Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 4193–4197, June 1990
Biochemistry

Yeast structural gene (*APN1*) for the major apurinic endonuclease: Homology to *Escherichia coli* endonuclease IV

(DNA repair/oxidative damage/alkylation damage/*Saccharomyces cerevisiae*)

SONYA C. POPOFF, ALEXANDER I. SPIRA, ARLEN W. JOHNSON*, AND BRUCE DEMPLE†

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138

Communicated by James C. Wang, March 26, 1990

ABSTRACT **DNA damage generated by oxygen radicals includes base-free apurinic/apyrimidinic (AP) sites and strand breaks that bear deoxyribose fragments. The yeast *Saccharomyces cerevisiae* repairs such DNA lesions by using a single major enzyme. We have cloned the yeast structural gene (*APN1*) encoding this AP endonuclease/3′-repair diesterase by immunological screening of a yeast genomic DNA expression library in λgt11. Gene disruption experiments confirm that the Apn1 protein accounts for ≥97% of both AP endonuclease and DNA 3′-repair diesterase activities in yeast cell-free extracts. The DNA and predicted amino acid sequences for the *APN1* gene are homologous to those for the *nfo* gene encoding DNA endonuclease IV of *Escherichia coli*. This conservation of structure between a eukaryotic enzyme and its prokaryotic counterpart underscores the fundamental nature of their roles in DNA repair.**

Active oxygen species—namely, superoxide radical (O_2^-), hydroxyl radical ($HO^\cdot$), and H_2O_2 can directly or indirectly damage cellular DNA. These species are produced during normal aerobic metabolism (1) or from exposure to exogenous agents such as ionizing radiation (2) or H_2O_2 (3). Prominent products of free radical attack on DNA are single-strand breaks with blocked 3′ termini that are refractory to DNA repair synthesis (4). These agents and the antitumor drugs bleomycin and neocarzinostatin also cause base loss to form apurinic/apyrimidinic (AP) sites in DNA (2, 5, 6). AP sites in DNA also originate from spontaneous or alkylation-induced depurination and from the action of DNA glycosylases, which remove a variety of damaged bases (7).

Enzymes that repair these types of DNA damage have been identified from both prokaryotic and eukaryotic cells (8, 9). Exonuclease III (encoded by the *xth* gene) (10) and endonuclease IV (the *nfo* gene product) (11) of *Escherichia coli* have been characterized extensively. Exonuclease III comprises ≈90% of the total AP endonuclease and 3′-repair diesterase activity in *E. coli* cell-free extracts, while endonuclease IV accounts for ≈5% of the total under normal growth conditions (9, 12). Endonuclease IV can be induced up to 10-fold by agents that generate intracellular superoxide (13). Both of these enzymes have relatively broad substrate specificity and initiate the repair of both DNA strand breaks with 3′ blocking groups and AP sites (4, 9, 12, 14, 15). Strains of *E. coli* bearing mutations in the *xth* locus are hypersensitive to H_2O_2 (16), while *nfo* mutants are hypersensitive to bleomycin and to *t*-butylhydroperoxide (11).

The major DNA 3′-repair diesterase and AP endonuclease from the yeast *Saccharomyces cerevisiae* has been purified and characterized (14, 15) and this eukaryotic enzyme has many biochemical properties in common with endonuclease IV from *E. coli*. Although AP endonucleases from other eukaryotes have been identified, the physiological role of these proteins has not been established, owing to a lack of enzyme-deficient mutants. We report here the isolation and characterization of the yeast structural gene encoding the major AP endonuclease/3′-repair diesterase, which we call *APN1* (for AP endonuclease).‡ The predicted Apn1 polypeptide is remarkably homologous to its bacterial counterpart, DNA endonuclease IV.

MATERIALS AND METHODS

Preparation of Antisera. The yeast AP endonuclease/3′-repair diesterase was purified from the yeast strain EJ2169 by a modification of a published procedure (14). Briefly, EDTA and dithiothreitol were omitted from all buffers, and protease inhibitors were added (aprotinin, leupeptin, pepstatin, each to 1 μg/ml; benzamide to 1 mM). These changes increased the stability and yield of the enzyme and eliminated the need for Affi-Gel Blue chromatography. For antibody production, the purified protein (50 μg) was emulsified in complete Freund's adjuvant, split into two aliquots, and injected subcutaneously into two female New Zealand rabbits. Four weeks after the primary injection, each rabbit was given two booster injections at 2-week intervals, each time with 12 μg of the purified yeast protein emulsified in incomplete Freund's adjuvant.

