ptashne, 1980). The plasmid pMN206 which confers the ability to complement the tsl mutation could not render the cell sensitive to UV irradiation nor STS-cell to make the turbid plaque with λ phage infection as in the case of plasmids containing the wild-type lexA gene. Since pMN206 deletes the coding region for 17 amino acids at carboxyl-terminus of lexA product, it is suggested that the product originated from pMN206 may result in the incomplete repression of the host recA gene. Plasmid pMN400 which deletes the promoter and the coding region for 5 amino acids at the amino-terminus of lexA gene and the product directed by the deleted gene was not identified, while pMN400 confers the ability to complement the tsl mutation. One possible explanation is that the region coding for lexA product is transcribed by the promoter of the β-lactamase gene of pBR322 and is expressed at a low level, although it is not ruled out the possibility that pMN400 confers tsl^+ by marker rescue.

Structual informations of lexA gene

(1). Initiation of transcription.

Teh well defined characteristics common to the promoter regions of E. coli are the "Pribnow box" (Pribnow, 1975) and the "-35 region" (Gilbert, 1976), which are located about 10 and 35 bp upstream from the RNA start site, respectively. In the case of the lexA gene, there is the sequence

TATACTC at the -10 region which matches the first four and the important sixth of "Pribnow box" (Fig. 24). The "-35 region" of this gene contains a common sequence, and is similar to that region of *lacI* gene of *E. coli* (Caros, 1978).

(2) Initiation of transcription

About 10 nucleotides upstream from the proposed translation initiation codon, there is a probable ribosome binding sequence AGGGGG (Shine and Dalgarno, 1975). In other frames starting from the dispersed AUG or GUG triads on the sequence, more than a hundred amino acids could not be arranged in a string in both directions. This strongly suggests that the coding region described above is that of the *lexA* gene. Initiated at 29 nucleotides downstream from a site from which transcription can initiate *in vitro*, a protein of 202 amino acids can be encoded in this region.

(3) Characteristics in the promoter region

As described in Section 1, the promoter region of the *recA* gene has the symmetrical structure. In the promoter region of the *lexA* gene, the structure homologous to the symmetrical region of the *recA* gene was found (boxed in Fig. 25). Amoung 22 nucleotides of the boxed regions, 14 nucleotides are identical. Those are -T-CTGTAT---C--ACAG-AT. Another characteristics common to the both promoter regions are the homologous sequence of the 9 nucleotides, ATAC-GTAT, which has a small dyad symmetry in it. Especially in the region from the position +4 to the position +36 of *recA* gene, 5 sets of the 9 nucleotides structures appear (Fig. 25)

Regulation of expression of the colicin E2 gene

The plasmid pGA2 seems not to confer the ability to repress the *recA* gene (Table 4), on the other hand suppresses production of colicin E2 in contrast to the plasmids pGA1 and pGA3. Also, pMN206, whose product complement the *tsl* mutation but seems not to repress the *recA* gene completely, confers

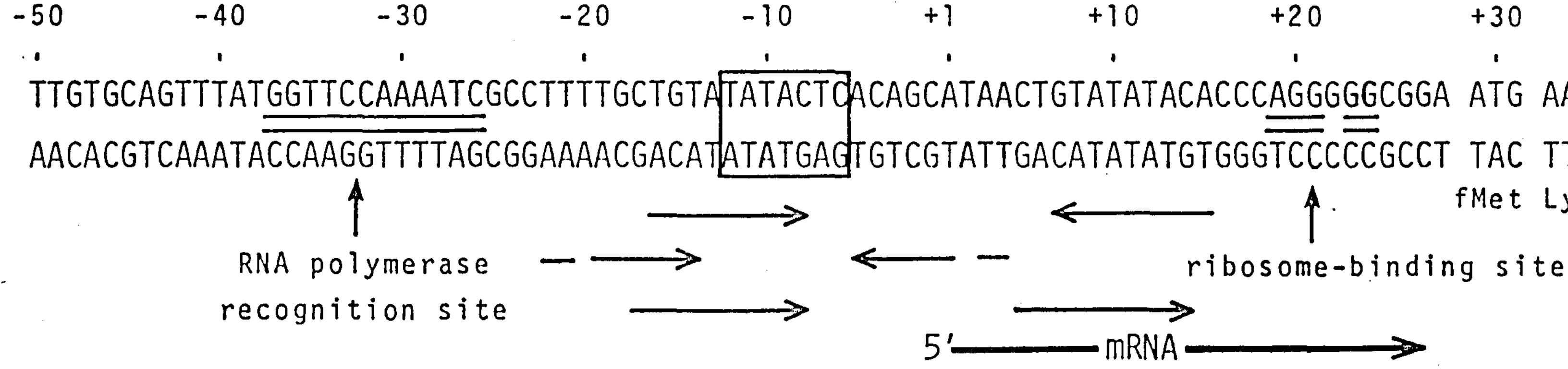


Fig. 24. Region of transcription-initiation of lexA gene.

