THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 20, pp. 13939–13948, May 19, 2006

Aprataxin Forms a Discrete Branch in the HIT (Histidine Triad) Superfamily of Proteins with Both DNA/RNA Binding and Nucleotide Hydrolase Activities*

Received for publication, July 21, 2005, and in revised form, February 24, 2006 Published, JBC Papers in Press, March 16, 2006, DOI 10.1074/jbc.M507946200

Amanda W. Kijas[‡], Janelle L. Harris[‡§], Jonathan M. Harris[§], and Martin F. Lavin[‡¶1]

From the [‡]Queensland Institute of Medical Research, P. O. Box Royal Brisbane Hospital, Queensland, Australia 4029, [§]School of Life Sciences, Queensland University of Technology, Brisbane, Australia 4000, and [¶]Central Clinical School, University of Queensland, Brisbane, Australia 4029

Ataxia with oculomotor apraxia type 1 (AOA1) is an early onset autosomal recessive spinocerebellar ataxia with a defect in the protein Aprataxin, implicated in the response of cells to DNA damage. We describe here the expression of a recombinant form of Aprataxin and show that it has dual DNA binding and nucleotide hydrolase activities. This protein binds to double-stranded DNA with high affinity but is also capable of binding double-stranded RNA and single-strand DNA, with increased affinity for hairpin structures. No increased binding was observed with a variety of DNA structures mimicking intermediates in DNA repair. The DNA binding observed here was not dependent on zinc, and the addition of exogenous zinc abolished DNA binding. We also demonstrate that Aprataxin hydrolyzes with similar efficiency the model histidine triad nucleotide-binding protein substrate, $AMPNH_2$, and the Fragile histidine triad protein substrate, Ap_4A. These activities were significantly reduced in the presence of duplex DNA and to a lesser extent in the presence of single-strand DNA, and removal of the N-terminal Forkhead associated domain did not alter activity. Finally, comparison of sequence relationships between the histidine triad superfamily members shows that Aprataxin forms a distinct branch in this superfamily. In addition to its capacity for nucleotide binding and hydrolysis, the observation that it also binds DNA and RNA adds a new dimension to this superfamily of proteins and provides further support for a role for Aprataxin in the cellular response to DNA damage.

Ataxia with oculomotor apraxia type1 (AOA1)[2] is an early onset neurodegenerative disorder that is characterized by cerebellar ataxia, peripheral neuropathy, and hypoalbuminemia (1). Date *et al.* (2) and Moreira *et al.* (3) identified the gene defective in this disorder, *APTX*, using linkage analysis. Two different size transcripts were identified for Aprataxin by Northern blot analysis and sequencing and were predicted to encode short (19 kDa) and long (39 kDa) forms of the protein. A series of mutations has been described in *APTX*, primarily localizing to the histidine triad (HIT) domain centrally located in the molecule (1, 4). Aprataxin also contains an N-terminal Forkhead-associated (FHA) domain that binds phosphoproteins and a C-terminal putative zinc finger DNA binding domain (2, 3, 5, 6). This protein localizes to the nucleus and nucleolus (6) and has also been shown to interact with several proteins involved in the cellular response to DNA damage (2, 6–9). These proteins include x-ray cross-complementing protein 1 (XRCC1), XRCC4, poly(ADP-ribose) polymerase 1, and p53 (2, 6–9). Aprataxin has also been shown to form part of a novel single-strand break repair complex that includes XRCC1 and DNA ligase 3 (9). This complex is distinct from the single-strand break repair complex that includes XRCC1, DNA ligase 3, polynucleotide kinase, and polymerase β. The N-terminal 20 amino acids of the Aprataxin FHA domain was found to interact with XRCC1 and was dependent on the phosphorylation of XRCC1 (5, 7). A role for Aprataxin in the cellular response to DNA damage is further substantiated by enhanced sensitivity of AOA1 cells to a variety of genotoxic agents. AOA1 cells are hypersensitive to agents that cause single-strand breaks in DNA, including hydrogen peroxide and methyl methanesulfonate (6). These cells also display higher levels of induced chromosomal aberrations in response to these agents compared with control cells. No increased sensitivity was observed after mitomycin C or ionizing radiation exposure (6, 7). Together these observations point to a role for Aprataxin in the cellular response to DNA single-strand breaks. Evidence for such a role was recently provided by Mosesso *et al.* (10), who demonstrated an inability of AOA1 cells to process camptothecin-induced DNA damage using the comet assay.

It remains unclear as to what the exact role of Aprataxin is in the cellular response to DNA damage. The examination of the putative functional domains within Aprataxin provides evidence as to what this role might be. The presence of a HIT domain in the central region of the molecule suggests that it has nucleotide hydrolase activity. The HIT domain superfamily of proteins is characterized by the presence of the HIT motif, *HXHXHXX* where *X* is a hydrophobic amino acid. This superfamily consists of four branches including histidine triad nucleotide-binding protein (HINT), fragile histidine triad (FHIT), galactose-1-phosphate uridylyltransferase, and DcpS, a scavenger pyrophosphatase (11, 12).

Phylogenetic tree analysis by Date *et al.* (2) and Hirano *et al.* (13) showed that Aprataxin branches more closely with FHIT family proteins. The FHIT family has been shown to hydrolyze diadenosine polyphosphate substrates. Contrary to this, Aprataxin has also been classified as a HINT protein (11, 14) based on its 31% identity with the rabbit and human HINT sequences over an 86-amino acid region (11). The recent reports on the catalytic activity of bacterially expressed recombinant Aprataxin have also been conflicting (13, 15). Hirano *et al.* (13) reported that due to intramolecular inhibition by the FHA domain, there was no significant hydrolysis on GpppBODIPY, a known fluoro-

* The costs of publication of this article were 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.

[1] To whom correspondence should be addressed: Queensland Institute of Medical Research, P. O. Box Royal Brisbane Hospital, Herston, Queensland, Australia 4029. Tel.: 61-7-3362-0341; Fax: 61-7-33620106; E-mail: Martin.Lavin@qimr.edu.au.

[2] The abbreviations used are: AOA1, Ataxia with oculomotor apraxia type 1; HINT, histidine triad nucleotide-binding protein; HIT, histidine triad; FHA, Forkhead-associated; XRCC1, x-ray cross-complementing protein 1; FHIT, fragile histidine triad; PMSF, phenylmethylsulfonyl fluoride; APTX, Aprataxin; DS, double strand; SS, single strand; Ap_4A, diadenosine tetraphosphate; $AMPNH_2$, adenosine-5′-monophosphoramidate.

genic substrate for FHIT family proteins (16). When they added the N-terminal FHA domain separately from the C-terminal portion of the Aprataxin protein, they found low levels of activity, $K_{cat} = 1.78 \times 10^{-3}s^{-1}$, and mutations in the HIT domain abrogated this activity. In contrast, Seidle *et al.* (15) obtained similar low levels of activity ($K_{cat} = 0.4 \times 10^{-3}s^{-1}$) using the full-length protein on the same substrate but observed no apparent inhibition by the FHA domain. They also demonstrated activity ($K_{cat} = 2.8 \times 10^{-3}s^{-1}$) on the HINT model substrate, *tert*-butoxycarbonyl-L-lysine methylcoumarinamide, an AMP molecule linked to lysine.

Here we describe the expression of a yeast recombinant human Aprataxin protein and its purification to a single detectable band after SDS-PAGE. We demonstrate for the first time that Aprataxin possesses both DNA and RNA binding activities as well as a capacity to hydrolyze nucleotide-based substrates. The presence of DNA significantly reduced nucleotide hydrolase activity.

MATERIALS AND METHODS

Expression Construct—Aprataxin was cloned into the yeast overexpression plasmid pMH1, which has been adapted for the use of the self-cleaving affinity technology of New England Biolabs IMPACT-CN system (17). The entire coding region of human *APTX* (NP_778243) lacking the stop codon was amplified using primers 5′-TTTGTCGACATGATGCGGGTGTGCTGGTTG-3′ and 5′-CTGTGTCCAGTGCTTCCTGAGAT G-3′, designed to introduce a SalI site at the 5′ end of the PCR product. This was digested with SalI and cloned into pMH1, which had been digested with SalI-SmaI to remove the *MLH1* gene, creating pMHA. This cloning strategy results is the addition of a glycine to the C terminus of the human Aprataxin protein after intein cleavage. The Aprataxin V263G mutant was cloned as above and had been previously generated, as described in Gueven *et al.* (6). Bacterial expression constructs of full-length, N-terminal deleted, C-terminal deleted, and N-and C-terminal deleted forms of Aprataxin were prepared using PCR with the following primer sets: 5′-CGGGATCCATGATGCGGGTGTGCTGGTTG-3′ and 5′-CCGCTCGAGTCACTGTGTCCAGTGCTTCCT-3′; 5′-CGGGATCCTCAGGCAACAGTGATTCTATAGA-3′ and 5′-CCGCTCGAGTCACTGTGTCCAGTGCTTCCT-3′; 5′-CGGGATCCA TGATGATGCGGGTGTGCTGGTTGG-3′ and 5′-CCGCTCGAGTCAACGAAGGGGCAGCTTCA-3′; 5′-CGGGATCCTCAGGCAACAGTGATTCTATAGA-3′ and 5′-CCGCTCGAGTCAACGAAGGGGCAGCTTCA-3′, respectively. These amplifications introduced BamHI and XhoI sites for cloning into pGEX-6P-1 (Amersham Biosciences). This cloning strategy leads to the addition of five amino acids (GPLGS) to the N terminus after cleavage with PreScission protease.