Polyclonal antiserum from one rabbit was purified by using an affinity column of the AP endonuclease/3′-repair diesterase. Purified protein (25 μg) was coupled to Affi-Gel 10 (Bio-Rad) or CNBr-activated Sepharose (Pharmacia) according to the conditions recommended by the supplier. Antiserum was applied to the affinity column and eluted according to the conditions reported by Dake *et al.* (17).

Yeast Extracts and Enzyme Assays. Yeast cell-free extracts were prepared by pelleting cells from logarithmic-phase cultures grown in 5 ml of YPD medium (18), washing the cells in cold extraction buffer [30 mM KCl/50 mM Tris·HCl, pH 7.5/10% (vol/vol) glycerol/1 mM phenylmethylsulfonyl fluoride/1 mM benzamide/aprotinin, leupeptin, and pepstatin A (1 μg/ml each)], and resuspending the cell pellets in 0.5 ml of extraction buffer. Cells were disrupted by mixing in a Vortex with glass beads (diameter, 420–600 μm) at top setting (10 × 15 sec). Cell debris and glass beads were removed by centrifugation at 12,000 × *g* for 10 min at 4°C. Protein concentrations were determined by the method of Bradford (19).

Yeast crude extracts were assayed for 3′-phosphoglycolaldehyde (3′-PGA) diesterase and AP endonuclease activities

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: AP, apurinic/apyrimidinic; 3′-PGA, 3′-phosphoglycolaldehyde.

*Present address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115.
†To whom reprint requests should be addressed.

‡The sequence reported in this paper has been deposited in the GenBank data base (accession no. M33667).

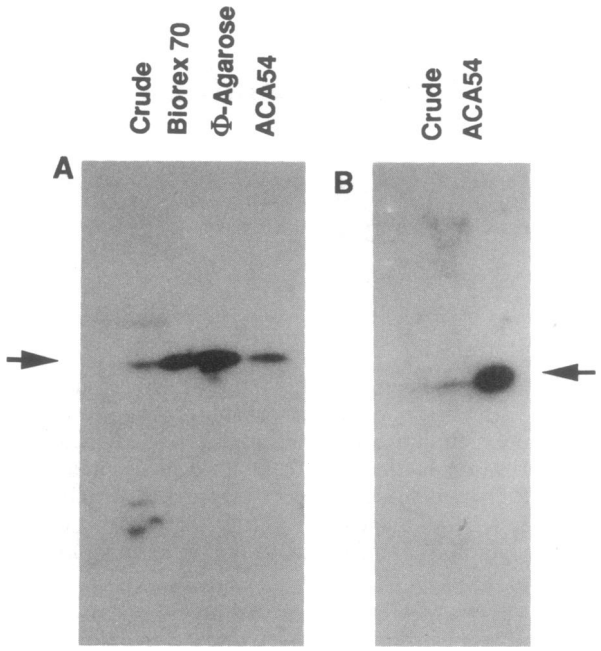

FIG. 1. Immunoblot analysis. Protein samples from either yeast crude extracts or purified fractions (as designated) were transferred from SDS/polyacrylamide gels to nitrocellulose membranes and were probed with affinity-purified (*A*) or phage-purified (*B*) sera. Bound antibody was detected with ^{125}I-labeled protein A and autoradiography. Arrows indicate M_r 40,500.

as described (14, 15). One unit of enzyme releases 1 pmol of product per min at 37°C. The 3′-PGA diesterase activity was determined after paper chromatographic analysis (14), unless noted otherwise.

Immunoprecipitation of enzymatic activities was performed by incubating yeast crude extracts with antibody preparations in a total vol of 20 μl in a buffer containing 50 mM Hepes (pH 7.6), 50 mM KCl, and bovine serum albumin (100 μg/ml). The mixtures were incubated on ice for a minimum of 4 hr, after which 10 μl of protein A agarose (Bio-Rad) was added. The samples were mixed and incubated on ice for ≥2 hr and centrifuged in a microcentrifuge at 4°C for 10 min. Aliquots of the supernatants were assayed for enzyme activities.

Hybridizations. The yeast genomic library in phage λgt11 and the host strain Y1090 (20) were obtained from Clontech (Palo Alto, CA). Phage from infected cells were transferred to nitrocellulose (Schleicher & Schuell) and probed with affinity-purified antiserum by standard procedures (21).

Table 1. Immunoprecipitation of DNA repair activities by specific antisera

Serum	3′-PGA diesterase, units/mg	AP endonuclease, units/mg
None	4.0	11.3
Preimmune		
5 μl	4.3	11.7
Immune		
1 μl	1.0	0.03
5 μl	0.47	0.07
Purified		
1 μl	1.10	ND
5 μl	0.3	ND

Cell-free extracts were from the yeast strain EJ2169, which was the same strain used in purification of the enzyme. 3′-PGA diesterase activity was determined without paper chromatography. ND, not determined.

