A Nick-sensing DNA 3′-Repair Enzyme from Arabidopsis*

Received for publication, February 11, 2002, and in revised form, April 2, 2002
Published, JBC Papers in Press, April 10, 2002, DOI 10.1074/jbc.M201411200

Stefania Petrucco‡, Giorgia Volpi, Angelo Bolchi, Claudio Rivetti, and Simone Ottonello
From the Department of Biochemistry and Molecular Biology, University of Parma, Parma 43100, Italy

DNA single-strand breaks, a major cause of genome instability, often produce unconventional end groups that must be processed to restore terminal moieties suitable for reparative DNA gap filling or ligation. Here, we describe a bifunctional repair enzyme from Arabidopsis (named AtZDP) that recognizes DNA strand breaks and catalyzes the removal of 3′-end-blocking lesions. The isolated C-terminal domain of AtZDP is by itself competent for 3′-end processing, but not for strand break recognition. The N-terminal domain instead contains three Cys3-His zinc fingers and recognizes various kinds of damaged double-stranded DNA. Gapped DNA molecules are preferential targets of AtZDP, which bends them by –73° upon binding, as measured by atomic force microscopy. Potential partners of AtZDP were identified in the Arabidopsis genome using the human single-strand break repairosome as a reference. These data identify a novel pathway for single-strand break repair in which a DNA-binding 3′-phosphoesterase acts as a “nick sensor” for damage recognition, as the catalyst of one repair step, and possibly as a nucleation center for the assembly of a fully competent repair complex.

DNA single-strand breaks (SSBs) caused by endogenously produced reactive oxygen species as well as by a number of oxidizing agents, ionizing radiation, or radiomimetic chemicals are a major source of genomic instability (1). Similar damage can also arise during DNA base excision repair, recombination, and other DNA transactions such as those controlled by eukaryotic DNA topoisomerases (2). Such lesions are usually rapidly cured by dedicated DNA polymerases and ligases. Gap filling and ligation by these enzymes can only occur if 3′-hydroxyl and 5′-phosphate termini are available at the boundaries of the damaged site. However, DNA strand breaks often bear unconventional end groups such as 3′-phosphate or 3′-phosphoglycolate as well as 5′-hydroxyl terminal moieties that must be removed or phosphorylated prior to reparative polymerization and/or ligation (1). DNA repair enzymes specialized in this task have been identified in a number of species ranging from bacteria to mammals (3–5). In eukaryotes, the paradigm of this class of enzymes is human polynucleotide kinase 3′-phosphatase, a bifunctional protein that acts as a phosphatase in the removal of 3′-phosphate-blocking lesions as well as a kinase in restoring 5′-phosphate termini (6, 7). Although fully competent in the restoration of repair-prone termini, human polynucleotide kinase 3′-phosphatase is by itself unable to locate damaged DNA sites. In fact, the recruitment of human polynucleotide kinase 3′-phosphatase, along with DNA ligase III and DNA polymerase β, onto damaged DNA is thought to be mediated by the repair complex assembly protein XRCC1, which in turn interacts with the nick-sensing enzyme poly-(ADP-ribose) polymerase-1 (PARP-1) (8–11). The latter is a multifunctional nuclear protein that, in addition to interacting with XRCC1, recognizes SSBs and induces a constrained bent configuration on its target DNA. Once activated by DNA binding, PARP-1 catalyzes the addition of negatively charged ADP-ribose moieties both to itself and to surrounding nuclear proteins (11–13). SSB recognition by PARP is mediated by an N-terminal DNA-binding domain containing two unusually long Cys3-His zinc fingers, separated by a 68-amino acid long spacer (13, 14). Interestingly, a single PARP-like zinc finger (enabling binding to nicked DNA) is also present in the repair enzyme DNA ligase III (15).

A structural and functional homolog of human polynucleotide kinase 3′-phosphatase (endowed with both 3′-phosphatase and 5′-kinase activities) has been recently identified in fission yeast (16). A related DNA 3′-phosphatase has also been described in Saccharomyces cerevisiae, which by contrast does not include 5′-kinase activity (17, 18).

Only limited information is presently available on the enzymology of DNA SSB repair (SSBR) in plants (19), organisms that are particularly exposed to a number of potentially genotoxic agents such as ozone, UV light, and environmental pollutants. Genes encoding PARP-homologous proteins with a structural organization resembling that of animal PARPs have been described in higher plants (20, 21). More recently, the first plant enzyme catalyzing the repair of DNA 3′-blocking lesions has been identified in maize (22). This enzyme (named ZmDP2) shares homology with polynucleotide kinase 3′-phosphatases, but, similarly to the S. cerevisiae enzyme, is devoid of an associated 5′-kinase activity. A conceptually assembled cDNA sequence from the Arabidopsis genome encodes the closest homolog of ZmDP2 (22). Relative to all known DNA 3′-repair proteins, the predicted Arabidopsis polypeptide includes a long N-terminal extension (532 amino acids) with sequence features suggestive of a role in DNA binding. To verify such a prediction and to gain new insight into SSBR in plants, we have isolated and functionally characterized this novel, putative 3′-end-processing enzyme from Arabidopsis. This enzyme (named AtZDP for Arabidopsis thaliana zinc finger DNA 3′-phosphoesterase) is a multifunctional modular protein with a C-terminal 3′-phosphoesterase domain and an N-terminal DNA-binding domain containing three PARP-like zinc fingers. AtZDP specifically binds to gap sites and sharply bends its target DNA. This
structural distortion, which marks damaged DNA, may also act as a nucleation center for the subsequent assembly of a fully competent DNA repair complex. Based on this view, on the peculiar structural organization of AtZDP, and on its unique ability to act as a nick-sensing 3'-phosphoesterase, we also carried out a whole genome analysis aimed to identify all the other putative components that are required to build up a fully competent SSBR complex in plants.

