Further Characterization of *Escherichia coli* Endonuclease V

MECHANISM OF RECOGNITION FOR DEOXYINOSINE, DEOXYURIDINE, AND BASE MISMATCHES IN DNA*

(Received for publication, May 5, 1997, and in revised form, August 22, 1997)

Min Yao and Yoke Wah Kow‡

From the Division of Cancer Biology, Department of Radiation Oncology Emory University School of Medicine, Atlanta, Georgia 30335

Endonuclease V from *Escherichia coli* has a wide substrate spectrum. In addition to deoxyinosine-containing DNA, the enzyme cleaves DNA containing urea residues, AP sites, base mismatches, insertion/deletion mismatches, flaps, and pseudo-Y structures. The gene coding for the enzyme was identified to be orf 225 or nfi (endonuclease fi{ve}). Using enzyme purified from an overproducing strain, the deoxyinosine- and mismatch-specific activities of endonuclease V was found to have different divalent metal requirements. The affinity of the enzyme is greater than 20-fold higher for DNA containing deoxyinosine than deoxyribonuclease or base mismatches. Under optimal cleavage conditions, endonuclease V forms two stable complexes with DNA containing deoxyinosine, but not with DNA containing base mismatches or deoxyribonuclease, suggesting that the 6-keto group of hypoxanthine in DNA is critical for stable interactions with the protein. The enzyme recognizes deoxyuridine in DNA but exhibits a much lower affinity to DNA containing deoxyuridine compared with DNA containing deoxyinosine. Interestingly, deoxyuridine-specific endonuclease activity of endonuclease V has a divalent metal requirement similar to the mismatch activity. A model for the mechanism of substrate recognition is proposed to explain these different activities.

Recently, we and others (1, 2) have shown that deoxyinosine 3‘-endonuclease is identical to endonuclease V, previously characterized as an enzyme with endonuclease activity toward DNA treated with acid, alkali, OsO₄, or 7-bromomethyl-benz[a]anthracen (3, 4). The gene coding for endonuclease V (nfi, endonuclease fi{ve}) was found to be identical to orf 225, located at 90 minutes of the *Escherichia coli* genome (GenBank™ accession no. U00006) (5). In addition to the less defined substrates, we have shown that homogeneous preparations of endonuclease V purified from the wild type *E. coli* cells recognize several well defined DNA lesions, including deoxyinosine (6), urea residues (6), AP sites (6), base mismatches (7), loops and hairpins (2), flaps, and pseudo-Y DNA structures (2). The ability of endonuclease V to recognize mismatches and abnormal replicative DNA structures suggests that the enzyme might play an important role in DNA metabolism (2).

Unlike DNA N-glycosylases, endonuclease V cleaves the DNA strand containing lesions at the second phosphodiester bond 3‘ to the lesion, leaving a nick with 3‘-hydroxyl and 5‘-phosphoryl groups. However, the lesion is not removed from the DNA by the enzyme. Endonuclease V forms stable complexes with DNA containing deoxyinosine both before and after cleavage, showing similar affinity to both the substrate and the product (8). Based on these results, we have proposed earlier (8) that, besides its endonuclease activity, the enzyme might function to target other repair protein(s) to the lesion. Therefore, it is possible that endonuclease V could initiate a novel repair pathway. Endonuclease V cleaves DNA containing base mismatches in a strand-specific manner; it cleaves the DNA strand containing mismatch closer to the 5‘ terminus (7). Furthermore, the mismatch-specific activity of the enzyme is reduced when the mismatch is flanked by GC pairs; however, its deoxyinosine-specific activity is not influenced by the sequence context. These results suggest that endonuclease V might have different modes of interaction between DNA containing deoxyinosine and mismatches. It is, therefore, important to understand the mechanism governing the recognition of various DNA lesions by this small protein (24,900 Da). To understand the possible mechanism that is involved in the substrate recognition by endonuclease V, we cloned the gene coding for the enzyme (2) so that a large quantity of the protein can be obtained rapidly for further biochemical and physical characterization.