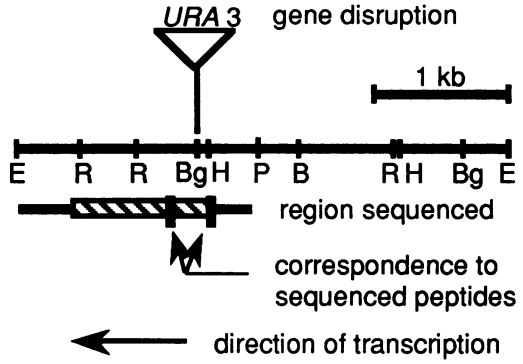

FIG. 2. Restriction map of the 3.8-kb *Eco*RI insert from phage λ6-1 DNA. *E*, *Eco*RI; R, *Eco*RV; Bg, *Bgl* II; H, *Hin*dIII; P, *Pvu* II; B, *Bam*HI. Hatched bar corresponds to the *APN1* coding region. A 3.8-kb *Bam*HI/*Bgl* II fragment (not shown to scale) containing the *URA3* gene was inserted at the *Bgl* II site.

Subcloning and DNA Sequencing. A 3.8-kilobase (kb) *Eco*RI fragment from one of the positive clones (λ6-1) was subcloned into the *Eco*RI site of plasmid pUC13 to generate pSCP8. Plasmid pSCP8 was sequenced with the Sequenase kit (United States Biochemical). Synthetic oligonucleotides (Operon Technologies, San Pablo, CA) corresponding to known peptide sequences from the yeast diesterase were used as sequencing primers. The peptides (single-letter code) and corresponding mixed oligonucleotide sequences were as follows: KYKFGAHM, 5′-AAI TA(T/C) AAI TT(C/T) GG(T/C/I) GC(T/C) CA(C/T) ATG-3′; MDDLNR, 5′-C(G/T) (A/G)TT (A/G/T/C)A(AG) (A/G)/TC (A/G)TC CAT (antisense strand); letters in parentheses indicate mixed nucleotides at those positions. Three additional oligonucleotides were synthesized (M. Byrne, Tufts University) as the DNA sequence was obtained; these corresponded to nucleotides 316–332, 604–620, and 827–843 of the DNA sequence shown in this paper. Portions of the *APN1* coding sequences and 5′ and 3′ flanking regions were sequenced by subcloning smaller DNA fragments into pUC13 and using pUC13/M13 forward and reverse universal primers. The entire DNA sequence was determined on both strands. Computer analysis of DNA sequences was done with Genetics Computer Group sequence analysis software (22).

A 2.0-kb *Pvu* II fragment containing the entire *APN1* coding region and *APN1* and vector flanking sequences was isolated from pSCP8 and subcloned into the *Sma* I site of the multicopy yeast vector yEp351 (23). The desired recombinants (yEpAPN1) were then used to transform the yeast strain NKY757 (*MATα, leu2, pep4-3, prb1-1122, trp1*) to Leu$^+$, using the LiOAc procedure described by Ito *et al.* (24).

Targeted Gene Disruptions. A 3.8-kb *Bam*HI/*Bgl* II fragment containing the yeast *URA3* gene was isolated from plasmid pNKY51 (25) and inserted into the *Bgl* II site of pSCP8 to disrupt the *APN1* gene. Linear DNA molecules resulting from *Eco*RI/*Bam*HI digestion were used to transform the yeast strain DBY747 (*MAT***a**, ura3, leu2, trp1, his3) by selection for growth in the absence of uracil. DNA isolated

Table 2. Dependence of DNA repair activities on the *APN1* gene

Strain	3′-PGA diesterase, units/mg	AP endonuclease, units/mg
DBY747 (*APN1*)	2.1	52
SCP762 (*apn1*::*URA3*)	0.07	1
NKY757 (yEp351)	1.3	31
NKY757 (yEpAPN1)	28.4	870

Activities were determined for cell-free extracts as described.

from Ura$^+$ transformants was analyzed on Southern blots to confirm the disruption in the yeast genome.

RESULTS

Antisera Specificity. The specificities of the antisera raised against the purified yeast AP endonuclease/3′-repair diesterase were examined by immunoblotting. Affinity-purified antiserum cross-reacted specifically with a protein in yeast crude extracts and purified fractions of M_r 40,500, the size of the purified polypeptide (Fig. 1*A*). This serum also cross-reacted weakly with other polypeptides that do not copurify with the endonuclease (Fig. 1*A*).

