JOURNAL OF BACTERIOLOGY, Nov. 1994, p. 6439–6448
0021-9193/94/$04.00+0

Vol. 176, No. 21

Purification and Characterization of the SegA Protein of Bacteriophage T4, an Endonuclease Related to Proteins Encoded by Group I Introns

MRIDULA SHARMA AND DEBORAH M. HINTON*

Section on Nucleic Acid Biochemistry, Laboratory of Molecular and Cellular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland 20892

Received 13 June 1994/Accepted 18 August 1994

Although not encoded by an intron, the bacteriophage T4 SegA protein shares common amino acid motifs with a family of proteins found within mobile group I introns present in fungi and phage. Each of these intron-encoded proteins is thought to initiate the homing of its own intron by cleaving the intronless DNA at or near the site of insertion. Previously, we have found that SegA also cleaves DNA. In this report, we have purified the SegA protein and characterized this endonuclease activity extensively. SegA protein cleaved circular and linear plasmids, DNA containing unmodified cytosines, and wild-type T4 DNA containing hydroxymethylated, glucosylated cytosines. In all cases, certain sites on the DNA were highly preferred for cleavage, but with increasing protein concentration or time of incubation, cleavage occurred at many sites. SegA cleaving activity was stimulated by the presence of ATP or ATPγS. Sequence analysis of three highly preferred cleavage sites did not reveal a simple consensus sequence, suggesting that even among highly preferred sites, SegA tolerates many different sequences. A T4 *segA* amber mutant that we constructed had no phenotype, and PCR analyses indicated that several T-even-related phages lack the *segA* gene. Taken together, our results show that SegA is an endonuclease with a hierarchy of site specificity, and these results are consistent with the insertion of *segA* DNA into the T4 genome some time after the divergence of the closely related T-even phages.

The bacteriophage T4 *segA* to *segE* genes (similar to endonucleases of group I introns) comprise five open reading frames (ORFs) of unknown function that are highly similar in their first 100 amino acids (51). Although there is no evidence that these ORFs are located within introns, they contain sequence motifs that are common to the GIY-YIG family of proteins encoded by group I introns in bacteriophage and in the mitochondria of *Neurospora crassa*, *Podospora anserina*, and *Saccharomyces douglasii* (7, 11, 13, 37, 51, 54). The I-TevI protein, a member of this family found within the group I intron of the *td* gene of bacteriophage T4, is a site-specific endonuclease which cleaves a *td* allele lacking the intron (8, 9, 45). Double-strand break repair results in the transfer of the intron to the intronless site, converting the gene that lacks an intron to an intron-containing allele (3, 5, 45, 52). This highly efficient process, called homing, has been observed for many group I introns, each of which encodes its own site-specific endonuclease that cleaves its respective intronless DNA at or near the insertion site (for reviews, see references 3, 4, 17, 35, and 44). The GIY-YIG family contains a subset of these intron endonucleases. Other intron-encoded proteins such as the highly characterized I-SceI, present in the intron in the 21S rRNA gene of *Saccharomyces cerevisiae* mitochondria (16, 30, 36), are functionally similar but share different common motifs (10, 15, 17, 23, 55, 57).

In this paper, we report the purification of the SegA protein and our characterization of its endonuclease activity. We find that although some sequences are highly preferred for cleavage, as the concentration of SegA is increased, any DNA is digested. This type of activity is reminiscent of some of the intron endonucleases which show great specificity for cleavage around the intron insertion site but also are very tolerant of base changes within this recognition sequence (6, 18, 58). We have also found that many T-even-related phages are missing the *segA* gene. Likewise, the presence of homing introns and the intron-encoded ORFs varies even among highly related species. These results support the proposed relationship between the *seg* genes and the intron-encoded endonucleases.

* Corresponding author. Mailing address: Building 8, Room 2A-13, National Institutes of Health, Bethesda, MD 20892. Phone: (301) 496-9885. Fax: (301) 402-0240.

MATERIALS AND METHODS

Strains and DNA. Unless otherwise indicated, construction of new plasmids, isolation and purification of DNA, and dideoxy sequencing reactions (49) were performed by previously published procedures (25, 26, 51).

Wild-type T4 DNA and T4 dC-DNA, containing unmodified cytosines, were obtained as previously described (28). A collection of T-even-related phages (T4, T2, T6, RB3, RB15, RB26, RB69, Tu1A, and LZ11 [14, 48, 60]) was obtained from S. Eddy and L. Gold (University of Colorado, Boulder), and the DNA was isolated by the procedure of Quirk et al. (45). Phage lambda DNA was from New England Biolabs. The plasmid pDH428 (27, 28) and the plasmids pRE201 and pMS216, which contain the *segA* gene downstream of a good ribosome binding site (RBS) and either the *lac* promoter of *Escherichia coli* (p_{lac}) or the leftward promoter of lambda (p_L), respectively, have been described (51) and are shown schematically in Fig. 1A. The vectors pTZ18U and pTZ19U were from U. S. Biochemical Corp. pCM7, which has the 1,795-bp *Xho*I-*Xba*I fragment from bacteriophage T4 containing gene 16 and part of gene 17 (T4 map units 94.779 to 96.567 [34]), and its parent vector, pKS (a pBR322 derivative), were the

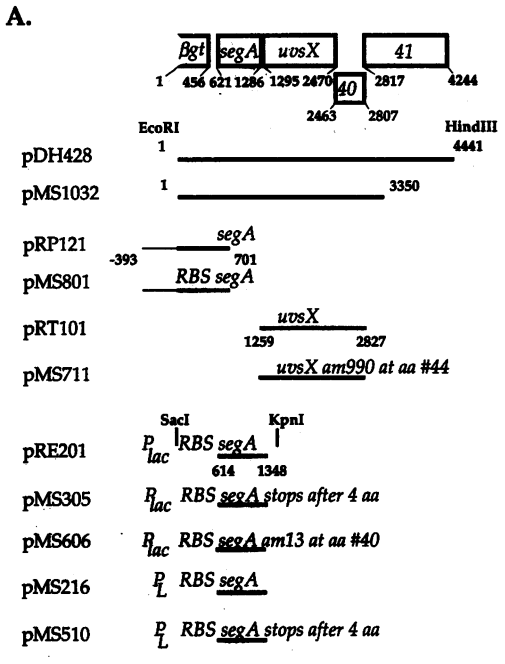

FIG. 1. T4 genome from the end of βgt through gene 41 and summary of plasmids containing inserts from this region. (A) T4 map from the *Eco*RI site at the end of βgt (position 1, T4 map unit 24.526 [34]) to the end of gene 41 (position 4244). The locations of the end of βgt and the genes *segA*, *uvsX*, 40, and 41 are indicated. Plasmids containing T4 DNA from this region are shown underneath. Positive numbers and boldface lines designate the inserted T4 sequences. In pRP121 and pMS801, positions −1 to −393 (designated by the thinner line) come from the 393 bp of plasmid DNA upstream of the T4 insert in pDH428. For pRE201, pMS305, pMS606, pMS216, and pMS510, the promoter, RBS, and restriction sites, *Sac*I and *Kpn*I, present in the vector are also indicated. All plasmids have the pBR322 origin, except pMS1032, which has the origin of pACYC177. (B) Sequences of *segA* and *uvsX* mutations. Relevant portions of the wild-type T4 DNA sequence at the start of the *segA* ORF are shown in regular type; numbers refer to the positions of the sequences relative to the map shown in panel A. Sequence changes and names of mutations are given in boldface type.

gifts of H. Wu and L. Black, University of Maryland School of Medicine, Baltimore. pRP121 (Fig. 1A) was obtained by inserting the 1,094-bp *Xmn*I fragment from pDH428 (from 393 bp upstream of the T4 insert in pDH428 to within the *segA* gene) into the *Sma*I site of pTZ19U. pRT101 (Fig. 1A) was created by inserting the 1,569-bp *Xmn*I fragment from pDH428, which contains the *uvsX* gene, into the *Sma*I site of the vector pTZ18U. In each case, the T4 insert is in the opposite orientation from that of the T7 promoter.