MATERIALS AND METHODS

All reagents were of analytical grade of purity. Restriction enzymes were purchased from New England Biolabs. General molecular biology procedures used were adapted from the protocols described in Sam broak *et al.* (9).

Cloning, Overexpression, and Purification of Endonuclease V—Homogeneous endonuclease V was purified from 400 g of *E. coli* paste according to the procedures described previously (6). Fraction V, which was judged to be homogeneous by silver staining, was blotted onto a polyvinylidene difluoride membrane and sent to the University of Texas Medical Branch at Galveston for N-terminal amino acid acid sequencing. The first 24 amino acid residues obtained for the N terminus of endonuclease V were determined to be MDLASLRAQQIELASSVIREDLD. When sequence homology for this amino acid sequence was searched through the GenBank™ using the Blast program (10), the N-terminal amino acid sequence was found to be identical to the N-terminal sequence of a hypothetical protein (24.9 kDa) translated from orf 225 (GenBank™ accession no. U00006) (5). The only difference was that the hypothetical protein had two extra amino acid residues at the N terminus, that is, the translated N-terminal amino acid sequence derived from orf 225 was MIMDLASLRAQQIELASSVREDRLD.

To confirm that orf 225 is the gene coding for endonuclease V, orf 225 was amplified from *E. coli* genomic DNA by the polymerase chain reaction using flanking primers derived from the genomic DNA sequences obtained from the GenBank™. The 5‘ primers and the 3‘ primer contained a NdeI restriction site (underlined) and an AvaI restriction site (underlined), respectively. Since the genomic sequence of orf 225 obtained from the GenBank™ (access no. U00006) (5) revealed two potential start codons for orf 225, two different 5‘ primers were used: 5‘ primer 1, CTGGGAAATCCATAGTATTAGGATCTCGC-

* This work is supported by National Institutes of Health Grant GM 37216 (to Y. W. K.). The cost of publication of this article was defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Division of Cancer Biology, Dept. of Radiation Oncology, Emory University School of Medicine, 145 Edgewood Ave., S.E., Atlanta, GA 30335. Tel.: 404-616-6951; Fax: 404-616-5689.
GTCA, 5′-primer 2, CTGGGAATTTCTAGGATCTGCCTGATATTAGCGC. The sequence for the 3′ primer was CGAGCATTCCCTGAGCTGGA-TAGGCGCCGTTGC. Polymerase chain reactions were carried out using Taq DNA polymerase (Boehringer Mannheim). The polymerase chain reaction fragment amplified with 5′ primer 1 and the 3′ primer was cloned into the Ndel and the Awas I sites of pET22b (+) vector (Novagen) to generate the plasmid pETI-1. The polymerase chain reaction fragment amplified with 5′ primer 2 and the 3′ primer was cloned into pET22b (+), yielding the plasmid pETI-2. Both the pETI-1 and the pETI-2 plasmids were transformed into E. coli strain BL21(DE3) by electroporation. When BL21(DE3) carrying plasmid pETI-1 or pETI-2 was grown and induced with isopropyl-1-thio-galactoside, the overexpression of a 25-kDa protein was observed to exhibit activity identical to endonuclease V (data not shown).

Based on the above observations, endonuclease V was then purified from an overexpressing host BL21(DE3) carrying pETI-2. The protein produced by the gene cloned in pETI-2 has the same N-terminal sequences as determined for endonuclease V. Four liters of BL21(DE3) carrying pETI-2 were grown in Luria broth containing 50 μg/ml ampicillin to an absorbance at 600 nm of 0.4. Isopropyl-1-thio-galactoside was then added to the culture at a final concentration of 0.34 mM MnCl2. Binding reactions were analyzed by an electrophoresis mobility shift assay as described under “Materials and Methods.” The protein concentrations used for the binding reactions were ranged from 0.82 to 130 nM.