The "Pribnow box" is boxed and the presumed RNA polymerase recognition site and ribosome-binding site are underlined. The regions of dyad symmetries and direct repeat are indicated by arrows.

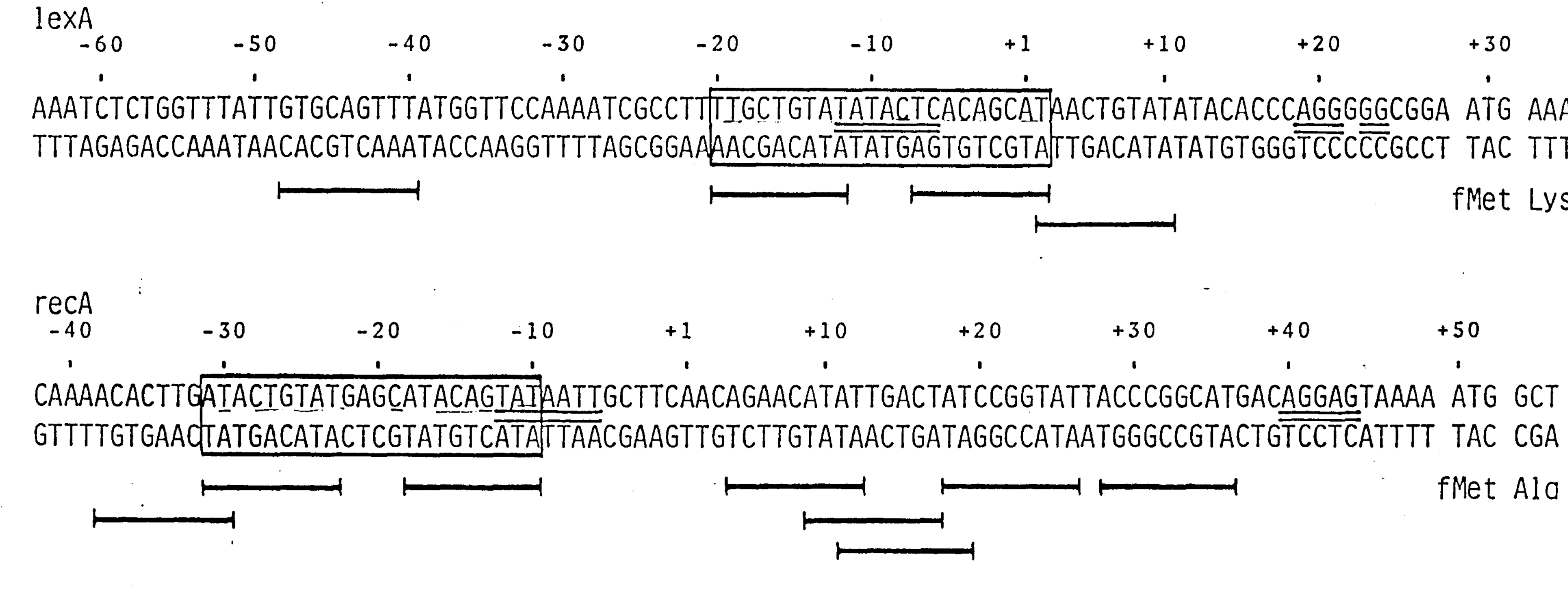


Fig. 25. Promoter regions of the lexA and recA. The homologous sequences comprised of 22 base pairs are boxed (see the text). The 9 nucleotides sequences which have more than 5 identical nucleotides amoung 8 nucleotides of the consensus sequence ATAC-GTAT are denoted by the bold lines. The RNA polymerase and ribosome-binding sites are underlined.

the ability to suppress the colicin E2 production. These indicate that the regions responsible for tsl complementation and suppression of colicin E2 production are overlapping. Therefore, it is likely that the lexA protein is involved in the regulation of the production of colicin E2 protein.

