Volume 13 Number 20 1985 Nucleic Acids Research

Sequence of the T4 recombination gene, *uvsX*, and its comparison with that of the *recA* gene of *Escherichia coli*

Hisao Fujisawa, Tetsuro Yonesaki and Teiichi Minagawa

Department of Botany, Faculty of Science, Kyoto University, Kyoto 606, Japan

Received 15 August 1985; Accepted 23 September 1985

ABSTRACT

We have determined the nucleotide sequence of the uvsX gene of bacteriophage T4 which is involved in DNA recombination and damage repair, and whose product catalyzes in vitro reactions related to recombination process in analogous manners to E. coli recA gene product. The coding region consisted of 1170 nucleotides directing the synthesis of a polypeptide of 390 amino acids in length with a calculated molecular weight of 43,760. Amino acid composition, the sequence of seven NH_2-terminal amino acids and molecular weight of the protein deduced from the nucleotide sequence were consistent with the data from the analysis of the purified uvsX protein. The nucleotide sequence and the deduced amino acid sequence were compared with those of the recA gene. Although a significant homology was not found in the nucleotide sequences, the amino acid sequences included 23% of identical and 15% of conservatively substituted residues.

INTRODUCTION

The uvsX gene is dispensable for T4 phage multiplication in ordinary laboratory strains of E. coli, but plays important roles in genetic recombination, DNA repair and replication (1-3). We have purified uvsX protein to near homogeneity. The purified gene product catalyzed in vitro reactions related to general recombination (4) as recA protein of E. coli (5-10). These include single stranded DNA dependent hydrolysis of ATP, stimulation of pairing of complementary single strands, D-loop formation and branch migration with a polarity. However, these proteins differ each other in some respects (4); for instance, uvsX protein had no proteolytic activity on E. coli lexA protein, which is cleaved by recA protein at a specific site and controls SOS-function of the cell (11), and there was no immunological cross-reaction between these proteins (4). In addition, electron microscopic observation has revealed that these proteins cooperatively bind to double stranded DNA, but uvsX protein does not extend the DNA as recA protein does (12).

We have determined the nucleotide sequences of the uvsX gene and its

flanking regions, and deduced amino acid sequence of the gene product. We also compared these sequences with those of the recA gene and its product.

MATERIALS AND METHODS

Bacteria, phages and plasmids

E. coli ED8689 (hsdR$^-$, su$^-$) and T4dC mutant (42$^-$ 56$^-$ denB unf) were received from Dr. H. Takahashi (Tokyo Univ.) (13). An amber mutant in uvsX gene (amxb) was described previously (14). Amber mutants in genes 40 (amF12) and 41 (amN81) were from Dr. L. Black (Univ. Maryl. Sch. Med.) and our laboratory stock, respectively. A plasmid, pBR325 (15) was provided by Dr. A. Oka (Res. Fac. Nucl. Acid) and the expression vectors, pNT204 and pNT45 (16) were a gift from Dr. M. Imai (Inst. Virus Res.)

Preparation and manipulation of DNA

Purification of plasmids and M13 RFI DNA, gel electrophoresis and other recombinant DNA techniques were done as described by Maniatis et al. (17). Restriction enzymes were obtained from Takara Co. and used as described by the supplier.

Marker rescue experiment

The method is described in a previous paper (18). Briefly, suspensions of candidate clones were transferred onto plate which had been seeded with a mixture of ED cells and UV-killed amxb.

DNA sequencing

DNA sequencing was done by the dideoxy method of Sanger et al. (19) on fragments cloned into M13mp18 or M13mp19 (20) using Takara's DNA sequencing kit.

Purification of uvsX protein

The purification of uvsX protein has been described (4).

SDS-polyacrylamide gel electrophoresis

Cells in 0.2 ml of culture were pelleted and dissolved in 40 µl of SDS-gel electrophoresis sample buffer and boiled for 2 min. Proteins were separated by electrophoresis on SDS-polyacrylamide slab gels (21) and the gels were stained with Coomassie brilliant blue.

