Journal of Biochemistry and Molecular Biology, Vol. 32, No. 5, September 1999, pp. 436–439

JOURNAL OF
Biochemistry
and
Molecular Biology

Cloning and Nucleotide Sequence of the *recA* Gene from *Shigella sonnei* KNIH104S Isolated in Korea

Yong-Chjun Park, Hee-Jung Shin, and Young-Chang Kim*
School of Life Sciences, Chungbuk National University, Cheongju 361-763, Korea

Received 8 April 1999, Accepted 18 May 1999

***Shigella sonnei* is an important cause of human enteric infections. *S. sonnei* KNIH104S was previously reported to be isolated from Korean shigellosis patients. We cloned a 2.8-kb *Kpn*I fragment containing the *recA* gene encoding a recombinase from the chromosomal DNA of *S. sonnei* KNIH104S. This recombinant plasmid was named pRAK28. *E. coli* HB101, a *recA* mutant, cannot grow on Luria-Bertani medium in the presence of the alkylating agent methylmethane sulfonate, however, *E. coli* HB101 harboring pRAK28 was found to grow on this medium. As far as we know, we are the first to sequence the *recA* gene from *S. sonnei*. This gene is composed of 1062 base pairs with an ATG initiation codon and a TAA termination codon. Nucleotide sequence comparison of the *S. sonnei recA* gene exhibited 99.7% and 99.5% identity with those of *S. flexneri* and *E. coli*, respectively.**

Keywords: Cloning, *recA* gene, Sequence, *S. sonnei* KNIH104S.

Introduction

Shigellosis is a common and serious disease that occurs throughout the world and is particularly prevalent in developing countries (Karnell *et al.*, 1991). *Shigella* spp. are Gram-negative microorganisms that harbor a high-molecular-weight plasmid encoding proteins essential for virulence (Sansonetti *et al.*, 1982). Genetic analysis of the plasmid has identified a complicated system coding for products essential for the internalization and spread of the bacteria within and between host cells, the secretion of these internalization proteins to the surface of the bacterium, and the regulation of virulence gene expression. The expression of these virulence genes is environmentally regulated at the level of transcription, and regulatory proteins controlling expression are encoded by genes located on the plasmid and on the chromosome (Hale, 1991).

Although *S. dysenteriae* is the cause for more illnesses internationally (Taylor *et al.*, 1989), *S. sonnei* is more dominant in Korea (Kim *et al.*, 1993) with the 50% infective dose being as low as 10^2–10^3 bacteria (DuPont *et al.*, 1972; Levine *et al.*, 1973). *Shigella* spp. are very fatal to humans, and children under the age of five are frequently the most severely infected. Recently, the annual global incidence of infection was estimated to exceed 100 million episodes, with a fatal outcome for ~600,000 individuals (Lindberg *et al.*, 1990; Lindberg and Pal, 1993). Moreover, high proportions of isolates are resistant to conventional antibiotics (Kim *et al.*, 1993; Lee *et al.*, 1996). Efficacious vaccines are thus urgently needed. Several promising candidate vaccines are at various levels of development, but, to-date, no vaccines are applicable for public health purposes. Since *Shigella* spp. are pathogenic only to humans and subhuman primates, *in vivo* testing must necessarily involve one or both of these two species.

Homologous recombination has been detected in a wide variety of organisms, from simple bacteriophages to complex eukaryotic cells. Of these, the RecA enzyme promotes fundamental and probably early steps in the processes leading to cross-over. The *recA* gene performs another equally important role in cell metabolism by controlling the expression of a group of unlinked genes that aid in the recovery of cells after exposure to DNA-damaging agents. This response, termed the SOS response, involves genes that participate in the repair of DNA damage, mutagenesis, and coordination of cell division events (Witkin, 1976).

In order to further the development of an *S. sonnei* vaccine against shigellosis, we are currently attempting to

* To whom correspondence should be addressed.
Tel: 82-431-261-2302; Fax: 82-431-268-2538
E-mail: youngkim@cbucc.chungbuk.ac.kr

make a carrier host of *S. sonnei*. For the construction of a stable vector system containing multi-antigenic genes in *S. sonnei*, we cloned the *recA* gene encoding a recombinase from the chromosomal DNA of the *S. sonnei* KNIH104S strain isolated in Korea. We report the cloning, sequencing, and analysis of the *recA* gene from *S. sonnei* KNIH104S chromosomal DNA.