Overexpression and Purification—The pMHA construct was introduced into *Saccharomyces cerevisiae* strain, BJ2168, using the standard lithium acetate transformation method followed by selection on tryptophan drop out minimal media. Subsequent growth and overexpression was also performed in this medium at 25–30 °C. Transformants were used to inoculate overnight starter cultures with 2% glucose as the carbon source. These saturated cultures were diluted 1/100 and regrown in media using 2% glycerol and 2% lactate as carbon sources to an A_{600} of 0.9. Cells were harvested and washed once in phosphate-buffered saline, then cell pellets were resuspended in an equal volume of chitin 500 buffer (20 mM Tris, pH 8, 1 mM EDTA, 1 mM PMSF, 10% glycerol, 500 mM NaCl) before freezing small beads on dry ice. Frozen cells were ground in dry ice using a coffee grinder. Ground cells were defrosted on ice and resuspended in 5 volumes of chitin 500 buffer with 0.01 mM Tris(2-carboxyethyl)phosphine. This lysate was centrifuged at 20,000 × *g* for 30 min at 4 °C. All the subsequent purification steps were performed at 4 °C. The clarified lysate was loaded at 0.5 ml/min onto a 1-ml chitin column equilibrated with chitin 500 buffer. The column was washed with 10 volumes of chitin 500 buffer, 5 volumes of chitin 1000 buffer (20 mM Tris, pH 8, 1 mM EDTA, 1 mM PMSF, 10% glycerol, 1000 mM NaCl), 5 volumes of cleavage buffer (20 mM Tris, pH 8, 500 mM NaCl, 10% glycerol, 1 mM EDTA pH8, 50 mM dithiothreitol, 1 mM PMSF) without dithiothreitol, then with 5 volumes of cleavage buffer with dithiothreitol and clamped to enable cleavage (16 h at 4 °C). Cleaved protein was eluted, concentrated, and diluted to a final concentration of 200 mM NaCl with water then loaded at 0.3 ml/min onto a 1-ml Biorex 70 cation exchange resin (Bio-Rad) equilibrated with buffer A200 (10 mM Tris, pH 7.5, 0.5 mM, EDTA, 1 mM PMSF, 10% glycerol, 200 mM NaCl). The column was washed with 20 volumes of buffer A200 before elution with 10 volumes of buffer A380 (10 mM Tris, pH 7.5, 0.5 mM, EDTA, 1 mM PMSF, 10% glycerol, 380 mM NaCl), collecting 1-ml fractions. Peak fractions were pooled and concentrated then diluted with an equal volume of water before snap-freezing on dry ice and subsequent storage at −70 °C. Purity was assessed by 15% SDS-polyacrylamide gel electrophoresis and subsequent Coomassie R250 staining as well as Western blot analysis using a rabbit anti-Aprataxin antibody raised against the full-length yeast recombinant protein. For experiments comparing the effect of EDTA on the recombinant protein, 2 protein preparations were made in parallel as above either with 1 mM EDTA in all buffers or was left out for control no EDTA preparation.

The mutant V263G protein was purified as outlined above for the chitin affinity chromatography strategy, with the inclusion of an ATP wash (5 volumes, 20 mM Tris, pH 8, 200 mM KCl, 10% glycerol, 3 mM $MgCl_2$, and 200 μM ATP) incorporated after the chitin 1000 wash. This was included to remove a 70-kDa co-purifying protein that appears to be a chaperone protein (HSP70) based on previous reports (17). Also, cleavage was performed at a higher temperature, 16 °C and for longer (24 h), to enhance the poor cleavage. In experiments where the V263G protein was used, wild type protein purified in the same manner was employed. The bacterial pGEX-6P-1 constructs were transformed into BL21 *E. coli* for overexpression at 25 °C. Isopropyl 1-thio-β-D-galactopyranoside (0.1 mM) was added for 2.5 h to induce protein expression. Cells were lysed by sonication and clarified by centrifugation (20,000 $g \times 40$ min). The supernatant was incubated with glutathione-Sepharose resin (Amersham Biosciences), glutathione *S*-transferase fusion protein was cleaved with an overnight incubation with PreScission protease, and proteins were recovered in the eluate. These proteins were further purified on a Biorex 70 cation exchange resin.

DNA Binding Assays—DNA binding was investigated using electrophoretic mobility shift assay using different size single-strand and duplex oligonucleotides. Oligonucleotides were radiolabeled by polynucleotide kinase (Invitrogen) and annealed to create duplexes. Sequences are based on those used by Chi and Kolodner (18), which are designed to minimize secondary structure. A common 21-base bottom oligo, 5′-TCAGCAGATAGGAACCATACT-3′, was annealed to the top sequences to create the various DNA structures: homoduplex, 5′-AGTATGGTTCCTATCTGCTGA-3′; mismatch, (G-A), 5′-AGTATGGTTCCGATCTGCTGA-3′; flap, 5′-AGTATGGTTCCTTAGACGACT-3′; 6-base pair hairpin, 5′-AGTATGGTTCGAGTACGTACTCCTATCTGCTGA-3′; nick, 5′-ATCTGCTGA-3′ and 5′-AGTATGGTTCCT-3′; one-base gap, 5′-ATCTGCTGA-3′ and 5′-AGTATGGTTCC-3′. RNA binding was investigated using a 21-base pair diced RNA sequence (DNA-PKcs) kindly provided by Dr. Rick Woods. The 11- and 37-bp duplexes were created as outlined above by annealing sequences: 5′-AGTATGGTTCC-3′ and 5′-GGAACCATACT-3′; 5′-ATGTGAATCAGTATGGTTCCTATCTGCTGAAGGAAAT-3′ and 5′-ATTTCCT-

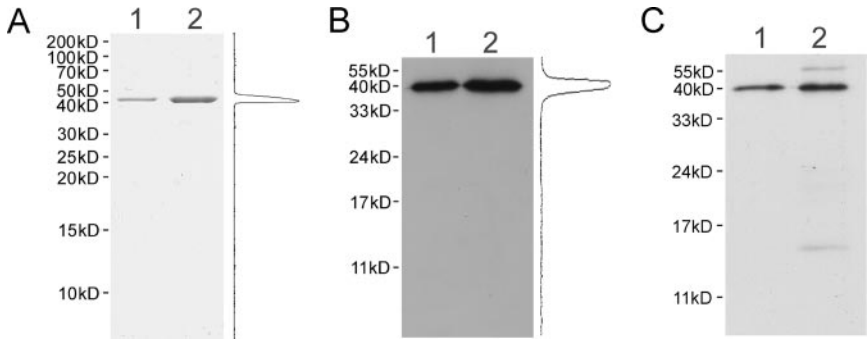

FIGURE 1. *A*, Coomassie R250-stained 15% SDS-polyacrylamide gel with 1.5 pmol (*lane 1*) and 4.5 pmol (*lane 2*) of purified Aprataxin (40 kDa). A histogram of the central region of the 4.5-pmol lane was generated by the ImageQuant program (Amersham Biosciences) and was aligned to the *left* of the *image*. *B*, Western blot of purified Aprataxin with 1.5 pmol (*lane 1*) and 4.5 pmol (*lane 2*) probed with a rabbit anti full-length Aprataxin antibody. A histogram of the central region of the 4.5-pmol lane was generated by ImageQuant program (Amersham Biosciences) and was aligned to the *left* of the image. C, Western blot of purified wild type Aprataxin (*lane 1*) and V263G (*lane 2*) probed with a rabbit anti-full-length Aprataxin antibody, Showing a less pure preparation.

TCAGCAGATAGGAACCATACTGATTCACAT-3′. A nick was introduced into the 37-bp sequence after nucleotide 16 by annealing the two corresponding oligonucleotides. Base 17 was left out, and the same strategy was used to create a gapped substrate.

Binding reactions were performed on ice for 15 min in 10 mM HEPES, pH 7, 20 ng/μl acetylated bovine serum albumin, 100 mM NaCl (unless otherwise stated), and 8% sucrose. The nucleic acid substrates were present at 1 pmol per reaction unless otherwise stated, and Aprataxin was included as indicated. Zinc sulfate was included where indicated by preincubating Aprataxin with zinc for 15 min before the addition of radiolabeled DNA duplex. Samples were run in 6% non-denaturing polyacrylamide gels containing 0.5× Tris borate-EDTA buffer at 20 mA for 25–35 min at 5 °C. Gels were dried on Whatman No. 3MM and visualized by x-ray film. Quantitation was performed by scanning films and using ImageQuant software (Amersham Biosciences). The percent bound was estimated as bound/(bound + unbound).

Catalytic Activity Assays—Initial experiments using pH 6, 7, and 8 HEPES buffer established that maximum catalytic activity was obtained at pH 7, which was used for subsequent investigations. Reactions were performed at 30 °C for 20 min to 3 h in 10 mM HEPES, pH 7, 30 mM NaCl, 1 μM Aprataxin with nucleotide substrates, ATP (Sigma), dATP (Fisher), $AMPNH_2$ (Sigma), and Ap_4A (Sigma) at 0.02–1.5 mM in 10-μl reactions. DNA duplex or single-strand oligonucleotide was added as indicated.

