Azotobacter vinelandii mutS: Nucleotide Sequence and Mutant Analysis

OSCAR LE, BINGHUI SHEN, SIIRI E. ISMAA,† AND BARBARA K. BURGESS*
Department of Molecular Biology and Biochemistry, University of California, Irvine, California 92717

Received 28 July 1993/Accepted 4 October 1993

An Azotobacter vinelandii homolog to the Salmonella typhimurium mutS gene was discovered upstream of the fdxA gene. The product of this gene is much more similar to S. typhimurium MutS than either is to the HexA protein of Streptococcus pneumoniae. An A. vinelandii ΔmutS mutant strain was shown to have a spontaneous mutation frequency 65-fold greater than that of the wild type.

Errors in DNA biosynthesis, chemical damage to DNA, and homologous recombination events can all give rise to mismatches in DNA base pairing. All organisms must therefore have some system for repairing those mismatches. One well-characterized repair system is the methyl-directed long-patch mismatch correction system of Escherichia coli (5, 14, 15). This system relies, in part, on the fact that DNA methylation is a postreplication event such that the newly synthesized strand (where the repair occurs, if necessary) is unmethylated for a short time. The first component of the system is a 97-kDa protein, MutS, that recognizes and binds to DNA at the site of the mismatch (23, 24).

In this study, we identified a mutS homolog in the free-living, obligately aerobic diazotroph Azotobacter vinelandii. This organism is increasingly being used for biochemical and genetic characterization of a variety of biological phenomena, in large part because it has an excellent transformation system (17, 18) and undergoes high-frequency reciprocal recombination (21). Studies of DNA replication or repair have not, however, previously been reported for this organism. Here we report the sequence of the A. vinelandii mutS gene and show that deletion of this gene from the A. vinelandii chromosome results in a Mut− phenotype.

DNA sequence analysis. We have previously reported the cloning of a 4.8-kb A. vinelandii DNA fragment that contains the fdxA gene and the subcloning of that fragment into pUC9 to form a plasmid designated pML1 (16). Here we determined the nucleotide sequence of the portion of pML1 upstream of the fdxA gene. Four M13mp18 derivatives were successfully constructed for sequencing, and the sequencing strategy is shown in Fig. 1. A total of 2,748 bp were sequenced, and the sequences of all fragments were overlapped, in both directions, to minimize sequencing errors. This sequencing strategy revealed a 2,605-bp open reading frame. As shown in Fig. 2, the predicted protein product encoded by this open reading frame is similar to that predicted by the DNA sequence of the Salmonella typhimurium mutS gene (6) and the Streptococcus pneumoniae hexA gene (20). On the basis of this amino acid sequence similarity, which is discussed in more detail below, this open reading frame has been designated the mutS gene of A. vinelandii.

As revealed by the DNA sequence, there is a stop codon (TGA) immediately in front of the start codon (ATG) of the mutS gene. This situation is not uncommon in A. vinelandii and occurs, for example, in the nif gene cluster when two genes with related functions are part of the same transcriptional unit (2, 9). In this case, it suggests that there is likely to be an open reading frame upstream of the mutS gene whose function may be related to that of mutS and which may be cotranscribed with mutS. Downstream of the mutS stop codon (TGA), there are inverted complementary sequences that are predicted to form a stem-loop structure which may act as a transcription termination signal. This would also be consistent with our finding, reported elsewhere (27), that an fdxA promoter is located between the stop site of the mutS gene and the start site of the fdxA gene.

Homology of A. vinelandii MutS with S. typhimurium MutS and S. pneumoniae HexA. Figure 2 compares the sequence of the putative A. vinelandii mutS gene product to those of the S. typhimurium MutS and S. pneumoniae HexA proteins. The mutS sequences from the two gram-negative organisms are much more similar to each other than either of them is to the HexA protein. These similarities occur over the entire length of the corresponding proteins to yield proteins with similar molecular weights. When the A. vinelandii and S. typhimurium MutS amino acid sequences are optimally aligned, 58% of the amino acids are identical and an additional 29% correspond to conservative changes. The similarity of these two proteins to S. pneumoniae HexA is also highly significant; ca. 36% of the amino acids in HexA are identical to either S. typhimurium MutS (7) or A. vinelandii MutS, and an additional 22% are conservative changes.

A previous study has compared the sequences of S. typhimurium MutS and S. pneumoniae HexA and discussed their relevance to the DNA mismatch repair systems used by those organisms (6). That study identified two regions where the sequences have an especially high degree of conservation. Region I is a 109-amino-acid stretch, enclosed by the thin vertical arrows in Fig. 2, close to the amino terminus of each protein. Comparison of the A. vinelandii MutS sequence to the S. typhimurium MutS sequence in this region shows that 80% of the amino acids are identical and that 18 of the remaining 24 amino acids represent conservative substitutions (Fig. 2). These findings further support the idea that region I has important functional significance (6).

