Order Number 1356814

Isolation and characterization of high copy number suppressors of the SOS system in *Escherichia coli*

Hunter, Rosalyn Alexandria, M.S.

The University of Arizona, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

ISOLATION AND CHARACTERIZATION OF HIGH COPY NUMBER SUPPRESSORS OF THE SOS SYSTEM IN *ESCHERICHIA COLI*

by

Rosalyn Alexandria Hunter

A Thesis submitted to the Faculty of the

DEPARTMENT OF MOLECULAR AND CELLULAR BIOLOGY

In Partial Fulfillment of the Requirements
for the Degree of

MASTER OF SCIENCE IN MOLECULAR AND CELLULAR BIOLOGY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 3

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgement the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Rosalyn A Hunter

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

David Mount

Dr. David Mount
Professor of Molecular and Cellular Biology

12/10/93

Date

ACKNOWLEDGEMENTS

I would like to give credit to the people who made this work possible. Thanks to Leslie Gregg-Jolly who always was there to help, to Ingrid Irbinskas who knew where everything was and who is a master in the manipulation of bacteria, to the other members of the lab who helped me including Greg Harlow, Mike Jenkins, Julie Mustard, and Keven Lewis, to Dr. Mount for having a good lab, and last and best, to my husband who helped me every step of the way. Thanks.

TABLE OF CONTENTS

TABLE OF CONTENTS - *Continued*

LIST OF ILLUSTRATIONS

LIST OF TABLES

ABSTRACT

In an attempt to find new genes that are involved in the induction of the SOS system of *Escherichia coli*, a plasmid library of *Escherichia coli* K-12 DNA cut with *Eco*RI was created in a pUC plasmid. The plasmids were transformed into the *Escherichia coli* strain AT492 containing a *sulA::lacZ* fusion. Colonies which did not show *lacZ* expression when the SOS systems had been induced by Mitomycin C were isolated. Four plasmids were found to suppress SOS induction when they were highly expressed. Kohara phage hybridization and restriction mapping suggest that these plasmids contain the genes for *lexA* and a truncated *recA* gene ending at the *Eco*RI site at basepair 1016.

INTRODUCTION

I. The SOS system in *E. coli*

The SOS system in *E. coli* is currently the largest known cellular stress response network encompassing more than 22 genes (Lewis, 1991). In response to DNA damage, an inducing signal is produced that activates the SOS system, a group of genes involved in genetic recombination, mutagenesis, and DNA repair (Little and Mount, 1982). These genes function to repair the DNA and to allow replication to continue by genetic recombination, excision repair, and by increasing mutagenesis during DNA replication.

The SOS system is regulated by the products of two genes, *recA* and *lexA*. The LexA protein is a repressor of all of the SOS genes. In fact, the genes of the SOS system are defined by the presence of a consensus LexA binding site near their promoters. The promoters of the SOS genes are normally bound by the LexA protein which prevents their expression. In the presence of DNA damage, the RecA protein changes conformation to an active form that can cleave LexA. Once cleaved, the LexA fragments can no longer bind to the promoters of the genes in the SOS system, thus derepressing the SOS system (see Figure 1). When the damage is repaired, RecA goes back to its non-active state and stops cleaving LexA, allowing the levels of LexA to rise so that repression of SOS genes is re-established.

The SOS system is believed to be a last-ditch repair mechanism that is induced when DNA damage becomes too great for the normal repair mechanisms to

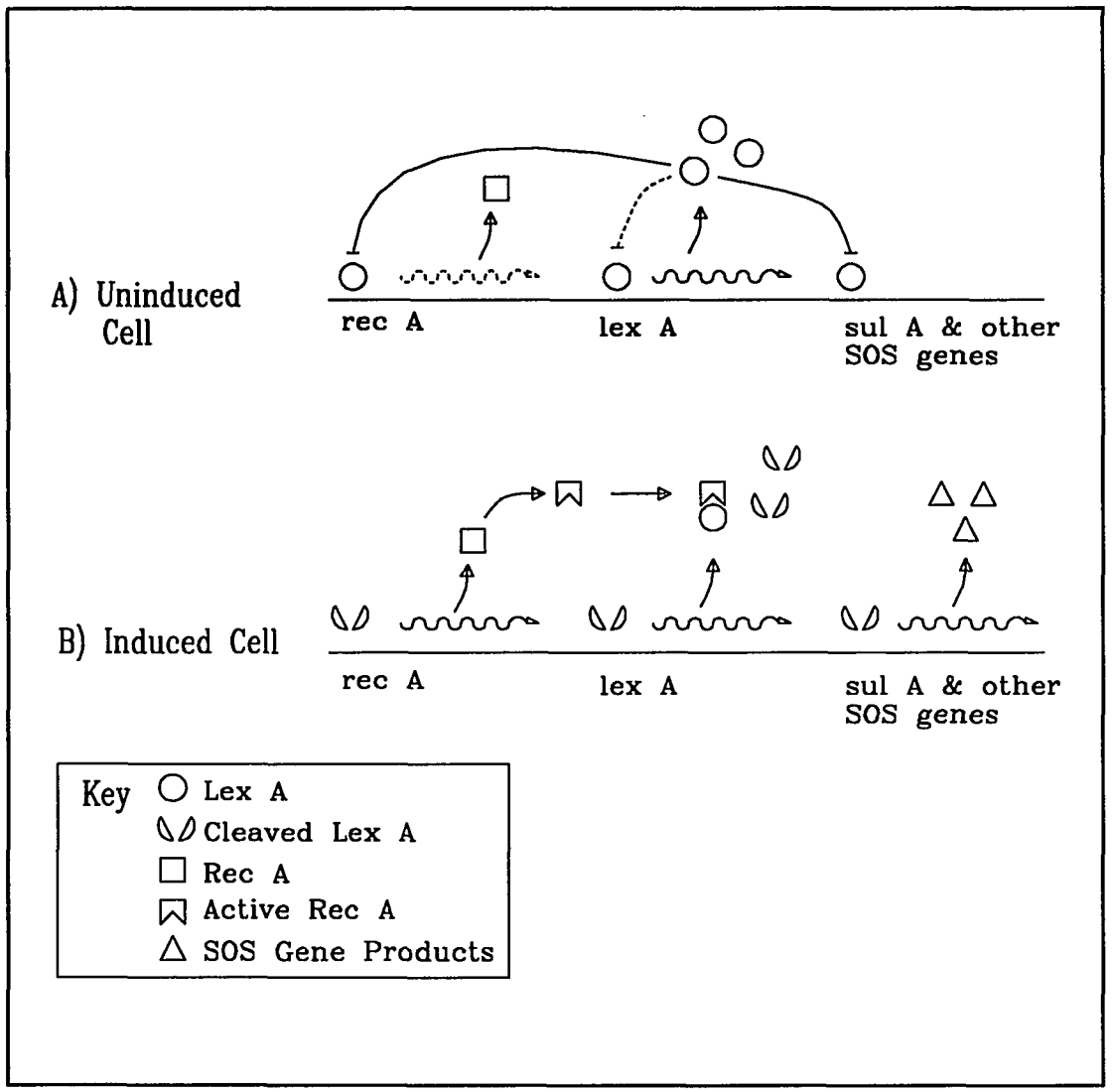

FIGURE 1 - Induction of the SOS system in *Escherichia Coli.*

A) When no DNA damage is present, *LexA* protein binds to the operators of the genes in the SOS system, preventing their expression. *LexA* binds to its own promoter when over-expressed and thus regulates its own expression. A small amount of *RecA* is produced.

B) When DNA damage is present *RecA* becomes activated. The Activated *RecA* cleaves *LexA* protein allowing *sulA* and the other SOS genes to be transcribed. When the DNA damage has been repaired, *RecA* is no longer found in its active state and the system reverts to A.

correct. The genes of the SOS system can repair damaged DNA in a way that produces many errors, yet allows replication to continue. Bacterial cells can more easily tolerate random mutations in genes than they can tolerate fragmented chromosomes. Therefore, those cells that have the ability to induce these repair mechanisms in emergencies are able to survive situations that would kill other cells. Presumably the genes are not always on, because under normal circumstances the housekeeping genes are more reliable at making accurate repairs and faithful replications of the chromosome than those genes repressed by LexA.