Construction of lac-lexA fusion gene

Amoung the constructed plasmids in which the lexA coding region was fused to the promoter region of lacZ gene by in vitro gene manipulation, one, pTH227, which rendered the host bacteria extremely sensitive to UV irradiation, was selected. This plasmid seems to express the function of the lexA gene efficiently. DNA sequence analysis of the fused region of pTH227 showed that the 95 bp DNA fragment containing a lac promoter was inserted in front of the ribosome-binding site (Shine-Dalgarno sequence) of the lexA gene. In consideration of an example of the gene expression from lac-cro fusion gene (Roberts et al., 1979), this joining is confirmed to be a fairly efficient one for the expression of lexA gene. As expected, the cells carrying a plasmid pTH227 produced the lexA protein about 1.5 % of the total cellular proteins when cells were cultured in a broth containing lactose.

Purification of the lexA protein

Using bacterial strain harboring this plasmid, the lexA protein was induced and purified more than 96 %. The amino acid composition and the amino-terminal sequence of the purified protein agreed with the prediction from nucleotide sequence analysis of lexA gene. Therefore, it was concluded that the product from the lac-lexA fusion gene is that of lexA gene.

The genetical and biochemical analyses suggested that the lexA product is a repressor for recA gene (Mount, 1977), lexA gene itself (Little and Mount,

1979; Brent and Ptashne, 1980) and colicin E2 gene (this work). Whereas, the specific binding activity of purified lexA protein to the DNA fragment bearing the entire recA gene could not be detected by a membrane binding method described by Bourgeois (1971). One of the probable cause is that the interaction between lexA protein and operator sequence is not so firm to be detectable on a membrane binding method. Another possibility is that the unknown component(s) is required for the tight binding to the operator in conjunction with the lexA protein. It cannot, however, be excluded the possibility that the binding activity of the lexA protein is inactivated during purification procedures. The interaction of lexA protein with the specific operator DNA is now under investivation.

Proteolytic cleavage of lexA protein with recA protein

Using tif-mutant form of recA protein, Craig and Roberts showed that the purified recA protein promotes the specific cleavage of λ repressor (1980). Also, it was reported that the lexA protein was cleaved with recA protein (Little et al., 1980). Using purified lexA protein and wild-type recA protein, it was showed that the lexA protein was cleaved with recA protein specifically into two peptides with molecular weights of about 13,500 and 9,000 in the presence of ATP and single strand DNA as described for λ repressor cleavage. It is very likely that the specific cleavage events of phage repressor and lexA protein with recA protein are the triggering events for induction of some "SOS-functions".

Site of specific cleavage

The analysis of amino acid sequence of the cleaved products showed that the cleavage was occurred between Ala^{84} and Gly^{85} residues. The comparison of amino acid sequence around the cleavage site of lexA protein with that of the λ repressor, in which it was suggested to be cleaved

at the junction of amino acid residues Ala^{111} and Gly^{112} (Pabo et al., 1979), showed that the region Gly^{75} to Gly^{85} of the lexA protein is similar to the region from Glu^{102} to Gly^{112} residues of λ repressor (Fig. 26): four amino acids are identical in the sequence of 11 amino acids and the other amino acids mimic each other in their hydrophobicities (Tanford, 1980). This suggests that the recA protein recognizes some amino acid residues and/or the some secondary structure of amino acid sequence around the cleavage site and promotes the specific cleavage.