**EXPERIMENTAL PROCEDURES**

**RNA Analyses**—Seed sterilization, germination, and hydroponic culture of *A. thaliana* seedlings (ecotype Columbia) were carried out following the protocols recommended by the Arabidopsis Biological Resource Center. Total RNA was isolated from 14-day-old Arabidopsis seedlings as described previously (23). For the RNase protection assay, the hybridization probe used was a [32P]-labeled antisense RNA transcribed in vitro with T7 RNA polymerase (Amersham Biosciences). Digestion of the pGEM-T-Easy vector (Promega, Madison, WI), carrying a 326-bp genomic fragment of AtZDP (from positions −242 to +84 with respect to the ATG initiator codon), with the MboII restriction enzyme was used to produce a truncated template for in vitro transcription reactions, yielding a riboprobe of 269 nucleotides, including 62 nucleotides of plasmid-derived sequence. For hybridization reactions, Arabidopsis total RNA (15 μg/assay) and control yeast RNA (4 μg/assay) were incubated overnight at 42 °C with the radiolabeled probe. Hybridization and RNase digestion conditions were as previously described (22). RNase protection assays were performed by precipitation with an equal volume of 5 M ammonium acetate and 2 volumes of absolute ethanol at −20 °C for 30 min, followed by centrifugation at 4 °C for 1 h. The RNase-protected probes were recovered by propanol precipitation after centrifugation. The RNA samples were denatured with sodium hydroxide and 8 M urea in 1 M Tris·HCl (pH 7.5) and precipitated with ethanol. For the RNase protection assay, nuclear RNA was isolated from Arabidopsis seedlings as described previously (23). For the RNase protection assay, the hybridization probe used was a 32P-labeled antisense RNA transcribed in vitro with T7 RNA polymerase (Amersham Biosciences). Digestion of the pGEM-T-Easy vector (Promega, Madison, WI), carrying a 326-bp genomic fragment of AtZDP (from positions −242 to +84 with respect to the ATG initiator codon), with the MboII restriction enzyme was used to produce a truncated template for in vitro transcription reactions, yielding a riboprobe of 269 nucleotides, including 62 nucleotides of plasmid-derived sequence. For hybridization reactions, Arabidopsis total RNA (15 μg/assay) and control yeast RNA (4 μg/assay) were incubated overnight at 42 °C with the radiolabeled probe. Hybridization and RNase digestion conditions were as previously described (22). RNase protection assays were performed by precipitation with an equal volume of 5 M ammonium acetate and 2 volumes of absolute ethanol at −20 °C for 30 min, followed by centrifugation at 4 °C for 1 h. The RNase-protected probes were recovered by propanol precipitation after centrifugation. The RNA samples were denatured with sodium hydroxide and 8 M urea in 1 M Tris·HCl (pH 7.5) and precipitated with ethanol.

**Sequence Analysis**—DNA sequencing was performed with the Thermo-Sequenase cycle sequencing kit (Amersham Biosciences). Nucleotide and amino acid sequence analyses were conducted at the Baylor College of Medicine Search Launcher web interface and at the Expert Protein Analysis System Proteomics Server of the Swiss Institute of Bioinformatics, respectively.

**Expression and Purification of Recombinant AtZDP.**—The full-length protein coding region of the AtZDP cDNA (1917 bp) was PCR-amplified (30 cycles) using 100 ng of an Arabidopsis cDNA library (25) as template and a high-fidelity thermophilic DNA polymerase (Vent, New England Biolabs Inc., Beverly, MA). The sequence-specific CpoI-tailed (+) and (−) primers used were ACGCGGTCCGATGCCGGTGGTTGC- TACTTG (primer 1) and CTCCGGACCGCTAAGTCCCTGGCGATG- TACTTG (primer 2) (translation start and stop codons are underlined), respectively. For expression of the isolated N-terminal domain (DBD), the AtZDP (+)-primer 1 was used in combination with the sequence-specific CpoI-tailed primer 3 (GAGCGGACCTATCTCCTCATCCTATCATT). The C-terminal domain (CD) encoding cDNA was amplified with the CpoI-tailed primer 4 (ACCGGGTCGATGAACTTGAGTGA- GTCACTTCTCAGGTC) and the AtZDP (−)-primer 2. Restriction fragments resulting from CpoI digestions of PCR products were then ligated in-frame into the CpoI site created at position 237 of the pET28b (+) expression vector (Novagen, Madison, WI). Clones carrying inserts of the expected size, in a correct orientation relative to the T7 promoter, were identified by restriction mapping. The pET-AtZDP constructs were sequenced and used for the transformation of BL21-Codon Plus (Novagen, Madison, WI) cells (Novagen). Protein expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside, followed by a 4-h incubation at 30 °C. After cell lysis, recombinant proteins were loaded onto a Talon metal affinity column (CLONTech, Palo Alto, CA) equilibrated with 10% glycerol, 300 mM NaCl, 50 mM Tris·HCl (pH 7.5), 0.1 mM benzamidine, and 0.1 mM phenylmethylene sulfonyl fluoride. The column was washed until the A280 of the flow-through was <0.05, and the protein was eluted with 250 mM imidazole in the same buffer. Protein concentration was determined following the Bradford method (26) using bovine serum albumin as a standard. The composition and purity of protein fractions were assessed by SDS-10% polyacrylamide gel electrophoresis (27). For immunoblot analysis, a polyclonal antibody raised against the purified maize ZmDP2 protein (22) and a conjugated anti-rabbit secondary antibody were utilized as described in the instructions included with the ECL Western blotting system (Amersham Biosciences).

**Phosphoesterase Assays.**—For DNA 3'-phosphatase activity, a 1-nucleotide gapped 45-bp duplex, either with or without a 3'-phosphate at the gap site, was prepared as described in detail elsewhere (22). A 3'-phosphorylate-modified DNA was prepared as described (28). As for the RNase protection assay, to run at 37 °C for 15-μl reaction mixtures containing 100 mM KCl, 50 mM Tris·HCl (pH 7.5), 10% glycerol, 5 mM dithiothreitol, and 100 μM ZnCl2 plus the substrate (specific activity of 0.25 Ci/pmol) and enzyme concentrations specified below. Where indicated, 30 ng of supercoiled plasmid DNA was included in the reaction mixture. After blocking with denaturing loading dye, reaction products were resolved on 8% sequencing gels. Phosphorimages of dried gels were acquired with a Personal Imager FX Bio-Rad and analyzed using the Multi-Analyset PC software (Bio-Rad). Nonlinear regression analysis of phosphorimaging data was performed with SigmaPlot (SPSS Inc., Chicago, IL).