II. Properties of *lexA* and the LexA protein

The LexA protein is 202 amino acids long. It is believed to bind DNA as a dimer. Although the structure of LexA has not yet been solved, It is belived that the N-terminal part of the protein encodes a DNA binding domain and the C-terminal end encodes a dimerization domain (Little and Mount, 1982).

Mutant forms of the *lexA* gene have helped us to understand the mechanism of SOS regulation. The mutant gene *lexA(ind⁻)*, which encodes a LexA protein that can not be cleaved, blocks the expression of all SOS genes and makes the cells very sensitive to DNA damaging agents; whereas in the mutant *lexA(def)*, which encodes a gene that is deleted or completely inactive, the SOS genes are constitutively expressed. Also a temperature sensitive allele was isolated, *lexA(ts)*, where the SOS system was expressed constitutively at high temperatures (Little and Mount, 1982).

The LexA protein is functionally and probably structurally similar to the lamda phage C1 repressor. Both proteins prevent expression by binding to the DNA of the operator in the major groove and preventing RNA polymerase from binding to the DNA. Also, both proteins are cleaved by active RecA, and both proteins are believed to bind DNA as dimers. LexA cleavage is believed to remove the dimerization domain of the protein and cause protein inactivation (Little and Mount, 1982).

The decrease in the in the pool of intact LexA protein leads to increased expression of the SOS genes (Walker, 1987). Therefore, if *lexA* is overexpressed, the level of LexA protein might become high enough to prevent the SOS response.

III. Properties of RecA protein and RecA activation

RecA protein has several functions in the cell: (1) homologous recombination in *E. coli*, (2) a DNA dependent ATPase, (3) a co-protease of LexA and lambda cI repressor cleavage and induction, (4) renaturation of single-stranded DNA, and (5) SOS mutagenesis (Kowalczykowski, 1991). The RecA protein is essential for certain types of DNA repair. Homologous recombination is the only mechanism that can safely repair DNA crosslinks *in vivo*. Bacteria which lack RecA are extremely sensitive to DNA damage. Even in the absence of other DNA damaging agents, oxidative damage caused by aerobically growing cells can kill half of the cells in a $RecA^-$ culture (Cox, 1993).

Although RecA has been extensively studied, the nature of its conversion to

an active state is poorly understood. The inactive form of RecA is believed to bind single-stranded DNA which may be formed when the chromosome is damaged. This binding along with other unknown factors, which may include the formation of RecA polymers, somehow converts the RecA protein to its active form which cleaves the LexA protein and induces the SOS system. Cleavage of LexA protein by RecA has been shown to require ATP or an analogue, and denatured DNA (Little and Mount, 1982). One theory suggests that conformational changes are induced by the hydrolysis of RecA bound ATP to ADP. (Kowalczykowski, 1991).

Current studies suggest that RecA forms a nucleoprotein filament of thousands of monomers. The RecA monomers begin binding at a region of single-stranded DNA, or in an underwound region of double-stranded DNA in the presence of ATP, and extend along the strand at a rate of 1000 monomers per minute. The RecA filament contains a grove that can bind up to three strands of DNA to facilitate homologous recombination. Filaments bound to ATP have a different structure than filaments not bound to ATP, and when ATP is hydrolyzed, it is hydrolyzed uniformly throughout the filament (Cox, 1993).

The structures of the RecA monomer and the RecA filament which binds ADP have been solved (Story et. al., 1992A and Story and Steitz, 1992B). However, the ATP binding forms and the RecA active state structures are not known. RecA protein is 352 amino acids long (see Figure 2). The N-terminal end of the protein (including an alpha-helix called A and the beta sheet called 0, which ends at amino acid 28) is believed to be involved in stabilizing the RecA filament. The regions

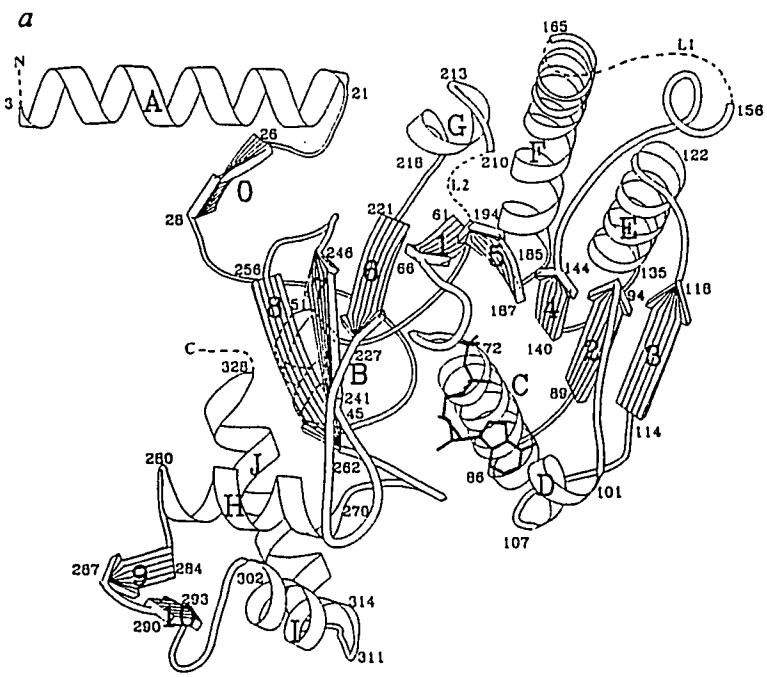

FIGURE 2 - Structure of RecA protein

The structure of RecA protein in shown. The image is from the paper Story et. al. 1992A.

L1 and L2 are thought to be DNA binding domains, but their structures are not yet known. ADP binds in the region near Alpha helices C and D. The C-terminal loop, which begins at amino acid 270 and includes alpha helices H and J and beta sheets 9 and 10, is belived to be involved in the formation of bundles of RecA fibers. The RecA protein is thought to form filaments, active filaments, and bundles *in vivo*. These structures are theorized to be important in RecA function (Story et. al., 1992A).

IV. Properties of other SOS genes

Most of the SOS genes encode enzymes that enact DNA repair, such as helicases and polymerases (see Table 1), but some of the SOS genes have less obvious functions. One of these genes, *sulA*, prevents cell division from occurring while the SOS system is induced. Only after the chromosome has been repaired and the SOS system is repressed again, will cell division occur. In this way, *sulA* prevents division before DNA replication can finish. This gene was discovered when mutant cells which expressed *sulA* constitutively were observed to form elongated or filamentous cells that did not divide (George et. al., 1975). The *sulA* gene would kill the cells if continually expressed, therefore often the *sulA* gene is deleted in SOS induction studies. In many studies where the SOS system is to be induced, the bacteria used contain a *sulA::lacZ* fusion. This fusion is very sensitive to changes in *lexA* binding and it shows when the SOS system is induced.

TABLE 1 - SOS system genes (*lexA* regulated genes)

Gene	Map site (min)	Function
lexA	91	SOS repressor
recA	58	General recombination, lexA repressor cleavage, SOS mutagenesis, recF pathway of recombination and DNA repair.
recN	58	recF dependent recombination, double-strand break repair
recQ	85	recF dependent recombination, resistance to thymine-less death
ruv	41	recF dependent recombination, daughter-strand gap repair
sulA(sfiA)	22	cell division inhibitor
umuC	25	SOS mutagenesis
umuD	25	SOS mutagenesis
uvrA	92	excision repair
uvrB	17	excision repair
uvrD	85	DNA helicase II, recF dependent recombination, excision repair, methyl-directed mismatch repair, DNA replication.
himA	38	Host factor component for lambda integration
dinA	2	DNA polymerase II. (Bonner et. al., 1990)
dinB	8	unknown
dinD	80-85	unknown - ATP binding site found
dinF	91	unknown
dinG	17.8	DNA Helicase (Lewis and Mount, 1992; Koonin 1993)

data from Peterson et. al., 1988.