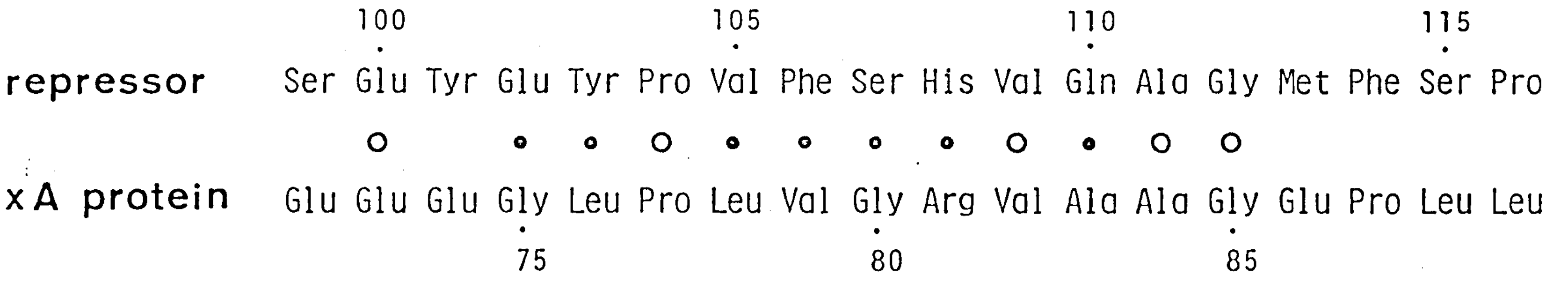


Fig. 26. Amino acid sequence of the homologous region of lexA protein and λ repressor.

Amino acid sequence of lexA product is numbered from Met residue (Fig. 17.) as no. 1 position. The smino acid sequence of λ repressor was determined by Sauer and Anderegg (1978). Circle represents amino acid identical in both proteins and dot shows amino acids which have similar hydrophobicity.

DISCUSSION

In consideration of new structural informations obtained here, the regulatory mechanism for the expression of "SOS-functions" is presented in Fig.27.

The comparison of the regulatory regions of the transcription for recA and lexA genes revealed the presence of the common sequence in both regions (Fig. 25). Recently, this common sequence was found also in the regulatory region of colicin E1 gene (A. Nakazawa, personal communication). This common sequence is probably a binding site for a single repressor. Genetical studies suggested that the product of lexA gene represses the recA gene (Mount, 1977; Gudas and Mount, 1977; McEntee, 1977), lexA gene itself (Little and Harper, 1979; Brent and Ptashne, 1980), colicin E1 gene (Tessman et al., 1978) and colicin E2 gene (this work). Therefore the lexA protein is a strong candidate for this hypothetical common repressor. The experiments using Mu(lac::Tn3) phage showed that the uvrA gene (Kenyon and Walker, 1980) and uvrB gene (Fogliano and Schendel, 1981) are transcriptionally controled by the recA-lexA regulatory system. However, it is unclear that the lexA protein represses directly these genes. It is possible that the unknown repressor which is also sensitive to the cleavage with recA protein is repressing these genes. Because in this case, these genes would appear to be under the control of lexA gene.

Two alternative models may be considered for the regulation of the coordinate expression of SOS-functions: One is that a single repressor (lexA protein) controls the various genes involved in the SOS-functions, another is that each gene concernes SOS-functions is controlled by own repressor which is inactivated by a common protease (recA protein).