Materials and Methods

Bacterial strains, plasmid, and culture conditions The bacterial strains and plasmids used or constructed in this study are presented in Table 1. Recombinant plasmid pRAK28 contains the *recA* gene, which is responsible for the recombination pathway or participates in the repair of DNA damage. Recombinant plasmid pRAK2801, pRAK2802, pRAK2803, and pRAK2804 were constructed from pRAK28 as shown in Table 1. All the strains carrying various recombinant plasmids were selected on Luria-Bertani (LB) agar medium containing ampicillin (50 μg/ml), tetracycline (15 μg/ml), isopropyl thio-β-D-galactoside (IPTG), and 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal).

Growth media for cloning Selection for ampicillin-resistant transformants was performed on LB agar containing 50 mg/L ampicillin. $RecA^+$ transformants were selected on LB agar containing 50 mg/L ampicillin and 2% methylmethane sulfonate (MMS).

Preparation and *in vitro* manipulation of plasmid DNA The isolation of plasmid DNA, transformation, restriction endonuclease digestion, ligation, agarose gel electrophoresis, and other standard recombinant DNA techniques were performed as described by Sambrook *et al.* (1989).

Cloning the *recA* gene in *S. sonnei* KNIH104S Chromosomal DNA and pBluescript SK(+) plasmid DNA were restricted with *Kpn*I, using conditions recommended by the manufacturer. The linearized pBluescript SK(+) DNA was treated with bacterial alkaline phosphatase for 60 min at 68°C, extracted with phenol and chloroform, and then precipitated with two volumes of cold (−20°C) absolute ethanol. The restricted chromosomal DNA was similarity treated with phenol and chloroform and precipitated by ethanol. Restricted plasmid and chromosomal DNAs were mixed with 1 to 5 U of T4 DNA ligase, 10 mM dithiothreitol, and 1 mM ATP. Ligation was performed for 12 to 18 h at 12°C. The ligation mixture was used to transform competent HB101 cells, and ampicillin-resistant transformants were selected. Ampicillin-resistant colonies were tested for MMS resistance by replica plating to drug plates. Colonies that grew in the presence of 2% MMS were recloned, and their plasmid DNAs were isolated.

SDS-PAGE Cells were harvested by centrifugation at 12,000 × *g* for 5 min and then suspended in 2× loading buffer. The suspended samples were boiled for 10 min and centrifuged at 12,000 × *g* for 10 min. Purification was achieved by SDS-PAGE using 12% separating and 3% stacking gels. After electrophoresis, the proteins in the gel were stained with Coomassie Brilliant Blue R-250.

Nucleotide sequencing Nucleotide sequences were determined directly from plasmids with an Applied Biosystems automated DNA sequencer (Foster City, USA). The derivatives of 2.8 kb KpnI fragment were generated with various restriction enzymes, and then subcloned into the polycloning site of the pBluescript SK(+) sequencing vector. Plasmid DNAs for the sequencing reaction were purified by standard procedures using a Qiagen Plasmid Kit (Qiagen Co., Chatsworth, USA). Both strands of the cloned DNA was sequenced.

Sequence analysis The sequence of the *recA* gene has been deposited in the GenBank database under the accession number AF101227. The nucleotide sequence and the deduced amino acids sequence were analyzed using DNASIS/PROSIS (Hitachi v. 7.0). Multiple alignments were carried out on computer using the Clustal_W algorithm (Thompson *et al.*, 1997), and then, the alignment was fine-tuned manually.

Table 1. Bacterial strains and plasmids used in this study.

Bacterial strain and plasmid	Relevant characteristics	Source or ref.
Strains		
S. sonnei KNIH104S	Sm^r, Tc^r	Lee *et al.*
E. coli XL1-Blue	*supE44hsdR17recA1endA1gyrA46thirelA1lac*$^-$ *F'[proAB*$^+$ *lacI*^{q} *lacZ*△M15Tn10(*tet*)]	
E. coli HB101	*supE44hsdS20*($r^-_B m^-_B$)*recA13ara-14 proA2 lacY1galK2rpsL20xyl-5mtl-1*	
Plasmids		
pBluescript SK(+)	Ap^r, multiple cloning site in *lacZa*;	Stratagene Co.
pRAK28	2.8 kb *Kpn*I fragment from *S. sonnei* KNIH104S inserted into SK(+)	This study
pRAK2801	Self-ligated large fragment of pRAK28 digested with *Pst*I	This study
pRAK2802	1.8 kb *Pst*I fragment from pRAK28 inserted into SK(+)	This study
pRAK2803	0.5 kb *Eco*RV-*Dra*I fragment from pRAK28 inserted into SK(+)	This study
pRAK2804	0.8 kb *Eco*RV-*Dra*I fragment from pRAK28 inserted into SK(+)	This study

Results and Discussion

Isolation of the *S. sonnei* KNIH104S *recA* gene A plasmid genomic DNA library of *S. sonnei* KNIH104S was constructed using pBluescript SK(+), and then transformed to *E. coli* strain HB101, a strain that carries the *recA* allele and is, therefore, deficient in homologous recombination and DNA repair functions. Since such repair is required for growth in the presence of the alkylating agent MMS, clones that expressed interspecific complementation of the *recA* phenotype were identified by growth on LB plates containing MMS. Of 500 clones screened in this manner, six displayed MMS resistance. The extracted plasmids from these MMS-resistant clones had the same restriction enzyme pattern and one of them was named as pRAK28.