The results of immunoprecipitation experiments were consistent with the presence of a single major AP endonuclease/3′-repair diesterase in *S. cerevisiae*. Incubation of yeast cell-free extracts with either crude or affinity-purified antisera immunoprecipitated up to 92% of the 3′-repair diesterase activity (assayed as 3′-PGA diesterase) and up to 98% of the AP endonuclease activities (Table 1). Sera from two different rabbits were equally effective in these experiments (data not shown). These data provide independent confirmation of biochemical experiments showing that the 40.5-kDa polypeptide is the major DNA 3′-repair diesterase and AP endonuclease of yeast (14, 15).

Gene Cloning. Immunopurified antiserum against the yeast AP endonuclease/3′-repair diesterase was used to probe a yeast genomic expression library in λgt11. Six positive clones were identified among 9×10^5 plaques. The DNA inserts in these six phages were from 3 to 3.8 kb long. Restriction mapping and Southern hybridizations revealed that all six clones shared a common yeast DNA fragment (data not shown). A physical map of a 3.8-kb *Eco*RI insert from one of these phages (λ6-1) is shown in Fig. 2. Synthetic oligonucleotide probes, corresponding to amino acid sequences of the N terminus and an internal peptide of the protein, hybridized specifically to the yeast DNA fragments in these phages (data not shown). We have named this cloned yeast gene *APN1* (for AP endonuclease).

Additional evidence that the *APN1* gene was contained on the selected recombinant phages was supplied by immunological experiments. Filters impregnated with the proteins expressed by one phage (λ6-1) were used to adsorb antibodies from the crude antiserum against the diesterase. When these antibodies were eluted from the filters and used in Western blots, a single 40.5-kDa species was identified in both crude extracts and a purified enzyme fraction (Fig. 1*B*). Thus, the cloned gene encodes the antigenic determinants of the AP endonuclease/3′-repair diesterase.