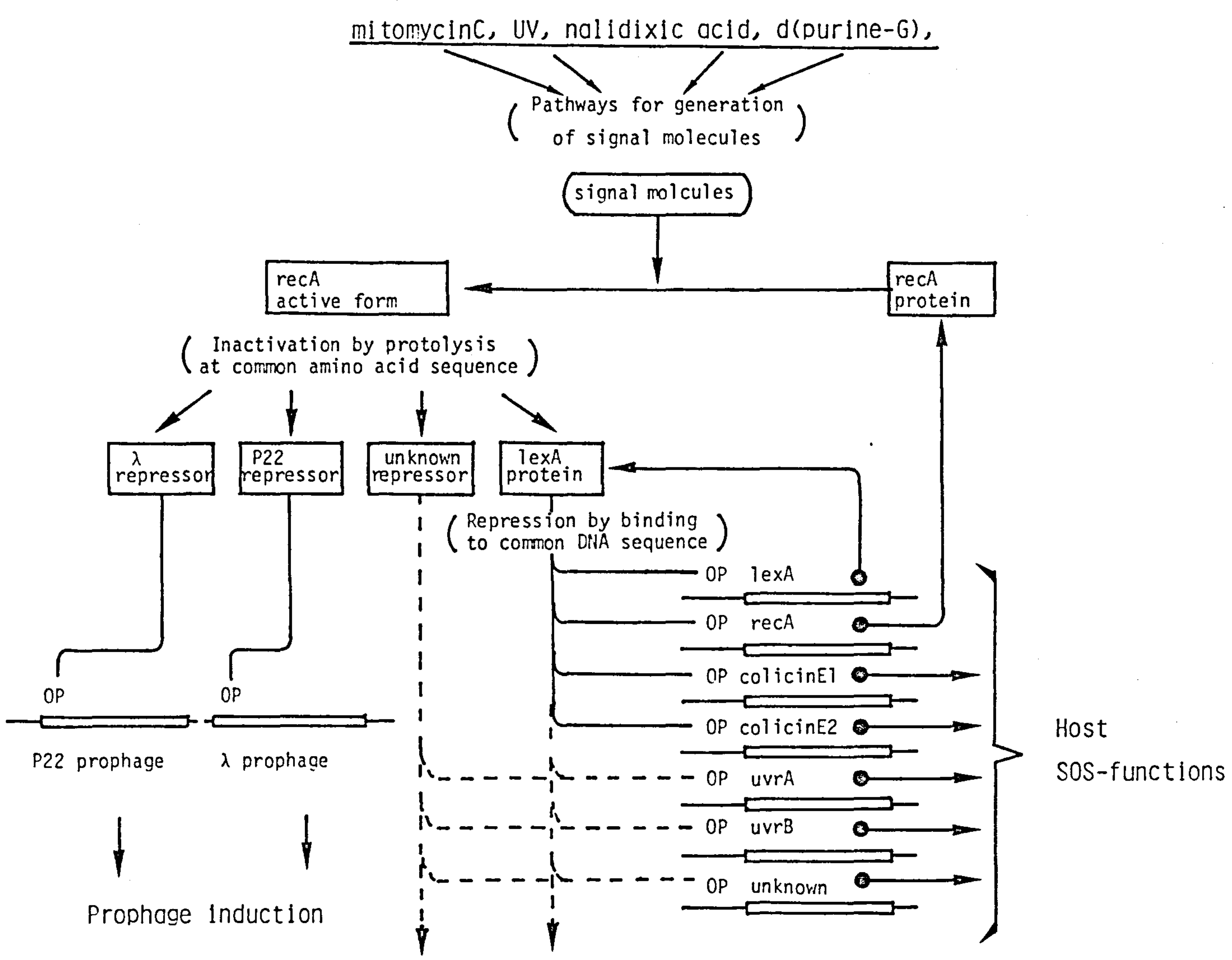


Fig. 27. Schematic representation of the regulatory mechanis for the expression of "SOS-functions".

As presented in Fig. 26, lexA protein and λ repressor contain a homologous amino acid sequence, in which a specific cleavage by recA protein is occured. It is plausible that this homologous amino acid sequence is the common sequence for the repressors of the genes involved in SOS-functions.

One of the remaining problems to be clarified in the SOS regulation is about the generation of signal molecule(s) which activates recA protein to the protease. This step may involve a variety of cell nucleases and may occur by different pathways depending on the nature of the initial damage. The recB mutation which is missing exonuclease V, is deficient in induction of recA protein after cells are treated with nalidixic acid, but proficient after cells are treated with bleomycin,which causes non-enzymatic degradation of DNA (Gudas and Pardee, 1975). The mutation in recF gene, whose function is not specified yet, decreases the inductions of λ prophage and recA protein by ultraviolet irradiation (Horii and Clark, 1973; Clark, 1973). These facts suggest that at least some divergence exists in the pathways for creation of signal molecule(s). Oishi and his collaborators have investigated signal molecules and found small proteins (nucleases) and oligonucleotides, in particular, dinucleotide d(purine-G) induce prophage ϕ80 in permiabilized cells (Irbe et al., 1981). In the presence of such dinucleotides recB function is no longer necessary. Whereas, such small oligonucleotides are not competent for the stimulation of the in vitro proteolysis by recA protein (unpublished result). The recA protein requires oligonucleotides more than hexamer in length for proteolytic activity. Thus a gap lying between the signal, d(purine-G) and the activation of recA protein must be filled up.