**DNA Binding Assay.**—For electrophoretic mobility shift assays, a 45-mer oligonucleotide (6) was end-labeled with polynucleotide kinase (Amersham Biosciences) and [γ-32P]ATP to a specific activity of 0.12 μCi/pmol and annealed with the two complementary 25- and 21-mer oligonucleotides to produce a 5'-labeled 1-nucleotide gapped 45-bp duplex. To generate a 45-bp duplex with a blocked 3'-end at the gap site, a 21-mer with a 3'-phosphate group was used in the annealing reaction along with the 23-mer and the labeled 45-mer. For the labeled 45-bp intact duplex, the 45-mer was hybridized with the complementary 25-mer oligonucleotide. Radiolabeled duplexes (0.1 pmol) were incubated on ice with 25 μl of electrophoretic mobility shift assay buffer (100 mM KCl, 50 mM Tris·HCl (pH 7.5), 10% glycerol, 10 mM MgCl2, 1 mM dithiothreitol, and 100 μM ZnCl2) containing 60 μg/ml bovine serum albumin and varying amounts (1–8 pmol) of the different AtZDP species. Unless otherwise indicated, 30 ng of Hoechst III-digested pBluescript plasmid DNA was included in the reaction to prevent the formation of large nonspecific protein-DNA aggregates. Competition experiment assays were carried out in the presence of 15, 30, or 150 ng of unlabeled 45-bp duplexes prepared as described above or in the presence of a nicked 43-bp duplex obtained from the annealing of a 43-mer (ATTG- ACGGGATCTCTAGAATTGCAGATTGTACCT) with the complementary 24-mer (AATTCCTTAGAGCATCTGCTTAAATT) and 19- mer (AATGACACCTGCGGATCCG). After 30 min, samples were loaded onto nondenaturing 5% polyacrylamide gels prerun at 100 V for 30 min at 4 °C. Electrophoresis was carried out at 150 V for 4–6 h at 4 °C in 1× Tris borate/EDTA. Gels were dried and subjected to phosphorimaging analysis. Zinc-free AtZDP was obtained by incubating (30 min at 20 °C) the protein with 2 mM EDTA, followed by a 20-h dialysis against electrophoretic mobility shift assay buffer to remove the excess EDTA. After dialysis, the oligonucleotide was removed by gel filtration. The 657-bp DNA template harboring a 1-nucleotide gap at 320 bp from one end was constructed as previously described (29). AtZDP-DNA complexes were assembled in 10 mM KCl, 5 mM Tris·HCl (pH 7.5), and 10 μM ZnCl2 and incubated for 30 min at room temperature prior to AFM imaging. Reactions were then diluted with 4 mM Hepes (pH 7.4), 10 mM NaCl, and 14 mM MgCl2 to obtain a final DNA concentration of ~1 ng and immediately deposited onto freshly cleaved ruby mica (Ted Pella Inc., Redding, CA). Samples were incubated for 2 min, rinsed with water, and blown dry on an AFM. AFM images were collected in air with a NanoScope III microscope (Digital Instruments Inc., Santa Barbara, CA) operating in tapping mode and equipped with a type E scanner. Olympus silicon nitride tips were used. All 512 × 512 pixel images were collected with a scan size of 2 μm at a scan rate varying between 2 and 4 scan lines/s. AFM images were analyzed using the ALEX software (30). Image integer values of the NanoScope file were converted to nanometers using the relation supplied with the NanoScope III documentation. Images were flattened by subtracting from each scan line a least-square fitted polynomial. No additional filters were applied to the images.

**DNA Analyses**—DNA sequencing was performed with the dyeoxy chain termination method using the Thermo-Sequence cycle sequencing kit (Amersham Biosciences). Nucleotide sequence analyses were conducted at the Baylor College of Medicine Search Launcher server and at the Expert Protein Analysis System Proteomics Server of the Swiss Institute of Bioinformatics, respectively. Ara-
A Zinc Finger DNA 3′-Phosphoesterase

Fig. 1. Mapping the AtZDP transcription initiation site. A, schematic representation of the AtZDP sequence surrounding the putative transcription initiation site. The positions of the antisense RNA probe utilized for RNase protection analysis and of the reverse transcription oligonucleotide primer are indicated by arrows below the sequence. The ATG translation start site, used as reference point (+1), and the TAG stop codon are shown. B, AtZDP mRNA 5′-end mapping. Total RNA samples derived from A. thaliana seedlings were analyzed by RNase protection (upper panel) and primer extension (lower panel) assays using the 32P-labeled antisense probes described for A. The AtZDP genomic sequence upstream of the ATG initiator codon is shown between the two panels. The experimentally determined location of the transcription start site is underlined. Possible minor sites for transcription initiation around position –50 are also indicated.

RESULTS

Isolation of a DNA 3′-Phosphoesterase cDNA from Arabidopsis—We have recently identified ZmDP2, the first plant DNA 3′-phosphoesterase capable of converting blocked 3′ termini into priming sites for reparative DNA polymerization. The ZmDP2 polypeptide was found to be similar to the C-terminal part of a predicted polypeptide conceptually derived from the Arabidopsis genome (32). Sequence analysis of this putative cDNA (AtZDP) revealed the presence of an in-frame ATG codon, 1273 bp upstream of the DNA 3′-phosphoesterase-homologous region, lying in an optimal sequence context according to translation initiation rules in plants (31). In keeping with such a prediction, a high-scoring transcription start site, 65 bp upstream of this putative initiation codon, was identified in the sequence of the AtZDP gene. To obtain reliable reference points for the isolation of the full-length AtZDP cDNA, we mapped the 5′-end of the corresponding mRNA by both RNase protection and primer extension using the hybridization probes reported in Fig. 1A. The combined use of these two approaches was meant to minimize possible RNA degradation or reverse transcription artifacts. As shown in Fig. 1B, both mapping experiments yielded an identical major transcription start site, 73 bp upstream of the putative ATG initiator codon. Although some minor transcription start sites are also apparent, they all precede this particular ATG codon, which thus appears to be the actual start site for AtZDP transcription initiation. Based on this finding, we designed the oligonucleotides reported in Fig. 2A and used them, along with total plasmid DNA from an Arabidopsis seedling cDNA library, to PCR-amplify the full-length AtZDP cDNA.