**RESULTS**

**Optimal conditions for the endonuclease activities of endonuclease V**—We have shown previously (6) that MnCl2 can replace MgCl2 and support the endonuclease activities of endonuclease V. An optimal MnCl2 concentration (0.34 mM) supports about 60–70% of deoxyinosine-specific endonuclease activity of endonuclease V at optimal MgCl2 (6). In contrast, the mismatch-specific activity of the enzyme is more than 20-fold higher in the presence of Mn2+ (0.34 mM) than Mg2+ (2 mM) (data not shown), showing sharp differences in the divalent ion requirements for the deoxyinosine-specific activity and mismatch-specific activity.

Endonuclease V was shown to have an optimal pH above 9.0 by earlier investigators using partially purified enzyme preparations (3, 4). However, highly purified endonuclease V prepared from an overproducing E. coli strain did not agree with these findings. Consistent with our results published earlier (6), endonuclease V cleaves deoxyinosine-containing DNA at a pH range from 6.0 to 9.5, with an optimal pH between 7.0 to 7.5. Less than 30% of deoxyinosine-specific activity remained when the pH was higher than 9.0. Similarly, the enzyme has a
sharp pH optimum of 8.0 for the mismatch-specific endonuclease activity. Only 50% of mismatch-specific activity remained at either pH 7.5 or 8.5, and less than 10% of the optimal activity when the pH of the reaction was higher than 9.0.

**Binding of Endonuclease V to DNA Containing Deoxyinosine and Base Mismatches**—We have showed previously (8) that endonuclease V purified from the wild type *E. coli* strain forms two complexes with DNA containing deoxyinosine with observed *Kd* for complexes I and II of 4 and 400 nM, respectively. In agreement with the earlier results, endonuclease V purified from the overproducing strain formed two stable complexes with oligodeoxy nucleotide containing deoxyinosine (Fig. 1). The *Kd* for complexes I and II (2 nM and 60 nM, respectively, data not shown) were lower than those observed previously using endonuclease V purified from wild type *E. coli* cells (8). This is probably due to a higher amount of active fraction in the enzyme preparation prepared from the overproducing strain.

Under optimal cleavage conditions (i.e. in the presence of 2 mM MgCl₂, 37 °C) most of the di-containing oligonucleotides remains bound with endonuclease V as protein-DNA complexes, with less than 10% of total radioactive material migrated as cleaved oligonucleotide (8) (Fig. 1). In contrast, DNA-protein complexes formed between endonuclease V and mismatch-containing DNA were relatively unstable, either in the presence of Mg²⁺ or Mn²⁺. In the presence of Mn²⁺, the majority of the radioactive material migrated as cleaved oligonucleotides in an EMSA when duplex A/A was used as DNA substrate (Fig. 1). Fig. 2 showed the quantitative results of an EMSA for endonuclease V with duplex A/A (with a 5'-end-labeled top strand) in the presence of 0.34 mM MnCl₂. Binding reactions were analyzed by an electrophoresis mobility shift assay as described under “Materials and Methods.” Panel B, the gel was dried under vacuum and quantified by a Fuji Bio-Imaging analyzer. The percentage of the radioactive signal for each band was plotted against the total enzyme concentration.

**Endonucleolytic Activities of Endonuclease V on DNA Containing Deoxyuridine**—In our earlier studies, we found that endonuclease V could not cleave deoxyuridine-containing oligodeoxynucleotides (6). Since partially purified endonuclease V preparations from earlier investigators were shown to recognize deoxyuridine-containing DNA (1, 3, 4), we decided to re-examine whether a highly purified endonuclease V preparation could recognize DNA containing deoxyuridine in various sequence contexts. Oligonucleotides U24 and U21 were annealed to SK(−) DNA to form the following duplexes which are identical to duplexes F and H used in earlier studies (6):

\[
\text{5'}-\text{GGCCGCGCTCTTGAACTAGTGG}-3' \\
\text{3'}-\text{TACCCGCGTACCTGTACCATGCAGGGSK-Sk(−)}
\]

**Duplex F**

\[
\text{5'}-\text{CCCCCGGCTCTAGGAAATTGACG}-3' \\
\text{3'}-\text{TAGGGGGCGCGAATTTTAAGCTA-Sk(−)}
\]