Another obscure point in the SOS regulation is the conceptual active form of recA protein. At a normal growth state, recA protein is

produced at a low level in bacterial cells. However, when cellular DNA is damaged, the syntheses of recA protein and proteins involved in SOS-functions start at a high rate. Before this induced synthesis, the recA protein must acquire the protease activity to inactivate the repressor of own gene. The recA protein cleaves λ repressor (Craig and Roberts, 1980), P22 repressor (Phizicky and Roberts, 1980) and lexA protein (Little et al., 1980; this work) into two peptides in the presence of ATP and single strand DNA. Whereas, it is not known whether the active protease form of recA protein is a complex bound with single strand DNA and/or nucleotide (ATP or ADP), or recA protein converts to an active structure catalytically with the aid of single strand DNA and ATP. Craig and Roberts (1980) presumed that binding of recA protein to polynucleotide causes a conformational change with recA protein, and the recA protein changed in form by binding to polynucleotide is active to cleave repressors. Because the efficient cleavage of repressor occurs whenever a constant ratio of recA protein to single strand DNA is provided for various concentrations of recA protein. Even if, recA protein – single strand DNA complex is the active form, in this case single strand DNA should be the signal molecule, the molecular nature of single strand DNA is still obscure. That is, which single strand DNA is a major activator for the recA protein, released DNA fragments from the chromosome by degradation, or the single stranded chromosomal DNA processed by nucleases? In latter case, proteolytic activity of recA protein will localize on the chromosomal DNA.

REFERENCES

Adhya, S. and Gottesman, M. (1978). Ann. Rev. Biochem. 47, 967-996.

Ambler, R.P. (1972). In Method in Enzymology, 25, eds. Hirs, C.H.W. and Timasheff, S.N. (New York: Academic Press), pp.262-272.

Anderson, E.H. (1946). Proc. Nat. Acad. Sci., U.S.A. 32, 120-128.

Appleyard, R.K. (1954). Genetics, 39, 429-439.

Backman, K., Ptashne, M. and Gilbert, W. (1976). Proc. Nat. Acad. Sci. U.S.A. 73, 4174-4178.

Bourgeois, S. (1971). In Methods in Enzymology, 21, eds. Grossman, L. and Moldave, K. (New York: Academic Press), pp.491-500.

Brooks, K and Clark, A.J. (1967). J. Viol. 1, 283-293.

Burgess, R.R., Travers, A.A., Dunn, J.J. and Bautz, E.K.F. (1969). Nature (London), 221, 43-46.

Calos, M.P. (1978). Nature (London), 274, 762-765.

Castellazzi, M., George, J. and Buttin, G. (1972). Mol. Gen. Genet. 119, 139-152.

Chothia, C. (1974). Nature (London), 248, 338-339.

Clark, A.J. (1973). Ann. Rev. Genet. 7, 67-86.

Clark, A.J. and Marguries, A.D. (1965). Proc. Nat. Acad. Sci. U.S.A. 53, 451-459.

Clark, A.J., Chamberlin, M., Boyce, R.P. and Howard-Flanders, P. (1966). J. Mol. Biol. 19, 442-454.

Clewell, D.B. and Helinski, D.R. (1969). Proc. Nat. Acad. Sci. U.S.A. 62, 1159-1166.

Clewell, D.B. and Helinski, D.R. (1970). Biochemistry, 9, 4428-4440.

Cowlishaw, J. and Ginoza, W. (1970). Virology, 41, 244-255.

Dugaiczyk, A., Boyer, H.W. and Goodman, H.M. (1975). J. Mol. Biol. 96 171-184.

Emmerson, P.T. and West, S.C. (1977). Mol. Gen. Genet. 155, 77-85.

Fogliano, M. and Schendel, P.F. (1981). Nature (London), 289, 196-198.

Gilbert, W. (1976). In RNA Polymerase, eds. Losik, R. and Chamberlin, M. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp.193-205.

Gudas, L.J. (1976). J. Mol. Biol. 104, 567-587.

Gudas, L.J. and Pardee, A.B. (1975). Proc. Nat. Acad. Sci. U.S.A. 72 2330-2334.

Gudas, L.J. and Pardee, A.B. (1976). J. Mol. Biol. 101, 459-477.

Gudas, L.J. and Mount, D.W. (1977). Proc. Nat. Acad. Sci. U.S.A. 74 5280-5284.

Hase, T., Wakabayashi, S., Matsubara, H., Kerscher, L., Oesterhelt, D., Rao, K.K. and Hall, D.O. (1978). J. Biochem. 83, 1657-1670.

Hertman, I. and Luria, S.E. (1967). J. Mol. Biol. 23, 117-133.

Horii, Z.I., and Clark, A.J. (1973). J. Mol. Biol. 80, 327-344.

Horii, T., Ogawa, T. and Ogawa, H. (1980). Proc. Nat. Acad. Sci. U.S.A. 77, 313-317.