Functional *APN1* Expression. Further confirmation that *APN1* encodes the main AP endonuclease/3′-repair diesterase of yeast was provided by gene disruption experiments. An *apn1*::*URA3* allele was constructed and used to replace the wild-type gene. The viability of haploid strains bearing the *apn1*::*URA3* allele demonstrates that Apn1 does not perform any essential function under normal growth conditions. The haploid *apn1*::*URA3* strains were highly deficient in both 3′-repair diesterase and AP endonuclease activities (≤3% of wild type; Table 2). Thus, the Apn1 protein is the main (if not the sole) AP endonuclease/3′-repair

```
-340  CCCGAGCACAAGAAAATAAAAAATTGAACAAAAATAAGAATAAAGTCACAAAGGGAAGAAGATGAGGGGGTTCTTTTTCCAGACTTTCTTCTTTTGGTCCTTCACCTTGGATATATGAAC  -221

-220  ATAGTAGTATTATTGAACACGAACTTTTTTCATCTAATACTTCTCTTTTTTTGCCAGTTTTTCTGGGAACTTGGAAACTGGAATTTTTCATTCTTACCTGGTTTCTTTTTCTTCATTTAT  -101

-100  TTACATGTTGGAAATCATTGTAAATATAAGATGACAAACTCCGAATAAGAAACACAAAACGCAACATTAATAAGCTTTTGGCATATCGGAACCATCGTAATGCCTTCGACACCTAGCTTT    21
                                                                                                                  M  P  S  T  P  S  F

  22  GTTAGATCTGCTGTCTCGAAATACAAATTTGGTGCGCACATGTCAGGTGCCGGTGGAATTTCTAATAGTGTAACTAATGCATTTAACACTGGCTGTAATTCGTTTGCCATGTTTTTAAAA  141
      V  R  S  A  V  S  K  Y  K  F  G  A  H  M  S  G  A  G  G  I  S  N  S  V  T  N  A  F  N  T  G  C  N  S  F  A  M  F  L  K

 142  TCTCCAAGAAAGTGGGTTTCTCCGCAGTATACACAGGAGGAAATAGATAAATTTAAGAAAAACTGTGCAACTTACAATTATAACCCATTGACGGACGTCTTACCGCATGGCCAATACTTT  261
      S  P  R  K  W  V  S  P  Q  Y  T  Q  E  E  I  D  K  F  K  K  N  C  A  T  Y  N  Y  N  P  L  T  D  V  L  P  H  G  Q  Y  F

 262  ATTAATTTGGCTAATCCGGATAGGGAAAAGGCAGAAAAGAGTTATGAGTCATTCATGGACGATCTGAACAGGTGTGAGCAATTAGGGATAGGCCTGTATAACTTACATCCAGGTTCTACT  381
      I  N  L  A  N  P  D  R  E  K  A  E  K  S  Y  E  S  F  M  D  D  L  N  R  C  E  Q  L  G  I  G  L  Y  N  L  H  P  G  S  T

 382  TTGAAAGGAGACCATCAGTTGCAGTTAAAACAACTAGCCTCATATTTGAACAAAGCAATTAAGGAAACGAAATTTGTTAAAATTGTATTAGAAAATATGGCTGGTACTGGAAATTTAGTA  501
      L  K  G  D  H  Q  L  Q  L  K  Q  L  A  S  Y  L  N  K  A  I  K  E  T  K  F  V  K  I  V  L  E  N  M  A  G  T  G  N  L  V

 502  GGAAGTTCTCTGGTAGATTTGAAGGAGGTTATTGGAATGATTGAAGATAAATCAAGAATCGGCGTTTGCATAGATACATGCCATACATTTGCAGCAGGCTACGATATCAGCACCACTGAA  621
      G  S  S  L  V  D  L  K  E  V  I  G  M  I  E  D  K  S  R  I  G  V  C  I  D  T  C  H  T  F  A  A  G  Y  D  I  S  T  T  E

 622  ACGTTTAACAATTTTTGGAAAGAGTTCAATGATGTAATAGGGTTCAAATATCTAAGTGCTGTTCACTTAAATGATTCCAAGGCTCCTTTAGGAAGCAATAGAGATTTGCATGAACGCTTG  741
      T  F  N  N  F  W  K  E  F  N  D  V  I  G  F  K  Y  L  S  A  V  H  L  N  D  S  K  A  P  L  G  S  N  R  D  L  H  E  R  L

 742  GGTCAAGGTTATTTGGGTATAGATGTGTTTAGAATGATCGCGCACTCTGAATACCTGCAGGGTATCCCTATTGTCTTGGAAACTCCATACGAAAATGATGAAGGCTATGGTAATGAAATC  861
      G  Q  G  Y  L  G  I  D  V  F  R  M  I  A  H  S  E  Y  L  Q  G  I  P  I  V  L  E  T  P  Y  E  N  D  E  G  Y  G  N  E  I

 862  AAACTTATGGAGTGGTTGGAATCAAAGAGTGAGAGCGAATTGTTAGAAGACAAGGAGTATAAAGAAAAAAATGATACGCTGCAGAAATTAGGTGCTAAATCGCGTAAGGAACAGTTAGAC  981
      K  L  M  E  W  L  E  S  K  S  E  S  E  L  L  E  D  K  E  Y  K  E  K  N  D  T  L  Q  K  L  G  A  K  S  R  K  E  Q  L  D

 982  AAGTTTGAGGTTAAACAAAAGAAGCGAGCTGGGGGCACCAAGAGGAAGAAAGCCACTGCAGAGCCAAGCGATAATGATATCTTGTCACAAATGACAAAGAAGAGGAAGACTAAGAAAGAA  1101
      K  F  E  V  K  Q  K  K  R  A  G  G  T  K  R  K  K  A  T  A  E  P  S  D  N  D  I  L  S  Q  M  T  K  K  R  K  T  K  K  E

1102  TAATGAGAAGCGAGAAGAATTTTAAATACGTAATCAATTTTTGTAGATTATCTCAACGTACGTAACTAAATGCTACCCTTCGTCATCCAAGATGCTGTTGCCCATTTCATAGAAACTACT  1221
      *

1222  TACTTGGTTTATAGATGTTTGCGTTGCGTTACAGAATAAGGCAGCCTTTTTTTTTTTTAGATTTTCATTTTTCACATACTGATGGATTAGCAGATTATAAAAGGGTAAAGTTAAAATCAT  1341

1342  TCCAAACTTGAACTAAGGAAGGCGCTAGACCATCCTCTTATGGAACGTTATTCCTAAGCATAGAAAGTTTTTTACTTGTAATCTTGAACGCGGGGTGTGTTCTTTATCTTTTTTTTCTTT  1461

1462  TTTACACAATTAGAC  1476
```

FIG. 3. *APN1* coding sequence with 5′ and 3′ flanking regions. Numbers correspond to nucleotide number with base pair 1 as the first nucleotide of the coding sequence. The translated amino acid sequence appears under each codon in single-letter code, with the asterisk indicating the predicted termination codon. Underlined amino acids correspond to those matching known Apn1 peptide sequences. Basic residues in the last 80 amino acids of the C terminus of Apn1 are double underlined. Possible TATA box and polyadenylylation sites are indicated with dotted underlines.

diesterase of yeast, consistent with the biochemical (14, 15) and immunological experiments (Table 1) cited above. These experiments underscore the conclusion that Apn1 combines 3′-repair diesterase and AP endonuclease functions in a single protein. Indeed, both activities are overproduced to a similar degree in strains bearing *APN1* on a multicopy plasmid (Table 2). Immunoblot experiments (data not shown) show that these overproducing strains contain high levels of a 40.5-kDa polypeptide.