Nucleotide sequence and analysis of the *recA* gene In order to determine the sequence of the *recA* region, subclones were constructed by cloning 1.8-kb *Pst*I (pRAK2802), 0.5-kb *Eco*RV-*Dra*I (pRAK2803), and 0.8-kb *Eco*RV-*Dra*I (pRAK2804) fragments from pRAK28 into a pBluescript SK(+) vector. The restriction maps of the constructed recombinant plasmids are shown in Fig. 1. The entire nucleotide sequence and the deduced amino acids sequence are shown in Fig. 2. The calculated M_r of the RecA polypeptide, based on the deduced sequence of 353 amino acids, is 37,928, which agrees with the estimated value by SDS-polyacrylamide gel electrophoresis (Fig. 3). As shown in Fig. 3, the *E. coli* XL1-Blue harboring pRAK28 produced a 38 kDa protein of RecA in the presence or absent of IPTG. This may be due to the presence of a self-promoter around the cloned *recA* gene.

According to analysis using the Blast program (Altschul *et al.*, 1997), the *recA* gene of *S. sonnei* KNIH104S (GenBank accession No. AF101227) is highly related to those of *E. coli* (GenBank accession No. X55552) and *S. flexneri* (GenBank accession No. X55553), being 99.5% and 99.7% identical, respectively. The amino acids sequences of RecA from *E. coli* and *S. flexneri* are the same and the RecA of *S. sonnei* KNIH104S is different at