**Duplex H**

The endonucleolytic activity of endonuclease V on these two DNA duplexes were compared with that on duplex I/SK which was formed by annealing oligo I to single-stranded SK(−)/T/A (see Equation 4).

\[
\text{5'}-\text{GTCGACTTAGAAGGATCCCCTTACGCGTAC}-3' \\
\text{3'}-\text{GGCCGCGACGTCCAGCTAGTGCCTCTTAGGGCCATGGTAAAGGC-}
\]

**SK(−) T/A**

(Eq. 4)
Fig. 4 shows that in the presence of 2 mM MgCl₂, the endonucleolytic activity of endonuclease V on deoxyuridine-containing DNA was much lower compared with that on deoxyinosine-containing DNA. The enzyme cleaved deoxyuridine-containing DNA only at a high enzyme concentration, and the activity did not appear to increase proportionally with an increase in enzyme concentration. Since duplex F and duplex H contained U/T and U/C mispairs, respectively, cleavage activity on these duplexes could be due to the mismatch-specific activity of endonuclease V. We have shown earlier that the mismatch-specific activity of endonuclease V was much higher in the presence of Mn²⁺ than Mg²⁺, and we thus examined whether endonuclease V had increased endonucleolytic activity on duplexes F and H when reactions were performed in the presence of Mn²⁺. In the presence of Mn²⁺, endonuclease V had a much higher activity on deoxyuridine-containing DNA, about 20-fold higher than in the presence of Mg²⁺ at most enzyme concentrations tested (Fig. 5). Interestingly, in the presence of Mn²⁺, the endonucleolytic activity of endonuclease V on duplexes F and H increased proportionally to the increase in the amount of enzyme. In contrast to the earlier observations (7, 8), the deoxyuridine-specific endonucleolytic activity of endonuclease V was found to have an optimal pH between 8.0 and 8.5, which is similar to the optimal pH observed for mismatch- and deoxyinosine-specific activity (data not shown). However, the deoxyuridine endonuclease activity of endonuclease V was less sensitive to higher pH values; 60–70% of the deoxyuridine endonuclease activity remained when the pH of the reaction was increased to 9.5. At this pH, both the mismatch- and deoxyinosine-specific activity of the enzyme was inhibited to less than 20% of the optimum activity (data not shown).

Endonuclease V Exhibits Specific Recognition of Deoxyuridine in DNA—Since duplex F and duplex H contain a U/T and U/C mispair, respectively, it is uncertain whether the endonucleolytic activity observed on these duplexes is due to the specific recognition of dU or to the base mismatch. To determine whether endonuclease V has a specific recognition for deoxyuridine, we prepared the following oligonucleotide duplexes:
E. coli Endonuclease V

mismatch-specific endonuclease activity. V exhibit features that is distinct from the deoxyinosine- and
Thus, the deoxyuridine endonuclease activity of endonuclease
when duplex U/A was used as DNA substrate (data not shown).

The deoxyuridine-endonuclease activity of endonuclease V on U21/SK and U24/SK were
assayed either in the presence of Mg$^{2+}$ (solid bar) or in the presence of Mn$^{2+}$
(open bar) with the indicated amount of endonuclease V. When performed in the
presence of Mg$^{2+}$, the reactions followed the standard endonuclease assay condi-
tions for deoxyinosine-specific activity. When performed in the presence of Mn$^{2+}$,
the reactions followed the standard endonuclease assay conditions for mismatch-
specific activity. The protein concentrations used for the cleavage reactions
ranged from 0.65 to 130 nM. The percentage of substrate cleavage was quantified and plotted against the enzyme
concentration.

5'-GGTCGACTUAGGGAGATCCCCCGGTAC-3'
3'-ACGTCCAGCTGAATCTCTCTAGGGCC-5'

**DUPEX U/X**

(Eq. 5)

5'-GGTCGACTXAGGGAGATCCCCCGGTAC-3'
3'-ACGTCCAGCTGAATCTCTCTAGGGCC-5'

**DUPEX X/U**

(Eq. 6)

In these duplexes, X stands for either dA, dT, dC, or dG. Thus, Duplex U/X can contain either U/A pair or U/T, U/C, or
U/G mispair and duplex X/U can contain either an A/U pair or a T/U, C/U, or G/U mispair.