Site-directed mutagenesis, using the method of Kunkel et al. (33) modified by adding the T4 polymerase accessory proteins, was used to construct the following plasmids (changes from the wild-type sequence are given in boldface italics; see Fig. 1B): pMS305 from pRE201 and oligomer X9 (*segA* stops after 4 amino acids (aa); 5′ATAATTATATTTCTTTTCTTTA**AGA**T CTTTTCATGTC3′), pMS606 from pRE201 and oligomer X10 [*segA*(Am13) at aa 40; 5′GTGCTAATAAAGTTCC**CT** ATCCCATATAACC3′], pMS801 from pRP121 and oligomer X18 (good *segA* RBS; 5′GAGAAGTTCATCGAATAG**TAA GGAGG**TAATGACATGAAAAGAC3′), and pMS711 from pRT101 and oligomer X14 [*uvsX*(Am990) at aa 44; 5′TGAA TATTGCACTTT**AG**GGTGAAATTACTGG3′]. The desired changes were confirmed by DNA sequencing. To construct pMS510, the 773-bp *Sac*I-*Kpn*I fragment from pMS305 was inserted into the vector pMS114 (51), which had been digested with *Sac*I and *Kpn*I. pMS510 is thus identical to pMS216, except that the *segA* gene truncates after four aa (Fig. 1B).

pMS1032 (Fig. 1A) was constructed as follows. The 3,017-bp *Bam*HI-*Pst*I fragment was first obtained from pGP1-2, a pACYC177-based plasmid (53), and then ligated in the presence of the following linker:

*Bam*HI *Pvu*II *Mlu*I *Eco*RI *Eco*RV *Pst*I
5′GATCCCAGCTGACGCGTGAATTCGATATCCTGCA3′
3′GTCGACTGCGCACTTAAGCTATAGG5′

The resulting vector, pMS1005, was then digested with *Eco*RI and *Eco*RV and ligated to the 3,917-bp *Eco*RI-*Eco*RV fragment from pDH4426 (27), to give pMS1032. Thus, pMS1032 contains the T4 region from the end of βgt to the beginning of gene 41 (positions 1 to 3350 in Fig. 1A) in a plasmid with the pACYC177 replication origin.

Isolation of T4 mutants. A *segA* amber mutation was recombined back into T4 phage as follows: a mid-log-phase culture of *E. coli* MV1190 (21) containing pMS606 was infected with T4D$^+$ phage (multiplicity of infection, 0.2). After 8 min without aeration, the culture was diluted 50-fold with L broth and then incubated at 37°C for 75 min with aeration. *E. coli* NapIV *supD* (43), which suppresses the amber codon with serine (the wild-type amino acid), was then used as plating bacteria with the phage. Phage plaques were lifted onto nitrocellulose (Schleicher & Schuell), and phage DNA was fixed to the filter, using the procedure from Schleicher and Schuell. The DNA was hybridized to the ^{32}P-labeled oligonucleotide X10, the oligomer used to make pMS606, and the filters were washed by the procedure of Ripley et al. (46). After washing at 68°C, only 13 of approximately 2,000 plaques remained bound to X10. One such phage was designated *segA*(Am13). A similar procedure was used to put a good RBS sequence upstream of *segA*(Am13) [by infecting pMS801 (good RBS)/MV1190 with T4 *segA*(Am13)] and to obtain T4 *uvsX*(Am990) and the double mutant T4 *segA*(Am13) *uvsX*(Am990) {by infecting pMS711 [uvsX(Am990)]/NapIV *supD* with T4D$^+$ and T4 *segA*(Am13), respectively}. However, in the case of pMS801, the nonsuppressing NapIV strain (43) was used as plating bacteria in the isolation of mutants. Sequencing of phage RNA was used to confirm the expected sequence changes in these mutations.

Purification of SegA from cells containing the cloned *segA* gene. (All steps were performed at or below 8°C.) Fifty liters of *E. coli* N4830, containing the temperature-sensitive lambda repressor *c*I857 (1) and pMS216, were grown at 29°C in L broth plus 25 μg of ampicillin per ml. At mid-log phase of growth, the temperature was raised to 42°C for 10 min and then lowered to 37°C. Cells, harvested after 2.5 h, were kept frozen at −80°C until resuspension in sonication buffer (20 mM Tris-HCl [pH 8.0], 1 mM EDTA, 10% glycerol, 1 mM β-mercaptoethanol, and 1 mM phenylmethylsulfonyl fluoride [PMSF]) plus 50 mM NaCl. (Final concentrations varied from

0.1 to 0.2 g of cells per ml.) After sonication on ice, the extract was centrifuged at 7,700 × *g* for 15 min, and then the resulting low-speed supernatant was centrifuged at 123,000 × *g* for 2 h.

The high-speed supernatant (15 ml at 9.2 mg/ml, representing 2.5 g of cells) was mixed with 30 ml of phosphocellulose resin (Whatman P11) equilibrated in sonication buffer plus 50 mM NaCl, and the suspension was rotated for 45 min. The resin, collected after centrifugation, was used to pack a 15-ml column (1.5 by 8.5 cm) which was washed with 35 ml of sonication buffer plus 50 mM NaCl and then 40 ml of sonication buffer plus 0.4 M NaCl at a flow rate of 20 ml/h. Proteins were eluted with a gradient of 0.4 to 1.0 M NaCl in sonication buffer (200 ml total, 5-ml fractions). The peak fraction of SegA, which eluted with 0.52 M NaCl, and the next fraction were pooled (total, 1 mg of protein).

An aliquot of the phosphocellulose fraction (7.5 ml) was diluted to a conductivity equivalent to that of sonication buffer plus 0.2 M NaCl and then applied to a double-stranded DNA cellulose column (2) (0.8 by 1.0 cm; 0.5 ml), previously equilibrated with that buffer. The column was washed with 5 ml of sonication buffer plus 0.2 M NaCl and 5 ml of buffer plus 0.25 M NaCl. A linear gradient from 0.25 to 0.8 M NaCl in sonication buffer (10 ml total in 1 h) was used to elute proteins, and 0.5-ml fractions were collected. The peak of SegA (0.5 ml; 0.23 mg of protein) eluted with 0.36 M NaCl. Although the SegA protein was highly purified after this step (see Fig. 2A), several low-molecular-weight contaminants were still present. A portion (0.45 ml) of this fraction was then diluted to a conductivity equivalent to that of sonication buffer plus 0.2 M NaCl and reapplied to the same double-stranded DNA cellulose column, equilibrated in buffer plus 0.2 M NaCl. The column was washed with 5 ml of buffer plus 0.2 M NaCl and then 5 ml of buffer plus 0.3 M NaCl, and the SegA-containing fraction (2.2 ml, 40 μg/ml) was eluted with buffer plus 0.6 M NaCl.

A portion of the SegA fraction (2.1 ml) was loaded on a Sephacryl S-200 column (Pharmacia; 0.7 by 20 cm, 8 ml) equilibrated in buffer plus 0.2 M NaCl, without the PMSF. One-milliliter fractions were collected at a flow rate of 4 ml/h. SegA protein was obtained in 2 ml at 40 μg/ml. Aliquots were stored at −80°C. The final yield was 50 μg of homogeneous protein per g of cells.

For some of the experiments, a partially purified fraction of SegA and the equivalent fraction from cells containing the plasmid with a truncation of the *segA* gene after 4 aa (pMS510; see above) were used. In this case, the 7,700 × *g* (low-speed) supernatant was obtained as described above. Protein was then obtained after a single round of phosphocellulose chromatography, using a gradient of 0.3 to 2.0 M NaCl in sonication buffer (PP fraction I) or after passing the extract through a DEAE-cellulose column in sonication buffer plus 0.05 M NaCl, followed by phosphocellulose chromatography with a gradient of 0.3 to 2.0 M NaCl in sonication buffer (PP fraction II). These final PP fractions contained 0.2 to 0.3 mg of total protein per ml; the SegA$^+$ fractions were approximately 20% pure.

In vitro endonuclease assay. Unless otherwise indicated, standard endonuclease reactions were performed as follows: supercoiled plasmid DNA or phage DNA was incubated with the indicated proteins for 30 min at 30°C in a reaction mixture containing 20 mM Tris-HCl (pH 8.0), 10% glycerol, 1 mM β-mercaptoethanol, 2 mM ATP, 5 mM $MgCl_2$, and NaCl (if indicated). (Amounts of DNA and proteins are given in the figure legends and are for an 8-μl reaction mixture.) Reaction products were separated on 0.8% agarose gels.

Sequencing of SegA cleavage sites. To determine the major SegA cleavage site in pDH428 and pCM7, the DNA (8 or 5.5 μg, respectively) was incubated with either the SegA$^+$ fraction (PP fraction II) or the equivalent fraction from cells containing pMS510 (240 to 400 ng of total protein) in a reaction mixture (160 μl) as described above. Linearized plasmid was the major product of each reaction after digestion with SegA; no digestion of the plasmid was detected after incubation with the SegA$^-$ control fraction. After phenol extraction and ethanol precipitation, the pDH428 product DNA was sequenced with the primers X16 (5′GCTTATTTCGTATCGTAACTC3′, positions 1815 to 1835 [Fig. 1A]) and X17 (5′CATTTAGAACAAATTGATACCCCTG3′, positions 2024 to 2000 [Fig. 1A]) which anneal upstream and downstream, respectively, of the cleavage site in *uvsX*, and the pCM7 product was sequenced with the primers 17-1 (5′GGCTCAATGTCTGCGGAAG3′) and 17-2 (5′CCACGAACTGCGTCAGGAGAGG3′), which anneal 46 bases upstream and 113 bases downstream, respectively, of the major cleavage site in gene 17.

To ask whether T4 DNA was cut at the *uvsX* site in vitro, a reaction (320 μl) was performed as described above, using 10.8 μg of T4 DNA and SegA (240 ng of PP fraction II) (NaCl concentration, 50 mM). After phenol extraction and ethanol precipitation, the DNA was sequenced by using the primers X16 and X17.