V. Properties of RecA fragments

RecA protein has many complex functions. One way that researchers study these functions is by creating *recA* mutants which contain partial activity. Most of these studies make deletions in the sequence of *recA* and create a protein with missing sections. In this way, areas important for the different functions of RecA may be studied. Since RecA has many functions, the results are often difficult to interpret (see Table 2).

Because the RecA protein is required for many important functions in bacterial cells, most of the RecA studies have used strains that contained both wild-type RecA and a mutant RecA expressed on a plasmid. In these studies, the interaction between the wild-type RecA and the mutant RecA was observed. Some RecA mutants can interfere with the activity of wild-type RecA *in vivo* (Larminat and Defais, 1989; Sedgewick and Yarranton, 1982). The mutant RecA335 was able to negatively complement wild-type RecA protein. Larminat and Defais theorized that this truncated protein may lead to poorer interaction between RecA molecules by creating mixed multimers of RexA affecting bacterial survival (Larminat and Defais, 1989).

One of the most comprehensive studies of RecA fragments was performed at Osaka University in Japan (Horii et. al., 1992). The researchers studied several C-terminal and N-terminal deleted RecA fragments and reached the following conclusions: A) Most RecA fragments whose C-terminal sequences had been deleted interfered with UV sensitization of $RecA^+$ bacteria as well as recombination, but did

TABLE 2 - Selected RecA Fragments and their Properties

NAME	%RECA[1]	DEL.	TERM END[2]	PROPERTIES	SOURCE[3]
RecA6337	95%	-15aa	N	UV sensitive, Severe interference w/ wild type	H
RecA6324	92%	-28aa	N	"	H
RecA6311	88%	-41aa	N	Little effect on wild type	H
RecA6293	83%	-59aa	N	"	H
RecA5290	82%	-67aa	C	UV sensitive, interfered with recombination but not SOS induction	H
RecA5275	78%	-77aa	C	"	H
RecA5259	73%	-93aa	C	"	H
RecA5203	57%	-149aa	C	"	H
RecA5327	92%	-26aa	C	SOS system always induced	H
RecA214	61%	-138aa	C	no interference with wild type	L
RecA335	95%	-17aa	C	Acts like wild type when alone. Interferes with cell survival, sos induction, and weigle reactivation	L
pMH27	79%	--	C	interferes with post-replicative repair but not recA activation	Y
pMH29	69%	--	C	"	Y
pMH21	60%	--	C	„	Y
pMH22	11%	--	C	"	Y
pMH28	8%	--	C	„	Y
pDR1461	73%	--	C	inhibits post replicative repair, prevents induction of SOS and reactivation of lambda phage.	S
pMH1	22%	--	C	"	S

1. The data of Yarranton and Sedgewick were measured in % of RecA. The percentages of the other RecA fragments was included for ease of comparison. 2. C means that the C terminal end of the protein was deleted. N means that the N terminal end was removed. 3. The source papers for the data were H (Horii et. al., 1992), L (Larminat and Defais, 1989), Y (Yarranton and Sedgewick, 1982), S (Sedgewick and Yarranton, 1982).

not interfere with induction of lambda prophage. B) Most N-terminal deleted fragments had little effect on wild-type RecA activity.

Two N-terminal deleted proteins did not follow the overall trends. The proteins, RecA6337 and RecA6324 which were deleted by 15 and 28 amino acids respectfully in the N-terminal region of RecA, severely inhibited wild-type RecA function although they comprised only about 5% as much protein as the wild-type RecA in the cell.

The C-terminal deleted protein RecA5327 which was missing only 26 amino acids from the end of the RecA gene, had a phenotype which suggested that it was always active. Cells containing these plasmids constitutively showed SOS induction. The RecA5327 protein cleaved LexA and the cells containing it showed filamentation (Horii et. al., 1992). The authors suggest that the last 25 amino acid residues at the C-terminal end may function as a regulator for DNA binding. They also theorize that the N-terminal region (specifically amino acids 41-59) may contain a site for RecA-RecA interaction (Horii et. al., 1992).

Larminat and Defais studied two C-terminal deleted RecA proteins, RecA335 missing the last 17 amino acids and ending at base pair 1058, and RecA214 missing 138 amino acids and ending at basepair 693. They stated that RecA335 had all of the functions of wild-type RecA when expressed in the absence of wild-type RecA, and yet in the presence of wild-type RecA it caused decreased SOS induction, cell survival, and Weigle reactivation. RecA214 could not cause SOS induction and had no effect on the wild-type RecA (Larminat and Defais, 1989).

Sedgewich and Yarranton observed that plasmids encoding an apparently inactive product caused both radiosensitization and reduced inducibility of some SOS functions even though they did not affect chromosomal RecA protein activity in undamaged cells (Sedgewick and Yarranton, 1982). They also found that plasmids encoding more than 22% of RecA can reduce the ability of cells to carry out post-replicative repair. However, these truncated RecA proteins did not interfere with wild-type RecA activation (Yarranton and Sedgewick, 1982). In the RecA structure paper, Story et. al. 1992A, the authors suggest that the findings of Yarranton and Sedgewick agree with their structure data showing that alpha helix A and beta strand 0 are involved in fiber formation. They suggest that the RecA fragments containing only these N-terminal regions, poison filament growth by adding to the growing end and preventing further polymerization (Story et. al., 1992A).

The above studies suggest that the altered activity of the SOS system is due to the actions of the truncated RecA proteins; however the work of Sedliakova et al. suggests that overproduction of wild-type RecA by a high-copy number plasmid is sufficient to partially inhibit SOS induction perhaps by lowering the activity of RecA protease (Sedliakova et. al., 1989). The authors looked at the effect of a high copy number plasmid carrying *recA* on a *sulA::lacZ* fusion. They observed that the plasmid containing wild-type *recA* inhibited induction of this fusion. They theorize that this is due to rapid masking of pyrimidine dimers by the RecA proteins and that these masked dimers are less efficiently distinguished by nucleases. Horii et al. (1992) also show this effect of wild-type RecA overexpression.

RecA fragment data is often confusing. Since the changes in RecA were

formed by modifying DNA sequences and introducing them on plasmids, it is not known if these RecA fragment proteins are properly folded or not. Some of the inconsistencies in the data between fragments of similar sizes may be due to improper folding of all or part of the protein. By combining the information found in these studies with new information about RecA structure, regions that are important to RecA function may be located, but the complexity of RecA function makes the mechanism of RecA activation difficult to understand.

VI. Previous studies of SOS induction

Previous studies have looked for constitutive mutants in the SOS system. They usually involve screening for mutants which continually turn on the SOS system. Mutations in *dam, dnaB, dnaE, dnaG, lig, polA,* and *uvrD* all can derepress the SOS system to various extents (Craig et. al., 1984; Casaregola et. al., 1982; Schuster et. al., 1973; Blanco and Pomes, 1977; Ossanna, 1988). This type of mutant search is less useful than our search in understanding the mechanisms of SOS induction because any mutation which will cause DNA damage to accumulate will cause induction of the SOS system. This problem can be avoided by looking not for genes which can induce the SOS system but for genes which prevent the induction of this system. Theoretically, such a gene may interact with the regulators of the SOS system, LexA and RecA. By finding such genes, we would be able to gain new insights on the mechanism of activation of the RecA protein.

This is not the first study to observe high copy number supressors of SOS

induction. Bagdasarian et al. (1986) report that a plasmid, R6-5, contained a locus whose product inhibited induction of *sfiA* (*sulA*) in a *recA*$^+$ host after treatment with nalidixic acid. The locus, called *psi*, was found on conjugating plasmids and it is theorized to protect single-stranded DNA and prevent SOS induction in order to avoid turning on the SOS system during conjugation.