Sequence Analysis and Structural Dissection of the AtZDP Protein—The AtZDP gene is located on chromosome 3 and is interrupted by 16 introns spanning a total of 4241 bp in the Arabidopsis genome. The AtZDP cDNA codes for an acidic protein of 638 amino acids with a predicted molecular mass of 71 kDa and an isoelectric point of 5.5. As revealed by alignment with homologous plant and animal sequences, AtZDP is a modular protein, with conserved residues clustered into separate blocks likely corresponding to distinct functional domains. Three PARP-like zinc finger motifs (CX2,CX2,CX2,HC), known to be involved in DNA strand break recognition in animal systems, are present in the N-terminal region, followed by a C-terminal ZmDP2-homologous region (starting at position 425), which exhibits typical sequence features of DNA 3′-phosphotransferases (18). A closer inspection of the N-terminal part of AtZDP further showed that all three zinc finger motifs bear a net positive charge (+5, +1, and +5, respectively). Basic residues, which are important for DNA binding (32), are conserved between the three fingers and in fingers I and III are predicted to lie on the same face of putative α-helices (Fig. 2B). Sequences downstream of the zinc fingers display a large prevalence of negatively charged amino acids. An isolated KRK sequence motif (amino acids 381–383), the only putative nuclear localization signal (33) we could identify in the AtZDP polypeptide, is found right after the third finger.

The full-length AtZDP cDNA and the two regions identified by sequence analysis as those corresponding to the putative DBD (amino acids 1–366) and the CD (catalytic; amino acids 412–638) of the AtZDP protein were individually subcloned and overexpressed in Escherichia coli. All cDNAs were inserted into the expression vector pET28 as in-frame fusions with vector sequences coding for an N-terminal hexahistidine tag. As shown in Fig. 3A, His6-tagged polypeptides of the expected sizes (73, 43, and 28 kDa for the full-length protein and the N- and C-terminal portions, respectively) became detectable upon isopropyl-β-D-thiogalactopyranoside induction and were purified to near homogeneity by metal affinity chromatography. Fig. 3B further shows that polyclonal antibodies previously raised against maize ZmDP2 recognize the full-length (73 kDa) and CD (28 kDa) polypeptides, thus conclusively demonstrating that AtZDP is indeed the Arabidopsis homolog of the ZmDP2 3′-phosphoesterase.

Available at www.mips.gsf.de/proj/thal/.

baidopsis sequences were obtained from the Arabidopsis thaliana Database of the Munich Information Center for Protein Sequences.7
**A** AtZDP modular structure. A, conceptual translation of the AtZDP cDNA. Amino acid residues included in PARP-like zinc finger motifs are shown in *boldface*, with zinc-coordinating cysteines and histidines *underlined*. The DNA 3'-phosphatase motifs DDDD-1–4 and A and B are enclosed in *brackets*. The positions of PCR primers utilized for amplification of AtZDP cDNAs encoding the full-length protein (primers 1 and 2), the catalytic domain (primers 4 and 2), and the DNA-binding domain (primers 1 and 3) are marked with *arrows* above the amino acid sequence.

**B** Predicted structure of the AtZDP DNA-binding domain. The DNA-binding domain is drawn to show three zinc-coordinated fingers (ZnFI, ZnFII, and ZnFIII). Repeated amino acid blocks preceding and following each finger are shown as *open* and *shaded* ovals. Schematic representations of the putative α-helices of zinc fingers I and III are also reported, with charged amino acid residues in *boldface*.

**Fig. 2.** AtZDP modular structure. A, conceptual translation of the AtZDP cDNA. Amino acid residues included in PARP-like zinc finger motifs are shown in *boldface*, with zinc-coordinating cysteines and histidines *underlined*. The DNA 3'-phosphatase motifs DDDD-1–4 and A and B are enclosed in *brackets*. The positions of PCR primers utilized for amplification of AtZDP cDNAs encoding the full-length protein (primers 1 and 2), the catalytic domain (primers 4 and 2), and the DNA-binding domain (primers 1 and 3) are marked with *arrows* above the amino acid sequence. B, predicted structure of the AtZDP DNA-binding domain. The DNA-binding domain is drawn to show three zinc-coordinated fingers (ZnFI, ZnFII, and ZnFIII). Repeated amino acid blocks preceding and following each finger are shown as *open* and *shaded* ovals. Schematic representations of the putative α-helices of zinc fingers I and III are also reported, with charged amino acid residues in *boldface*. 
DNA 3’-Phosphoesterase Activity Is Associated with the C-terminal Domain of AtZDP—Full-length AtZDP and its isolated domains were assayed for DNA 3’-phosphoesterase activity using 3’-end-blocked, 5’-labeled synthetic oligonucleotides as substrates. The conversion of a 3’-phosphate end into the corresponding 3’-hydroxyl species (as schematized in Fig. 4A) was supported by both the full-length protein and the isolated C-terminal domain, but not by the N-terminal domain (Fig. 4B). A much weaker 3’-end-processing activity was also observed with a phosphoglycolate (rather than phosphate)-blocked oligonucleotide (data not shown). The quantitative recovery of 5’-end-associated radioactivity observed in all of these assays guaranteed for the absolute 3’-specificity of the DNA phosphoesterase activity. DNA 3’-phosphoesterification proceeded with a nearly identical efficiency when the 3’-phosphate group was part of a single-stranded oligonucleotide or at the gap site of a double-stranded oligonucleotide. In fact, the apparent \( K_m \) and \( V_{max} \) values (derived from the curves reported in Fig. 4C) were 22 or 35 \( \mu M \) and 35 or 46 pmol/min/pmol of enzyme, respectively. Fig. 4C further shows that a significant enhancement of 3’-phosphate hydrolysis (with an apparent \( V_{max} \) of 151 pmol/min/pmol of enzyme) was observed in reaction mixtures containing the single-strand substrate and supplemented with supercoiled plasmid DNA. Quite curiously, a similarly enhanced rate of 3’-phosphate hydrolysis was also observed when using CD-AZDP (instead of full-length AZDP) as a source of enzyme (data not shown). These findings indicate that different structural forms of 3’-blocked DNA serve as substrates for AZDP and that the catalytic conversion can be activated by DNA. The N-terminal zinc finger region appears to be dispensable for 3’-phosphoesterase activity as well as for DNA-dependent enzyme activation.