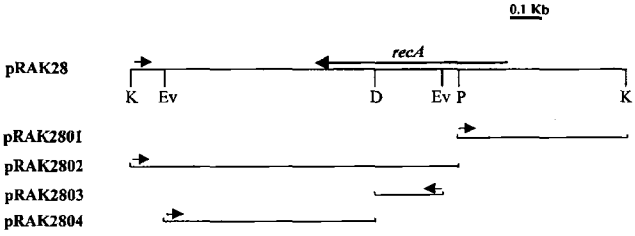

Fig. 1. Physical and genetic map of the cloned plasmid pRAK28 and its derivatives for DNA sequencing. The arrows indicate the direction of promoter in each vector. Cleavage sites for the enzymes are designated as follows: K, *Kpn*I; Ev, *Eco*RV; D, *Dra*I; P, *Pst*I.

```
5'CGACCGTGATGCGGTGCGTCGTCAGGCTACTGCGTATGCATTGCAGACCTTGTGGCAACA   60
ATTTCTACAAAACACTTGATACTGTATGAGTATACAGTATAATTGCTTCAACAGAACATA  120
TTGACTATCCGGTATTACCCCGGCATGACAGGAGTAAAAATGGCTATCGACGAAAACAAA  180
                                recA M A I D E N K

CAGAAAGCGTTGGCGGCAGCACTGGGCCAGATTGAGAAACAATTTGGTAAAGGCTCCATC  240
Q K A L A A A L G Q I E K Q F G K G S I

ATGCGCCTGGGTGAAGACCGTTCCATGGATGTGGAAACCATCTCTACCGGTTCGCTTTCA  300
M R L G E D R S M D V E T I S T G S L S

CTGGATATCGCGCTTGGGGCAGGTGGTCTGCCGATGGGCCGTATCGTCGAAATCTACGGA  360
L D I A L G A G G L P M G R I V E I Y G

CCGGAATCTTCCGGTAAAACCACGCTGACGCTGCAGGTGATCGCCGCAGCGCAGCTTGAA  420
P E S S G K T T L T L Q V I A A A Q L E

GGTAAAACCTGTGCGTTTATCGATGCTGAACACGCGCTGGACCCAATCTACGCACGTAAA  480
G K T C A F I D A E H A L D P I Y A R K

CTGGGCGTCGATATCGATAACCTGCTGTGCTCCCAGCCGGACACCGCGAGCAGGCACTG  540
L G V D I D N L L C S Q P D T G E Q A L

GAAATCTGTGACGCCCTGGCGCGTTCTGGCGCAGTAGACGTTATCGTCGTTGACTCCGTG  600
E I C D A L A R S G A V D V I V V D S V

GCGGCACTGACGCCGAAAGCGGAAATCGAAGGCGAAATCGGCGACTCTCACATGGGCCTT  660
A A L T P K A E I E G E I G D S H M G L

GCGGCACGTATGATGAGCCAGGCGATGCGTAAGCTGGCGGGTAACCTGAAGCAGTCCAAC  720
A A R M M S Q A M R K L A G N L K Q S N

ACGCTGCTGATCTTCATCAACCAGATCCGTATGAAAATTGGTGTGATGTTCGGTAACCCG  780
T L L I F I N Q I R M K I G V M F G N P

GAAACCACTACCGGTGGTAACGCGCTGAAATTCTACGCCTCTGTTCGTCTCGACATCCGT  840
E T T T G G N A L K F Y A S V R L D I R

CGTATCGGCGCGGTGAAAGAGGGCGAAAACGTGGTGGGTAGCGAAACCCGCGTGAAAGTG  900
R I G A V K E G E N V V G S E T R V K V

GTGAAGAACAAAATCGCTGCACCGTTTAAACAGGCTGAATTTCAGATCCTCTACGGCGAA  960
V K N K I A A P F K Q A E F Q I L Y G E

GGTATCAACTTCTACGGCGAACTGGTTGACCTGGGCGTAAAAGAGAAGCTGATCGAGAAA 1020
G I N F Y G E L V D L G V K E K L I E K

GCAGGCGCGTGGTACAGCTACAAAGGTGAGAAGATCGGTCAGGGTAAAGCGAATGCAACT 1080
A G A W Y S Y K G E K I G Q G K A N A T

GCCTGGCTGAAAGATAACCCGGAAACCGCGAAAGAGATCGAGAAGAAAGTACGTGAGTTG 1140
A W L K D N P E T A K E I E K K V R E L

CTGCTGAGCAACCCGAACTCAACGCCGGATTTCTCTGTAGATGACAGCGAAGGCGTAGCA 1200
L L S N P N S T P D F S V D D S E G V A

GAAACTAACGAAGATTTTTAATCGTTCGTTATTGATGATAATAGATAATGTGTGCACGGA 1260
E T N E D F *

CGGTCCCTCGCCCCTTTGGGGAGAGGGTTAGGGTGAGGGGAACAGGCCAACACTGGCGTG 1320

AAAAATTACTCTCACC 3'                                            1336
```

Fig. 2. Nucleotide sequence of the *recA* gene (GenBank accession No. AF101227). The amino acids sequences are also shown in one-letter codes beneath the corresponding codons and the stop codon is marked with an asterisk.

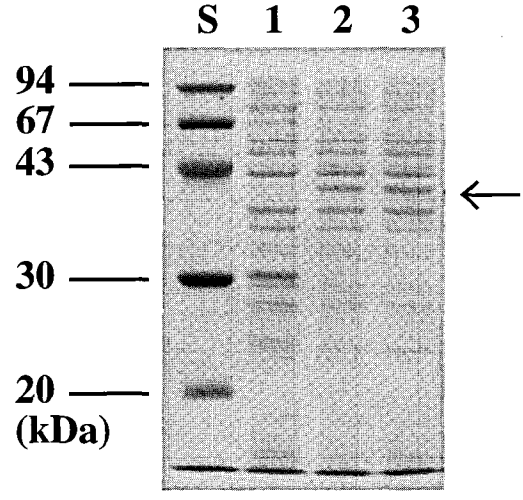

Fig. 3. SDS-PAGE profile of the cell extract of *E. coli* harboring the cloned *recA* gene from from *S. sonnei* KNIH104S. Lane 1, XL1-Blue (pBluescript SK+); lane 2, XL1-Blue (pRAK28); lane 3, XL1-Blue (pRAK28)-induced; lane S, size marker. The arrow indicates the *recA* protein.

only one site; the Arg-86 site of the RecA polypeptide present in both *E. coli* and *S. flexneri* is replaced by Leu in *S. sonnei* KNIH104S. In the Block database (Henikoff *et al.*, 1991), Lys (73%), Arg (15.9%), Ala (7.3%), Gln (2.4%), and Ser (1.2%) were found at this site among the 82 sequenced RecA proteins. This data indicates that the Arg-86 site is not strictly conserved.