RESULTS AND DISCUSSION

Cloning of T4 uvsX gene

Comparing the T4 physical map (22) and the genetic map around the uvsX gene (2, 23), we assumed the gene would be located on a fragment of about 3kb EcoRI digest. T4dC DNA was digested with EcoRI and electrophoresed on a

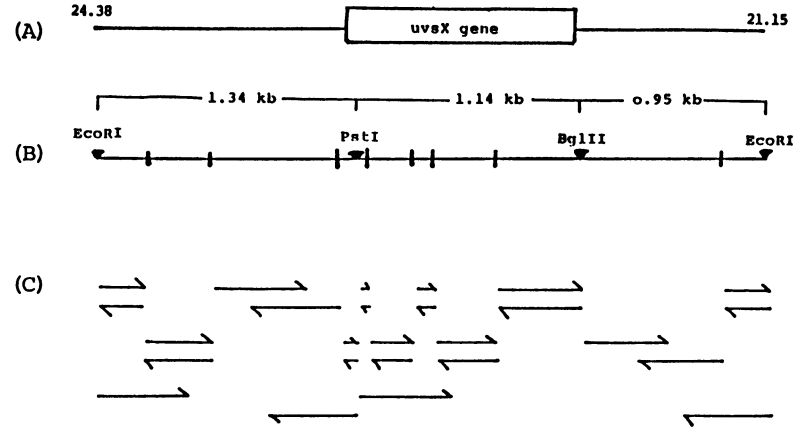

Fig. 1. Cleavage maps and sequencing strategy of the uvsX gene and its flanking regions. (A) The position of uvsX gene (boxed) in the T4 physical map (22). (B) The restriction map of the 3 kb Eco fragment. TaqI sites are indicated by vertical lines. (C) The sequenced fragments. The directions and the extents of the sequence determinations are shown by arrows.

1% agarose gel. The bands of about 3 kb were eluted and cloned into the EcoR1 site on the Cm gene of pBR325. Amp^R Cm^S clones were collected and screened by a marker rescue test with UV-irradiated amxb as described in Materials and Methods. pBSK101, a plasmid thus isolated, contained an EcoRI fragment with an estimated size of 3.43 kb (3 kb Eco fagment). The plasmid also rescued amF12 (40^-) but not amN81 (41^-). Mattson et al. (24) reported another EcoRI fragment which rescued tsL84 (40^-), amN57, amA461 and amH601 (41^-). These observations suggest that an EcoRI site exists in gene 40, which is split , upon EcoRI digestion, into two fragments; the 3kb Eco fragment and the fragment reported by Mattson et al. (24). Fig. 1 shows a physical map of the 3 kb Eco fragment, corresponding to 21.15 through 24.38 units in the T4 map (22).

Confirmation of uvsX gene on the 3kb Eco fragment

We have proved the presence of the entire uvsX gene sequence in the 3 kb Eco fragment by cloning the fragment into the expression vector pNT45, a derivative of pBR322 which carries the major leftward promoter (P_L) of lambda. In bacterial strains carrying pNT204, which harbors the temperature sensitive lambda repressor gene, cI857, temperature dependent expression of genes inserted downstream of the P_L promoter has been shown to occur (16). Amp^R transformants were isolated, followed by screening with marker rescue