DNA Strand Break Recognition by the N-terminal Domain of AtZDP—Based on the notion that the PARP finger domain mediates DNA SSB recognition, we focused on the DNA-binding properties of AtZDP using electrophoretic mobility shifts assays. Binding reactions were assembled by incubating a 1-nucleotide gapped 45-bp duplex with purified AtZDP. Two retarded species were reproducibly observed following electrophoresis (c1 and c2 in Fig. 5A), and a third, faster migrating species was sometimes detected (c3 in Fig. 5A). To define the binding specificity of these complexes, we performed competition experiments using various unlabeled versions of the ligand DNA molecule. As shown in Fig. 5A, a gapped 45-bp duplex, either with or without a blocked 3’ terminus at the site of the gap, was a more effective competitor of complex c2 formation than the intact double-stranded oligonucleotide. However, a 50-fold molar excess of either single-stranded DNA efficiently removed complexes c1 and c2. Competition was not affected by the sequence context of the ligand DNA because it was also observed when an unrelated duplex bearing a single-strand nick was employed (data not shown). By contrast, single-stranded DNA was ineffective in competing for these complexes, possibly having a limited effect on complex c3 only. It thus appears that AZDP is an enzyme that specifically binds strand breaks in duplex DNA molecules in a sequence-independent fashion and with a higher affinity for nicked or gapped templates. Fig. 5B finally shows that the isolated N-terminal

---

**Fig. 3. Expression and purification of recombinant AtZDP.** A, Coomassie Blue-stained SDS-8% polyacrylamide gel of total lysates from isopropyl-β-D-thiogalactopyranoside-induced bacterial cells transformed with the empty pET28b vector (lane 1), pET-AZDP (lane 2), pET-DBD (lane 4), or pET-CD (lane 6). Highly purified fractions (0.5 μg) of the corresponding histidine-tagged recombinant proteins are shown in lanes 3, 5, and 7, respectively. The migration positions of molecular mass markers are indicated on the right. B, immunoblot analysis of the protein samples shown in A. An anti-ZmDP2 polyclonal antibody was utilized for immunodetection. The loading order and electrophoresis conditions were the same as described for A.

**Fig. 4. DNA repair activity of AtZDP.** A, schematic representation of the substrate utilized for DNA 3-phosphatase assays (see “Experimental Procedures” for details). Asterisks indicate the position of the 32P label. B, phosphorimaging of a denaturing 8% polyacrylamide gel showing the unconverted 5’-labeled P21P substrate (lane 1) and the dephosphorylated P21 product (lane 2). For DNA 3'-phosphoesterase reactions, a fixed concentration (13 nM) of the P21P substrate was incubated in the presence of increasing amounts of full-length AZDP (lanes 3–5), the isolated CD (lanes 6–8), or the N-terminal domain (DBD; lanes 9–11). C, substrate concentration dependence of AtZDP activity. Reaction velocity is plotted against substrate concentration. Increasing amounts of labeled single-stranded P21P (○ and □) or of the gapped 45-bp duplex (○) were incubated in the presence of a fixed concentration (5 nM) of full-length AZDP, either with (△) or without (○ and □) supercoiled plasmid DNA. Data are the average of at least two independent experiments (performed in duplicate) that differed by <10% of the mean.
AtZDP Bends DNA upon Binding to Single-strand Breaks

Further insight into DNA binding by AtZDP was obtained by AFM. The full-length protein and a 657-bp duplex with a 1-nucleotide gap in the middle (so to allow the unambiguous identification of protein bound to the gap region) were utilized for this analysis. Two types of AtZDP-DNA complexes, with the protein bound to either the gap site (Fig. 6A) or the ends of the duplex (Fig. 6B), were thus visualized. In accordance with electrophoretic mobility shift assay results, the frequency of binding to the ends was on average three times less than that of binding to the gap site region (approximately one complex every 15 DNA molecules). Moreover, a significant distortion of the bound DNA (a magnified view of which is shown in Fig. 6C) was apparent in most of the latter complexes in the region surrounding the protein-binding site. This indicates that AtZDP binding to single-strand lesions induces a bend in the target DNA. The extent of such bending was directly measured from AFM images by drawing tangents to the entry and exit points of the DNA at the protein-binding site. An average bend angle of 70°, with a mode value distribution of 60°–80°, was thus measured (Fig. 6D). It should be noted, however, that because of increased flexibility at the joint gap site, SSBs within duplex molecules can produce V-shaped DNA conformations even in the absence of protein binding. This effect must be taken into account to quantify the extent of protein-induced bending. Based on the notion that bends or gaps reduce the end-to-end distance of DNA molecules, we utilized polymer chain statistics to measure the mean square end-to-end distance \( (R^2) \) and the corresponding bend angle values of free and AtZDP-bound molecules of the 657-bp duplex (29). As shown in Table I, a static bend angle value of 76° for the AtZDP-gapped DNA complex was derived from such analysis. This value, which compares fairly well with that produced by the “direct” tangent method, is considerably larger than the 17° bend angle value estimated for unbound DNA. It thus appears that AtZDP binding bends gapped DNA by ~73° (average of the values produced by the two methods), a structural distortion that may significantly contribute to the efficiency of DNA repair in vivo.