DNA Sequence of the *APN1* Gene. Further verification that the cloned DNA encoded the AP endonuclease/3′-repair diesterase was obtained from DNA sequencing. Double-stranded sequencing was primed by using synthetic oligonucleotides that corresponded to N-terminal and internal amino acid sequences of the protein (see *Materials and Methods*). The resulting DNA sequences predicted peptides that matched the known amino acid sequences. This information provided the direction of transcription within the cloned DNA and the physical distance between these peptides (Fig. 2).

The entire DNA sequence of *APN1* and its flanking regions was obtained by a combination of subcloning and synthesis of additional DNA primers (Fig. 3). There is a single long open reading frame that would encode a polypeptide of M_r 41,430, close to the M_r 40,500 measured for the purified protein (14).

Homology to *E. coli* Endonuclease IV. When the *APN1* DNA sequence and predicted Apn1 polypeptide were compared with the GenBank DNA and National Biomedical Research Foundation protein data bases, respectively, no striking similarities were found. However, the *APN1* gene and predicted Apn1 protein are homologous to, respectively, the *E. coli nfo* gene and its product, endonuclease IV (26).§ The enzymatic features of the Apn1 and endonuclease IV proteins are similar (12, 14, 15). A one to one comparison revealed 55% DNA sequence identity and 41% identity for the predicted polypeptides for the Apn1 protein compared to the Nfo protein (Fig. 4 *Upper*). This homology occurs along nearly the entire length of the endonuclease IV polypeptide, interrupted by a relatively short region of poorer homology at residues 115–125 in Nfo (Fig. 4 *Lower*). The Apn1 polypeptide contains an additional 80 residues at the C terminus compared to Nfo. This C-terminal region is notable for its high density of basic amino acids (30% of the last 80 residues are arginine or lysine; Fig. 3).

```
 17 FGAHMSGAGGISNSVTNAFNTGCNSFAMFLKSPRKWVSPQYTQEEIDKFK 66
    ||| | ||| |   |       || | |  | |     |   || ||
  4 IGAHVSRAGGLANAAIRAAEIDATAFALFTKNQRQWRAAPLTTQTIDEFK 53

 67 KNCATYNYNPLTDVLPHGQYFINLANPDREKAEKSYESFMDDLNRCEQLG 116
    |  | |      |||  | |||  |  |  |||   | |   ||||||
 54 AACEKYHYTS.AQILPHDSYLINLGHPVTEALEKSRDAFIDEMQRCEQLG 102

117 IGLYNLHPGSTLKGDHQLQ.LKQLASYLNKAIKETKFVKIVLENMAGTGN 165
    | | |||| |         |   |  | |   |  |  | || || |
103 LSLLNFHPGSHLMQISEEDCLARIAESINIALDKTQGVTAVIENTAGQGS 152

166 LVGSSLVDLKEVIGMIEDKSRIGVCIDTCHTFAAGYDISTTETFNNFWKE 215
    |    |  |  ||||| |||||||| ||||||  |
153 NLGFKFEHLAAIIDGVEDKSRVGVCIDTCHAFAAGYDLRTPAECEKTFAD 202

216 FNDVIGFKYLSAVHLNDSKAPLGSNRDLHERLGQGYLGIDVFRMIAHSEY 265
    |   |||||  |||| |  ||  | |  || |  | | || |
203 FARTVGFKYLRGMHLNDAKSTFGSRVDRHHSLGEGNIGHDAFRWIMQDDR 252

266 LQGIPIVLETPYENDEGYGNEIKLMEWLESKSESE 300 Apn1
    |||  |||  |      ||
253 FDGIPLILET..INPDIWAEEIAWLKAQQTEKAVA 285 Nfo
```