The second most similar region, designated region II, is located close to the carboxyl terminus of each protein and is enclosed by thick vertical arrows in Fig. 2. In this region, comparison of the A. vinelandii MutS sequence to the S. typhimurium MutS sequence reveals that 80% of the amino acids are identical and that 20 of the remaining 25 amino acids

* Corresponding author.
† Present address: The Heart Research Institute, Camperdown, New South Wales 2050, Australia.
represent conservative changes. Region II contains an ATP-binding site (7), and the consensus sequence for this site (29) is shown in parentheses in Fig. 2. This region has also recently been shown to be homologous to an analogous predicted product of the RepI gene in humans and the DicI gene of mice (3, 12). In addition to these two major regions of similarity, a putative helix-turn-helix DNA-binding motif has been identified in the S. typhimurium MutS and S. pneumoniae HexA sequences (6). As shown in Fig. 2, this motif is also conserved in the A. vinelandii MutS sequence, in which it spans amino acid residues 771 to 790.

Construction of an A. vinelandii mutant strain that does not synthesize MutS. Taken together, the above comparisons indicate that, at least insofar as MutS is concerned, the DNA mismatch repair system of A. vinelandii is likely to be structurally and functionally analogous to that of S. typhimurium. To determine whether or not MutS is active in A. vinelandii, we therefore constructed a mutant strain that cannot synthesize MutS. As shown in Fig. 3, this was accomplished by first constructing a plasmid designated pOL1, in which the first 1,270-bp ClaI-EcoRV fragment of the mutS gene was deleted and replaced by a kanamycin resistance cartridge (28). The deletion alone would remove the first 417 amino acid residues from the NH2-terminal end of the MutS protein, and the kanamycin resistance cartridge insertion additionally resulted in a 1-bp frame shift that would affect the remaining mutS sequence. Thus, disruption of mutS in this manner is expected to eliminate any possible MutS function.

Purified, linearized A. vinelandii DNA, derived from pOL1, was then introduced into competent cells of wild-type A. vinelandii (17) and allowed to undergo homologous recombination, as described previously for the construction of several A. vinelandii nif deletion mutant strains (1, 21), except that the selection in this case was for kanamycin resistance. The resulting kanamycin-resistant ΔmutS mutant strain is designated JG131 and should be identical at the DNA level to wild-type A. vinelandii except that the chromosomal mutS gene has been disrupted. As expected, JG131 has a stable kanamycin resistance phenotype in the absence of selective pressure.

To confirm the mutation at the DNA level, genomic DNA was isolated from the wild-type and JG131 strains of A. vinelandii and restricted with BglI and EcoRI. Figure 3 shows the locations of these restriction sites relative to the mutS gene. As shown in Fig. 4, when the wild-type DNA fragments were run on an agarose gel and then probed with a radiolabeled 1.27-kb ClaI-EcoRV fragment specific to mutS, a single band of the size predicted from the restriction sites of the mutS gene shown in Fig. 3 was observed. This demonstrates that wild-type A. vinelandii contains only one copy of the mutS homolog per chromosome. When JG131 DNA was similarly probed with the ClaI-EcoRV fragment, no bands were observed, demonstrating that this fragment had been successfully removed from the chromosome (Fig. 4). The presence of the kanamycin resistance cartridge at the appropriate location in the A. vinelandii JG131 genome was also demonstrated by Southern hybridization as shown in Fig. 4. Western (immunoblot) analysis with antibodies raised against purified A. vinelandii FdI demonstrated that the fdxA gene located downstream of the mutS gene was not disrupted in JG131 (data not shown).

Deletion of mutS results in a Mut− phenotype. To confirm that the A. vinelandii mutS gene is a mutator gene in that organism, we compared the spontaneous mutation rate of wild-type A. vinelandii to that of our ΔmutS mutant strain JG131 by measuring their frequencies of mutation to rifampin

FIG. 1. Sequencing strategy for a portion of A. vinelandii chromosomal DNA that includes the mutS gene. Locations of relevant restriction enzyme sites are indicated above the upper bar. The arrows below the upper bar show the coding regions and the direction of transcription of mutS and fdxA. The subcloned A. vinelandii chromosomal DNA restriction fragments are denoted in the lower half of the figure by thick solid lines. pBS300 and pBS303 were sequenced 5' to 3', while pBS301 and pBS304 were sequenced 3' to 5'. The short arrows at the bottom represent start points and directions of sequencing. The dots indicate where synthetic oligonucleotides were used as primers for sequencing instead of the universal M13 primer. The following restriction enzymes were used: S, Sall; Sm, SmaI; P, PstI; A, AccI; Pu, PvuII; E, EcoU; Ap, Apal; X, XhoI; C, ClaI; EV, EcoRV; B, BstHII.