### DISCUSSION

This work identifies AtZDP as the first nick-sensing DNA 3'-phosphoesterase thus far described in any organism. AtZDP is a multifunctional DNA-binding enzyme that recognizes and bends damaged DNA carrying SSBs and that catalyzes the repair of 3'-blocking terminal lesions. Based on its characteristic zinc finger domain structure and peculiar functional capabilities, we suggest that AtZDP acts as a nick-sensing and 3'-end-restoring component of a multiprotein DNA repair complex in plants.

**AtZDP Identifies a Novel Family of Nick-sensing DNA 3'-Phosphoesterases**—To our knowledge, AtZDP is the first DNA repair enzyme embodying three \( \text{Cy}_{58}-\text{His} \) zinc fingers. An AtZDP homolog (ZmDP2) has been previously isolated in maize based on its ability to complement diphosphonucleoside phosphatase mutants. This maize protein appeared to be endowed with only the 3'-phosphoesterase function (22). More recent experiments indicate, however, the presence of a nick-sensing DNA-binding domain in the N-terminal part of ZmDP2. AtZDP belongs to the DNA 3'-phosphatase family, and conserved motifs in its sequence suggests that it covalently binds substrate DNA via the first aspartate of the motif DDDD-1 and that catalyzes the 3'-phosphoglycolate-blocked terminal lesion. This work identifies AtZDP as the first DNA-binding enzyme that recognizes and bends damaged DNA carrying SSBs and that catalyzes the repair of 3'-blocking terminal lesions. Based on its characteristic zinc finger domain structure and peculiar functional capabilities, we suggest that AtZDP acts as a nick-sensing and 3'-end-restoring component of a multiprotein DNA repair complex in plants.

**AtZDP Bends DNA upon Binding to Single-strand Breaks**—Further insight into DNA binding by AtZDP was obtained by AFM. The full-length protein and a 657-bp duplex with a 1-nucleotide gap in the middle (so to allow the unambiguous identification of protein bound to the gap region) were utilized for this analysis. Two types of AtZDP-DNA complexes, with the protein bound to either the gap site (Fig. 6A) or the ends of the duplex (Fig. 6B), were thus visualized. In accordance with electrophoretic mobility shift assay results, the frequency of binding to the ends was on average three times less than that of binding to the gap site region (approximately one complex every 15 DNA molecules). Moreover, a significant distortion of the bound DNA (a magnified view of which is shown in Fig. 6C) was apparent in most of the latter complexes in the region surrounding the protein-binding site. This indicates that AtZDP binding to single-strand lesions induces a bend in the target DNA. The extent of such bending was directly measured from AFM images by drawing tangents to the entry and exit points of the DNA at the protein-binding site. An average bend angle of 70°, with a mode value distribution of 60°–80°, was thus measured (Fig. 6D). It should be noted, however, that because of increased flexibility at the joint gap site, SSBs within duplex molecules can produce V-shaped DNA conformations even in the absence of protein binding. This effect must be taken into account to quantify the extent of protein-induced bending. Based on the notion that bends or gaps reduce the end-to-end distance of DNA molecules, we utilized polymer chain statistics to measure the mean square end-to-end distance \( (R^2) \) and the corresponding bend angle values of free and AtZDP-bound molecules of the 657-bp duplex (29). As shown in Table I, a static bend angle value of 76° for the AtZDP-gapped DNA complex was derived from such analysis. This value, which compares fairly well with that produced by the “direct” tangent method, is considerably larger than the 17° bend angle value estimated for unbound DNA. It thus appears that AtZDP binding bends gapped DNA by ~73° (average of the values produced by the two methods), a structural distortion that may significantly contribute to the efficiency of DNA repair in vivo.

---

S. Petrucco, unpublished data.
recognizable nick-sensing domain, is not capable of 3'-phosphoglycolate end processing, and strongly discriminates against single-stranded 3'-phosphate-blocked substrates (18). It has been proposed that a substantial functional redundancy exists for the repair of 3'-terminal lesions, with different enzymes bearing distinct, yet partially overlapping substrate specificities. This is clearly the case in yeast, where in addition to Tpp1, two additional enzymes, the apurinic/apyrimidinic endonucleases Apn1 and Apn2, have been shown to catalyze the removal of various 3'-blocking lesions, including those bearing a 3'-phosphoglycolate group. An apurinic endonuclease redox protein (Arp), as yet uncharacterized with respect to its DNA repair capabilities, has previously been identified in Arabidopsis (35). It will thus be interesting to find out whether a similar redundancy, especially regarding blocking lesions other than a 3'-phosphate, also exists in plants.

**FIG. 6.** AtZDP is a DNA strand break sensor that induces DNA bending. **A,** AFM images of gap-bound AtZDP. The protein is visible as a white dot in a central position of the 657-bp DNA, corresponding to the 1-nucleotide gap site. The V-shaped conformation induced upon protein binding is clearly visible. **B,** AFM images of end-bound AtZDP on the same DNA template as in **A.** **C,** three-dimensional view of one of the complexes shown in **A.** **D,** frequency distribution of DNA bend angles measured with the tangent method. The average bend angle is 70° ± 31°. The number of bins is the square root of the sample size. The images shown in **A** and **B** have a real size of 250 × 250 nm.

Damaged DNA Recognition by AtZDP Zinc Fingers—AtZDP zinc fingers act as a “nick sensor” and sharply bend target DNA upon binding to single-strand lesions. Interestingly, a similar degree of bending has been reported for PARP binding to nicked DNA sites (12), as if a bent conformation of target DNA were a sort of general prerequisite for the correct positioning of downstream-acting DNA repair components. In fact, it has been proposed that the human repair complex binds to the inside of DNA bends generated by PARP at SSB sites. In this way, it would serve as a “docking” platform for repair components, while still allowing access to SSB termini on the outside of the bend. Considering that one PARP-like finger is sufficient for nick sensing (15), it is conceivable that multiple fingers, as in AtZDP and PARP, may serve additional roles in the control of genome stability. Indeed, the specific binding of AtZDP to single- and double-stranded DNA ends may suggest that this enzyme is involved in the assembly of multipurpose strand break repair complexes.