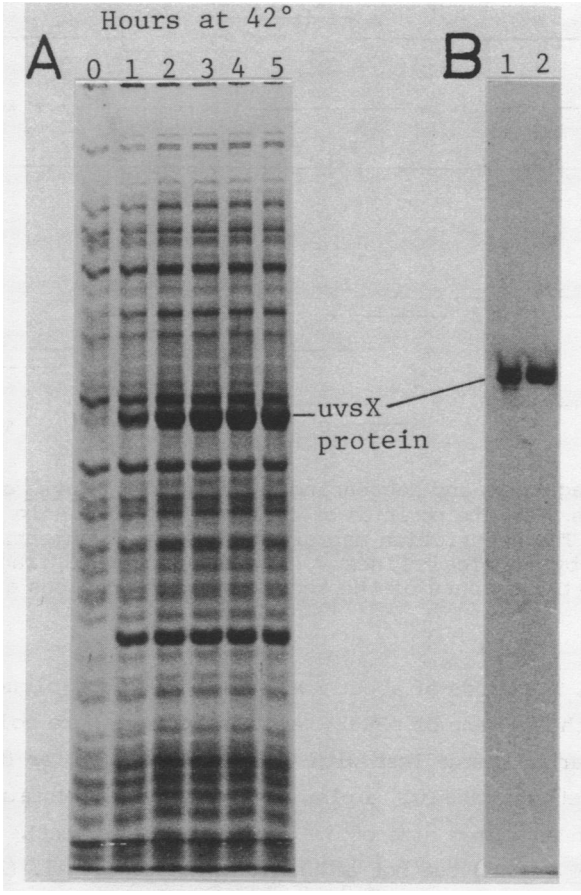

Fig. 2. Overexpression of uvsX gene. (A) Cells were grown at 30°C and shifted to 42°C. Samples were taken at indicated times after the shift, pelleted and dissolved in SDS-sample buffer, followed by electrophoresis in SDS-polyacrylamide gels as described in Materials and Methods. (B) Purified uvsX protein from phage infected cells (lane 1) and induced cells (lane 2) were subjected to SDS-polyacrylamide gel electrophoresis.

of a UV-irradiated T4 uvsX mutant. Several candidate clones carrying the uvsX gene were transferred to 42°C to derepress the P_L promoter. At various times after the temperature shift, cells were collected and total cell proteins were analyzed by SDS-polyacrylamide gel electrophoresis (Fig. 2A). About a half of clones showed a dramatic increase in the amount of proteins in two bands, the larger one of which had a molecular weight identical to that previuosly determined for uvsX protein (4). The overproduced protein was purified to near homogeneity (Fig. 2B). The overall yield was 10 mg from 1 g (wet weight) cells . The purified protein

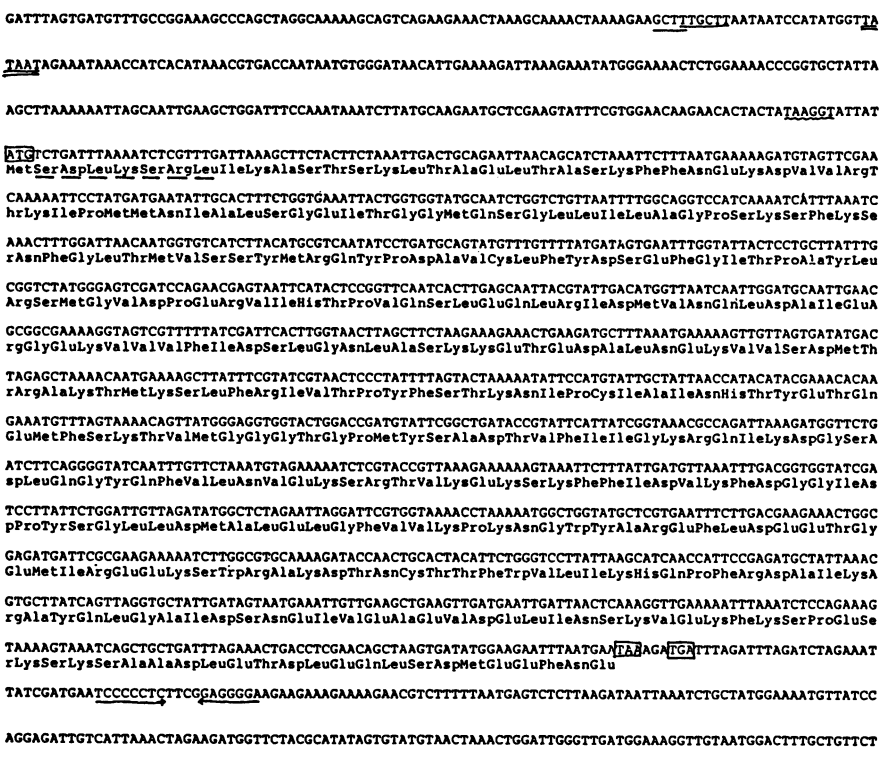

Fig. 3. Nucleotide sequences of the uvsX gene and its flanking regions. The initiation codon (ATG) and two termination codons (TAA, TGA) are boxed. The Shine-Dalgarno sequence is indicated by a wavy line. The putative -10 and -35 promoter sequences are indicated by a double line and an underline, respectively. A palindromic sequence is indicated by arrows. The predicted amino acid sequence is shown under the nucleotide sequence. A dashed underline indicates the amino acid sequence confirmed by protein sequencing.

showed the same specific activity in hydrolysis of ATP to ADP + Pi (T. Shibata, personal communication), D-loop formation and branch migration as uvsX protein purified from phage-infected cells (data not shown).