The products of hydrolysis were resolved and quantitated using anion exchange high performance liquid chromatography on a Hypersil SAX column 250 × 4.6 mm (Phenomenex). Parent substrates used here and products of the hydrolysis reactions, AMP and ATP, were identified by retention times and co-elution with standards and quantified based on their absorbance at 259 nm. The reaction products were resolved using a gradient starting with 1 mM ammonium phosphate buffer, pH 2.8, 0–40%, run against 750 mM ammonium phosphate buffer, pH 3.7, for all reactions except $AMPNH_2$. These reaction products were resolved using a gradient starting with water 0–40% run against 750 mM ammonium phosphate buffer, pH 3.7, where AMP was the only detectable product. Peak areas were used to generate a standard curve for AMP over the range 50–1000 pmol ($R^2 = 0.99$), indicating both a linear and sensitive detection of the primary product of hydrolysis. Determinations of kinetic parameters were made by estimating the best fit curve to the Michaelis equation using an Excel macro available online at University of Paisley, Department of Biological Sciences.

Phylogenetic Tree Analysis—Human sequences (NP_778243, NP_005331, Q9BX68, NP_612638, AAL40394, NP_000146, NP_054745, AAH32336, including Hint2 and Hint3) were used from each branch Hint1, FHIT, galactose-1-phosphate uridylyltransferase, DcpS, and Aprataxin to identify orthologs in mouse (NP_079821, NP_032274, NP_057867, NP_081306, AAC24117) bony fish (NP_999894, AAH81526, CAF92838, AAP47149, NP_957034) and yeast (NP_014901, NP_010158, NP_009574, NP_014816, NP_010591) by pBAST. These were then aligned by on-line ClustalW multiple sequence alignment program using the default parameters at the European Bioinformatics Institute web site. The alignment was imported into Molecular Evolutionary Genetic Analysis (MEGA Version 2.1) to build a Neighbor-Joining tree with pairwise deletion and Poisson correction using human Chk2 (NP_001005735) as the root. The reliability of the inferred tree was analyzed by the bootstrap test, 1000 replications.

RESULTS

Cloning and Expression of Recombinant Aprataxin—Full-length Aprataxin cDNA was cloned into a self-cleaving affinity expression construct as described by Hall and Kunkel (17). This yeast expression system is driven by a GAL1 promoter inducing overexpression of an Aprataxin fusion protein. Purification of the fusion protein was achieved by affinity chromatography utilizing the chitin binding domain followed by dithiothreitol-induced self-cleavage. Further purification was carried out using cation exchange chromatography. This approach generated a single detectable protein by Coomassie R250 staining (Fig. 1*A*). An antibody generated against the full-length Aprataxin protein did not detect any degradation products by Western blot analysis (Fig. 1*B*).

This expression system was also used to produce the V263G mutant form of Aprataxin (Fig. 1*C*). This mutation is located within the HIT motif of Aprataxin and was identified in a Japanese AOA1 patient (2). The recovery of mutant protein was 95-fold lower than that obtained for the wild type protein after elution from the chitin resin and less pure as is evident from Fig. 1*C*. This was not due to reduced levels of the induced fusion protein in yeast cell extracts, which were comparable by Western blot analysis (data not shown), but due to reduced binding to the chitin resin and extremely inefficient cleavage even with increased temperature and length of the cleavage compared with the wild type protein.

DNA Binding Activity—AOA1 cells are hypersensitive to DNA-damaging agents and are defective in repair of DNA single-strand breaks (6, 10). Aprataxin has also been shown to bind to several proteins involved in DNA repair and localizes to the site of DNA damage. Accordingly, we investigated whether Aprataxin was capable of binding directly to DNA. Using gel electrophoretic mobility shift assays, we demonstrated that recombinant Aprataxin was able to bind to a 21-base pair double-stranded DNA molecule (Fig. 2*A*, *lanes 5* and *6*). A single discrete binding species was observed at low concentrations of Aprataxin. Increasing

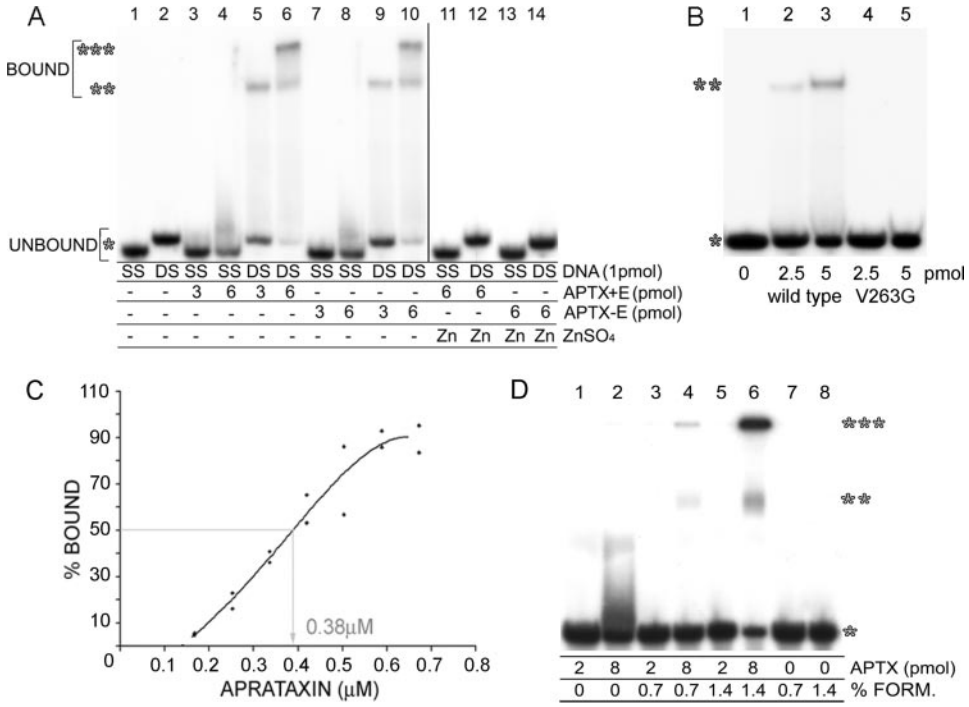

FIGURE 2. *A*, binding of Aprataxin (*APTX*) to a 21bp DNA duplex (*DS*) and single-stranded oligo (*SS*) as detected by electrophoretic mobility shift assays. *Lanes 1* and *2* are control reactions lacking Aprataxin; *lanes 3–6* show Aprataxin purified in the presence of EDTA (*APTX+E*) binding to SS and DS DNA (1 pmol) with increasing protein; *lanes 7–10* show Aprataxin purified in the absence of EDTA (*APTX-E*) binding to SS and DS DNA (1 pmol) with increasing protein; *lanes 11–14* show the effect of zinc sulfate (200 μM) on Aprataxin DNA binding. A *single asterisk* indicates unbound DNA, *two asterisks* indicate the Aprataxin-DNA complex, and *three asterisks* indicate more than one molecule of Aprataxin binding to one DNA molecule. *B*, the affinity of wild type and mutant (V263G) Aprataxin for duplex DNA. The Aprataxin concentration required to bind 50% of the duplex was estimated from the plot of duplicate reactions with increasing protein concentration. *C*, the role of the HIT domain in DNA binding activity. The wild type control (*lanes 2* and *3*) or the HIT motif mutant protein, V263G (*lanes 4* and *5*) were incubated with 1.5 pmol of homoduplex DNA. A *single star* indicates unbound DNA, and *two asterisks* indicates the Aprataxin-DNA complex. *D*, formaldehyde (*FORM.*) cross-linking of Aprataxin to single-strand DNA. *Lanes 1* and *2* show the unstable binding of Aprataxin to single-stranded DNA (2 pmol) with increasing protein. *Lanes 3–6* are the same conditions with the addition of formaldehyde. *Lanes 3* and *4*, 0.74% formaldehyde; *lanes 5* and *6*, 1.6% formaldehyde. *Lanes 7* and *8* are control reactions to ensure that the cross-linked species observed are specific to the addition of Aprataxin. Binding reactions were performed as described under "Materials and Methods" but after binding the samples were incubated at room temperature for 5 min (for cross-linking) then returned to ice just before loading. A *single asterisk* indicates unbound DNA, *two asterisks* indicate the Aprataxin-DNA complex, and *three asterisks* indicate aggregates of cross-linked Aprataxin-DNA complex in the well.

the concentration of Aprataxin gave rise to a second and more retarded band, indicating more than one molecule of Aprataxin binding to each duplex (Fig. 2*A*, *lane 6*). Under the conditions used Aprataxin failed to form a stable complex with single-stranded DNA, but there was evidence of some nonspecific binding or unstable complex formation at higher Aprataxin concentrations (Fig. 2*A*, *lane 4*).

One of the domains identified in Aprataxin by sequence homology was a C_2H_2-type zinc finger (2, 3). The formation of the zinc finger structure depends on a zinc ion that is coordinated by the two cysteines and two histidines, enabling binding to DNA or RNA. Because Aprataxin contains a putative zinc finger domain, protein was purified in the presence and absence of EDTA. This chelator failed to alter the binding to either single or double-strand DNA (Fig. 2*A*, *lanes 3–6* (+EDTA), *lanes 7–10* (−EDTA)). Unexpectedly, the addition of zinc under both conditions abolished DNA binding (Fig. 2*A*, *lanes 11–14*). The other domain that could be responsible for this DNA binding activity is the HIT domain, which is known to bind and hydrolyze nucleotide-based substrates. To investigate the possible role of this domain we generated a mutation in the HIT motif that removes one of the key hydrophobic amino acids in the consensus sequence, V263G. When used in the electrophoretic mobility shift assays in parallel with wild type protein under the same conditions the V263G mutant form failed to bind duplex DNA (Fig. 2*B*). To determine the affinity of Aprataxin for duplex DNA, binding was performed with increasing Aprataxin concentration. When 50% of duplex DNA is bound by Aprataxin, a K_d value of 0.38 μM was obtained (Fig. 2*C*). To further investigate binding to single-stranded DNA, formaldehyde cross-linking was employed. Under these conditions a distinct binding species was observed (Fig. 2*D*, *lane 4*). As the concentration of formaldehyde was increased, the amount of this stabilized complex increased (*lane 6*). It is also evident that some cross-linked material remained in the well.