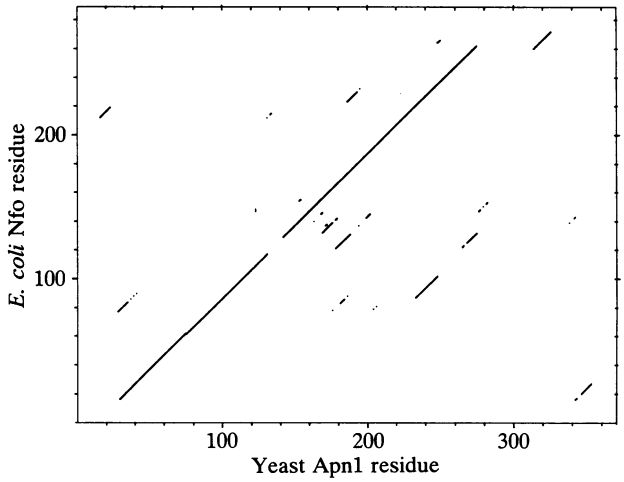

FIG. 4. Comparison of Apn1 and endonuclease IV sequences. (*Upper*) Detailed comparison of selected Apn1 (top strand) and Nfo (bottom strand) amino acid sequences. Sequences were aligned by using the program BESTFIT (22). Exact matches are indicated with lines and gaps are indicated by dots. (*Lower*) Dot plot of amino acid homology between the yeast Apn1 and *E. coli* Nfo proteins. The entire translated peptides of each sequence (amino acids 1–368 for Apn1 and amino acids 1–286 for Nfo) were compared. This was generated by using the Genetics Computer Group programs DOTPLOT (22) (window, 30 residues; stringency, 14 matches minimum).

DISCUSSION

We used an immunological screen to clone the yeast structural gene (*APN1*) encoding the major DNA AP endonuclease/3′-repair diesterase. Engineered disruptions of *APN1* caused the simultaneous loss of >97% of both AP endonuclease and DNA 3′-repair diesterase activities in cell-free extracts, which confirms that both activities reside in a single protein. The predicted Apn1 polypeptide exhibits substantial identity to the amino acid sequence of the bacterial DNA repair enzyme endonuclease IV. This homology is reflected in the action of these two enzymes on the same range of damaged DNA substrates.

The *APN1* gene is distinct from known yeast genes. First, the *APN1* sequence has no identity with yeast sequences in the GenBank/EMBL data bases. The *APN1* gene is on chromosome XI, as determined by hybridization to yeast chromosomes separated electrophoretically (unpublished data). No *RAD* gene or other DNA repair gene has been mapped to chromosome XI (27). These results are consistent with previous work showing that various *rad* mutants have normal levels of the 3′-PGA diesterase (14).

The homology between the yeast Apn1 and bacterial endonuclease IV sequences is important for several reasons. This structural conservation provides for the enzymatic functions common to these two proteins (12, 14, 15). In turn, the evolutionary retention of these DNA repair functions indicates that they fulfill fundamental biological roles.

The yeast and bacterial enzymes are both transition metalloproteins (J. D. Levin and B.D., unpublished data) (12, 14, 15). Some of the homologous sequences might therefore constitute metal-binding sites within the enzymes. The mechanism by which these endonucleases identify specific damages (e.g., AP sites) within long stretches of undamaged DNA is of general interest and must be reflected in their homologous structures. Finally, both Apn1 and endonuclease IV are members of a rare class of enzymes that contain both 3′-phosphomonoesterase and 3′-phosphodiesterase activities (12, 15). The structure(s) that accounts for this unusual feature must also reside in polypeptide sequences shared by Apn1 and its *E. coli* counterpart.

The possible function of the highly basic C-terminal region of Apn1 is unclear. Since this sequence is absent from

§At the time of the most recent search (January 1990) the *nfo* gene sequence was not in the data base and was kindly provided to us by R. Cunningham (State University of New York, Albany, NY).

endonuclease IV, it appears unnecessary for damage recognition and phosphoester cleavage of DNA *in vitro*. Perhaps such a presumptively cationic region interacts with DNA in eukaryotic chromatin in a way that allows access of the enzyme to DNA damages.

Both the Apn1 enzyme and endonuclease IV repair lethal DNA damages that are caused either by oxygen radicals or by alkylating agents. One key difference between the yeast enzyme and endonuclease IV is at the level of gene expression. Apn1 endonuclease activity is produced at a fairly constant level under a variety of conditions (ref. 14; S.C.P., unpublished data), while endonuclease IV is inducible by agents that generate intracellular superoxide (13). The major constitutive AP endonuclease/3′-repair diesterase of *E. coli* is exonuclease III (4, 11), a polypeptide whose sequence (10) does not bear homology to endonuclease IV or Apn1. The inducibility of endonuclease IV probably provides for increased repair of DNA lesions that are attacked inefficiently by exonuclease III.