Nucleotide sequences of the uvsX gene and its flanking regions

The sequencing strategy is shown in Fig. 1C. The 3kb Eco fragment was digested with BglII and PstI, and the resulting three fragments were separated by agarose gel electrophoresis. Each fragment was digested with TaqI and cloned into M13mp18 or M13mp19. DNA sequences of the cloned fragments were determined by the dideoxy method of Sanger et al. (19). Nucleotide sequences of an open reading frame coding for a polypeptide of 390 amino acid residues, corresponding to the uvsX gene as described below,

Table 1 Amino Acid Composition of uvsX protein

Amino acid	Residues observed	predicted
Ala	25	22
Arg	17	16
Asn	-	14
Asp	-	25
Asx	41	39
Cys	-	3
Gln	-	12
Glu	-	33
Glx	49	45
Gly	27	24
His	3	3
Ile	23	25
Leu	31	31
Lys	33	34
Met	15	15
Phe	21	21
Pro	13	13
Ser	35	34
Thr	25	24
Trp	-	3
Tyr	10	11
Val	24	26

and its flanking regions are presented in Fig. 3.

To show that the open reading frame encodes the uvsX gene, the amino acid sequence of the seven residues from the NH_2 terminus of the uvsX protein isolated from T4-infected cells was determined. The sequence was NH_2-Ser-Asp-Leu-Lys-Ser-Arg-Leu, in perfect agreement with the sequence predicted from the nucleotide sequence, except for the absence of the Met residue at the NH_2 terminus. The amino acid composition of the purified uvsX protein is also in good agreement with the composition of the protein predicted from the nucleotide sequence (Table 1). The molecular weight of uvsX protein deduced from the nucleotide sequence is 43,760 in the absence of the initiating Met residue. This is in reasonable agreement with the molecular weight of 39,000 which was estimated by SDS-polyacrylamide gel electrophoresis (4). From these facts, we conclude the sequence codes for uvsX gene.

The initiation ATG codon (boxed) of the uvsX gene is preceeded by a Shine-Dalgarno sequence (wavily lined), a ribosome binding site, at an appropriate distance. Consensus promoter sequences, the putative -35 (underlined) and -10 (doubly lined) sequences were found at about 200 bp upstream from the ATG codon. The putative -35 sequence is consistent to the middle promoter of T4, which requires the mot gene product of T4 for

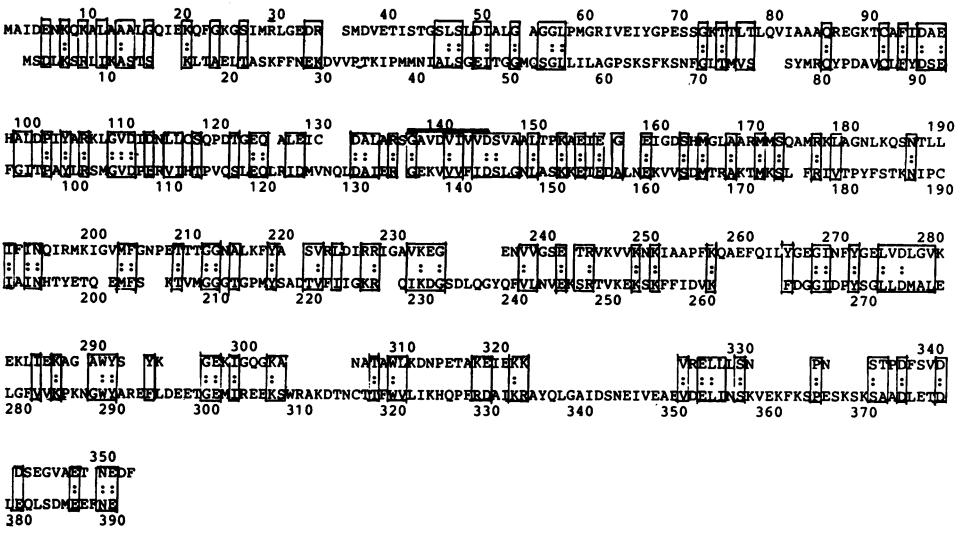

Fig. 4. Amino acid sequences of uvsX and recA proteins. The recA protein sequence (upper line) is described by Horii et al. (32) and Sancar et al. (33); the uvsX protein sequence (lower line) is determined in this paper. Identities(dotted) and conservative substitutions are boxed. The common sequence in ATP binding sites is indicated by a double line.