FIG. 2. Comparison of homologous DNA mismatch repair proteins A. vinelandii (Av) MutS, S. typhimurium (S) MutS, and S. pneumoniae (Sp) HexA. Residues identical to those of A. vinelandii MutS are boxed. The two regions with significantly higher conservation (8) are enclosed by thin vertical arrows (region I) and thick vertical arrows (region II). A putative nucleotide-binding site is emphasized by including the consensus sequence in parentheses above the sequence.
Anabaena

VOL.

vector,

was

where for

with

cleaved

polymerase. Similarly, pUC-4K, the ClaI

gene

transferase
determined).

functions

of

mutS

and

I

magnitude

elects

gene

and

that of the

in

that

known

in

DNA-binding [Fe-S] protein (11). Most important

involvement

of whether

question

recently

protein,

FIG. 3.

mutS and

ClaI

gene

and

HindIII

function is unknown

in

that

the

element-binding

proposed

to

RNA

procedures

The

observed for MutS

A.

vinelandii.

[Fe-S]

proteins

4, 19).

In

and

those reported else-

or

or

and

mutS

-?

FdxA-

FdxA+

of Rif'

No.

Spontaneous

mutation

from

elsewhere

mutS

mutS

and

FdxA

related? The finding (Fig. 1) that the

mutS and fdxA genes are adjacent on the A. vinelandii chromo-

some and transcribed in the same direction brings up the

question of whether these locations are coincidental or the

functions of the two proteins are somehow related. The fdxA

gene encodes a small [Fe-S] protein designated A. vinelandii

ferredoxin I (AvFdxI) whose function is unknown (16). The

involvement of [Fe-S] proteins in DNA repair has a precedent

in that endonuclease III (10) and MutY (13, 26) are both

thought to be [Fe-S] proteins. In addition, another putative

[Fe-S] protein, the iron-regulatory element-binding protein, is

known to bind RNA (22), and the patB gene product of

Anabaena strain 7120 has recently been predicted to be a

DNA-binding [Fe-S] protein (11). Most important for this

study is the fact that the fdxA gene product, AvFdxI, has also

recently been proposed to be a DNA-binding protein (25). We

therefore examined a mutant strain that does not synthesize

FdxI to determine whether it too has a mutator phenotype.

The FdxI- strain used in this study, designated JG120, was

constructed by site-directed mutagenesis (8). It has a stop
codon at position 24 of the FdxI amino acid sequence

and contains no material that cross-reacts with FdxI antibodies.

Table 1 shows that this strain has the same spontaneous

mutation rate as wild-type A. vinelandii. Thus, removal of FdxI

in vivo does not result in a mutator phenotype.

Nucleotide sequence accession number. The nucleotide se-

cquence described here has been deposited in the GenBank/EMBL
data base under accession no. M63007.

FIG. 4. Southern analysis of wild-type A. vinelandii OP and ΔmutS

mutant strain JG131 genomic DNAs digested by EcoRI and BglI.

Lanes 1 and 3 contained digested wild-type A. vinelandii genomic

DNA, while lanes 2 and 4 contained digested ΔmutS mutant strain

genomic DNA. Lanes 1 and 2 were probed with a radiolabeled 1.27-kb

EcoRI-ClaI mutS-specific fragment, and lanes 3 and 4 were probed

with a radiolabeled 1.28-kb PstI fragment of pUC4K which is specific

for the kanamycin resistance cartridge. The positions of the restriction

sites relative to mutS are shown in Fig. 3.

Table 1. Spontaneous mutation frequencies of MutS+, FdxA+,

MutS-, and FdxA- strains

<table>
<thead>
<tr>
<th>A. vinelandii strain</th>
<th>Mean no. of Rif'</th>
<th>No. of independent</th>
<th>Relative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mutants/10(^8)</td>
<td>cultures</td>
<td>mutation</td>
</tr>
<tr>
<td></td>
<td>cells ± SD</td>
<td></td>
<td>frequency</td>
</tr>
<tr>
<td>Wild type (MutS+ FdxA+)</td>
<td>1.2 ± 1.0</td>
<td>39</td>
<td>1</td>
</tr>
<tr>
<td>JG131 (MutS- FdxA+)</td>
<td>78 ± 35</td>
<td>38</td>
<td>65</td>
</tr>
<tr>
<td>JG120 (MutS+ FdxA-)</td>
<td>1.4 ± 1.7</td>
<td>21</td>
<td>1</td>
</tr>
</tbody>
</table>

* Frequencies were calculated from the numbers of independent cultures. A few mutational "jackpots," which were defined as cultures having more than five times the mean number of mutants, were excluded from the calculation of the mean and standard deviation, as described elsewhere (4, 19). Data obtained for wild-type strains OP (the most commonly used wild-type strain) and Trans (the parent for JG131 and JG120) were indistinguishable and were therefore combined.
This work was supported by National Institutes of Health grant RO1-GM45209.

REFERENCES