To investigate the kinetics of binding, a time course experiment was performed. Because this was found to be rapid, incubations were performed on ice. Within 1 min post-incubation, Aprataxin DNA binding was evident (Fig. 3*A*). Competition experiments with unlabeled duplex DNA were used to investigate the dynamics of binding to DNA and whether single-stranded DNA could disrupt the Aprataxin-DNA complex. The results in Fig. 3*B* reveal a linear decrease of 57% (S.D. = 9) with each fold increase in unlabeled competitor over the range 1–8-fold. To investigate the dynamics of the competition with either duplex or single-stranded DNA, unlabeled competitor DNA was added at 8-fold excess for 1–20 min. The duplex DNA at 1 min (Fig. 3*C*, *lane 3*) was able to quickly compete for Aprataxin to the same level as that observed at the 20 min time point (*lane 5*). When an 8-fold excess of single-stranded DNA was added to the Aprataxin-DNA complex, it failed to disrupt the complex at 1, 10, and 20 min (Fig. 3*C*, *lanes 6–8*), which is in agreement with the reduced affinity of Aprataxin for single-strand DNA. To investigate the footprint size of Aprataxin on duplex DNA, we employed 11-, 21-, and 37-bp substrates. The results for Fig. 3*D* show that the affinity of Aprataxin for the 11-bp fragment is significantly less than that for the other 2, and only a single retarded fragment was observed. Two distinct binding species were observed with the 21-bp substrate and up to 4 binding species with the 37-bp fragment, suggesting a footprint size of ~10 bp. In view of the increased affinity of Aprataxin for the 37-bp

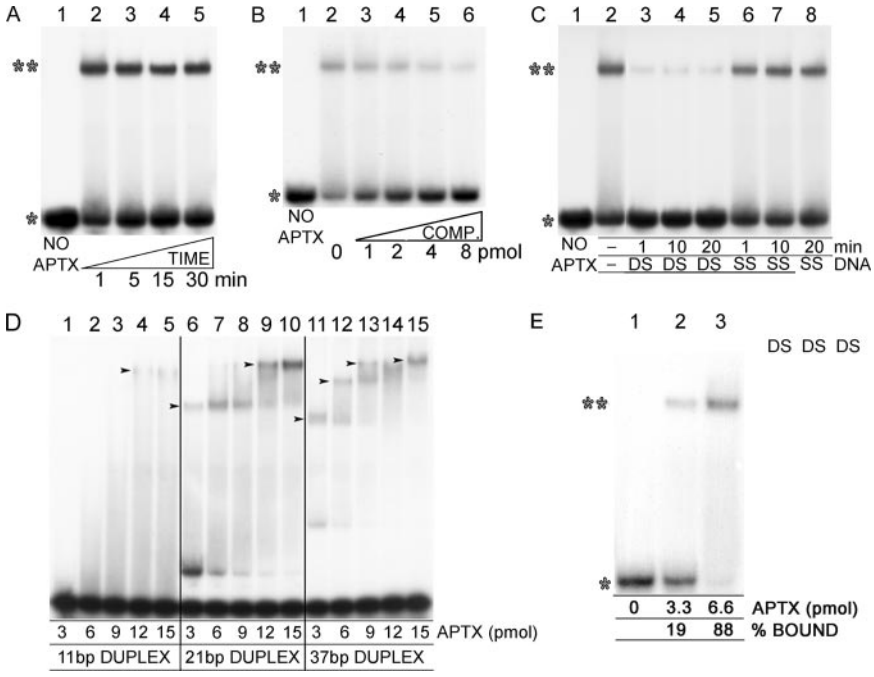

FIGURE 3. **DNA binding characteristics of Aprataxin.** *A*, time course for stable Aprataxin-DNA complex formation. Aprataxin (3pmol) was incubated with 1 pmol of DS DNA duplex for 1–30 min before electrophoresis. A *single asterisk* indicates unbound DNA, and *two asterisks* indicate the Aprataxin-DNA complex. *B*, stability of the Aprataxin-DNA complex. Preformed Aprataxin-DNA complex was used in competition experiments with increasing amounts of unlabeled DNA (*COMP.*). Aprataxin (3pmol) was incubated with 1 pmol of DS DNA duplex for 15 min before the addition of 0–8 pmol of unlabeled competitor duplex 5 min before electrophoresis. A single *asterisk* indicates unbound DNA, and *two asterisks* indicate the Aprataxin-DNA complex. *C*, stability of the Aprataxin-DNA complex in the presence of an 8-fold excess cold DS DNA or SS DNA over time. Aprataxin (3pmol) was incubated with 1 pmol of DS DNA duplex for 15 min (*lane 2*). An 8-fold excess (8pmol) of unlabeled DS duplex was added to this preformed Aprataxin-DNA complex for 1, 10, and 20 min (*lane 3–5*) or unlabeled SS DNA for 1, 10, and 20 min (*lanes 6-8*). A *single asterisk* indicates unbound DNA, and *two asterisks* indicate the Aprataxin-DNA complex. *D*, binding of Aprataxin to an 11-bp (*lanes 1–5*), 21-bp (*lanes 6–10*), and 37-bp (*lanes 11–15*) duplex with increasing protein concentration (3–15 pmol) to estimate the binding footprint on DNA (1 pmol). The *arrowheads* indicate each of the discrete species detected. Binding to the 11-bp duplex gave rise to one discrete band at the highest protein concentration (*lanes 4* and *5*). Two discrete bands were observed on the 21-bp duplex (*lanes 9* and *10*) and 4 on the 37-bp duplex (*lane 15*). The very lower band across all lanes is the un-incorporated nucleotide after radiolabeling. This happens to co-migrate with the unbound 11-bp duplex. *E*, binding of Aprataxin to a 37-mer single-strand oligonucleotide (2 pmol). Binding was carried out with an increasing amount of Aprataxin (*lanes 2* and *3*). The percent found was estimated using ImageQuant program.

duplex fragment, we investigated the capacity of Aprataxin to bind to a larger single-strand oligonucleotide (37 nucleotides). In contrast to that observed with the 21-mer single-strand oligonucleotide, a stable Aprataxin-DNA complex was observed with the longer oligonucleotide in the absence of a cross-linking agent (Fig. 3*E*). No additional shifted binding species here observed with increasing Aprataxin concentration, indicating that no additional molecules of Aprataxin bound the oligonucleotide.

Aprataxin has been shown to interact with XRCC1 and poly(ADP-ribose) polymerase, both of which are involved in base excision repair and repair of single-strand breaks. To investigate whether Aprataxin displayed DNA structure specific binding, we employed a series of DNA structures that mimic various transition states in the process of repair. The introduction of a single mismatch (G-A) into the duplex sequence had no effect on efficiency of binding compared with the homoduplex (Fig. 4*A*). The efficiency of binding was reduced to between 28 and 72% when either a single flap, nick, or gap was introduced into the duplex. However, all of these structures effectively give rise to a shorter continuous duplex for which Aprataxin has a significantly reduced affinity (see Fig. 3*D*). The introduction of a nick or a gap into the larger 37-bp duplex did not lead to a significant increase in DNA binding affinity compared with the homoduplex (Fig. 4*B*).

Gueven *et al.* (6) have shown that Aprataxin is localized not only to the nucleus but also the nucleolus, where it associates with nucleolin. This nucleolar localization could mean it associates with RNA, as has been reported for nucleolin (19, 20). Because Aprataxin binds efficiently to duplex DNA, we investigated whether it could also bind double-stranded RNA. The results in Fig. 5*A* show that Aprataxin displays a

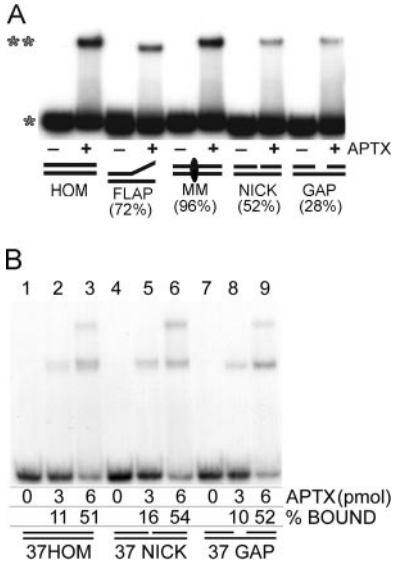

FIGURE 4. **Comparison of Aprataxin binding to various DNA structures mimicking possible transition states during repair.** *A*, structures investigated include a single flap, mismatch (G-A), nick, and gap as compared with the homoduplex (*HOM*) DNA substrate. Aprataxin (2 pmol) was incubated with 1 pmol of each radiolabeled 21-bp DNA substrate for 15 min before electrophoresis. A *single asterisk* indicates unbound DNA, and *two asterisks* indicate Aprataxin DNA complexes. *B*, comparison of Aprataxin (3 and 6 pmol) binding to a 37-bp homoduplex and nicked-homoduplex and single nucleotide gap homoduplex (2 pmol of each). Incubation was for 15 min. The percent bound as compared with the homoduplex control was estimated using the ImageQuant program (Amersham Biosciences) and is indicated in *parentheses*. A *single asterisk* indicates unbound DNA, and *two asterisks* indicate the Aprataxin-DNA complex.