expression (25). Interestingly, the uvsX gene has two overlapped -35 middle promoter sequences. There is no direct evidence that the synthesis of uvsX protein depends on the mot protein, but kinetic study of the incorporation of ^{3}H-leucine into proteins indicated that the synthesis was not immediate early but delayed early (14). There are two tandemly aligned termination codons, TAA and TGA (boxed), which are followed by a stem- and loop-sequence rich in GC (arrowed).

Structure of uvsX protein

The protein has a Tyr-rich region between residues 76 and 99 where four Tyr residues are occuring at 4 to 10 residue intervals, and its COOH-terminal region (residues 374 to 390) is rich in acidic residues. These features are common among single stranded DNA binding proteins from various sources (26-30). However, it should be noted that unlike other single stranded DNA binding proteins the single stranded DNA-uvsX protein complex is extremely unstable at physiological temperature (unpublished). Walker et al. (31) suggested sequences common to ATP binding sites among ATPases. One of them, Gly-X-X-X-hydrophobic-hydrophobic-hydrophobic-hydrophobic-Asp is present in uvsX protein (residues 135 to 143).

Comparison of the nucleotide and amino acid sequences of uvsX and those of recA.

As mentioned in Introduction, uvsX and recA proteins play a pivotal role in recombination and catalyze several common in vitro reactions. We compared the nucleotide sequences of genes uvsX and recA , the latter having been determined by Horii et al. (32) and Sancar et al. (33). No significant homology in nucleotide sequences was found between these genes. The amino acid sequences of the uvsX and recA proteins are aligned to get good matches as shown in Fig. 3, although any alignment of two proteins will show some accidental homology as pointed out by Doolittle (34). Approximately, 23 % of amino acid residues along the sequences is identical each other and a further 15 % is conservatively (35) substituted. The conserved sequence proposed for the ATP-binding site of ATPases is found at the similar position from the NH_2 terminus, starting at residue 135 of uvsX protein and at residue 137 of recA protein. In addition, the amino acid sequences of the flanking regions of the ATP-binding site are considerably conserved; 50% in 50 residues on the NH_2 terminal side and 53% in 14 residues on the COOH terminal side. The COOH-terminal regions of two proteins are rich in acidic residues as commonly observed in DNA-binding proteins. However, in recA protein there is no Tyr-rich region at the corresponding position of uvsX protein. Further analysis and comparison of the structure and functions of these proteins would provide information necessary to understand the essential structure of recombination protein.

ACKNOWLEDGMENTS

We give special thanks to Dr. K. Wada (Osaka Univ.) for determining the NH_2 terminal amino acid sequence and the amino acid composition of the uvsX protein. We thank Dr. M. Takanami for advices on DNA sequencing, Dr. Y. Shimura for critical reading of the manuscript and Dr. Y. Ryo for valuable discussion. This work was supported by a grant-in-aid for Scientific Research from the Ministry of Education, Science, and Culture of Japan.

REFERENCES

1 Hamlett, N. V. and Berger, H. (1975) Virology 63, 539-567.
2 Cunningham, R. P. and Berger, H. (1977) Virology 80, 67-82.
3 Bernstein, C. and Wallace, S. S. (1983) In Bacteriophage T4, C. K. Mathews, E. M. Kutter, G. Mosig, and P. B. Berget, eds., Amer. Soc. Microbiol. Washington D. C. pp.138-151.
4 Yonesaki, T., Ryo, Y., Minagawa, T., and Takahashi, H. (1985) Eur. J. Biochem. 148, 127-134.