Howard-Flanders, P. and Boyce, R.P. (1966). Radiat. Res. Suppl. 6, 156-184.

Itoh, T. and Tomizawa, J. (1980). Proc. Nat. Acad. Sci. U.S.A. 77, 2450-2454.

Irbe, R.M., Morin, L.M.E. and Oishi, M. (1981). Proc. Nat. Acad. Sci. U.S.A. 78, 138-142.

Jacob, F. and Monod, J. (1961). J. Mol. Biol. 3, 318-356.

Jenkins, S.T. and Bennett, P.M. (1976). J. Bacteriol. 125, 1214-1216.

Kenyon, C.J. and Walker, G.C. (1980). Proc. Nat. Acad. Sci. U.S.A. 77, 2819-2823.

Kessler, S.W. (1975). J. Immunol. 115, 1617-1624.

Kirby, E.P., Jacob, F. and Goldthwait, D.A. (1967). Proc. Nat. Acad. Sci. U.S.A. 58, 1903-1910.

Korn, D. and Weissbach, A. (1962). Biochim. Biophys. Acta, 61, 775-790.

Laemmli, U.K. (1970). Nature (London), 227, 680-685.

Lederberg, E.M. (1960). Symposium Soc. Gen. Microbiol. 10, 115-131.

Little, J.W. and Kleid, D.G. (1977). J. Biol. Chem. 252, 6251-6252.

Little, J.W. and Harper, J.E. (1979). Proc. Nat. Acad. Sci. U.S.A. 76, 6174-6151.

Little, J.W., Edmiston, S.H.M., Pacelli, L.Z. and Mount, D.W. (1980). Proc. Natl. Acad. Sci. U.S.A. 77, 3225-3229.

Lowry, O.H., Rosenbrough, N.J., Farr, A.L., and Randall, R.J. (1951). J. Biol. Chem. 193, 265-275.

Mandel, M. and Higa, A. (1970). J. Mol. Biol. 53, 159-162.

Maxam, A.M. and Gilbert, W. (1977). Proc. Nat. Acad. Sci. U.S.A. 74, 560-564.

McEntee, K. (1977). Proc. Nat. Acad. Sci. U.S.A. 74, 5275-5279.

McEntee, K., Weinstock, G.M. and Lehman, I.R. (1979). Proc. Nat. Acad. Sci. U.S.A. 76, 2615-2619.

McPartland, A., Green L. and Echols, H. (1980). Cell, 20, 731-737.

Melechen, N.E., and Skaar, P.D. (1962). Virology, 16, 21-29.

Miki, T., Ebina, Y., Kishi, F. and Nakazawa, A. (1981). Nucleic Acids Res. 9, 529-545.

Monk, M., and Gross, J. (1971). Mol. Gen. Genet. 110, 299-306.

Morita, M. and Oka, A. (1979). Eur. J. Biochem. 97, 435-443.

Mount, D.W. (1971). J. Bacteriol. 107, 388-389.

Mount, D.W., Walker, A.C. and Kosel, C. (1973). J. Bacteriol. 116, 950-956.

Mount, D.W. (1977). Proc. Nat. Acad. Sci. U.S.A. 74, 300-304.

Miura, A. and Tomizawa, J. (1968). Mol. Gen. Genet. 103, 1-10.

Nirenberg, M.W. (1963). In Method in Enzymology, 6, eds. Colowick, S.D. and Kaplan, N.O. (New York: Academic Press), pp.17-23.

Noack, D., and Klaus, S. (1972). Mol. Gen. Genet. 115, 216-224.

Oka, A. (1978). J. Bacteriol. 133, 916-924.

Ogawa, H., Shimada, K and Tomizawa, J. (1968). Mol. Gen. Genet. 101, 227-244.

Ogawa, T., Wabiko, H., Tsurimoto, T., Horii, T., Masukata, H. and Ogawa, H. (1978). Cold Spring Harbor Symp. Quant. Biol. 43, 909-915.

Otsuji, N., Sekiguchi, M. Iijima, T. and Takagi, Y. (1959). Nature (London), 184, 1079-1080.

Ozeki, H., Stocker, B.A.D. and Smith, S.M. (1962). J. Gen. Microbiol. 28, 671-687.

Pabo, C.O., Sauer, R.T., Proc. Nat. Acad. Sci. U.S.A. 76, 1608-1612.