Repair functions by PARP fingers seem to be independent from the associated enzyme activity. In keeping with this view, the DNA 3'-phosphoesterase activity of AtZDP does not require the nick-binding function, and conversely, strand break recognition by its zinc fingers does not discriminate between blocked or unmodified 3' termini. Furthermore, we found that double-
stranded DNA activates AtZDP phosphoesterase activity regardless of the presence of the zinc fingers, thus indicating that AtZDP fingers are not directly involved in either substrate recognition or DNA-dependent enzyme activation. It has been proposed that in addition to nick sensing, the zinc finger of DNA ligase III contributes to stabilization of substrate binding to the enzyme active site (15). In turn, PARP fingers have been shown to enhance the effect of activator DNA on PARP enzyme activity (36). It is thus tempting to speculate that also in the case of AtZDP, both substrate and activator DNA binding to the catalytic domain are similarly optimized by the three PARP-like zinc fingers.

**Single-strand Break Repair in Plants**—The human SSBR complex has recently been shown to be composed of five interacting modular proteins that appear to coordinate activity in DNA repair (8). Based on the observation that AtZDP is a modular polypeptide sharing structural motifs with animal SSBR proteins, we reasoned that a similar multiprotein complex might operate in plants. Having in mind the unique modular composition of AtZDP and thus the likely existence of differences in complex assembly and modular assortment, we searched the *Arabidopsis thaliana* Database for putative components of a plant SSBR repairosome. A list of *Arabidopsis* polypeptides that we believe may be part of such a complex is presented in Fig. 7, together with their predicted catalytic domains and protein-protein or protein-DNA interaction motifs. Candidate XRCC1-like and DNA polymerase β genes were already annotated in the *Arabidopsis thaliana* Database. Similar to its human homolog, the XRCC1-like polypeptide of *Arabidopsis* is predicted to contain multiple protein-protein interaction BRCT (BRCA1 C-terminal) modules. Interestingly, similar BRCT modules are also present in the putative DNA polymerase β as well as in a member of one of the two PARP families thus far described in plants (21). No human polynucleotide kinase 3'-phosphatase homolog other than AtZDP is encoded in the *Arabidopsis* genome. Accordingly, the only putative 5'-kinase of *Arabidopsis* (17) does not bear an associated 3'-phosphatase domain. Finally, various candidate sequences sharing strong similarity with the DNA ligase catalytic domain, but all lacking a PARP-like zinc finger, were obtained from a homology search of the *Arabidopsis thaliana* Database using human DNA ligase III as a query. By excluding a number of sequences belonging to either class I or IV DNA ligases, we propose the polypeptide encoded by the At1g66730 gene as the best candidate for a DNA ligase involved in SSBR in plants. In fact, in addition to a high-scoring DNA ligase-like domain, the At1g66730 polypeptide contains an N-terminal region that is conserved within a broad group of eukaryotic DNA repair components and is thought to target enzyme activity for nucleic acids in the context of various lesion-specific repair pathways (37, 38).

### Table I

| Bend angle analysis of AtZDP - DNA complexes |
|-----------------------------------------------|
| L<sup>a</sup> (R<sup>b</sup>)<sup>c</sup> | 204 nm (657 bp) | 192.6 nm (657 bp) | 204 nm (657 bp) |
| Bend angle | 23663 nm<sup>d</sup> | 15,932 nm<sup>d</sup> | 24,056 nm<sup>d</sup> |
| Tangents | 17 °<sup>d</sup> | 76 °<sup>d</sup> | 70 °<sup>d</sup> |
| No. of molecules | 482 | 54 | |

<sup>a</sup> L is the contour length of the complexes as derived from AFM images.

<sup>b</sup> R<sup>c</sup> is the mean square end-to-end distance.

<sup>c</sup> Value is the mean square end-to-end distance predicted for a 204-nm long DNA molecule with a persistence length of 53 nm.

<sup>d</sup> Value is the expected bend angle at the gap site of a 204-nm long DNA molecule with an (R<sup>c</sup>) value of 23,663 nm<sup>d</sup>.

A minimal set of five putative repair components potentially capable of cooperating with AtZDP in the assembly of a fully competent SSBR complex is encoded by the *Arabidopsis* genome. Despite the occurrence of common DNA recognition and protein-protein interaction modules, the detailed architecture of this complex is likely to be substantially different from that of the mammalian SSB repairosome. As exemplified by the unique domain organization of AtZDP, most of such diversity probably originates from the dissimilar assortment of functionally homologous modules among individual repair components. Therefore, following experimental validation of the actual functionality of the components identified by this whole genome analysis, future studies will have to address the mode and
order of recruitment of such components into a functional repair complex. Because of the PARP-like DNA damage recognition capacity of AtZDP, it will be especially interesting to determine whether this nick-sensing 3′-phosphoesterase can act as an autonomous nucleation center for SSBR complex assembly and how additional components are recruited into such a complex. Ultimately, the in vitro reconstitution of a fully competent plant repairesome will lead to understanding the physiological significance and far-reaching adaptive implications of the unique SSBR strategy operating in plants.

Acknowledgments—We thank Dr. Michele Minet for the Arabidopsis cDNA library. We gratefully acknowledge Ricardo Percudani for assistance with sequence analysis, Alessio Peracchi for helpful discussions, Roberto Tirindelli for critical reading of the manuscript, and Nicola Vannini for helping with AFM imaging. We are grateful to Gian Luigi Rossi for encouragement and continuous support.