5 Clark, A. J., and Margulies, A. D. (1965) Proc. Natl. Acad. Sci. USA 53, 451-459.
6 Ogawa, T., Wabiko, H., Tsurimoto, T., Horii, T., Masukata, H., and Ogawa, H.(1978) Cold Spring Harbor Symp. Quant. Biol. 43, 909-915.
7 Weinstock, G. M., McEntee, K., and Lehman, I. R. (1979) Proc. Natl. Acad. Sci. USA 76, 126-130.
8 Shibata, T., Das Gupta, C., Cunningham, R. P., and Radding, C. M. (1979) Proc. Natl. Acad. Sci. USA 76, 1638-1642.
9 McEntee, K., Weinstock, G. M., and Lehman, L. R. (1979) Proc. Natl. Acad. Sci. USA 76, 2615-2619.
10 Cox, M. M., and Lehman, I. R. (1981) Proc. Natl. Acad. Sci. USA 78, 3433-3437.
11 Horii, T., Ogawa, T., Nakatani, T., Hase, T., Matsubara, H. and Ogawa, H. (1981) Cell 27, 515-522.
12 Griffith, J. and Formosa, T. (1985) J. Biol. Chem. 260, 4484-4491.
13 Takahashi, H. and Saito, H. (1982) Virology 120, 122-129.
14 Yonesaki, T., Miyazaki, J., and Minagawa, T. (1984) Mem. Fac. Sci. Kyoto Univ. Ser. B. 9, 87-96.
15 Bolivar, F. (1978) Gene 4, 121-136.
16 Shigesada, K., Tsurushita, N., Matsumoto, Y., and Imai, M. (1984) Gene 29, 199-209.
17 Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning, Cold Harbor Lab. Cold Spring Habor N.Y.
18 Minagawa, T., Yonesaki, T., and Fujisawa, H. (1983) Plant. Cell. Physiol. 24, 247-253.
19 Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
20 Norrander, J., Kempe, T., and Messing, J. (1983) Gene 26, 101-106.
21 Laemmli, U. K. (1970) Nature (London) 227, 680-685.
22 Kutter, E., and Ruger, W. (1983) In Bacteriophage T4, C. K. Mathews, E. M. Kutter, G. Mosig, and P. B. Berget, eds. Amer. Soc. Microbiol. Washington D. C. pp.277-290.
23 Childs, J. D. (1983) Mutation Res. 71, 1-14.
24 Mattson, T., Van Houwe, G., Bolle, A., Selzer, G., and Epstein, R. (1977) Molec. gen. Genet. 154, 319-326.
25 Brody, E., Raussay, D., and Hall, D. H. (1983) In Bacteriophage T4, C. K. Mathews, E. M. Kutter, G. Mosig, and P. B. Berget, eds. Amer. Soc. Microbiol. Washington D.C. pp.174-183.
26 Coleman, J., and Oakley, J. (1980) Crit. Rev. Biochem. 7, 247-287.
27 Walker, J. M., Gooderman, K., Hastings, J. R. B., Mays, E., and Johns, E. (1980) FEBS Let. 122,262-270.
28 Williams, K. R., LoPresti, M., Setoguchi, M., and Konigsberg, W. H. (1980) Proc. Natl. Acad. Sci. USA 77, 4614-4617.
29 Dunn, J. J., and Studier, F. W. (1981) J. Mol. Biol. 148, 303-330.
30 Williams, K. R., Spicer, E. K., LoPresyti, M. B., Geggenmeimer, R., and Chase, J. (1983) J. Mol. Biol. 254, 6426-6432.
31 Walker, J. E., Saraste, M., Runswick, M. J., and Gay, M. J. (1982) EMBO. J. 1, 945-951.
32 Horii, T., Ogawa, T., and Ogawa, H. (1980) Proc. Natl. Acad. Sci. USA 77, 313-317.
33 Sancar, A., Stachelek, C., Konigsberg, W. and Rupp, W. D. (1980) Proc. Natl. Acad. Sci. USA 77, 2611-2615.
34 Doolittle, R. F. (1981) Science (Washington) 214, 149-158.
35 Dayhoff, M. O. (1978) in Atlas of Protein Sequence and structure, Vol.5, Suppl. 3, published by National Biomedical Research Foundation, Washington D. C. pp.356-358.