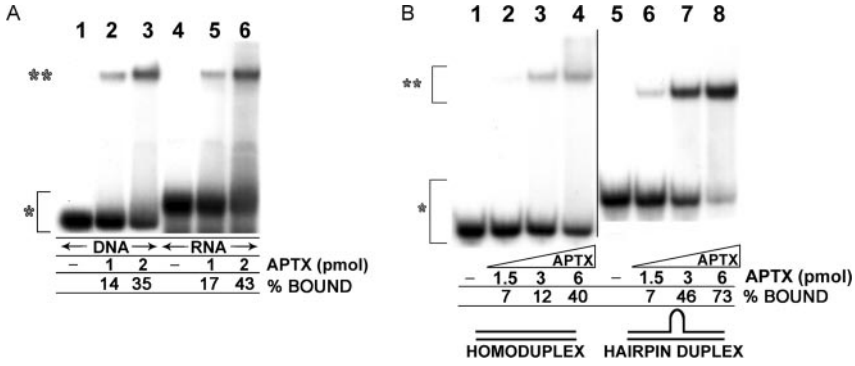

FIGURE 5. *A*, comparison of Aprataxin binding to DNA and RNA. Aprataxin was incubated with 0.4 pmol of radiolabeled DS DNA or RNA for 15 min with 0 pmol (*lanes 1* and *4*), 1 pmol (*lanes 2* and *5*), and 2 pmol (*lanes 3* and *6*) of Aprataxin. A *single asterisk* indicates unbound DNA or RNA, and *two asterisks* indicates the Aprataxin-DNA complex. *B*, comparative binding of Aprataxin to homoduplex (21 bp) hairpin duplex. To investigate differences in affinity to this secondary structure as compared with homoduplex DNA, Aprataxin (0, 1.5, 3, and 6 pmol) was bound to 2 pmol of radiolabeled homoduplex (*lanes 1–4*) and hairpin duplex (*lanes 5–8*) for 15 min. A *single asterisk* indicates unbound DNA, and *two asterisks* indicate the Aprataxin-DNA complex.

similar affinity for binding double-stranded RNA and DNA. Because it is not expected that RNA in the nucleolus will be double-stranded but capable of giving rise to secondary structures composed of double-stranded regions, such as hairpin loops, the binding we observe here may be explained by Aprataxin binding to RNA hairpin loops. This is supported by the results in Fig. 5*B* which show that Aprataxin binds to a duplex structure designed to create a centrally located 6-base pair hairpin loop. This binding occurs with greater affinity than with the homoduplex DNA.

Catalytic Activity of Aprataxin—A HIT domain has been shown to be centrally located in Aprataxin (3, 6). This domain characterizes a superfamily of proteins with nucleotide hydrolase and transferase activity (11). Aprataxin has been classified as a HINT protein within the HIT superfamily on the basis of sequence analyses (11, 14). We investigated whether Aprataxin has activity characteristic of a HIT or HINT protein using a high performance liquid chromatography-based assay to assess nucleotide cleavage. When $AMPNH_2$ (a model substrate for HINT proteins) was employed as a substrate for Aprataxin, the cleavage product generated, AMP, was identified by retention time and co-elution with an AMP standard. The catalytic activity of Aprataxin on $AMPNH_2$ substrate was $10.06 \times 10^{-3}s^{-1}$ (Fig. 6*A*). This activity was comparable with that obtained with the FHIT substrate Ap_4A with an activity of $10.79 \times 10^{-3}s^{-1}$, producing AMP and ATP as the products of hydrolysis (Fig. 6*A*). Very low activity was obtained when ATP ($0.26 \times 10^{-3}s^{-1}$) was used as the substrate, and there was no detectable hydrolysis of dATP over a time period of 3 h. To further investigate the catalytic activity of Aprataxin on Ap_4A (Fig. 6*B*) and $AMPNH_2$ (Fig. 6*C*) substrates, we estimated several kinetic parameters. The turnover values were comparable and low for both Ap_4A ($K_{cat} = 11.5 \times 10^{-3}s^{-1}$) and $AMPNH_2$ ($K_{cat} = 19.5 \times 10^{-3}s^{-1}$). The affinity of Aprataxin for these two substrates was vastly different, with a 46-fold higher affinity for Ap_4A ($K_m = 18.0\ \mu M$) than $AMPNH_2$ ($K_m = 837.5\ \mu M$). The V263G mutant displayed no detectable activity on the Ap4A substrate.

To determine whether the DNA binding by Aprataxin might influence hydrolytic activity, we monitored the catalytic activity in the presence of DNA. The addition of increasing concentrations of duplex DNA to a reaction mix containing the Ap_4A substrate significantly reduced catalytic activity. At 50% inhibition the K_i for duplex DNA was 0.05 μM (Fig. 7*A*). Single-stranded DNA was less effective, with a K_i of 3.8 μM at 50% inhibition (Fig. 7*B*). The kinetics of inhibition reflected the binding capacity of the different forms of DNA observed in the electrophoretic mobility shift assays. This inhibition might be explained by competition between the Ap_4A and DNA for the catalytic site. Alternatively, DNA might induce a conformational change in Aprataxin after binding, which would alter access by Ap_4A to the active site.

A previous report has shown that the enzymatic activity of Aprataxin is negatively regulated by an intramolecular interaction involving the N-terminal domain (13). They showed that an N-terminal truncated

A

SUBSTRATE (1.5 mM)	ACTIVITY pmols AMP/second/pmol Aprataxin
Ap_4A	10.79×10^{-3}(10.44×10^{-3};11.14×10^{-3})
$AMPNH_2$	10.06×10^{-3}(9.92×10^{-3};10.21×10^{-3})
ATP	0.26×10^{-3}(0.27×10^{-3};0.25×10^{-3})
dATP	no detectable hydrolysis

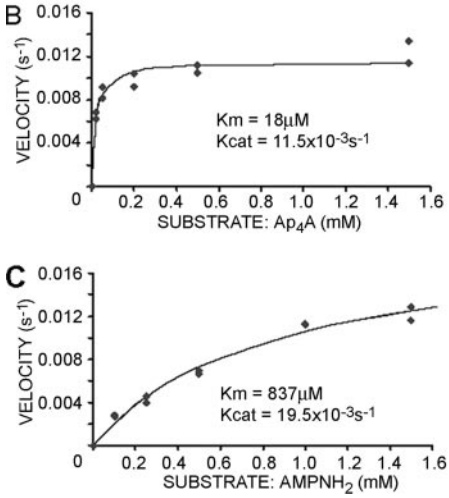

FIGURE 6. *A*, catalytic activity of Aprataxin on a range of nucleotide-based substrates. Aprataxin was incubated for 3 h with 1.5 mM concentrations of each putative substrate (Ap_4A, $AMPNH_2$, dATP, and ATP) at 30 °C. The amount of AMP produced over this time was used to calculate the reaction velocity. *B*, kinetic parameters of Aprataxin on the FHIT substrate, Ap_4A. A plot of the activity obtained in the presence of a range of Ap_4A substrate concentrations (0.02–1.5 mM) *versus* velocity (pmol of AMP produced/s). These data, performed in duplicate, were fitted to the Michaelis equation to calculate the K_m and K_{cat} (given under the curve). *C*, kinetic parameters of Aprataxin on the HINT substrate, $AMPNH_2$. A plot of the activity obtained in the presence of a range of $AMPNH_2$ substrate concentrations (0.1–1.5 mM) *versus* velocity (pmol of AMP produced/s). These data performed in duplicate were fitted to the Michaelis equation to calculate the K_m and K_{cat} (given under the curve).

form of the protein was substantially more active than the purified full-length form. This N-terminal region corresponds to the FHA domain which binds to a phosphorylated region of XRCC1 (7, 9). We investigated the role of the FHA and other functional domains on the nucleotide hydrolase and DNA binding activities of Aprataxin using a series of truncated recombinant forms (Fig. 8*A*). The purified proteins employed are shown in Fig. 8*B*. Use of the Ap_4A substrate demonstrated that full-length Aprataxin and an N-terminal-truncated form lacking the FHA domain had comparable hydrolytic activities with K_{cat} values of 17.1×10^{-3} and $17.6 \times 10^{-3}s^{-1}$, respectively (Fig. 8, *C* and *D*). The other truncated forms of Aprataxin had no detectable hydrolase activity (Fig. 8*A*). Furthermore, only the full-length and N-terminal-truncated forms of Aprataxin formed stable complexes with DNA (Fig. 8, *E* and *F*). In addition, the use of the XRCC1 triply-phosphorylated peptide that was previously shown to bind the FHA domain of Aprataxin (9) did not

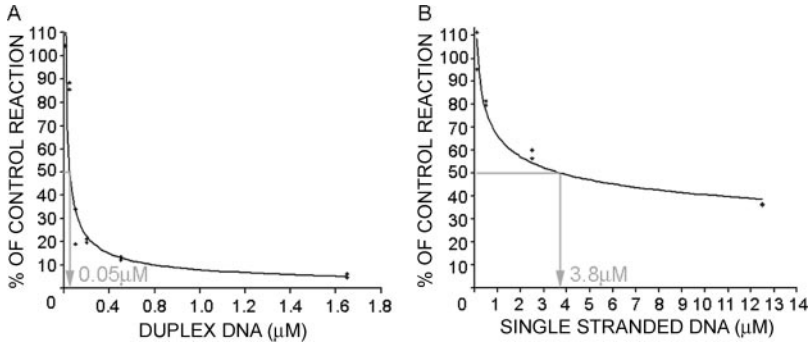

FIGURE 7. *A*, the effect of the 21-bp duplex DNA on the catalytic activity of Aprataxin (1 mM) using Ap_4A as a substrate. The velocity of the reaction is expressed as a percent of the control reaction, that has no DNA and is plotted against increasing DNA concentration in the presence of 500 μM Ap_4A. The concentration at which 50% inhibition occurs is 0.05 μM DNA. *B*, the effect of the 21-bp single-stranded DNA on the catalytic activity of Aprataxin (1 μM) using Ap_4A as a substrate. The velocity of the reaction is expressed as a percent of the control reaction, which has no DNA and is plotted against increasing DNA concentration in the presence of 500 μM Ap_4A. The concentration at which 50% inhibition occurs is 3.8 μM DNA.