To assay SegA cleavage at the *uvsX* site in vivo, *E. coli* N4830 cells containing pMS216 (*segA*$^+$ Amp^r) or pMS510 (mutant *segA* Amp^r) were transformed with the compatible Kan^r plasmid pMS1032, which contains the SegA cleavage site present in *uvsX* and a pACYC177 replication origin (Fig. 1A). Transformants were grown at 29°C to mid-log phase of growth in 25 ml of L broth with 25 μg of ampicillin per ml and 40 μg of kanamycin per ml. Cultures were induced at 42°C for 10 minutes and then incubated at 37°C for 2 h. Plasmid DNA was isolated from 5 ml of culture as previously described (12), and the region at the *uvsX* site was sequenced by using the primers X16 and X17.

Screening T-even genomes for the *segA* gene. Genomic DNAs from a collection of T-even-related phages were screened for the presence of the *segA* gene by hybridization to ^{32}P-labeled DNA oligonucleotides (containing sequences from 756 to 726, from 846 to 825, or from 999 to 975) or to a ^{32}P-labeled RNA probe containing most of the *segA* gene (sequences from 1258 to 702). (Sequence positions are as designated in Fig. 1). The presence of the *segA* gene was also assayed by PCR analysis. Primers derived from the end of the T4 gene 42 (42B, 5′TGGTTGATCATTGGTGGA3′) and the beginning of the T4 *uvsX* gene (X23, 5′GTGCAATATTCATCATAGG3′) (500 ng of each primer) were mixed with heat-denatured phage DNA (100 to 200 ng) in a reaction mixture (50 μl) containing 10 mM Tris · HCl (pH 8.3), 50 mM KCl, 3.4 mM $MgCl_2$, 0.001% gelatin, a 200 μM concentration of each deoxynucleoside triphosphate, and 1.25 U of ampli-*Taq* polymerase (Perkin-Elmer Cetus). The mixture was then subjected to 35 cycles of 95°C for 1.5 min, 40 or 55°C for 1.5 min, and 70°C for 3 min. Aliquots of 9 μl were used for electrophoresis on 0.8% agarose gels.

RESULTS

Purification of the bacteriophage T4 SegA protein and characterization of the SegA nuclease activity. We have previously cloned the T4 *segA* ORF downstream of a good RBS and the strong leftward promoter of lambda, p_L, and showed that heat induction of *E. coli* N4830 (*c*I857), harboring this SegA$^+$ plasmid, pMS216, results in a high level of the 25-kDa SegA protein (51). Because the biological function of SegA is

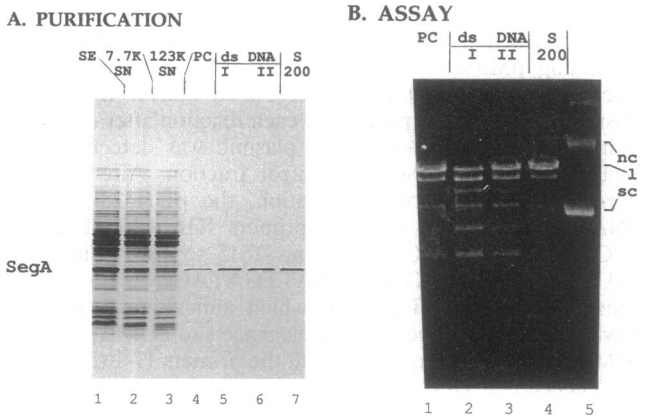

FIG. 2. Purification of the T4 SegA protein and assays of its endonuclease activity. (A) Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of protein fractions from cultures of *E. coli* N4830/pMS216 as the SegA protein was purified (see Materials and Methods). Lane 1, sonicated extract (SE; 8 μg of protein); lane 2, 7,700 × *g* supernatant (7.7K SN; 8 μg); lane 3, 123,000 × *g* supernatant (123K SN; 7 μg); lane 4, phosphocellulose fraction (PC; 0.6 μg); lane 5, ds DNA cellulose fraction I (ds DNA I, 0.5 μg); lane 6, ds DNA cellulose fraction II (ds DNA II, 0.3 μg); and lane 7, S-200 Sephacryl fraction (S 200; 0.3 μg). (B) Agarose (0.8%) gel of SegA reaction products. pDH428 supercoiled DNA (50 ng, lanes 1 to 4) was incubated with 10 ng (lane 1), 8 ng (lane 2), or 4 ng (lanes 3 and 4) of the indicated fractions, keeping the final NaCl concentration at 70 mM. Lane 5, control without protein. Locations of pDH428 supercoiled DNA (sc), linear DNA (l), or nicked circle DNA (nc) are shown.

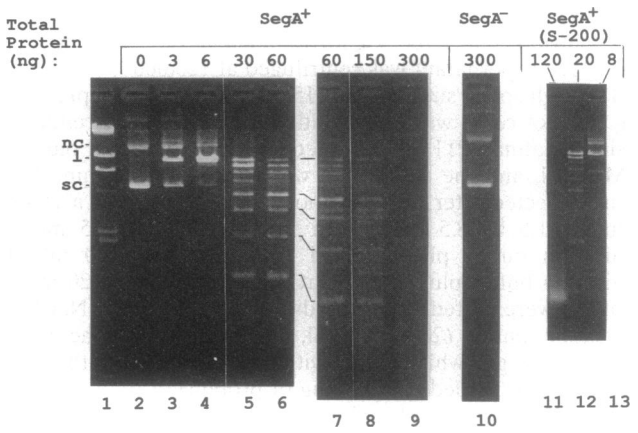

FIG. 3. The pattern of SegA digestion of pDH428 shows a hierarchy of cleavage sites. pDH428 (200 ng, lanes 3 to 10; 50 ng, lanes 11 to 13) was incubated with the indicated amount of the partially purified SegA$^+$ protein fraction (PP II, lanes 3 to 9) or the SegA$^-$ control (lane 10) or with highly purified SegA (S-200 fraction, lanes 11 to 13) as described in Materials and Methods, with 75 to 80 mM NaCl and the products separated on a 0.8% agarose gel. Lane 1, lambda DNA restricted with *Hin*dIII; fragments of 23, 9.4, 6.6, 4.4, 2.3, and 2.0 kbp can be seen. The positions of pDH428 nicked circle (nc), linear (l), and supercoiled (sc) DNAs are shown. Lines between lanes 6 and 7 indicate equivalent bands.

unclear (see below), we took two approaches in characterizing the activities of the protein in vitro. First, we purified SegA protein by using phosphocellulose, double-stranded DNA cellulose, and S-200 chromatography (see Materials and Methods) (Fig. 2A) and characterized the cleaving activity of the purified protein. Second, we compared the activity of a partially purified fraction of SegA expressed by pMS216/N4830 with that of an equivalent fraction from N4830 cells containing pMS510. This plasmid is identical to pMS216, except for a CA-to-TCT change which results in the truncation of SegA after 4 aa (see Materials and Methods) (Fig. 1).

The SegA protein was assayed for its activity on DNA, using both plasmids and phage genomes (see below) as substrates. SegA cleaved the plasmid pDH428 (Fig. 1A) (28), yielding linearized plasmid and smaller products. This activity was observed in a partially purified fraction containing the SegA protein (Fig. 3, lanes 3 to 9), but not in the corresponding fraction from cells with the SegA$^-$ plasmid pMS510 (Fig. 3, lane 10). The same endonuclease activity was observed as the protein was purified (Fig. 2B, lanes 1 to 4). For this set of reactions, equivalent amounts of SegA, as judged by protein gel analysis, were added to each reaction mixture, indicating that the SegA activity remained relatively constant through extensive purification of the protein.

Our characterization of the SegA endonuclease revealed that it has a hierarchy of site specificity and depending on conditions is highly specific or general in its cleavage of DNA. Low levels of the protein or short reaction times yielded primarily nicked circle and linearized plasmid (Fig. 3, lanes 3, 4, and 13; Fig. 4A, lane 5). This linearization occurred at a specific site in the plasmid, and we detail below our analysis of the sequence at this and two other preferred cut sites. However, as the amount of protein or reaction time was increased, different products were observed, ranging from a set of discrete, linear DNAs to smaller, less discrete species to a smear of small DNAs (Fig. 3, lanes 5 to 9, 11, and 12). Although these various reaction products could have arisen from a contaminating nuclease acting after SegA-generated linearization of the plasmid, our results suggest that all of the reaction products were generated by SegA. First, whereas incubation of SegA with either supercoiled pDH428 or linear DNA (see below) gave this pattern of digestion, our control fraction from pMS510 did not digest either DNA (Fig. 3, lane 10; and see below). Furthermore, incubation of pDH428 that had been previously linearized with SegA with the control fraction from pMS510 did not result in further digestion (data not shown). Thus, the products generated by SegA were not substrates for subsequent degradation, at least by a nuclease present in the SegA$^-$ control fraction. In addition, the same reaction products were observed when using the purified protein (Fig. 2B, lanes 1 to 4; Fig. 3, lanes 11 to 13). Finally, our preparation of SegA did not have any detectable single-stranded DNA nuclease activity (data not shown), and the SegA-generated products could be ligated in the presence of T4 DNA ligase (Fig. 4B). This argues that the compatible ends remaining after SegA digestion were not destroyed by a nonspecific nuclease. Taken together, these results suggest that all the cleavage activity we observed was due to SegA and that SegA has many possible cleavage sites, but some sites are highly preferred over others.