Fig. 6 shows that endonuclease V cleaved oligo U in these
duplexes efficiently in the presence of Mn$^{2+}$, no matter whether
dU was paired with dA or mispaired with other bases. We have
shown earlier (7) that if X/U or U/X (X being any of the four
bases) in these duplexes were replaced by mismatched pairs, endonuclease V cleaved only the top strand whose 5’ terminus
is closer to the mismatches. However, if U/X or X/U was
replaced with LX or XL, the enzyme cleaves the strand contain-
ing deoxyinosine whether it is on the top strand or the bottom
strand. Since endonuclease V cleaved the deoxyuridine-con-
taining strand in both duplex U/X and duplex X/U (Fig. 6, A
and B), the enzyme specifically recognizes deoxyuridine in a
manner similar to the recognition of deoxyinosine. In addition,
endonuclease V possesses a very weak activity on single-
stranded DNA containing deoxyuridine (data not shown), fur-
ther indicating that the enzyme recognizes deoxyuridine spe-
cifically in the DNA. However, it is interesting to note that the
deoxyuridine-specific activity of the enzyme had a similar di-
valent ion requirement as the mismatch-specific activity. Even
on Duplex U/A, which contains a U/A pair, the endonucleaseytic
activity of the enzyme was 5-fold more active in the presence of
Mn$^{2+}$ than in the presence of Mg$^{2+}$. Furthermore, in the pres-
ence of Mn$^{2+}$, although endonuclease V binds more tightly to
duplex U/A than to duplex A/A, more than 20% of radioactive
material migrated as cleaved oligonucleotides in an EMSA
when duplex U/A was used as DNA substrate (data not shown).
Thus, the deoxyuridine endonuclease activity of endonuclease
V exhibit features that is distinct from the deoxyinosine- and
mismatch-specific endonuclease activity.

**DISCUSSION**

To further understand the mechanism involved in the sub-
strate recognition by endonuclease V, the gene coding for the
enzyme was cloned into an overproducing pET plasmid and
large quantities were purified. In agreement with earlier stud-
ies (8), purified endonuclease V formed two stable complexes
with DNA containing deoxyinosine, both under optimal cleav-
age and noncleavage conditions. The estimated affinity of en-
donuclease V for dI-containing oligonucleotides in complexes I
and II were 2 and 60 nM, respectively. In contrast, the enzyme
exhibits a much weaker interaction with DNA containing base
mismatches. Under noncleavage conditions, i.e. in the presence
of NaCl, the $K_{dI}$ and $K_{dII}$ determined for complexes I and II
with mismatched DNA was 58 and 62 nM, respectively. Under
endonuclease V in complex I. Thus, agreement with our earlier studies (8) which suggest that complex II was formed at a high enzyme concentration (130 nM). Although less than 10% of complex I was formed between the enzyme and mismatch-containing DNA, close to 40% of complex II was formed at a high enzyme concentration (130 nM). These data suggest that the second molecule of the enzyme dissociates from mismatched DNA after cleavage. This is in sharp contrast to the interaction of the enzyme with deoxyinosine-containing DNA in which the enzyme remains bound to DNA containing deoxyinosine even after cleavage (Fig. 1) (8). The much larger Kd of the enzyme for DNA containing base mismatches further demonstrates a weaker interaction of endonuclease V with DNA containing mismatches. Interestingly, the Kd of the enzyme for DNA containing base matches is similar to that for DNA containing deoxyinosine. This is in agreement with our earlier studies (8) which suggest that complex I formed with different substrates were as follows, deoxyinosine >> deoxycytidine > deoxynebularine, single base mismatches. It is possible that this 6-keto group interacts specifically with positively charged group(s) of the protein through an ionic interaction. However, endonuclease V did not form stable complexes with DNA containing a guanine which possesses a 6-keto and a 2-amino group. It is possible that the 2-amino group of guanine extends outside the “binding pocket” of endonuclease V, thus providing steric hindrance to the interaction of endonuclease V with the 6-keto group. Furthermore, the ability of endonuclease V to recognize deoxyuridine in DNA, as demonstrated by the ability of endonuclease V to nick both single- and double-stranded DNA containing deoxyuridine, suggests that the 4-keto group of uracil can interact favorably with endonuclease V and deoxyuridine can also enter the binding pocket of the protein. Similar to guanine, the 5-methyl group of thymine might also provide steric hindrance to the unique interaction of endonuclease V with dU. The inability of endonuclease V to bind to thymine is thus reminiscent of the failure of uracil DNA N-glycosylase to recognize thymine in DNA (11, 12).