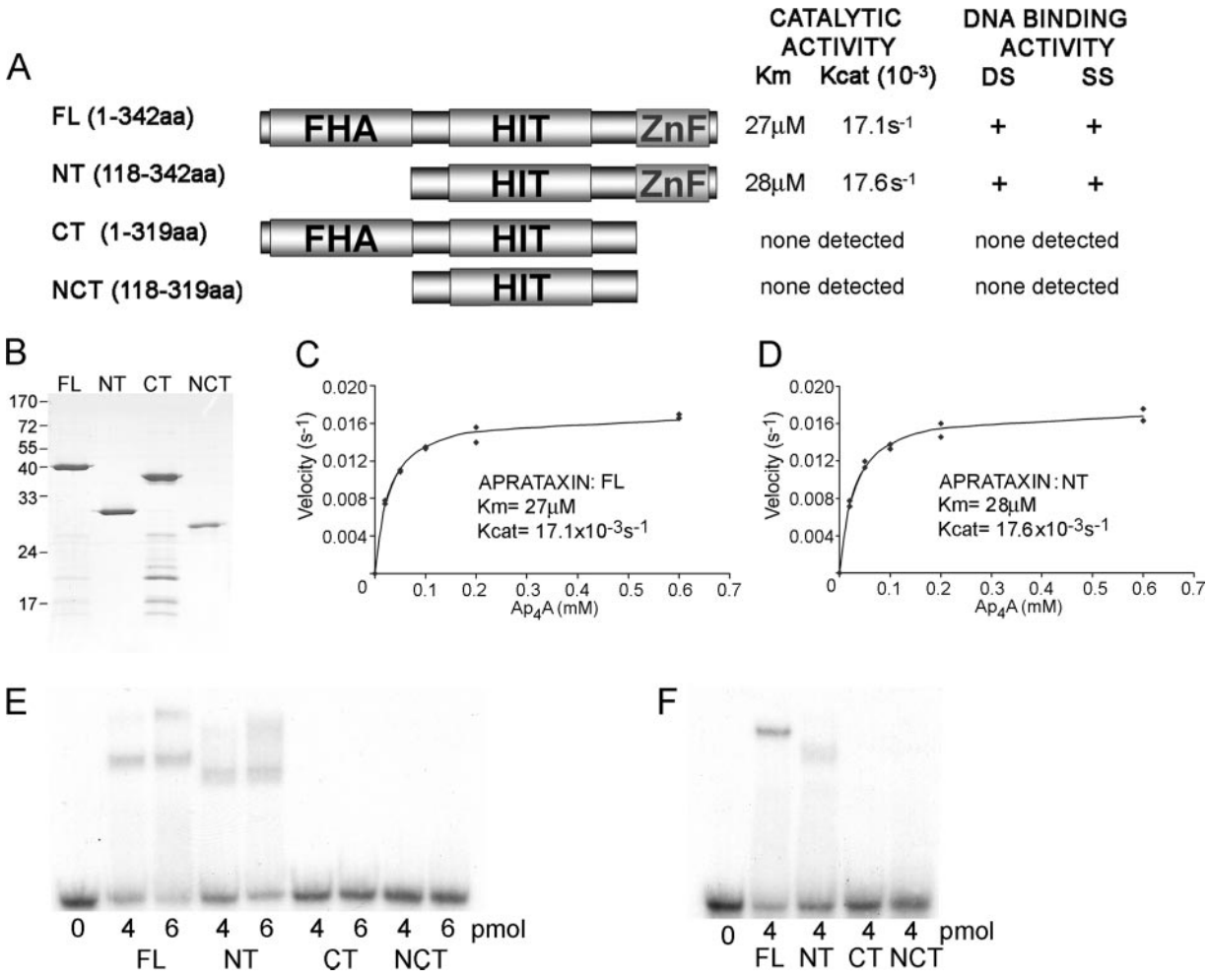

FIGURE 8. **Structure and activity of different recombinant forms of Aprataxin.** *A*, schematic representation of full-length Aprataxin and truncated forms and a summary of catalytic and DNA binding activities for these different forms. *FL*, full-length; *NT*, N-terminal truncation; *CT*, C-terminal truncation; *NCT*, N- and C-terminal truncated. *B*, Coomassie staining of purified recombinant Aprataxin and truncated forms (30 pmol of each). FL corresponds to a molecular of 40 kDa; NT, 26 kDa; CT, 37 kDa; NCT, 23 kDa. *C*, kinetic parameters for full-length Aprataxin using Ap_4A substrate. Activity was determined over the range of substrate concentration (0.02–0.6 mM). K_m and K_{cat} were calculated as described in the legend to Fig. 6. Velocity is expressed as pmol of AMP produced/pmol of Aprataxin/s. *D*, kinetic parameters for N-terminal-truncated Aprataxin using Ap_4A substrate. Velocity is expressed as in *C* above. *E*, binding of recombinant forms of Aprataxin to DS DNA (2pmol). *F*, binding of recombinant forms of Aprataxin to SS DNA (2pmol).

interfere with either the DNA binding or nucleotide hydrolase activities (results not shown). A summary of the results for catalytic and DNA binding activities for the different forms appears in Fig. 8*A*.

Sequence Analysis of the HIT Superfamily—There exists some disagreement on the relationship of Aprataxin to the other HIT superfamily members as to whether it is more closely related to the FHIT or HINT branches (2, 11, 13, 14). To address this we have performed a phylogenetic analysis with members from all branches and several orthologs from each to gain a more robust insight. Sequences from each branch for human, mouse, bony fish, and yeast were aligned by ClustalW and used to build a Neighbor Joining bootstrap consensus tree. When this analysis was applied, Aprataxin formed a discrete cluster like each of the other previously described branches of the HIT superfamily, HINT, FHIT, galactose-1-phosphate uridylyltransferase, and DcpS (Fig. 9). Each branch is deeply rooted with a yeast ortholog.

DISCUSSION

It is well established that Aprataxin plays a role in protecting cells against genotoxic stress. This evidence stems from its interaction with a number of proteins involved in DNA break repair, including poly(ADP-

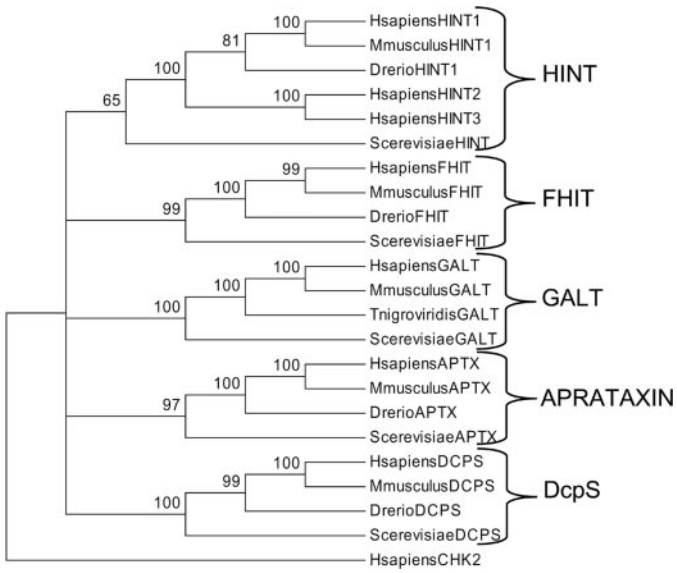

FIGURE 9. **Phylogenetic relationships of the HIT superfamily of proteins, with representatives from all branches including Aprataxin.** The protein sequences from each branch (indicated by the *bracket*) including Aprataxin were aligned by ClustalW then used to build a Neighbor Joining tree (with pairwise deletion) to investigate how these sequences relate to each other using human CHK2 protein as the root. The bootstrap test was applied with 1000 replicates, and the values are indicated on the *top* at each node. *Drerio*, *Danio rerio*; *Tnigroviridis*, *Tetraodon nigroviridis*.

ribose) polymerase 1, XRCC1, p53, and XRCC4 and the increased sensitivity of AOA1 cells to agents that cause single-strand breaks in DNA (6, 7). It also participates in a novel single-strand break repair complex together with XRCC1 and DNA ligase 3, and recent evidence shows that AOA1 cells have a reduced capacity to repair single-strand breaks in DNA (9, 10). Furthermore, exposure of cells expressing EGFP-Aprataxin to uranium ions gives rise to EGFP-Aprataxin foci distributed along particle trajectories as discreet tracts in the nucleus, demonstrating that it associates with chromatin in response to DNA damage (6). Complete co-localization with XRCC1 across the length of these tracts was demonstrated, indicating that both proteins were localized to the same sites of DNA damage. We demonstrated here that a purified recombinant full-length form of Aprataxin was able to bind duplex DNA. This was dynamic binding as it was readily competed off the duplex. The inability of an 8-fold excess of single-strand DNA to disrupt the Aprataxin double-stranded DNA complex supports a preference for duplex DNA. The affinity of binding for duplex DNA increases with increasing duplex length. Reduced affinity of binding was observed with the 11-bp duplex compared with the 21- and 37-bp fragments. Disruption of the 37-bp duplex with a nick or single nucleotide gap did not impact on the affinity. The estimated footprint size on duplex DNA represents approximately one turn of the helix (10 bp).