VII. Objective of the present work

In this study, *E. coli* plasmid libraries were used to screen for genes which serve as high copy number suppressors of SOS system induction. A DNA damaging agent (Mitomycin C) was used to induce the SOS system. Colonies in which the SOS system was not derepressed were isolated and afterwards the phenotype was shown to be plasmid linked. The genes on the plasmid were identified using Kohara phage hybridization and DNA sequencing.

RecA activation is an essential step required for the derepression of the SOS system. One goal of this study was to find factors involved in RecA activation. If RecA requires the presence of some other gene product for activation to occur, then it is possible that overexpression of this product or products might cause squelching of the essential components needed for RecA activation. This type of study would find different genes than a mutant search, because if the gene was essential, a mutation in this gene would be lethal.

One gene that we expected to find in this study was *lexA*. Since the LexA protein concentration determines the rate of SOS induction, overexpression of *lexA*

should create an excessive amount of LexA protein that would overwhelm the ability of RecA to cleave it.

The SOS system is one of the best known systems in *E. coli* and yet many important parts of it are poorly understood. This research was intended to discover how the early stages of this system work.

MATERIALS AND METHODS

I. Media

Strains were grown in Luria broth (L broth) containing 10g Bacto-tryptone (Difco), 5g yeast extract (Difco), 10g NaCl and 1 pellet NAOH (J.T. Baker) per liter. L plates also contained 15g Difco agar. MacConkey agar (Difco) was prepared using alpha-lactose (Sigma) as the sugar source, 40g agar, and 10g sugar. SOC broth was made with 5g of bactotryptone, 1.25g yeast extract, 2.5ml NaCl, 0.625ml 1M KCL, 2.5ml 2M MgSO4, 3.75ml 40% glucose, for a total of 250ml of filter sterilized broth.

The antibiotics were used in the following concentrations: 100µg/ml Ampicillin (Sigma), and 0.07µg/ml mitomycin C (Sigma) if not otherwise stated. During the initial isolation of the plasmids, 0.1µg/ml mitomycin C was used. In most cases, ampicillin was included in the agar before plating, whereas mitomycin C was spread on the plates the day before use. Other chemicals used were purchased from Sigma Chemical Co. (St. Louis, MO).

II. Kohara Mapping

The Kohara phage filter was purchased from Takara Biochemical. It contained the mini-set Kohara phage library and was used to localize the inserted DNA. The plasmids were digested with *Eco*RI, and the *E. coli* DNA fragment was

isolated by agarose gel electrophoresis and removed from the gel using Geneclean (Bio101). Random primers were used to make radioactive P^{32} labeled copies of the fragment. These labeled copies were used for Southern blotting. The random primer kit was purchased from Bethesda Research Laboratories (Gaithersburg, MD).

The positively-charged nylon membrane filter was prepared for hybridization and probing by using the technique listed by MagnaGraph. The radioactive insert DNA was hybridized with the filter for 19 hours and the film was exposed to the filter overnight.

III. Plasmids and Strains

The strains of *Escherichia coli* used in these experiments are listed in Table 3. The plasmids used and their derivations are listed in Table 4. AT489 is a precursor strain to AT492 and shows a slightly cleaner phenotype. Since the contrast in color is greater between colonies in the presence and absence of Mitomycin C, AT489 was used exclusively in later experiments.

IV. Plasmid isolation and Transformations

Plasmid isolation was done using the "Jimmy Crack Corn" method which is a modification of the Holmes and Quigley method cited in Maniatis (1982) p. 366 and in Biotechniques 6(5):460. The entire procedure was done in one 1.5ml tube so that large numbers could be done at once. The centrifugation step was reduced to

15 minutes and the viscous cellular pellet was removed using a sterile toothpick. The aqueous solution containing the plasmid was concentrated by precipitation with isopropanol and a wash with 70% ethanol and then vacuum dried and filled with water or TE buffer. The concentrated DNA was pure enough for most purposes. It contained some RNA but most chromosomal DNA was removed by the procedure.

Transformations were done using the procedure described by Inoue et al. with the exception that the heat shock was done for 45 seconds in a 37°C bath.

V. Restriction Mapping

Restriction maps were derived by 1% agarose gel electrophoresis of single and double digests of the plasmids with the enzymes *Bam*HI, *Pst*I, *Eco*RI, *Eco*RV, and *Hin*dIII. The restriction enzymes were purchased from Stratagene except for *Eco*RV which was obtained from Gibco. Maps were matched to *E. coli* physical maps of Kohara (Kohara et al. 1987) and Rudd (Rudd, 1992).

VI. Competent Cell production

The competent cells were made using the procedure of Inoue et al. 1990. SOC broth was used to grow the cells instead of SOB suggested in the paper.

TABLE 3 - Bacterial strains

Strain	Genetic Markers	Source
AT489	recA$^+$, sulA11, lacΔU169(argF-lacZYA), thr-1, leuB6, his-4, arg-E3, ilv(ts), galK2, rpsL31, supE44, srl::Tn10, λsulA::lac c1(ind$^-$).	A. Thliveris
AT492	as AT489 with F'lacI^Q, lacZΔM15, TN9.	A. Thliveris
DM1790	thr-1, ara-14, leuB6, Δ(gpt-proA)62, Δlac, tsx-33, Δpro, gal$^+$, sup$^+$, hisG4, rfbD1, mgl-51, rpsL31, kdgK51, xyl-5, mtl-1, argE3, thi-1.	Mount
KP245	λ c1(ind$^-$), recAop::Lac, sulA211, Lac Δ U169(argF-lacZYA),LexA71::TN5 malB45, zja505::TN10malB45, thr-1, leuB6, his-4, arg-E3, ilv(ts), galK2, rpsL31, supE44.	Kevin Peterson
DH5α	F$^-$, endA1, hsdR17 (r_k^-, m_k^-), supE44, thi-1, λ$^-$, recA1, gyrA96, relA1,lac ΔU169(argF-lacZYA), ϕ80dlac.	Bethesda Research Laboratories

TABLE 4 - Plasmids

Plasmid	Base plasmid and Relevant markers	Source
pUC8	lacZα donor,amp^r	Mount
pUC19	pUC8, expanded cloning site, lacZα donor, amp^r	Gibco
pSR4	pUC8, amp^r, random *Eco*RI E. Coli fragments	(This study) Rundlett 1988
pSR5	"	"
pSR8	"	"
pSR12	"	"
pATT100	pBR322, amp^r, lexA$^+$, tet^r	A. Thliveris

VII. DNA Sequencing

DNA for sequencing was isolated using a Quiagen kit. The DNA was then sequenced at the University of Arizona Macromolecular Structures Facility. The Facility uses an ABI 373A DNA sequencing system which uses automatic electrophoresis and detection of fluorescently labeled DNA fragments. The sequencing is done by dideoxysequencing using the Taq cycle sequencing method.

VIII. Sequence Analysis

The sequence obtained for pSR5 was entered into the sequence analysis software system GCG using SEQ.ED and compared to the GENBANK sequence database using the program BLASTN. The pSR5 insert sequence pulled out several *recA* sequences as well as the multiple cloning site of pUC8. The GCG program GAP was used to align the sequences.

RESULTS

I. Isolation of plasmids

In 1988, graduate student Steve Rundlett produced an *E. coli* plasmid library by digesting *E. coli* genomic DNA with the restriction enzyme *Eco*RI and ligating the fragments into the plasmid pUC8. These plasmids were transformed into the reporter strain AT492 which is deleted for *lacZ* but which contains a *sulA::lacZ* fusion so that colonies appear red on MacConkey Lactose (MacLac) plates only when the SOS system is induced.