Pacelli, L.Z., Edmisten, S.H. and Mount, D.W. (1979). J. Bacteriol. 137, 568-573.

Phizicky, E.M. and Roberts, J.W. (1980). J. Mol. Biol. 139, 319-328.

Pribnow, D. (1975). Proc. Nat. Acad. Sci. U.S.A. 72, 784-788.

Reznikoff, W.S. and Abelson, J.N. (1978). In the Operon, eds. Miller, J.H. and Reznikoff, W.S. (New York: Cold Spring Harbor Laboratory) pp.221-243.

Roberts, J.W., Roberts, C.W. and Craig, N.L. (1978). Proc. Nat. Acad. Sci. U.S.A. 75, 4714-4718.

Roberts, R.J., Breitmeyer, J.B., Takachinik, N.F. and Meyers, P.A. (1975). J. Mol. Biol. 91, 121-123.

Roberts, T.M., Kacich, R. and Ptashne, M. (1979). Proc. Nat. Acad. Sci. U.S.A. 76, 760-764.

Rosenberg, M., DeCrombrugghe, B. and Musso, R. (1976). Proc. Nat. Acad. Sci. 76, 717-721.

Rosenberg, M., Court, D., Shimatake, H., Brady, C. and Wulff, D.L. (1978). Nature, 414-422.

Rosenberg, M. and Court, D. (1980). Annu. Rev. Genet. 14, in press.

Sancar, A., Hack, A.M. and Rupp, W.D.(1979). J. Bacteriol. 137, 692-693.

Sauer, R.T. and Anderegg, R. (1978). Biochemistry, 17, 1092-1100.

Schaller, H., Gray, C. and Herrmann, K. (1974). Proc. Nat. Acad. Sci. U.S.A. 72, 737-741.

Schuster, H., Beyersmann, D., Mikolajczyk, M. and Schlicht, M. (1973). J. Virol. 11, 879-885.

Shibata, T., DasGupta, C., Cunningham, R.P. and Radding, C.M. (1979). Proc. Nat. Acad. Sci. U.S.A. 76, 1638-1642.

Shimada, K., Weisberg, R.A. and Gottesman, M.E. (1973). J. Mol. Biol. 80, 297-314.

Shine, S. and Dalgarno, L. (1975). Nature (London), 254, 34-38.

Simoncsitz, A., Brownlee, G.G., Brown, R.S., Rubin, J.R. and Guilley, H. (1977). Nature, 269, 833-836.

Slater, D.W. and Spiegelman, S. (1966). Proc. Nat. Acad. Sci. U.S.A. 56, 164-170.

Smith, S.M., Ozeki, H. and Stocker, B.A.D. (1963). J. Gen. Microbiol. 33, 231-242.

Sutcliffe, J.G. (1978). Cold Spring Harbor Symp. Quant. Biol. 43, 77-90.

Tanford, C. (1980). Proteins, (New York: John Wiley & sons Inc.), pp. 139-145.

Tessman, E., Gritzmacher, C. and Peterson, P. (1978). J. Bacteriol. 135, 29-38.

Thomas, M., White, R.L. and Davis, R.W. (1976). Proc. Nat. Acad. Sci. U.S.A. 73, 2294-2298.

Tomizawa, J. and Ogawa, H. (1968). Cold Spring Harbor Symp. Quant. Biol. 33, 243-251.

Weber, K., and Osborn, M. (1969). J. Biol. Chem. 244, 4406-4412.

Weinstock, G.M., McEntee, K. and Lehman, I.R. (1979). Proc. Nat. Acad. Sci. 76, 126-130.

Weiss, B. (1971). In Methods in Enzymology, eds. Grossman, L. and Moldave, K. (New York: Academic Press) 21, pp.319-326.

Wilson, G.A. and Young, F.E. (1975). J. Mol. Biol. 97, 123-125.

Witkin, E.M. (1976). Bacteriol. Rev. 40, 869-907.

Yoshimori, R.N. (1971). Dissertation (Univ. of California, San Francisco, CA.).

Youngs, D.A. and Smith, K.C. (1973). J. Bacteriol. 114, 121-127.

Young, I.G., Leppik, R.A., Hammilton, J.A. and Gibson, F. (1972). J. Bacteriol. 110, 18-25.

Zimmerman, C.L., Appella, E. and Pisano, J.J. (1977). Analytical Biochemistry, 77, 569-573.