Aprataxin failed to form a stable complex with a 21-mer single-strand oligonucleotide, but formaldehyde cross-linking gave rise to a stable complex. Increasing the length of the oligonucleotide to a 37-mer allowed for stable binding without the need for a cross-linker. Under these conditions only one binding species was observed, whereas 3–4 molecules of Aprataxin bound the 37-bp duplex. The stable binding to the 37-mer oligonucleotide compared with the 21-mer might be explained by the capacity to form a secondary structure. Analysis of secondary structure using OligoAnalyzer 3.0 (Integrated DNA Technologies) predicts 3 stem loop structures involving 3–7-nucleotide pairing for the 37-mer. Although these predicted structures have low ΔG values (0.98–1.25 kcal·mol^{-1}), it is possible that Aprataxin stabilizes the predicted duplex regions, leading to stable complex formation. None of these pairings is predicted for the 21-mer. Aprataxin joins a number of other DNA-binding proteins that display non-sequence-specific DNA binding characteristics and are involved in DNA repair, transcription, DNA replication, and chromosome maintenance. In this context a recent report by Weinberg *et al.* (21) demonstrated that the C terminus of p53, expressed in bacteria, binds to nonspecific short double-stranded DNA fragments as part of a scanning process before finding specific binding sites. This could well apply to Aprataxin, where it binds DNA in a nonspecific fashion before finding sites of damage where it may exert its activity as part of a repair complex.

The DNA binding was not dependent on the deoxyribose portion of the DNA duplex structure since Aprataxin also bound double-stranded RNA. This could mean that it has an additional role in the repair or processing of RNA. We have previously shown that Aprataxin localizes to both the nucleus and the nucleolus (6). In lymphoblastoid cells, up to 50% of Aprataxin is present in nucleoli, yet the role of Aprataxin in this subnuclear structure is not yet known. These subnuclear structures are multifunctional, being the sites for ribosome biogenesis, RNA processing, cell cycle control, and viral replication (22, 23). It is also of interest that nucleoli have been shown to be a storage site for proteins involved in DNA damage recognition and repair (23–25). This may also apply to Aprataxin, but failure to see any net movement of this protein from nucleoli in response to DNA damage does not support this possibility (6). Alternatively, it may play an additional role in some aspect of RNA processing. Its association with an RNA-binding protein, nucleolin in the nucleolus, supports such a role (6). The double-stranded RNA binding we observed here may represent the capacity of Aprataxin to bind secondary structure in mRNA or small nucleolar RNAs as exemplified by the higher affinity of Aprataxin for the 6-base pair hairpin structures that would mimic RNA secondary structure.

Three functional domains have been identified in Aprataxin on the basis of homology (3, 6). These include the FHA domain, which facilitates binding to phosphorylated proteins, a HIT domain for nucleotide binding and hydrolysis, and a putative C-terminal zinc finger that could allow binding to DNA and or RNA. Although Aprataxin contains only a single C_2H_2 type zinc finger domain, the presence of this domain could be sufficient for DNA binding, but most zinc fingers are usually present in multiple copies (26). Mendez-Vidal *et al.* (27) have shown that a single zinc finger in the mouse Wig1 protein is essential for binding double-stranded RNA, whereas the two additional zinc fingers are dispensable. Initial DNA binding studies were carried out in the absence of exogenous zinc, and preincubation with low concentrations of zinc completely abolished this binding, suggesting that it is unlikely that the zinc finger motif is responsible for the DNA binding observed here. Inclusion of EDTA during purification of zinc finger proteins has been shown to chelate any zinc ions from these proteins (28) and, thus, prevent them from forming the required structure to bind nucleic acid. In this study purification of recombinant Aprataxin in the presence or absence of EDTA did not impact on DNA binding, providing further evidence for lack of dependence on zinc. Interestingly, the HIT domain mutant, V263G, was unable to bind DNA, suggesting that this domain is likely to be involved in the DNA binding observed here. But the V263G protein co-purified with a heat shock protein, indicating the protein could be misfolded compared with the wild type protein as a result of the mutation or an artifact of overexpression. The same was found when this mutant was expressed in a bacterial system (data not shown).

The third functional domain, a centrally located HIT domain, has been predicted to have nucleotide binding and hydrolase activities (2, 3, 11). We demonstrated here that Aprataxin has hydrolase activity against HINT and FHIT model substrates, producing AMP as the prod-

uct of hydrolysis in both reactions, as has been observed for all of the HINT and FHIT branch proteins investigated to date. A recombinant form of the HIT domain alone failed to display hydrolase activity. Aprataxin displayed comparable levels of activity on both the HINT ($AMPNH_2$) and FHIT (Ap_4A) substrates used here. These activities were ~10-fold lower than either of the activities reported for rabbit Hint1 activity on $AMPNH_2$ (29) and human FHIT activity on Ap_4A (30). The more distinguishing kinetic parameter was Aprataxin affinity, which was 47-fold higher for the FHIT substrate, Ap_4A ($K_m = 18.0\ \mu M$), than for the HINT substrate, $AMPNH_2$ ($K_m = 837.5\ \mu M$). An even better HINT substrate has been identified; this is an AMP linked to a lysine residue. Rabbit Hint1 displays much higher turnover, $K_{cat} = 0.23 s^{-1}$, and also higher affinity, $K_m = 0.47\ \mu M$, for this substrate (31). Seidle *et al.* (15) recently reported catalytic activity of Aprataxin using an AMP-lysine-based substrate and found that Aprataxin displayed 100-fold lower affinity ($K_m = 47\ \mu M$) and turnover ($K_{cat} = 3 \times 10^{-3} s^{-1}$) than Rabbit Hint1, well below what would be expected for this branch of proteins. We observed that removal of the FHA domain from Aprataxin did not alter the K_m or K_{cat} for the Ap_4A substrate. This contrasts with two other studies that show either a reduced (15) or increased activity (13) when the N terminus of the protein is absent. These results might be explained by the differences in the length of the N-terminal sequence remaining in the different protein forms.

The presence of DNA duplex and single-stranded DNA to a lesser extent was able to effectively inhibit the catalytic activity observed here. This inhibition may be in part due to a conformational change that occurs in the presence of DNA but may also be due to competition for the HIT domain. The importance of the HIT domain for the two activities described here is further supported by disruption of activity by the V263G mutation in this domain and by retention of activity when the FHA domain is deleted, but with the remainder of the molecule intact. Loss of activity by deletion of the C terminus points to the importance of this end of the molecule in maintaining the proteins catalytic and DNA binding activities. DNA binding capacity may have evolved through functional diversification of the HIT domain where the low level of nucleotide hydrolase activity reported here and by others (13, 15) may represent residual activity of an ancestral function. Crystallographic studies in the presence of DNA will provide more definitive evidence for the HIT domain in Aprataxin representing a novel DNA and RNA binding domain.

The catalytic activities reported here and elsewhere (13, 15) together support the notion that none of the substrates assayed, either FHIT or HINT substrates, display levels of turnover that are likely to be physiologically relevant and that Aprataxin may represent a distinct family within the HIT superfamily, implying that a better substrate is yet to be identified. The proposal that Aprataxin belongs to or is more closely related to either the FHIT or HINT branches of the HIT superfamily of proteins is based on limited sequence and phylogenetic tree analysis (2, 11, 13, 14). The hydrolytic activities determined in this study on nucleotide-based substrates and in other studies suggest that Aprataxin has a unique activity profile distinct from the HINT and FHIT branches (13, 5). We propose that Aprataxin forms a separate branch in the HIT superfamily of proteins. To further substantiate this, we carried out a more thorough analysis of the sequence relationships within and between the HIT superfamily branches using several orthologs from each branch. Aprataxin was found to form a discrete cluster, like each of the other previously described branches in the HIT superfamily (HINT, FHIT, galactose-1-phosphate uridylyltransferase, and DcpS). This cluster like the others is deeply rooted, and the divergence appears to predate the radiation of the eukaryotes, indicating that each family is very