The AT492 colonies containing the plasmids were replica plated onto MacLac media containing mitomycin C, a DNA damaging agent that induces the SOS system. Those colonies which did not turn red in the presence of mitomycin C were grown and the plasmids were isolated. These plasmids were then retransformed into AT492 and the phenotype was confirmed to be plasmid-linked (see Figure 3).

Rundlett found four plasmids which reduced or prevented the induction of the SOS system in *E. coli*. These plasmids were designated pSR4, pSR5, pSR8, and pSR12 and are believed to contain *E. coli* genes which can act as suppressors of SOS induction when expressed on a high copy number plasmid. My goal was to discover which genes were encoded on these plasmids.

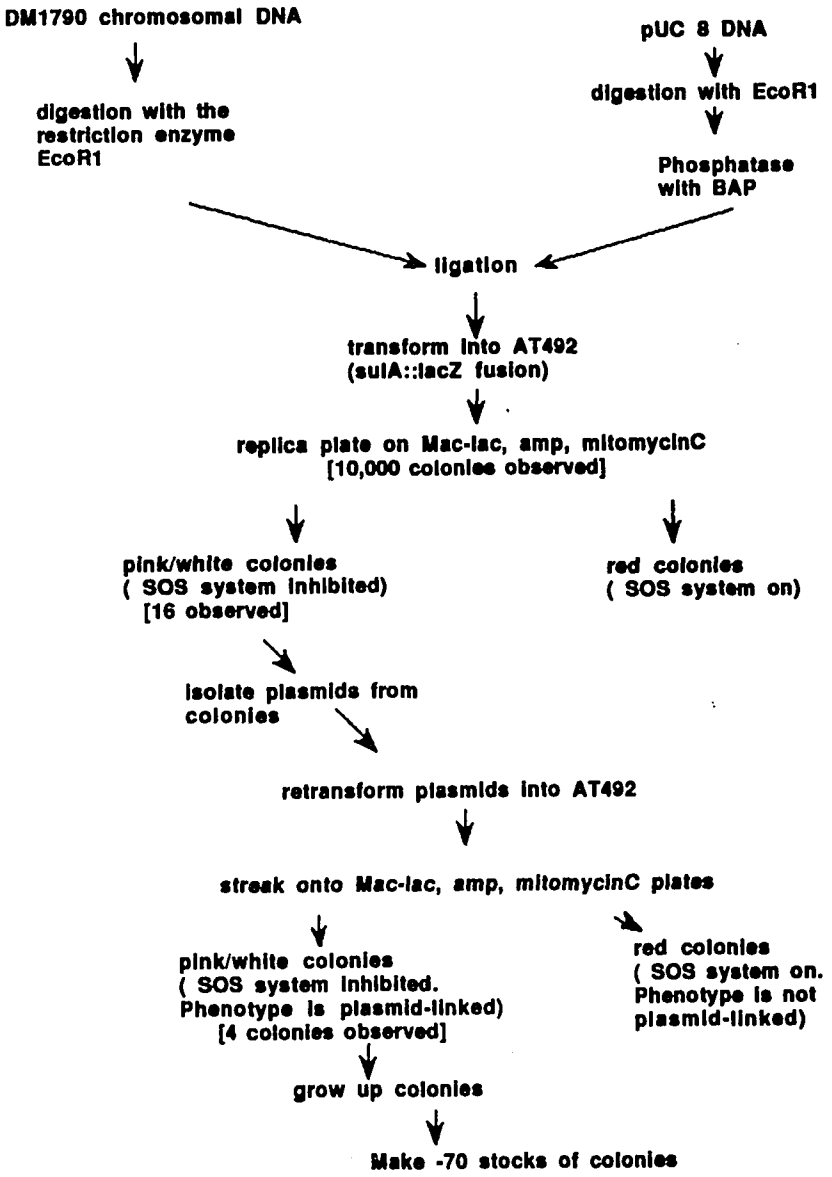

FIGURE 3 - Procedure for Isolating Genes that Suppress SOS induction.

Chromosomal DNA was cut with *Eco*RI and ligated to pUC8 plasmids. The plasmid library created was used to transform AT492 cells which contains a *sulA:lacZ* fusion. The colonies were screened for the induction inhibition phenotype and then checked to see that the phenotype was plasmid linked. Those colonies which showed the phenotype linked to the plasmid were preserved for future study.

II. LexA phenotype test

Because LexA was expected to be able to cause this phenotype, the first test that was done was a test for the presence of LexA. In this experiment, the cells were grown from -70°C stock, and the plasmids were isolated. The plasmids were then transformed into the strain KP245 which is LexA deficient, so that the SOS system is always induced. The strain, KP245, contains a reporter gene, *recAop::lac*, which makes the cells red on MacLac plates. The normal *lac* gene is deleted and *sulA* is defective in this strain. If LexA is expressed by the plasmids it can bind the operators of the SOS genes and turn off expression of the reporter gene.

The plasmids pSR4, pSR8 and pSR12 were successfully transformed into KP245 and grown on MacLac plates. The colonies were compared to KP245 colonies containing the control plasmids pUC8 and pATT100. The negative control plasmid, pUC8, appeared purple on MacLac plates, and the positive control plasmid, pATT100, which contains the gene *lexA* appeared white. The plasmids pSR8 and pSR4 were purple and red respectively, showing that they did not contain active *lexA* genes; while the colonies containing the plasmid pSR12 appeared white showing the presence of a functional *lexA* gene on the plasmid. The colonies containing pSR12 appeared as white as the control plasmid pATT100.

III. Kohara phage mapping of plasmids

The three plasmids not shown to contain a functional *lexA* were used to probe a Kohara phage library of *E. coli*. The plasmids (pSR4, pSR5, and pSR8) were hybridized to a filter which contained a Kohara phage library of *E. coli*. The library revealed that pSR8 hybridized to the Kohara fragments 634(1F8), 635(12B2), and 636(1E4). Inserts for both pSR4 and pSR5 hybridized to the Kohara fragment 446(23B7). An examination of the Kohara phage library revealed that fragments 634, 635, and 636 overlap in a region which contains the *lexA* gene and fragment 446 contains the *recA* gene (see Figure 4). These data suggested that the *lexA* and *recA* genes had been isolated. Both of these genes are already known to regulate the SOS system.

IV. Restriction Digest Mapping of Plasmids

To confirm that *recA* and *lexA* were present in the plasmids as suggested by the Kohara phage data, the plasmids were mapped by digestion with several restriction enzymes. These data were used to construct maps that were compared to the Rudd *E. coli* map at the locations suggested by the Kohara phage fragments (see Figure 4).

Restriction digest information suggested that pSR12 was similar to pSR8. The plasmid maps suggest that they contain inverted copies of the same *Eco*RI fragment (see Figure 5). These data along with the previous phenotype observed in