S. cerevisiae seems to be unusual in having such a dominant single activity to repair both AP sites and deoxyribose damages. *E. coli* and human cells (D. S. Chen and B.D., unpublished data) each have at least two major enzymes for these functions, while additional activities can be detected in those organisms if sensitive methods are used (9, 11, 28). This dependence on a single enzyme is consistent with, for example, the 10-fold increased sensitivity of *apn1::URA3* strains to methyl methanesulfonate and to hydrogen peroxide compared to *APN1* strains (unpublished data). These sensitivities indicate that the Apn1 enzyme is required to initiate multiple DNA repair pathways. It is possible, of course, that another such enzyme in yeast has escaped our detection. However, no additional AP endonuclease or 3′-repair diesterase activity was revealed in crude extracts when various supplements were added to the enzyme reactions (e.g., Mg^{2+}).

The availability of the cloned *APN1* gene now allows several fundamental issues to be addressed—for example, the ability of the yeast enzyme to substitute for the *E. coli* enzymes in the repair of both oxidative and alkylation damages and the role of Apn1 in limiting both spontaneous and mutagen-induced mutagenesis.

We thank R. P. Cunningham for providing us with the *nfo* DNA sequence, Nancy Kleckner and Edith Gralla for supplying yeast strains and plasmids, and Edith Gralla for helpful comments on the manuscript. This work was supported by a grant to B.D. from the National Institutes of Health (ES03926). B.D. is a Dreyfus Foundation Teacher-Scholar.

1. Halliwell, B. (1987) *FASEB J.* **1**, 358–364.
2. Hutchinson, F. (1985) *Prog. Nucleic Acid Res. Mol. Biol.* **32**, 115–154.
3. Imlay, J. A. & Linn, S. (1988) *Science* **240**, 1302–1309.
4. Demple, B., Johnson, A. & Fung, D. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7731–7735.
5. Rabow, L., Stubbe, J., Kozarich, J. W. & Gerlt, J. A. (1986) *J. Am. Chem. Soc.* **108**, 1730–1731.
6. Povirk, L. F. & Goldberg, I. R. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3182–3186.
7. Lindahl, T. (1982) *Annu. Rev. Biochem.* **51**, 61–87.
8. Wallace, S. S. (1988) *Environ. Mol. Mut.* **12**, 431–477.
9. Weiss, B. & Grossman, L. (1987) *Adv. Enzymol.* **60**, 1–34.
10. Saporito, S. M., Smith-White, B. J. & Cunningham, R. P. (1988) *J. Bacteriol.* **170**, 4542–4547.
11. Cunningham, R. P., Saporito, S., Spitzer, S. G. & Weiss, B. (1986) *J. Bacteriol.* **168**, 1120–1127.
12. Levin, J. D., Johnson, A. W. & Demple, B. (1988) *J. Biol. Chem.* **263**, 8066–8071.
13. Chan, E. & Weiss, B. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 3189–3193.
14. Johnson, A. W. & Demple, B. (1988) *J. Biol. Chem.* **263**, 18009–18016.
15. Johnson, A. W. & Demple, B. (1988) *J. Biol. Chem.* **263**, 18017–18022.
16. Demple, B., Halbrook, J. & Linn, S. (1983) *J. Bacteriol.* **153**, 1079–1082.
17. Dake, E., Hofmann, T. J., McIntire, S., Hudson, A. & Zassenhaus, H. P. (1988) *J. Biol. Chem.* **263**, 7691–7702.
18. Sherman, F., Fink, G. & Hicks, J. (1983) *Methods in Yeast Genetics* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
19. Bradford, M. M. (1976) *Anal. Biochem.* **76**, 248–254.
20. Young, R. A. & Davis, R. W. (1983) *Science* **222**, 778–782.
21. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1987) *Current Protocols in Molecular Biology* (Wiley, New York).
22. Devereux, J. R., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
23. Hill, J. E., Myers, A. M., Koerner, T. J. & Tzagoloff, A. (1986) *Yeast* **2**, 163–167.
24. Ito, H., Fukuda, Y., Kousaku, M. & Kimura, A. (1983) *J. Bacteriol.* **153**, 163–168.
25. Alani, E., Cao, L. & Kleckner, N. (1987) *Genetics* **116**, 541–545.
26. Saporito, S. M. & Cunningham, R. P. (1988) *J. Bacteriol.* **170**, 5141–5145.
27. Mortimer, R. K., Schild, D., Contopoulou, C. R. & Kans, J. A. (1989) *Yeast* **5**, 321–403.
28. Bernelot-Moens, C. & Demple, B. (1989) *Nucleic Acids Res.* **17**, 587–600.