REFERENCES

1. Friedberg, E. C. (1995) Trends Biochem. Sci. 20, 381
2. Poulist, J. J., Yao, K. C., Robertson, C. A., and Nash, H. A. (1999) Science 286, 552–555
3. Ward, J. E. (1998) in DNA Damage and Repair (Nickoloff, J. A., and Hoeckstra, M. F., eds) Vol. 2, pp. 65–84, Human Press Inc., Totowa, NJ
4. Demple, B., and Harrison, L. (1994) Annu. Rev. Biochem. 63, 915–948
5. Wilson, D. M., Engelward, B. P., and Samson, L. (1998) in DNA Damage and Repair (Nickoloff, J. A., and Hoeckstra, M. F., eds) Vol. 1, pp. 29–64, Human Press Inc., Totowa, NJ
6. Karimi-Busheri, F., Daly, G., Robin, P., Canas, B., Pappin, D. J., Sgouros, J., Miller, G. G., Foroni, C., Percudani, R., Rossi, G. L., and Ottonello, S. (1999) J. Biol. Chem. 274, 24187–24194
7. Jilani, A., Ramotar, D., Slack, C., Ong, C., Yang, X. M., Seherer, S. W., and Lasko, D. D. (1999) J. Biol. Chem. 274, 24176–24186
8. Whitehouse, C. J., Taylor, R. M., Thistlethwaite, A., Zhang, H., Karimi-Busheri, F., Lasko, D. D., Weinfeld, M., and Caldecott, K. W. (2001) Cell 104, 107–117
9. Masson, M., Niedergang, C., Schreiber, V., Muller, S., Menissier-de Murcia, J., and de Murcia, G. (1998) Mol. Cell. Biol. 18, 3563–3571
10. Dantzer, F., Schreiber, V., Niedergang, C., Trucco, C., Flatter, E., De La Rubia, G., Oliver, J., Rolli, V., Menissier-de Murcia, J., and de Murcia, G. (1999) Mol. Cell. Biol. 19, 2371–2373
11. Caldecott, K. W., Aoufouchi, S., Johnson, P., and Shatt, S. (1996) Nucleic Acids Res. 24, 4387–4394
12. Le Cam, E., Paek, P., Menissier-de Murcia, J., Cognet, J. A., Barbin, A., Sarantoglou, V., Reve, B., Delain, E., and de Murcia, G. (1994) J. Mol. Biol. 235, 1062–1071
13. Gradwohl, G., Menissier-de Murcia, J., Molinete, M., Simonin, F., Koken, M., Hoeijmakers, J. H., and de Murcia, G. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 2990–2994
14. Mazen, A., Menissier-de Murcia, J., Molinete, M., Simonin, F., Gradwohl, G., Poirier, G., and de Murcia, G. (1989) Nucleic Acids Res. 17, 4689–4698
15. Mackey, Z. B., Niedergang, C., Menissier-de Murcia, J., Leppard, J., Au, K., Chen, J., de Murcia, G., and Tomkinson, A. E. (1999) J. Biol. Chem. 274, 21679–21687
16. Meijer, M., Karimi-Bushe, F., Huang, T. Y., Weinfeld, M., and Young, D. (2002) J. Biol. Chem. 277, 4550–4555
17. Vance, J. R., and Wilson, T. E. (2001) J. Biol. Chem. 276, 15073–15081
18. Vance, J. R., and Wilson, T. E. (2001) Mol. Cell. Biol. 21, 7191–7198
19. Brit, A. B. (1999) Trends Plant Sci. 4, 20–25
20. Mahajan, P. B., and Zao, Z. (1999) Plant Physiol. 118, 895–905
21. Babiychuk, E., Cottrill, P. B., Storozhenko, S., Fangthong, M., Chen, Y., O’Farrell, M. K., Van Montagu, M., Inze, D., and Kushnir, S. (1998) Plant J. 15, 635–645
22. Betti, M., Petrucco, S., Bolchi, A., Dieci, G., and Ottonello, S. (2001) J. Biol. Chem. 276, 18038–18045
23. Petrucco, S., Bolchi, A., Foroni, C., Percudani, R., Rossi, G. L., and Ottonello, S. (1996) Plant Cell 8, 69–80
24. Sambrook, J., and Russell, D. W. (2001) Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
25. Minet, M., Dufour, M. E., and Lacroute, F. (1992) Plant J. 2, 417–422
26. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
27. Laemmli, U. K. (1970) Nature 227, 680–685
28. Sander, M., and Huang, S. M. (1995) Biochemistry 34, 1267–1274
29. Rivetti, C., Walker, C., and Bustamante, C. (1998) J. Mol. Biol. 280, 41–59
30. Rivetti, C., Guthold, M., and Bustamante, C. (1996) J. Mol. Biol. 264, 919–932
31. Lutke, H. A., Chere, K. C., Michel, F. S., Moss, K. A., Kern, H. F., and Scheele, G. A. (1987) EMBO J. 6, 43–48
32. Molinete, M., Vermeulen, W., Burke, A., Menissier-de Murcia, J., Kupper, J. H., Hoeijmakers, J. H., and de Murcia, G. (1995) EMBO J. 14, 2109–2117
33. Crags, G., and Kellie, S. (2001) J. Biol. Chem. 276, 23719–23725
34. Thaller, M. C., Schippa, S., and Rossolini, G. M. (1998) Protein Sci. 7, 1647–1652
35. Babiychuk, E., Kushnir, S., Van Montagu, M., and Inze, D. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 3299–3303
36. Ikejima, M., Noguchi, S., Yamashita, R., Ogura, T., Sugimura, T., Gill, D. M., and Miwa, M. (1990) J. Biol. Chem. 265, 21907–21913
37. Dronkert, M. L., van We, J., Boeke, M., Vasconcelos, M. L., van Steeg, H., Tan, T. L., Hoeijmakers, J. H., and Kanaar, R. (2000) Mol. Cell. Biol. 20, 4553–4561
38. Moshous, D., Callebaut, I., de Chasseval, R., Corneo, B., Cavazzana-Calvo, M., Le Deist, F., Texier, I., Sanal, O., Bertrand, Y., Philippe, N., Fischer, A., and de Villartay, J. P. (2001) Cell 105, 177–186