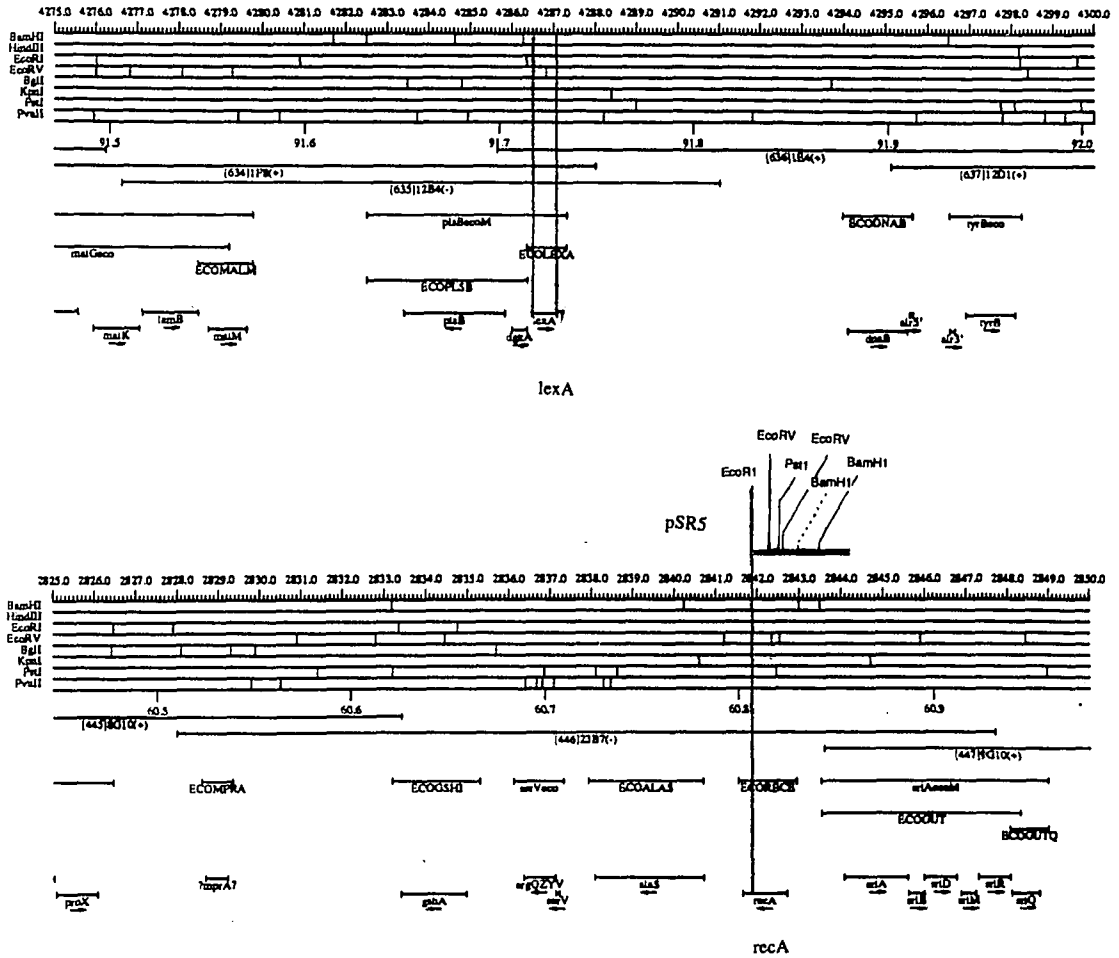

FIGURE 4 - Kohara Phage Maps showing the locations of *recA* and *lexA*.

A) The *lexA* gene is shown to lie within the overlap of the Kohara fragments 634, 635, and 636.

B) The *recA* sequence is shown to lie in a region of Kohara fragment 446 that does not overlap any other fragments. The location of the insert sequence from pSR5 is shown above.

pSR12 suggest that pSR8 and pSR12 contain an intact *lexA* gene, and that this gene is preventing SOS derepression by overexpressing LexA protein and hiding the effect of RecA activation. The fact that *lexA* is one of the only genes that lies within the overlap region of the three Kohara fragments strongly supports the fact that it is present on these plasmids. The restriction map data are puzzling however because both pSR8 and pSR12 show the presence of an extra *Eco*RI site as well as an extra *Eco*RV site. Because the sequence is found in both plasmids, and both plasmids were isolated seperately, the sequence difference may have been present in the strain used to isolate the DNA.

The SOS inhibiting effect of pSR5 and pSR4 was puzzling, because *recA* is a positive regulator of the SOS system. The *recA* gene on the plasmid pSR5 contains an internal *Eco*RI site suggesting that the entire *recA* gene may not have been present. A number of papers suggested that truncated RecA fragments can have an inhibitory effect on SOS induction (Larminat and Defais, 1989; Sedgewick and Yarranton, 1982). If the *Eco*RI site in pSR5 ends within the *recA* gene, a truncated RecA fragment would be created that could interfere with induction of the SOS system.

To confirm this theory, the plasmid was mapped by digestion with several restriction enzymes. Although map data suggested that the plasmid pSR5 contained a truncated *recA* gene, there were two pieces of evidence that did not match. First, one of the expected *Bam*HI sites was missing, and second, the size of the insert did not match the size of predicted the *Eco*RI fragment (see Figure 5). In order to make certain that *recA* was present, the plasmid was sequenced.

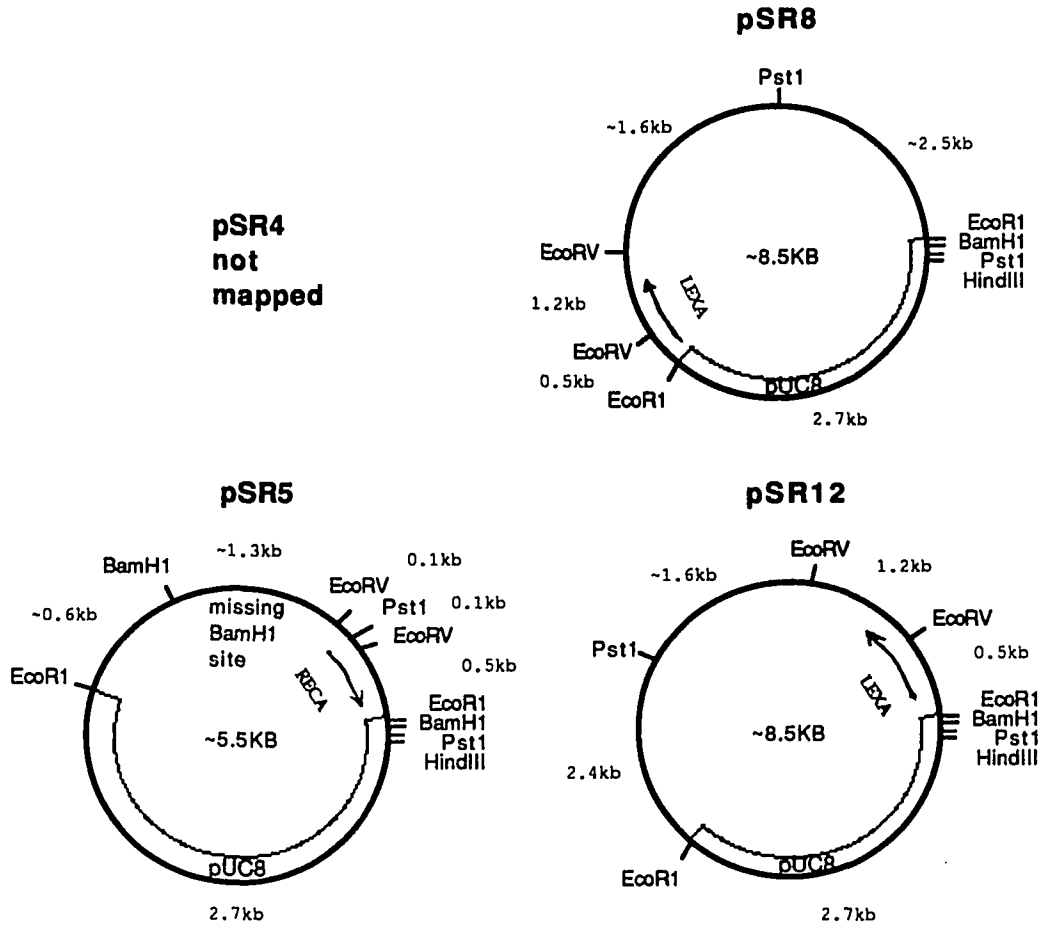

FIGURE 5 - Plasmid Maps made from Restriction Data for the Plasmids pSR5, pSR8, and pSR12.

These plasmid maps were made by analysis of plasmid digests. The maps for pSR8 and pSR12 are preliminary maps that were not completed. The map for pSR5 shows the location of the truncated *recA* sequence on the plasmid. (Maps not to scale).