Our previous work indicated that cleavage by SegA was absolutely dependent on Mg^{2+} and was stimulated by the presence of ATP (51). As seen in Fig. 4A, increasing concentrations of either ATP or ATPγS resulted in increased cleavage, and the amount of stimulation was similar in both cases (compare lanes 2 to 4 with lanes 6 to 8). Furthermore, no hydrolysis of [γ-^{32}P]ATP was observed in a reaction in which pDH428 was digested (data not shown). These results indicate that hydrolysis of ATP is not required for the stimulatory effect. Stimulation of SegA protein by ATP was also seen when the purified protein was used (Fig. 4A, lanes 9 and 10), strongly

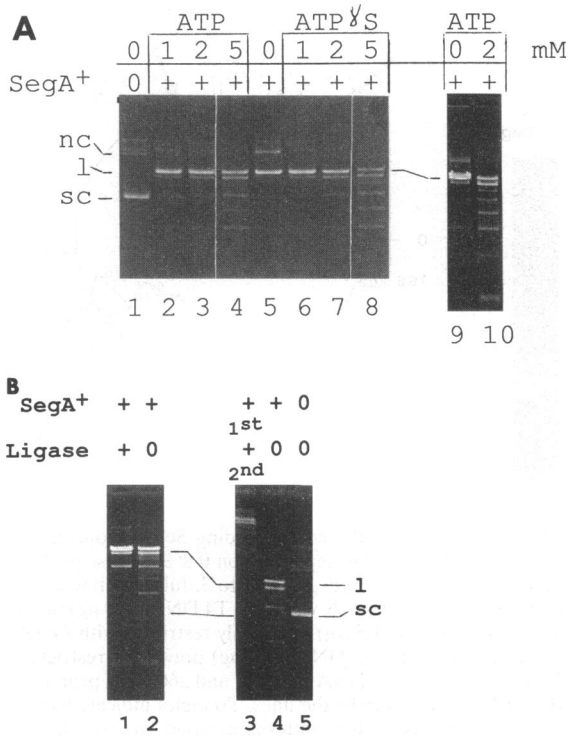

FIG. 4. Characterization of the SegA cleavage reaction in the presence of ATP, ATPγS, and T4 DNA ligase. (A) Stimulation of SegA cleavage by ATP and ATPγS. pDH428 (40 ng, lanes 2 to 8; 50 ng, lanes 9 and 10) was incubated with SegA (9 ng of PP I fraction, lanes 2 to 8, or 20 ng of S-200 fraction, lanes 9 and 10) in the presence of the indicated concentration of ATP (lanes 2 to 4, 9 and 10) or ATPγS (lanes 6 to 8) or without either (lane 5), as described in Materials and Methods. Reaction mixtures for lanes 9 and 10 had 75 mM NaCl. Lane 1, control reaction that lacked protein. The locations of nicked circle (nc), linear (l), and supercoiled (sc) plasmids are indicated. (B) SegA cleavage products can be ligated by T4 DNA ligase. Lanes 1 and 2, pDH428 (200 ng) digested with SegA (10 ng of PP fraction II) in the presence (lane 1) or absence (lane 2) of 0.01 Weiss unit of T4 DNA ligase. Lanes 3 and 4, pDH428 (160 ng) digested with SegA (40 ng of PP fraction I) in a 32-μl reaction mixture and one-fourth of the reaction mixture was then either loaded on the gel (lane 4) or incubated with 5 Weiss units of T4 DNA ligase for 20 min at 30°C before being loaded (lane 3). Lane 5, control reaction lacking any protein. The positions of linear (l) and supercoiled (sc) pDH428 are shown.

suggesting that ATP acts on the SegA protein itself rather than indirectly through another protein in the preparation. ATP appears simply to stimulate SegA cutting; no new cut sites are seen in the presence of ATP that are not detected with a greater amount of SegA (data not shown). Likewise, although SegA activity was inhibited as the NaCl concentration was increased above 100 mM, the salt concentration also did not affect the hierarchy of cut sites (i.e., it did not inhibit a star activity as is seen with some restriction enzymes) (data not shown).

Incubation of pDH428 with SegA in the presence of T4 DNA ligase resulted in less cleavage (Fig. 4B, lanes 1 and 2), and addition of ligase after digestion of pDH428 by SegA generated larger DNA species (Fig. 4B, lanes 3 and 4). These results indicated that SegA generates compatible 3′-hydroxyl and 5′-phosphate ends.

SegA protein cleaved wild-type T4 and T2 DNA (containing hydroxymethylated, glucosylated cytosines), T4 dC-DNA (containing unmodified cytosines), and lambda DNA (also unmodified), again with a pattern that was consistent with a hierarchy of specificity. Digestion of wild-type T4 DNA under conditions that gave limited digestion of pDH428 yielded discrete products that could be seen above a background smear (Fig. 5A, lane 3). This result was seen more clearly by using wild-type T4 DNA that had been previously digested with *Nde*I (lane 6) or T4 dC-DNA that had been previously digested with *Eco*RI (lane 9) or *Bam*HI plus *Pvu*I (lanes 11 to 13). A major cleavage site in the T4 DNA could be mapped by analyzing the products from these substrates. After limited incubation with SegA, the level of the doublet of *Nde*I fragments at 9.2 kbp decreased and a new fragment of 8.5 kbp was seen (marked by triangles in lane 6), while the level of the largest *Eco*RI fragment of 11.4 kbp decreased with the appearance of a fragment of 7.6 kbp (marked by triangles in lane 9). Taken together, these results position a major SegA cleavage site in the T4 DNA at approximately 50 map units on the T4 genome (Fig. 5B). In addition, the SegA-generated DNA fragment of 6.5 kbp, observed by using the T4 DNA digested with *Bam*HI plus *Pvu*I (marked by the arrow in Fig. 5A, lanes 11 to 13), was consistent with the SegA cleavage site we have mapped within the T4 gene *uvsX* (see below). Thus, there are preferred sites for SegA cleavage in T4. However, as with pDH428, incubation with high levels of SegA results in extensive digestion of the T4 genome (Fig. 5A, lane 15).

Sequence analysis of three highly preferred SegA cleavage sites. Previously, we had used restriction analysis to estimate the position of the major cleavage site in pDH428 to the middle of the *uvsX* gene (51). To determine the sequence at the cut, we performed dideoxy sequencing reactions with primers that annealed upstream and downstream of the deduced site (see Materials and Methods). As marked by arrows in Fig. 6A (left panel), two strong stops for Sequenase were seen upon extension of the primers annealed to pDH428 DNA that had been previously incubated with SegA. (Extension past these points was expected in this experiment because some uncut pDH428 still remained after the limited SegA digestion.) This analysis mapped two SegA cleavage sites, consistent with two cut sites that each have a single base 3′ overhang (Fig. 6B). A similar analysis with the products of limited SegA digestion of wild-type T4 DNA (digestion products shown in Fig. 5A, lane 3) indicated that the same cleavage sites were generated with this substrate (Fig. 6A, middle panel).

Because no physiological role has been determined for SegA (see below), we had no genetic evidence to confirm that SegA has an in vivo function consistent with the endonuclease activity we observed in vitro. To ask whether the protein acts as an endonuclease when present in the cell, we transformed cells with both a target plasmid, pMS1032 (bearing T4 sequences from the end of β*gt* to within gene 41 and the pACYC177 origin [Fig. 1A]), and either pMS216, which expresses high levels of SegA, or the control plasmid pMS510. Primer extension analysis with the isolated plasmid DNA revealed that in the cells expressing SegA, but not in the SegA$^-$ control cells, pMS1032 was cleaved within the *uvsX* gene at the same site observed in vitro (Fig. 6A, right panel). Thus, SegA is active as an endonuclease in vivo and it exhibits the same specificity at the *uvsX* site. However, as can be seen from the sequencing gel, the steady-state level of DNA cut at this site was low, perhaps reflecting the ability of DNA ligase to reseal nicked circular product.