The interaction of endonuclease V with DNA can be best illustrated in Scheme 1 in which deoxyuridine is shown overlying with deoxyinosine, deoxycytidine with deoxyadenosine, and thymidine with deoxyguanosine. Endonuclease V interacts specifically with the 6-keto group of hypoxanthine and 4-keto group of uracil, probably through an ionic interaction (Scheme 1A). The complexes formed by the enzyme with duplex U/A is much less stable than those formed with duplex I/A. The weaker interaction of the enzyme with DNA containing deoxyuridine can be explained by the fact that the keto group of uracil (4-keto) is located in a position toward the imidazole ring of hypoxanthine, perhaps further away from the group(s) in the protein that give rise to the strong interaction with the 6-keto group of hypoxanthine. When the keto groups are replaced by amino groups (Scheme 1B), no specific interaction of endonuclease V with the DNA was observed, demonstrating the importance of the keto groups for the specific interaction. For thymine and guanine (Scheme 1C), although both bases have the required keto group for specific interaction with endonuclease V, the presence of a 5-methyl (thymine) or a 2-amino group (guanine) seems to abolish this specific interaction, presumably due to steric hindrance. However, endonuclease V can still interact with these bases in a strand specific manner when they are present in a base mismatches, probably due to changes in the secondary structures of DNA. Based on this model, it is thus expected that mutation of some specific amino acid residues might affect only the deoxyinosine-specific endonuclease activity but not the mismatch-specific activities of endonuclease V.

Acknowledgments—We thank Drs. Susan Wallace, Paul Doetsch, and Lois Rabow for helpful discussion and Dr. Paula Vertino for critical reading of the manuscript. We also thank the Protein Chemistry Laboratory, the University of Texas Medical Branch at Galveston, for N-terminal sequencing of endonuclease V.

REFERENCES
1. Guo, G., Ding, Y., and Wei, B. (1997) J. Bacteriol. 179, 310–316
2. Yao, M., and Kow, Y. W. (1996) J. Biol. Chem. 271, 50672–50676
3. Gates, F. T., III, and Linn, S. (1977) J. Biol. Chem. 252, 1647–1653
4. Demple, B., and Linn, S. (1982) J. Biol. Chem. 257, 2848–2855
5. Blattner, F. R., Burland, V. D., Plunkett, G., III, Sofia, H. J., and Daniels, D. L. (1993) Nucleic Acids Res. 21, 5408–5417
6. Yao, M., Hatabet, Z., Melamede, R. J., and Kow, Y. W. (1994) J. Biol. Chem. 269, 16260–16268
7. Yao, M., and Kow, Y. W. (1994) J. Biol. Chem. 269, 31389–31396
8. Yao, M., and Kow, Y. W. (1995) J. Biol. Chem. 270, 28609–28616
9. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
10. Altschul, S. F., Gish, W., Miller, W., Meyers, E. W., and Lipman, D. J. (1990) J. Mol. Biol. 203, 403–410
11. Savva, R., McAuley-Hecht, K., Brown, T., and Pearl, L. (1995) Nature 373, 487–493
12. Mol, C. D., Arvai, A. S., Slupphaug, G., Kavli, B., Alseth, I., Krokan, H. E., and Tainer, J. A. (1995) Cell 80, 869–878