V. Sequencing and sequence analysis

Sequence analysis confirmed that pSR5 contained a truncated *recA* gene that was cut at the *Eco*RI site internal to *recA* at basepair 1016. The match in the region shown is exact except for a G to C mutation, and a single nucleotide gap in the sequence (see Figure 6). The gap would normally constitute a frameshift mutation. However, because the RecA fragment has some activity, and because the plasmid was sequenced only once, It seems likely that no frameshifts have occurred in the sequence, and the gap is due to an error in interpreting the sequence data.

```
        pSR5 insert x recA            Gaps:      1
        Percent Similarity: 99.728   Percent Identity: 99.455

           EcoR1
pSR5    35 ggaattcagcctgtttaaacggcgcagcgattttgttcttcaccactttc 84
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
recA  1022 ggaattcagcctgtttaaacggcgcagcgattttgttcttcaccactttc 973
           .         .         .         .         .
        85 acgcgggtttcgctacccaccacgttttcgccctctttcaccgcgccgat 134
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
       972 acgcgggtttcgctacccaccacgttttcgccctctttcaccgcgccgat 923
           .         .         .         .         .
       135 acgacggatgtcgagacgaacagaggcgtagaatttcagcgcgttaccac 184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
       922 acgacggatgtcgagacgaacagaggcgtagaatttcagcgcgttaccac 873
           .         .         .         .         .
       185 cggtagtggtttccgggttaccgaacatcacaccaattttcatacggatc 234
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
       872 cggtagtggtttccgggttaccgaacatcacaccaattttcatacggatc 823
           .         .         .         .         .
       235 tggttgatgaagatcagcagcgtgttggactgcttcaggttacccgccag 284
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
       822 tggttgatgaagatcagcagcgtgttggactgcttcaggttacccgccag 773
           .         .         .         .         .
       285 cttacgcatcgcctggctcatcatacgtgccgcaagggccatgtgagagt 334
           |||||||||||||||||||||||||||||||||| ||||||||||||
       772 cttacgcatcgcctggctcatcatacgtgccgcaaggcccatgtgagagt 723
           .         .         .         .         .
       335 cgccgatttcgNcttcgatttccgctttcggcgtcaagtgccgccacgga 384
           ||||||||||:||||||||||||||||||||| ||||||||||||||
       722 cgccgatttcgccttcgatttccgctttcggcgtc.agtgccgccacgga 674
                   .
       385 gtcaacgacgataacgtc. 402
           ||||||||||||||||||
       673 gtcaacgacgataacgtct 655
```

FIGURE 6 - Sequence comparison of pSR5 to *recA*

The insert sequence of pSR5 shows identity with the sequence of *recA*. The pSR5 insert sequence is shown starting at the *Eco*RI site in the pUC8 multiple cloning site. The *recA* gene shown here is transcribed right to left. The sequence begins to match at the *Eco*RI site internal to *recA*. This shows clearly that the *recA* sequence cloned onto the plasmid ends at the internal *Eco*RI site.

DISCUSSION

These experiments did not discover any new SOS regulatory elements. Because only one restriction enzyme was used, and because plasmids are biased in the size of fragments that they will incorporate, it is possible that other genes may be found whose products can prevent induction of the SOS system. This experiment was performed again using different restriction enzymes and a different *E. coli* strain as a source of DNA. However, in these new experiments no new genes were found.

The difference in size between the *Eco*RI fragment in pSR5 and the fragments shown in the Rudd map may be due to a mutation in the genomic DNA creating a new *Eco*RI site; or perhaps when the original genomic digest was made, the restriction enzyme had star activity and cut the genomic DNA at a site similar to an *Eco*RI site that when cloned into the pUC8 plasmid became an *Eco*RI site. It is clear that an *Eco*RI site now exists in pSR5, however this may not be so in pSR4. The plasmid pSR4 was not mapped because the plasmid could not be digested using normal methods. Perhaps the plasmid was heavily supercoiled by the isolation procedure that we used. The colonies that prevented SOS induction appeared white or light pink when incubated overnight on MacLac plates in the presence of Mitomycin C, however, the colonies stayed white for only a short time and turned pink or red after a few days. Mitomycin C is a crosslinking agent that joins the DNA strands and prevents replication. In order for us to see colonies at all, some of the cells must have been healthy enough to replicate their DNA. To

continue replication and division, the bacterial cells must eventually be able to overcome the repair inhibition caused by the introduced genes. The fact that many of the colonies were pink suggests that the prevention of derepression was leaky. The colonies overall had white cells but some had the SOS system induced. Those cells which induced the SOS system would in time dominate the colony causing the darkening of colonies over time.

It was always a concern in this study that the growth of these strains on mitomycin C and the repression of repair genes could cause mutations in the strain and plasmids that would affect our results. The lack of a *Bam*HI site in the pSR5 map may be due to proliferation of a plasmid which had a mutation destroying the restriction site. It is also possible that there are sequence differences between the strain used to make the Rudd map, and the strain that was used in this study. The other differences between our data and the published sequence may be due to such mutations.

The phenotypes of the plasmids pSR8 and pSR12 were caused by the presence of *lexA* on the plasmids. Overexpression of *lexA* prevents derepression by providing so many copies of LexA that activated RecA can not cleave them all and derepress the system. The fact that pSR8 did not show a positive phenotype in the LexA test may be due to a mutation obtained in the plasmid after its isolation which created a *lexA*(def) gene.

The phenotypes of the plasmids pSR4 and pSR5 were caused by the presence of the sequence for a truncated RecA protein. The *recA* fragment that was causing the high copy number suppression encodes a C-terminal deleted protein. The

sequence ends at the internal *Eco*RI site at 1016 in the *recA* gene. This RecA fragment may have been interfering with the wild-type RecA by hiding sites of DNA damage, or forming hybrid RecA polymers with reduced activity.

The truncated *recA* gene encoded by the plasmid pSR5 is most similar to the truncated *recA* gene on the plasmid pDR1461 described in the papers by Sedgewick and Yarranton (Sedgewick and Yarranton, 1982; Yarranton and Sedgewick, 1982). The authors reported that their plasmid, pDR1461, which contains 73% of the *recA* gene and ends at the same *Eco*RI site as the *recA* gene in pSR5, interfered with reactivation of lambda phage, increased the sensitivity of cells to UV, and affected SOS induction of cells containing wild-type RecA protein (Sedgewick and Yarranton, 1982). On the contrary, plasmids containing recA fragments of very similar sizes to pSR5 (pMH29 encoding 69% of *recA* and pMH27 encoding 79% of *recA*) did not interfere with the ability of wild-type RecA protein to induce the SOS system (Yarranton and Sedgewick, 1982).

Although the sequence data clearly shows the termination of the RecA sequence at basepair 1016, transcription should add 19 additional amino acids from the cloning vector that may have a role in the function of this protein.

Current evidence suggests that RecA forms a multimer of many RecA proteins that bind all along the length of a DNA molecule. The fibers work in concert in the hydrolysis of ATP (Cox, 1993). The presence of defective or partially active RecA monomers may have an inhibitory effect on the action of the RecA fiber. If the RecA protease depends on the concerted action of a collection of monomers, then monomers with losses in function are sure to affect activity.

If the active form of RecA is a multimer where most or all fragments must be active to cause induction, then in the presence of truncated RecA proteins we would form mixed multimers with little or no activity that would not be able to cleave LexA. This would prevent the activity of wild-type RecA because the concentration of pure active multimers in a cell over-expressing defective RecA would be very small.

The RecA fragment encoded by pSR5 should be 260 amino acids in size. It lacks 96 amino acids found in the C-terminal end of RecA and gains 19 amino acids from the plasmid (see Figure 7). The truncated RecA protein lacks the C-terminal loop that is thought to be necessary for bundling of fibers (Story et. al., 1992A). The plasmid, pSR5, was isolated as an inhibitor of SOS induction, this suggests that bundling of RecA may be important in SOS induction perhaps by regulating RecA activation.

In conclusion, we have isolated a C-terminal deleted RecA fragment that can inhibit induction of the SOS system when expressed from a high copy number plasmid. These findings support the work of other researchers and they may be of help in ultimately understanding the mechanism of RecA protein activation.