Another major cleavage site of SegA was found by using pCM7, a plasmid containing the T4 gene 16-17 region. Limited incubation with SegA linearized pCM7, but under these con-

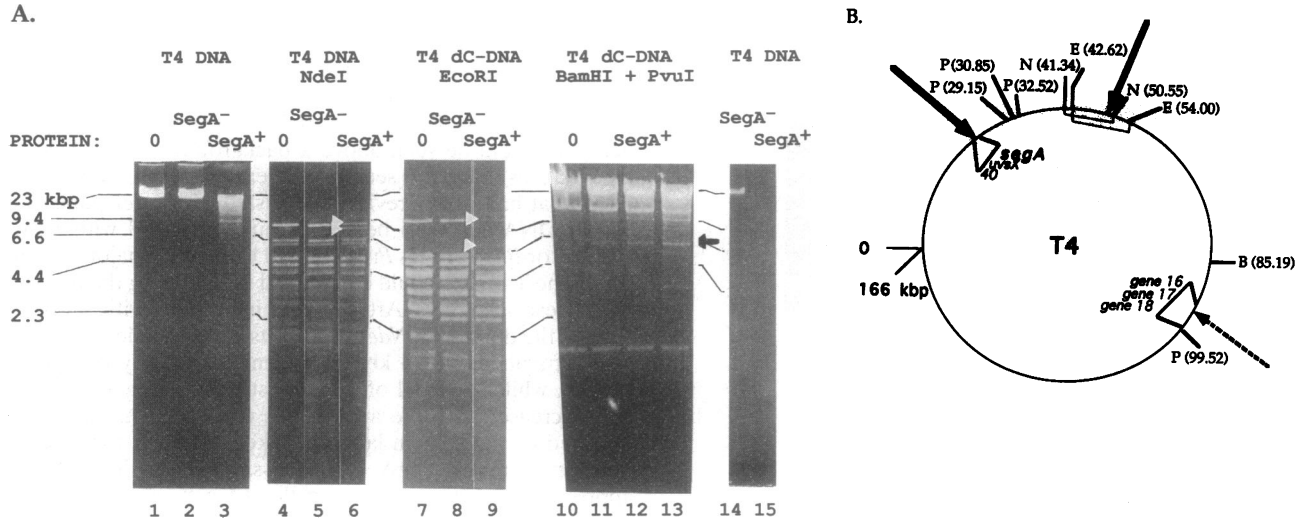

FIG. 5. SegA digestion of T4 DNA. (A) DNA was incubated with SegA protein (SegA$^+$), with the corresponding SegA$^-$ control fraction (SegA$^-$), or in the absence of protein (0) as described in Materials and Methods, and the products were separated on 0.8% agarose gels. SegA$^+$ reactions were performed with the PP II fraction, except for lane 15, in which the PP I fraction was used. Lanes 1 to 3, full-length wild-type T4 DNA (270 ng) and 0 ng (lane 1) or 6 ng (lanes 2 and 3) of protein (NaCl concentration, 50 mM); lanes 4 to 6, wild-type T4 DNA (225 ng) previously restricted with *Nde*I and 0 ng (lane 4) or 6 ng (lanes 5 and 6) of protein; lane 7 to 9, T4 dC-DNA (1.5 μg) previously restricted with *Eco*RI and 0 ng (lane 7) or 64 ng (lanes 8 and 9) of protein (NaCl concentration, 40 mM); lanes 10 to 13, T4 dC-DNA (500 ng) previously restricted with *Bam*HI and *Pvu*I and 0, 3, 6, or 10 ng of protein, respectively; lanes 14 and 15, full-length wild-type T4 DNA (180 ng) and 260 ng of protein (NaCl concentration, 67.5 mM; incubation time, 15 min). The locations of lambda *Hin*dIII markers are shown by the lines. Triangles indicate bands used to map the SegA cleavage site at 50 map units; the thick arrow denotes SegA-generated DNA consistent with the cleavage site within the *uvsX* gene (see text). (B) Map of T4 genome. Although T4 DNA is linear, it is shown as a circle because it is circularly permuted and terminally redundant. The locations and map units of the *Bam*HI (B) and *Pvu*I (P) sites and of the large *Eco*RI (E) and *Nde*I (N) fragments, which are cleaved by SegA, are indicated. Major cleavage sites for SegA within *uvsX* and at map unit 50 are denoted by the solid arrows. The cleavage site within gene 17 (determined from analysis of plasmid DNA containing this region of T4) is shown by the dashed arrow (see text for details).

ditions, the vector DNA was much less reactive (Fig. 7A). Sequence analysis revealed two cleavage sites, designated site gene 17 and site gene 17′, which lie within 35 bp of each other in gene 17 and give approximately 75 and 25% cleavage, respectively (Fig. 7B).

Comparison of the three deduced SegA cleavage sites revealed a 26 of 55 base match between the gene 17 and *uvsX* sites and a 21 of 55 base match between the gene 17 and 17′ sites (shown with the bars in Fig. 7B). However, there are only 11 matches among all three sites (bold bars in Fig. 7B). This analysis suggests either a relaxed consensus sequence or perhaps similarities from structure rather than sequence. In fact, the sequence **GAANTGNT**, which was cleaved in the sites *uvsX* and gene 17 (shown in boldface italics in Fig. 7B) is also found in site gene 17′ but was not cleaved. Thus, a preferred SegA-cut site cannot be predicted simply by a sequence match to these deduced sites.

Search for a biological role for *segA*. The T4 *segA* gene lies in a region of the T4 genome which has not been identified genetically, and to our knowledge, no mutations in *segA* have ever been isolated. We constructed a T4 *segA* mutant (Am13) with an amber mutation at aa 40 of the *segA* ORF (Fig. 1B). We found that T4 Am13 and its wild-type T4 parent plated with similar efficiencies on *E. coli* NapIV relative to NapIV *supD*, which suppresses with the wild-type amino acid, serine. This result suggests that under normal growth conditions, this *segA* mutant has no phenotype.

The T4 *uvsX* gene is a phage analog of *E. coli recA* and is needed for normal phage recombination, repair, and replication (20, 27, 31, 32, 38, 39). Because in some cases T4 genes that are involved in the same pathway are clustered together on the phage genome, we asked whether *segA*, like *uvsX*, was needed for UV repair. The efficiency of plating of T4 *segA*$^+$ or T4 *segA*(Am13) on nonsuppressing and suppressing strains was assayed after UV irradiation by John W. Drake (National Institute for Environmental Health Sciences, Research Triangle Park, N.C.). Survival curves for each phage in each strain were identical (data not shown). We also constructed a T4 phage with mutations in both *segA* and *uvsX* [T4 *segA*(Am13) *uvsX*(Am990)] and found that the efficiency of plating on NapIV relative to NapIV *supD* was identical to that of T4 containing the *uvsX*(Am990) mutation alone. Thus, we found no evidence to suggest that *segA* is involved in phage recombination or repair.

During infection, the *segA* gene is efficiently transcribed prereplicatively (24). However, because the region immediately upstream of the *segA* ORF shows little, if any, match to the consensus RBS sequence (Fig. 1B), we expect the level of SegA protein during T4 infection to be very low. To ask how higher levels of SegA would affect phage growth, we plated T4 *segA*$^+$ on cells containing the SegA$^+$ plasmid, pRE201 (Fig. 1A), the control plasmid pMS305 (SegA$^-$ [Fig. 1A]), or the vector. Titers were similar and did not differ in the presence or absence of isopropyl-β-D-thiogalactopyranoside (IPTG), an inducer for the p_{lac} promoter present on the plasmids. We also constructed a T4 phage containing the good RBS present in pMS216 upstream of the *segA*(Am13) mutation. We found that the efficiency of plating of this phage on NapIV relative to NapIV *supD* is similar to that of T4 *segA*(Am13) alone and to that of wild-type T4. Thus, under normal growth conditions, increased expression of *segA* did not appear to be either beneficial or detrimental to the phage.

Screening T-even-related phage for the *segA* gene. To screen for *segA* DNA in various T-even phage genomes, we hybridized

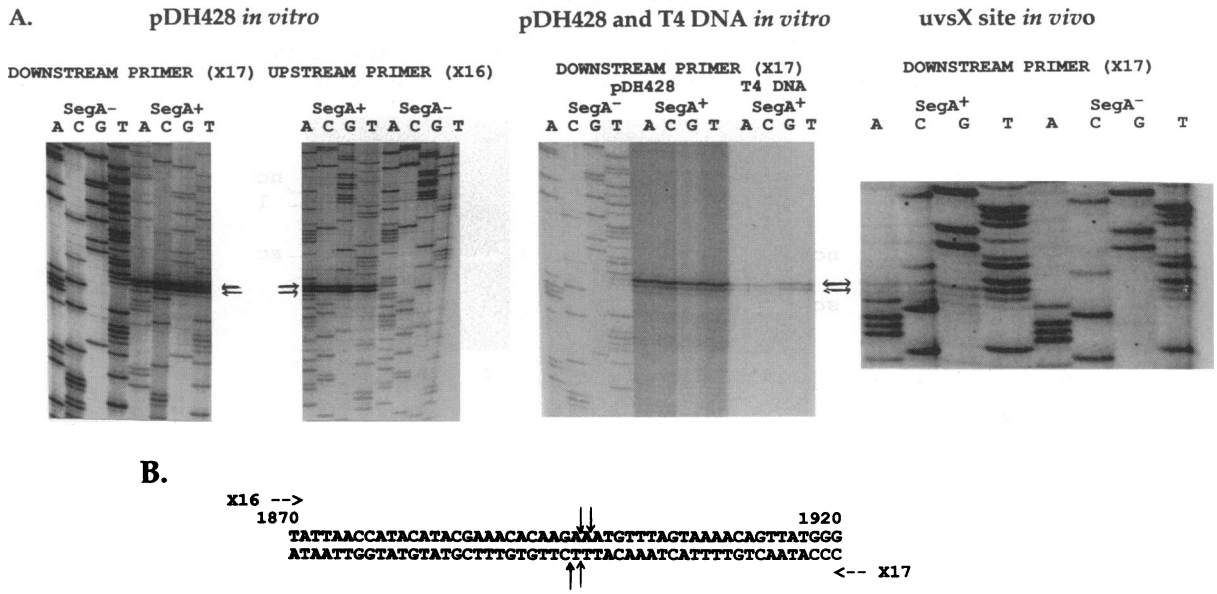

FIG. 6. Determination of the sequence of the SegA cleavage site within the *uvsX* gene in vitro and in vivo. (A) SegA cleavage at the *uvsX* site. Sequencing gels were obtained after extension of the downstream primer X17 or the upstream primer X16. (Left and middle panels) pDH428 or T4 DNA that had been digested in vitro with SegA$^+$ or the SegA$^-$ control fraction; (right panel) pMS1032 DNA (which contains *uvsX* [Fig. 1A]) isolated from cells expressing SegA (*E. coli* N4830/pMS216/pMS1032) or from SegA$^-$ control cells (*E. coli* N4830/pMS510/pMS1032) (see Materials and Methods for details). Arrows indicate the positions of SegA cleavage. (B) Sequence of DNA surrounding the SegA cleavage site in *uvsX* starting at position 1870 (*uvsX* ORF starts at 1295 [Fig. 1A]). Solid arrows indicate determined cut sites; dashed arrows show the directions of the primers used for sequencing.