For construction of the host carrier of *S. sonnei* vaccine against shigellosis, we cloned the *aroA* gene (Park *et al.*, 1997) for auxotrophs, the *asd* gene (Park *et al.*, 1999) for balanced-lethal vector system (Nakayama *et al.*, 1988), and we constructed genetically-defined *aroA* deletion mutants (Park *et al.*, 1998) of *S. sonnei* KNIH104S. Construction of *recA* deleted mutant of *S. sonnei* KNIH104S for a stable vector system is now on the way.

References

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.

DuPont, H. L., Hornick, R. B., Snyder, M. J., Libonati, J. P., Formal, S. B. and Gangarosa, E. J. (1972) Immunity in shigellosis. II. Protection induced by oral live vaccine or primary infection. *J. Infect. Dis.* **125**, 12–16.

Hale, T. L. (1991) Genetic basis of virulence in *Shigella* species. *Microbiol. Rev.* **55**, 206–224.

Henikoff, S. and Henikoff, J. G (1991) Automated assembly of protein blocks for database searching. *Nucleic Acids Res.* **19**, 6565–6572.

Karnell, A., Bruce, A. D. S., Shigehiro, K., Hanaa, S., Finn, P. R., Phung, D. C., Dang, D. T. and Lindberg, A. A. (1991) An auxotrophic live oral *Shigella flexneri* vaccine: Development and testing. *Rev. Infect. Dis.* **13**, S367–S361.

Kim, H. H., Lee, M. W., Shin, Y. H., Yoo, C. K., Kim, D. S., Kim, M. S., Cho, S. Y. and Kim, D. J. (1993) Epidemiological studies on the bacterial agents of gastrointestinal infection. *The Report of KNIH.* **30**, 3–8.

Lee, K. U., Ju, Y. R., Park, Y. M. A., Lee, K. J., Lee, Y. J., Kim, C. K., Hong, S. K., Park, Y. C., Hwang, S. Y., Lim, J. Y. and Kim, Y. C. (1996) Construction of allelic exchange mutants for a gene replacement of *Shigella sonnei*. *J. Korean Soc. Microbiol.* **31**, 35–43.

Levine, M. M., DuPont, H. L., Formal, S. B., Hornick, R. B., Takeuchi, A., Gangarosa, E. J., Snyder, M. J. and Libonati, J. P. (1973) Pathogenesis of *Shigella dysenteriae* 1 (Shiga) dysentery. *J. Infect. Dis.* **127**, 261–270.

Lindberg, A. A. and Pal, T. (1993) Strategies for development of potential candidate *Shigella* vaccine. *Vaccine* **11**, 168–179.

Lindberg, A. A., Karnell, A., Pal, T., Sweiha, H., Hultenby, K. and Stocker, B. A. (1990) Construction of an auxotrophic *Shigella flexneri* strain for use as a live vaccine. *Microbiol. Pathog.* **8**, 433–440.

Nakayama, K., Kelly, S. M. and Curtiss III, R. (1988) Construction of asd+ expression-cloning vector: Stable maintenance and high level expression of cloned genes in a Salmonella vaccine strain. *Bio/Technology* **6**, 693–697.

Park, Y. C., Shin, H. J. and Kim, Y. C. (1999) Cloning and nucleotide sequence of the *asd* gene from *Shigella sonnei* KNIH104S. *Kor. J. Microbiol.* **35**, 13–17.

Park, Y. C., Shin, H. J., Hong, S. G., Hwang, S. Y., Park, K. D., Ju, Y. R., Lee, K. J. and Kim, Y. C. (1997) Cloning and nucleotide sequence of the *aroA* gene from *Shigella sonnei* KNIH104S. *J. Korean Soc. Microbiol.* **32**, 151–159.

Park, Y. C., Hwang, S. Y., Shin, H. J., Lee, K. J., Ju, Y. R., Bae, Y. S. and Kim, Y. C. (1998) Construction of genetically defined *aroA* deletion mutants of *Shigella sonnei* KNIH104S. *J. Korean Soc. Microbiol.* **33**, 373–380.

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sansonetti, P. J., Kopecko, D. J. and Formal, S. B. (1982) Involvement of a plasmid in the invasive ability of *Shigella flexneri*. *Infect. Immun.* **35**, 852–860.

Taylor, D. N., Bodhidatta, L., Brown, J. E., Echeverria, P., Kunanusont, C., Naigowit, P., Hanchalay, S., Chatkaemorakot, A. and Lindberg, A. A. (1989) Introduction and spread of multiresistant *Shigella dysenteriae* 1 in Thailand. *Am. J. Trop. Med. Hyg.* **40**, 77–85.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876–4882.

Witkin, E. M. (1976) Ultraviolet mutagenesis and inducible DNA repair in *Escherichia coli*. *Bacteriol. Rev.* **40**, 869–907.