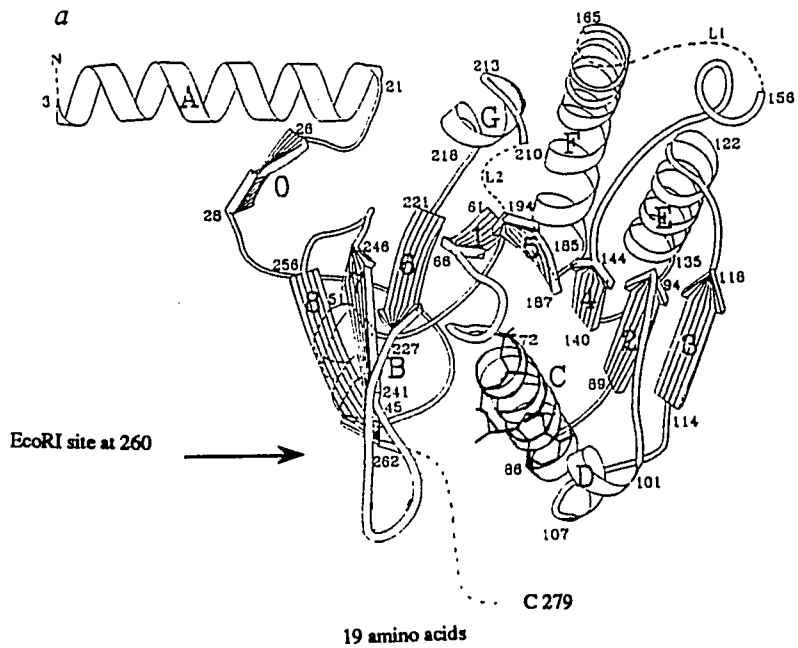

FIGURE 7 - Structure of RecA fragment from pSR5 *recA* gene

The probable structure of the RecA fragment encoded by the plasmid pSR5. Compare to Figure 2. Structure is modified form from structure by Story et al. 1992A.

REFERENCES

BAGDASARIAN, M., A. Bailone, M. M. Bagdasarian, P. A. Manning, R. Lurz, K. N. Timmis, and R. Devoret. 1986. An inhibitor of SOS induction, specified by a plasmid locus in Escherichia coli. PNAS U.S.A. 83:5723-5726.

BLANCO, M. and L. Pomes. 1977. Prophage induction in *Escherichia coli* K12 cells deficient in DNA polymerase I. Mol. Gen. Genet. 154:287-292.

BONNER, C. A., S. Hays, K. McEntee, and M. F. Goodman. 1990. DNA polymerase II is encoded by the DNA damage inducible *din*A gene of Escherichia coli. PNAS U.S.A. 87(19):7663-7667.

CASAREGOLA, S., R. D'Ari, and O. Huisman. 1982. Role of DNA replication in the induction and turn-off of the SOS response in *Escherichia coli*. Mol.Gen.Genet. 185:440-444.

COX, Michael M. 1993. Relating Biochemistry to Biology: How the recombinational repair function of RecA protein is manifested in its molecular properties. BIOESSAYS Vol. 15(9):617-623.

CRAIG, R. L., J. A. Arraj, and M. G. Marinus. 1984. Induction of damage inducible (SOS) repair in *dam* mutants of *Escherichia coli* exposed to 2-aminopurine. Mol. Gen. Genet. 194:539-540.

GEORGE, J., M. Castellazzi, and G. Buttin. 1975. Prophage induction and cell division in *E. coli*. III. Mutations *sfiA* and *sfiB* restore division in *tif* and *lon* strains and permit the expression of mutator properties of tig. Mol. Gen. Genet. 140:309-332.

HORII, Toshihiro, Naoki Ozawa, Tomoko Ogawa, and Hideyuki Ogawa. 1992. Inhibitory effects of the N- and C-terminal truncated *Escherichia coli recA* gene products on functions of the wild-type *recA* gene. J. Mol. Biol 223:105-114.

INOUE, H., H. Nojima, and H. Okayama. 1990. High efficiency transformation of *Escherichia coli* with plasmids. GENE 96:23-28.

KOHARA, Yuji, Kiyotaka Akiyama, and Katsumi Isono. 1987. The physical map of the whole *E. coli* chromosome: Application of a new strategy for rapid analysis and sorting of a large genomic library. CELL 50:495-508.

KOONIN, Eugene V. 1993. *Escherichia coli din*G gene encodes a putative DNA helicase related to a group of eukaryotic helicases including Rad3 protein. Nucleic Acid research 21(6):1497

KOWALCZYKOWSKI, Steven C. 1991. Biochemistry of genetic recombination: Energetics and mechanism of DNA strand exchange. Annu. Rev. Biophys. Chem. 20:539-575.

LARMINAT, Florence and Martine Defais. 1989. Modulation of the SOS response by truncated RecA proteins. Mol Gen Genet 216:106-112.

LEWIS, Lysle Kevin. 1991. Use of genetic and computer modeling approaches to identify DNA damage-inducible genes in Escherichia coli. PhD dissertation. The University of Arizona.

LEWIS, L. Kevin, and David W. Mount. 1992. Interaction of lexA repressor with the asymmetric *din*G operator and complete nucleotide sequence of the gene. J.Bac. 174(15):5110-5116.

LITTLE, John W. and David W. Mount. 1982. The SOS regulatory system of Escherichia coli. Cell 29:11-22.

MANIATIS, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning,Cold Spring Harbor Laboratory Press, Cold Spring Harbor,N.Y.

OSSANNA, Nina. 1988. Isolation and characterization of SOS constitutive mutations in Escherichia coli. PhD dissertation. The University of Arizona.

PETERSON, Kenneth R., Nina Ossanna, Andrew T. Thliveris, Don G. Ennis and David W. Mount. 1988. Derepression of specific genes promotes DNA repair mutagenesis in Escherichia coli. Journal of Bacteriology 170(1):1-4.

RUDD, Kenneth E. 1992. "Alignment of *E. coli* DNA sequences. A revised, integrated genomic restriction map " in A short course in Bacterial Genetics: A laboratory manual and handbook for *Escherichia coli* and related bacteria. by Jeffrey Miller. Cold Spring Harbor Press. Cold Spring Harbor, New York. pp. 2.3 - 2.43.

RUNDLETT, Steve. 1988. Unpublished data.

SCHUSTER, H., D. Beyersmann, M. Mikolajczyk and M. Schlicht. 1973. Prophage Induction by high temperature in thermosensitive *dna* mutants lysogenic for bacteriophage lambda. J. Virol. 11:879-885.

SEDGEWICK, Steven G. and Geoffrey T. Yarranton.1982. Cloned Truncated *recA* genes in *E. coli* I. Effect on radiosensitivity and *recA*+ dependent Processes. Mol Gen Genet 185:93-98.

SEDLIAKOVA, M., F. Masek, and K. Kleibl. 1989. In UV-irradiated Escherichia coli PQ35 overproducing the recA protein, expression of the sfiA gene and dimer excision are alleviated. Mol Gen Genet 217:427-429.

STORY, Randall M., Irene T. Weber, and Thomas A. Steitz. 1992A. The structure of *E. coli recA* protein monomer and polymer. NATURE. 355:318-324.

STORY, Randall M. and Thomas A. Steitz. 1992B. Structure of the *recA* protein-ADP complex. NATURE. 355:374-376.

WALKER, Graham C. 1987. "The SOS response of *Escherichia coli*" in *Escherichia Coli* and *Salmonella typhimurium* - cellular and molecular biology. vol2. American Society for Microbiolgy. Washington D.C.

YARRANTON, Geoffrey and Steven G. Sedgwick. 1982. Cloned Truncated *recA* genes in *E. coli* II. Effects of truncated gene products on in vivo *recA+* protein Activity. Mol Gen Genet 185:99-104.