the genomic DNAs of T4, T2, T6, RB3, RB15, RB26, RB69, LZ11, and Tu1A to ^{32}P-labeled oligonucleotides containing various portions of the *segA* sequence and to a ^{32}P-labeled RNA probe containing most of the *segA* gene (see Materials and Methods). We were unable to detect *segA* DNA in any phage genome other than that of T4. We also used PCR analysis to determine the distribution of the *segA* gene among phages that are related to T4. We selected oligomers that would anneal to the T4 DNA at the end of gene 42 and at the beginning of *uvsX*, shown schematically in Fig. 8A. (Both gene 42 and *uvsX* are highly conserved among T2, T4, and T6.) After amplification of this region with T4 DNA, a species of the expected size of 2.1 kbp was observed (Fig. 8B). However, with T2 DNA, a smaller DNA of 1.1 kbp was seen. Recent DNA sequence analysis indicates that T2 differs from T4 in this region; in T2, both the *segA* gene and the upstream β*gt* gene are missing, and a different gene, β-*glucosyl-HMC-α-glucosyl-transferase* is located here (59). Interestingly, although the T2 β-glucosyl-HMC-α-glucosyl transferase glucosylates DNA, it is not related by sequence to the T4 β*gt* ORF (59). Thus, the *segA* gene is missing in T2 DNA, but its absence is not a simple deletion of just this gene. An analysis of the PCR products obtained with several other T-even phage genomes (Fig. 8B) indicates that while some phage are similar to T2 and some yield a slightly smaller product, none appears to be like T4.

DISCUSSION

Although no genetic locus has ever been identified between the bacteriophage T4 *bgt* and *uvsX* genes, our DNA sequence analyses have revealed an ORF, designated *segA* (19, 51). To investigate the role of *segA*, we purified and characterized the SegA protein, and we generated a T4 *segA*(Am) mutant. Our biochemical analyses showed that SegA is a Mg^{2+}-dependent DNA endonuclease with the very interesting property that it recognizes a hierarchy of sites; certain sites are highly preferred, but with increasing protein concentration or time, cleavage will occur at many locations. Thus, depending on the assay, the protein can appear highly specific or very relaxed in its choice of sites. An unusual feature of this endonuclease is that its activity increases with increasing concentrations of ATP. By affecting the activity of the protein, the level of ATP influences whether the SegA cleavage is specific or relaxed.

The ability of SegA to cleave DNA at very specific sites indicates that it must recognize some particular sequence or structure of the DNA. However, our sequence analysis of three highly preferred sites suggests that even under conditions in which SegA cleavage is highly specific, no simple consensus sequence, at least in the immediate region of the cut site, is recognized by the protein. These results are consistent with a protein that can recognize a sequence that has more than one particular base at any of a number of positions or a protein whose cleavage site is far removed from its recognition site.

What is the role of the SegA endonuclease during T4 infection? Although T4 encodes multiple nucleases that digest the host chromosome (56), the activity of SegA on both unmodified and modified phage DNA would seem to eliminate such a function for this gene. Our results also suggest that *segA* is not an essential gene, and it is not required for UV repair. However, because the first 100 aa of the SegA protein are very similar to those of four other T4 ORFs, *segB* to *segE*, which have been identified only by sequence (51), our screening of the single *segA* mutant would have missed a phenotype if any one of the *seg* genes is sufficient for a particular function.

An intriguing possibility is that the *segA* gene does not have any biological role for the phage, but is simply descended from the insertion of mobile, endonuclease-encoding DNA into the T4 genome. We have previously shown that although none of the five T4 *seg* genes appears to be in an intron, they share amino acid motifs found in the GIY-YIG family of proteins

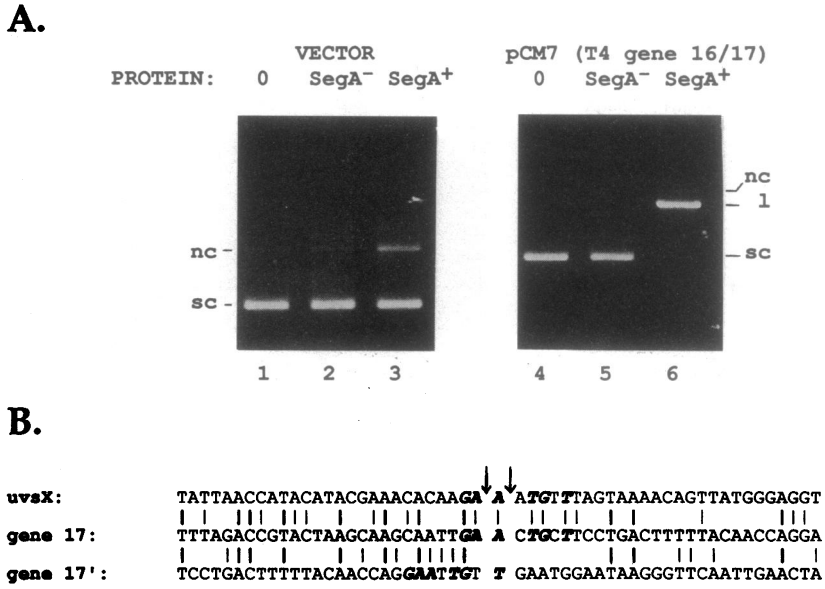

FIG. 7. SegA cleavage within T4 gene 17. (A) Agarose (0.8%) gel showing the products of incubation of vector (lanes 1 to 3) or pCM7, which contains the T4 gene 16 and part of gene 17 (lanes 4 to 6, 200 ng of each plasmid) with 10 ng of the SegA$^+$ (fraction PP II) or the SegA$^-$ control or no enzyme, as described in Materials and Methods. The locations of nicked circle (nc), linear (l), and supercoiled (sc) plasmid DNAs are denoted. (B) Comparison of the sequences of the SegA cleavage sites in gene 17 (designated gene 17 and gene 17′) and that in *uvsX*. Arrows show location of cuts on strand shown. (Cleavage sites on the complementary strand of the *uvsX* site are given in Fig. 6B; the same relative positions are cleaved in the 17 and 17′ sites.) Bases that are identical in all three sites are connected by boldface bars; identical bases between *uvsX* and gene 17 or gene 17 and gene 17′ are shown with thin bars. The sequence GAANTGNT, shown italicized and in boldface type, is cleaved in both the *uvsX* and gene 17 sites but not in the gene 17′ site (see text for further discussion).

encoded by group I introns of fungi and phage (51). Characterization of I-TevI, a member of this family, as well as group I-encoded endonucleases sharing different amino acid motifs, has shown that these endonucleases initiate the process of intron homing, the movement of their own intron DNA into an intronless allele. Each endonuclease cleaves at a specific site at or near the site of insertion. Double-strand break repair then converts that DNA into an intron-containing gene (for reviews, see references 3, 4, 17, 35, and 44). It has been suggested that it was the endonuclease-encoding DNA rather than the intron which was the original mobile element, in part because homing and splicing are independent processes and because the presence of these ORFs varies greatly in closely related species (for a recent review, see reference 40). Furthermore, the discovery of the VMA1-derived endonuclease of *S. cerevisiae*, which is encoded by a mobile spacer within a protein coding sequence rather than by an intron, indicates that introns are not an obligatory location for these ORFs (22, 29). Likewise, neither the yeast HO-endonuclease, a site-specific nuclease needed for mating type switching, nor the yeast endonuclease Endo.SceI, which has been implicated in recombination, is located within introns, yet both are related by sequence to a family of group I endonucleases distinct from the GIY-YIG family (41, 42, 47).

A number of observations are consistent with the idea that *segA* arose from the insertion of mobile endonuclease DNA of the GIY-YIG family into an intergenic location within the T4 genome. First, we have found that the *segA* gene is not present in the genomes of many T-even-related phages, including those of T2 and T6, which are highly homologous to T4. This result is consistent with the insertion of DNA containing the *segA* gene into the T4 genome some time after the divergence of these closely related phages. Second, within the GIY-YIG family, SegA shares the strongest relationship not to the other phage member, the T4-encoded endonuclease I-TevI, but rather to ORFs found within group I introns of *N. crassa* and *P. anserina* (51). This finding also supports the horizontal movement of endonuclease-encoding DNA from fungi to phage (or vice versa). Finally, the activity of the purified SegA protein shares features with some of the intron-encoded endonucleases. Recent characterization of the cleavage sites of I-TevI, a member of the GIY-YIG family, and I-SceII and I-CreI, members of another family, have shown that although each of these protein exhibits specificity for its insertion site, it tolerates many changes within its recognition sequence (6, 18, 50, 58). Our finding that SegA has a hierarchy of specificity supports the interesting idea that these endonucleases can be both highly specific (for insertion into a specific sequence) as well as more relaxed (for the infrequent transposition into a broad range of targets) (6, 40). Further information about the common properties of the GIY-YIG proteins awaits the purification and characterization of the fungal intron-encoded ORFs.

As yet, movement of *segA* DNA has not been detected. However, we would not expect to see transposition of the *segA*-containing DNA into the mapped cleavage sites in the T4 genes *uvsX* or 17, since both of these genes are required for normal phage growth and thus, disruption of the genes should be very deleterious for the phage. The effect of cleavage or movement into the site around 50 map units is unclear because the function of this region of the phage has not been determined. We have also not observed movement of the *segA* DNA into the T2 or LZ11 genomes, which lack the *segA* gene, upon infection by these phages of cells that express SegA endonuclease and contain *segA* DNA on a plasmid (data not shown). However, sequencing of the T2 DNA between 42 and *uvsX* indicates that its lack of *segA* is not a simple deletion, since other DNA has replaced the *segA* gene (59). Thus, at present it is unclear whether *segA* is an active mobile element, a fossil

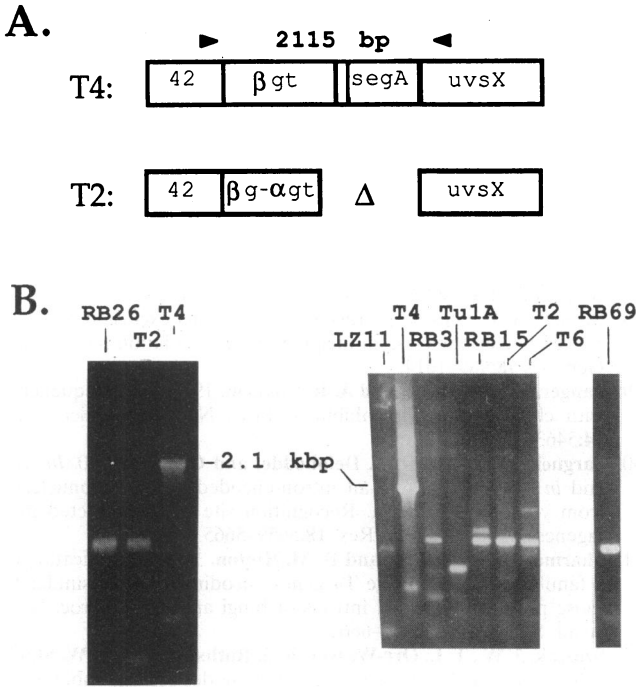

FIG. 8. Many T-even phage genomes lack the *segA* gene. (A) The region between genes 42 and *uvsX* in the phages T4 and T2. The genes 42 (hydroxymethylase), β*gt* (β-glucosyl transferase), *segA*, *uvsX*, and β*g*-α*gt* (β-glucosyl-HMC-α-glucosyl transferase), the locations of the primers used for PCR (shown as black triangles), and the 2,115-bp PCR product from T4 is indicated (see text for details). (B) DNA products obtained after PCR of the indicated T-even genomic DNAs. PCR was performed as described in Materials and Methods, and the products were run on 0.8% agarose gels. The 2.1-kbp DNA corresponding to the expected product from T4 DNA is indicated.

of a once-mobile gene that spread into a wide range of species, or a protein that is simply derived from the same primordial endonuclease as those now found in some group I introns.

ACKNOWLEDGMENTS

We thank H. Wu and L. Black for their gifts of plasmids; S. Eddy, L. Zwerdlinger, and L. Gold for T-even phages; and N. Nossal, M. Weitzman, M. Gellert, and R. Gumport for helpful discussions. We thank N. Nossal also for T4 polymerase, polymerase accessory proteins, and the protocol for using the accessory proteins in site-directed mutagenesis. We are especially grateful to J. Drake for assaying the survival of the *segA* mutant after irradiation.

REFERENCES

1. **Adhya, S., and M. Gottesman.** 1982. Promoter occlusion: transcription through a promoter may inhibit its activity. Cell **29:**939–944.
2. **Alberts, B., and G. Herrick.** 1971. DNA-cellulose chromatography. Methods Enzymol. **21:**198–217.
3. **Belfort, M.** 1990. Phage T4 introns: self splicing and mobility. Annu. Rev. Genet. **24:**363–385.
4. **Belfort, M.** 1991. Self-splicing introns in prokaryotes: migrant fossils. Cell **64:**9–11.
5. **Bell-Pedersen, D., S. Quirk, J. Clyman, and M. Belfort.** 1990. Intron mobility in phage T4 is dependent upon a distinctive class of endonucleases and independent of DNA sequences encoding the intron core: mechanistic and evolutionary implications. Nucleic Acids Res. **18:**3763–3770.
6. **Bryk, M., S. M. Quirk, J. E. Mueller, N. Loizos, C. Lawrence, and M. Belfort.** 1993. The td intron endonuclease I-TevI makes extensive sequence-tolerant contacts across the minor groove of its DNA target. EMBO J. **12:**2141–2149.
7. **Burger, G., and S. Werner.** 1985. The mitochondrial URFI gene in *Neurospora crassa* has an intron that contains a novel type of URF. J. Mol. Biol. **186:**231–242.
8. **Chu, F. K., G. F. Maley, F. Maley, and M. Belfort.** 1984. Intervening sequence in the *thymidylate synthase* gene of bacteriophage T4. Proc. Natl. Acad. Sci. USA **81:**3049–3053.
9. **Chu, F. K., G. F. Maley, D. K. West, M. Belfort, and F. Maley.** 1986. Characterization of the intron in the phage T4 *thymidylate synthase* gene and evidence for its self excision from the primary transcript. Cell **45:**157–166.
10. **Colleaux, L., M.-R. Michel-Wolwertz, R. F. Matagne, and B. Dujon.** 1990. The *apocytochrome b* gene of *Chlamydomonas smithii* contains a mobile intron related to both *Saccharomyces* and *Neurospora* introns. Mol. Gen. Genet. **223:**288–296.
11. **Collins, R. A., C. A. Reynolds, and J. Olive.** 1988. The self-splicing intron in the *Neurospora apocytochrome b* gene contains a long reading frame in frame with the upstream exon. Nucleic Acids Res. **16:**1125–1134.
12. **Crouse, G. F., A. Frischauf, and H. Lehrach.** 1983. An integrated and simplified approach to cloning into plasmids and single-stranded phages. Methods Enzymol. **101:**78–89.
13. **Cummings, D. J., J. M. Domenico, and F. Michel.** 1988. DNA sequence and organization of the mitochondrial *NDI* gene from *Podospora anserina*: analysis of alternate splice sites. Curr. Genet. **14:**253–264.
14. **Datta, D. B., B. Arden, and U. Henning.** 1977. Major proteins of the *Escherichia coli* outer cell envelope membrane as bacteriophage receptors. J. Bacteriol. **131:**821–829.
15. **Delahodde, A., V. Goguel, A. M. Becam, F. Creusot, J. Perea, J. Banroques, and C. Jacq.** 1989. Site-specific DNA endonuclease and RNA maturase activities of two homologous intron-encoded proteins from yeast mitochondria. Cell **56:**431–441.
16. **Dujon, B.** 1980. Sequence of the intron and flanking exons of the mitochondrial 21s *rRNA* gene of yeast strains having different alleles at the *omega* and *rib 1* loci. Cell **20:**185–197.
17. **Dujon, B.** 1989. Group I introns as mobile genetic elements: facts and mechanistic speculations—a review. Gene **82:**91–114.
18. **Dürrenberger, F., and J.-D. Rochaix.** 1993. Characterization of the cleavage site and the recognition sequence of the I-*Cre*I DNA endonuclease encoded by the chloroplast ribosomal intron of *Chlamydomonas reinhardtii*. Mol. Gen. Genet. **236:**409–414.
19. **Ellis, R. L., and D. M. Hinton.** 1989. Identification of a new bacteriophage T4 gene in the region of the phage genes *uvsX* (recombination protein) and 41 (primase-helicase). J. Cell. Biochem. **13D:**123.
20. **Formosa, T., and B. M. Alberts.** 1986. Purification and characterization of the T4 bacteriophage uvsX protein. J. Biol. Chem. **261:**6107–6118.
21. **Geisselsoder, J., F. Witney, and P. Yuckenberg.** 1987. Efficient site-directed in vitro mutagenesis using phagemid vectors. Biotechniques **5:**786–791.
22. **Gimble, F. S., and J. Thorner.** 1992. Homing of a DNA endonuclease gene by meiotic gene conversion in *Saccharomyces cerevisiae*. Nature (London) **357:**301–306.
23. **Hensgens, L. A. M., L. Bonen, M. de Haan, G. van der Horst, and L. A. Grivell.** 1983. Two intron sequences in yeast mitochondrial *coxI* gene: homology among URF-containing introns and strain-dependent variation in flanking exons. Cell **32:**379–389.
24. **Hinton, D. M.** 1989. Transcript analyses of the *uvsX*-40-41 region of bacteriophage T4. Changes in the RNA as infection proceeds. J. Biol. Chem. **264:**14432–14439.
25. **Hinton, D. M.** 1991. Transcription from a bacteriophage T4 middle promoter using T4 motA protein and phage-modified RNA polymerase. J. Biol. Chem. **266:**18034–18044.
26. **Hinton, D. M., and N. G. Nossal.** 1985. Bacteriophage T4 DNA replication protein 61. Cloning of the gene and purification of the expressed protein. J. Biol. Chem. **260:**12858–12865.
27. **Hinton, D. M., and N. G. Nossal.** 1986. Cloning of the bacteriophage T4 *uvsX* gene and purification and characterization of the T4 uvsX recombination protein. J. Biol. Chem. **261:**5663–5673.
28. **Hinton, D. M., L. L. Silver, and N. G. Nossal.** 1985. Bacteriophage

T4 DNA replication protein 41. Cloning of the gene and purification of the expressed protein. J. Biol. Chem. **260**:12851–12857.
29. **Hirata, R., Y. Ohsumi, A. Nakano, H. Kawasaki, K. Suzuki, and Y. Anraku.** 1990. Molecular structure of a gene, *vmaI*, encoding the catalytic subunit of H^+-translocating adenosine triphosphatase from vacuolar membranes of *Saccharomyces cerevisiae*. J. Biol. Chem. **265**:6726–6733.
30. **Jacquier, A., and B. Dujon.** 1985. An intron encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. Cell **41**:383–394.
31. **Kreuzer, K. N., and J. W. Drake.** 1994. Repair of lethal DNA damage in bacteriophage T4, p. 89–97. *In* J. D. Karam, J. W. Drake, K. N. Kreuzer, G. Mosig, D. H. Hall, F. A. Eiserling, L. W. Black, E. K. Spicer, E. Kutter, K. Carlson, and E. S. Miller (ed.), Molecular biology of bacteriophage T4. American Society for Microbiology, Washington, D.C.
32. **Kreuzer, K. N., and S. W. Morrical.** 1994. Initiation of T4 DNA replication, p. 28–42. *In* J. D. Karam, J. W. Drake, K. N. Kreuzer, G. Mosig, D. H. Hall, F. A. Eiserling, L. W. Black, E. K. Spicer, E. Kutter, K. Carlson, and E. S. Miller (ed.), Molecular biology of bacteriophage T4. American Society for Microbiology, Washington, D.C.
33. **Kunkel, T. A., J. D. Roberts, and R. A. Zakour.** 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
34. **Kutter, E., B. Guttman, G. Mosig, and W. Rüger.** 1990. Genomic map of bacteriophage T4, p. 1.24–1.51. *In* S. J. O'Brien (ed.), Genetic maps, vol. 5. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **Lambowitz, A. M.** 1989. Infectious introns. Cell **56**:323–326.
36. **Macreadie, I. G., R. M. Scott, A. R. Zinn, and R. A. Butow.** 1985. Transposition of an intron in yeast mitochondria requires a protein encoded by that intron. Cell **41**:395–402.
37. **Michel, F., and B. Dujon.** 1986. Genetic exchanges between bacteriophage T4 and filamentous fungi? Cell **46**:323.
38. **Mosig, G.** 1987. The essential role of recombination in phage T4 growth. Annu. Rev. Genet. **21**:347–371.
39. **Mosig, G.** 1994. Homologous recombination, p. 54–82. *In* J. D. Karam, J. W. Drake, K. N. Kreuzer, G. Mosig, D. H. Hall, F. A. Eiserling, L. W. Black, E. K. Spicer, E. Kutter, K. Carlson, and E. S. Miller (ed.), Molecular biology of bacteriophage T4. American Society for Microbiology, Washington, D.C.
40. **Mueller, J. E., M. Bryk, N. Loizos, and M. Belfort.** Homing endonucleases. *In* R. Roberts, S. Linn, and S. Lloyd. (ed.), Nucleases, 2nd ed., in press. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
41. **Nakagawa, K., N. Morishima, and T. Shibata.** 1991. A maturase-like subunit of the sequence-specific endonuclease Endo.SceI from yeast mitochondria. J. Biol. Chem. **266**:1977–1984.
42. **Nakagawa, K., N. Morishima, and T. Shibata.** 1992. An endonuclease with multiple cutting sites, Endo.SceI, initiates genetic recombination at its cutting site in yeast mitochondria. EMBO J. **11**:2707–2715.
43. **Nelson, M. A., M. Ericson, L. Gold, and J. F. Pulitzer.** 1982. The isolation and characterization of TabR bacteria: hosts that restrict bacteriophage T4*rII* mutants. Mol. Gen. Genet. **188**:60–68.
44. **Perlman, P. S., and R. A. Butow.** 1989. Mobile introns and intron-encoded proteins. Science **246**:1106–1109.
45. **Quirk, S. M., D. Bell-Pedersen, J. Tomaschewski, W. Rüger, and M. Belfort.** 1989. The inconsistent distribution of introns in the T-even phages indicates recent genetic exchanges. Nucleic Acids Res. **17**:301–315.
46. **Ripley, L. S., J. S. Dubins, J. G. deBoer, D. M. DeMarini, A. M. Bogerd, and K. N. Kreuzer.** 1988. Hotspot sites for acridine-induced frameshift mutations in bacteriophage T4 correspond to sites of action of the T4 type II topoisomerase. J. Mol. Biol. **200**:665–680.
47. **Russell, D. W., R. Jensen, M. J. Zoller, J. Burke, B. Errede, M. Smith, and I. Herskowitz.** 1986. Structure of the *Saccharomyces cerevisiae HO* gene and analysis of its upstream regulatory region. Mol. Cell. Biol. **6**:4281–4294.
48. **Russell, R. L., and R. J. Huskey.** 1974. Partial exclusion between T-even bacteriophages: an incipient genetic isolation mechanism. Genetics **78**:989–1014.
49. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
50. **Sargueil, B., D. Hatat, A. Delahodde, and C. Jacq.** 1990. *In vivo* and *in vitro* analyses of an intron-encoded DNA endonuclease from yeast mitochondria. Recognition site by site-directed mutagenesis. Nucleic Acids Res. **18**:5659–5665.
51. **Sharma, M., R. L. Ellis, and D. M. Hinton.** 1992. Identification of a family of bacteriophage T4 genes encoding proteins similar to those present in group I introns of fungi and phage. Proc. Natl. Acad. Sci. USA **89**:6658–6662.
52. **Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein, and F. W. Stahl.** 1983. The double-strand-break repair model for recombination. Cell **33**:25–35.
53. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82**:1074–1078.
54. **Tian, G.-L., F. Michel, C. Macadre, P. P. Slonimski, and J. Lazowska.** 1991. Incipient mitochondrial evolution in yeasts. II. The complete sequence of the gene coding for cytochrome b in *Saccharomyces douglasii* reveals the presence of both new and conserved introns and discloses major differences in the fixation of mutations in evolution. J. Mol. Biol. **218**:747–760.
55. **Turmel, M., J. Boulanger, M. N. Schnare, M. W. Gray, and C. Lemieux.** 1991. Six group I introns and three internal transcribed spacers in the chloroplast large subunit *ribosomal* RNA gene of the green alga *Chlamydomonas eugametos*. J. Mol. Biol. **218**:293–311.
56. **Warner, H. R., and D. P. Snustad.** 1983. T4 DNA nucleases, p. 103–109. *In* C. K. Mathews, E. M. Kutter, G. Mosig, and P. B. Berget (ed.), Bacteriophage T4. American Society for Microbiology, Washington, D.C.
57. **Wenzlau, J. M., R. J. Saldanha, R. A. Butow, and P. S. Perlman.** 1989. A latent intron-encoded maturase is also an endonuclease needed for intron mobility. Cell **56**:421–430.
58. **Wernette, C., R. Saldanha, D. Smith, D. Ming, P. S. Perlman, and R. A. Butow.** 1992. Complex recognition site for the group I intron-encoded endonuclease I-SceII. Mol. Cell. Biol. **12**:716–723.
59. **Winkler, M., and W. Rüger.** 1993. Cloning and sequencing of the genes of β-glucosyl-HMC-α-glucosyl-transferases of bacteriophages T2 and T6. Nucleic Acids Res. **21**:1500.
60. **Zwerdlinger, L., S. Eddy, and L. Gold.** Unpublished data.