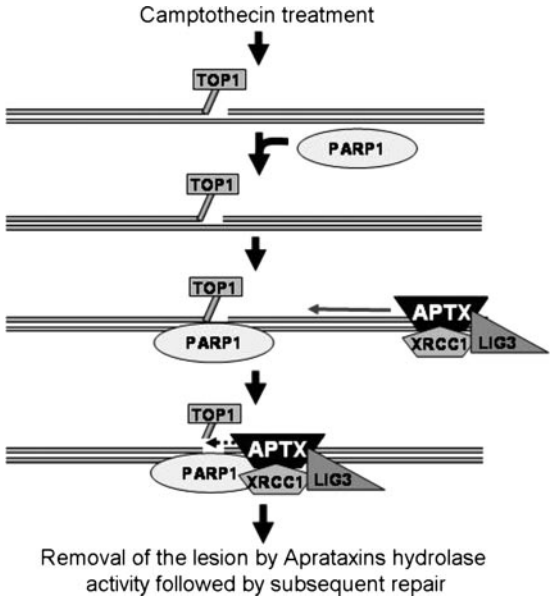

FIGURE 10. **Model of the Aprataxin role in the repair of single-strand breaks.** Camptothecin induces the formation of a trapped topoisomerase 1 (*TOP1*) cleavage complex, which leads to the recruitment of poly(ADP-ribose) polymerase (*PARP*) by direct binding to topoisomerase I as described recently (32). This in turn provides a signal for the Aprataxin/XRCCI/DNA ligase 3 to locate to the site of damage. This complex has been reported by Luo *et al.* (9). It is envisaged that the capacity of Aprataxin to bind duplex DNA facilitates recruitment of this complex to the site of damage. Formation of this complex may be enhanced by the known capacity of all of these molecules to interact (6–9). This would then allow Aprataxin to remove this lesion with its hydrolase activity. This would be followed by the involvement of the base excision repair machinery.

distantly related. This is also supported by the extremely low sequence identities observed here, for example the pairwise sequence identity of human Aprataxin and FHIT protein sequences was only 20%. Overall this suggests that some HIT superfamily members may have evolved by convergent evolution. This divergent nature of the HIT superfamily sequences was only recognized when the crystal structure for galactose-1-phosphate uridylyltransferase protein was compared with Hint1, allowing it to be classified into the HIT superfamily due to the structural similarities.

We have described here two functional activities for recombinant full-length Aprataxin, which are the capacity to bind RNA and DNA and the ability to hydrolyze nucleotide based substrates. Several previous studies using cellular-based investigations including, localization, interactions, and repair assays have implicated Aprataxin in DNA break repair (6, 9, 10). When combined with the observations made in this study, demonstrating *in vitro* binding of Aprataxin to DNA, we propose that its role may be to scan the genome through its capacity to bind double-stranded DNA in a sequence-independent manner (Fig. 10). In response to DNA damage, Aprataxin would localize to the lesion as part of a complex with DNA ligase 3 and XRCCI as described by Luo *et al.* (9). Once recruited, Aprataxin, through its nucleotide hydrolase activity, would process the lesion as part of a base excision repair mechanism (Fig. 10). The increased sensitivity to camptothecin observed in AOA1 cells implicates the topoisomerase 1 cleavage complex as a substrate for Aprataxin. The more general sensitivity to agents that cause single-stranded breaks in DNA suggests that Aprataxin may have several such substrates.

The analysis applied here confirms the classification of the HIT superfamily with four distinct branches and, in addition, demonstrates that Aprataxin forms a distinct branch. Furthermore, the data described here add a new function to the HIT superfamily of proteins, which is the capacity to bind nucleic acids in addition the ability to bind and hydro-

lyze nucleotide-based substrates. The challenge ahead is to define the role of Aprataxin in DNA repair and possibly RNA processing and identify the specific physiological substrates.

Acknowledgments—We thank the Dr. T. A. Kunkel laboratory (in particular, Molly), which provided the pMH1 vector and yeast strain used in this study, our laboratory colleague, Dr. Rick Woods, who provided the diced RNA products used in this study, and Dr. Steve Smerdon for the use of laboratory facilities. We also thank the Australian National Health and Medical Research Council for research support and Tracey Laing for typing the manuscript.

REFERENCES

1. Le Ber, I., Bouslam, N., Rivaud-Pechoux, S., Guimaraes, J., Benomar, A., Chamayou, C., Goizet, C., Moreira, M. C., Klur, S., Yahyaoui, M., Agid, Y., Koenig, M., Stevanin, G., Brice, A., and Durr, A. (2004) *Brain* **127,** 759–767
2. Date, H., Onodera, O., Tanaka, H., Iwabuchi, K., Uekawa, K., Igarashi, S., Koike, R., Hiroi, T., Yuasa, T., Awaya, Y., Sakai, T., Takahashi, T., Nagatomo, H., Sekijima, Y., Kawachi, I., Takiyama, Y., Nishizawa, M., Fukuhara, N., Saito, K., Sugano, S., and Tsuji, S. (2001) *Nat. Genet.* **29,** 184–188
3. Moreira, M., Barbot, C., Tachi, N., Kozuka, N., Uchida, E., Gibson, T., Mendonca, P., Costa, M., Barros, J., Yanagisawa, T., Watanabe, M., Ikeda, Y., Aoki, M., Nagata, T., Coutinho, P., Sequeiros, J., and Koenig, M. (2001) *Nat. Genet.* **29,** 189–193
4. Tranchant, C., Fleury, M., Moreira, M. C., Koenig, M., and Warter, J. M. (2003) *Neurology* **60,** 868–870
5. Date, H., Igarashi, S., Sano, Y., Takahashi, T., Takahashi, T., Takano, H., Tsuji, S., Nishizawa, M., and Onodera, O. (2004) *Biochem. Biophys. Res. Commun.* **325,** 1279–1285
6. Gueven, N., Becherel, O. J., Kijas, A. W., Chen, P., Howe, O., Rudolph, J. H., Gatti, R., Date, H., Onodera, O., Taucher-Scholz, G., and Lavin, M. F. (2004) *Hum. Mol. Genet.* **13,** 1081–1093
7. Clements, P. M., Breslin, C., Deeks, E. D., Byrd, P. J., Ju, L., Bieganowski, P., Brenner, C., Moreira, M., Taylor, A. M. R., and Caldecott, K. W. (2004) *DNA Repair* **3,** 1493–1502
8. Sano, Y., Date, H., Igarashi, S., Onodera, O., Oyake, M., Takahashi, T., Hayashi, S., Morimatsu, M., Takahashi, H., Makifuchi, T., Fukuhara, N., and Tsuji, S. (2004) *Ann. Neurol.* **55,** 241–249
9. Luo, H., Chan, D. W., Yang, T., Rodriguez, M., Chen, B. P., Leng, M., Mu, J., Chen, D., Songyang, Z., Wang, Y., and Qin, J. (2004) *Mol. Cell. Biol.* **24,** 8356–8365
10. Mosesso, P., Piane, M., Palitti, F., Pepe, G., Penna, S., and Chessa, L. (2005) *Cell. Mol. Life Sci.* **62,** 485–491
11. Brenner, C. (2002) *Biochemistry* **41,** 9003–9014
12. Liu, H., Rodger, N. D., Jiao, X., and Kiledjian, M. (2002) *EMBO J.* **21,** 4699–4708
13. Hirano, M., Furiya, Y., Kariya, S., Nishiwaki, T., and Ueno, S. (2004) *Biochem. Biophys. Res. Commun.* **322,** 380–386
14. Kwasnicka, D. A., Krakowiak, A., Thacker, C., Brenner, C., and Vincent, S. R. (2003) *J. Biol. Chem.* **278,** 39051–39058
15. Seidle, H. F., Bieganowski, P., and Brenner, C. (2005) *J. Biol. Chem.* **280,** 20927–20931
16. Draganescu, A., Hodawadekar, S. C., Gee, K. R., and Brenner, C. (2000) *J. Biol. Chem.* **275,** 4555–4560
17. Hall, M. C., and Kunkel, T. A. (2001) *Protein Expression Purif.* **21,** 333–342
18. Chi, N. W., and Kolodner, R. D. (1994) *J. Biol. Chem.* **269,** 29984–29992
19. Serin, G., Joseph, G., Faucher, C., Ghisolfi, L., Bouche, G., Amalric, F., and Bouvet, P. (1996) *Biochimie* (*Paris*) **78,** 530–538
20. Ghisolfi-Nieto, L., Joseph, G., Puvion-Dutilleul, F., Amalric, F., and Bouvet, P. (1996) *J. Mol. Biol.* **260,** 34–53
21. Weinberg, R. L., Freund, S. M., Veprintsev, D. B., Bycroft, M., and Fersht, A. R. (2004) *J. Mol. Biol.* **342,** 801–811
22. Shaw, P., and Doonan, J. (2005) *Cell Cycle* **4,** 102–105
23. Olson, M. O. (2004) *Sci. STKE 2004,* pe10
24. Rubbi, C. P., and Milner, J. (2003) *EMBO J.* **22,** 6068–6077
25. Wsierska-Gadek, J., and Horky, M. (2003) *Ann. N. Y. Acad. Sci.* **1010,** 266–272
26. Stefl, R., Skrisovska, L., and Allain, F. (2005) *EMBO J.* **6,** 33–38
27. Mendez-Vidal, C., Wilhelm, M. T., Hellborg, F., Qian, W., and Wiman, K. G. (2002) *Nucleic Acids Res.* **30,** 1991–1996
28. Matt, T., Martinez-Yamout, M. A., Dyson, H. J., and Wright, P. E. (2004) *Biochem. J.* **381,** 685–691
29. Bieganowski, P., Garrison, P. N., Hodawadekar, S. C., Faye, G., Barnes, L. D., and Brenner, C. (2001) *J. Biol. Chem.* **277,** 10852–10860
30. Lima, C. D., Klein, M. G., and Hendrickson, W. A. (1997) *Science* **278,** 286–290
31. Krakowiak, A., Pace, H. C., Blackburn, G. M., Adams, M., Mekhalfia, A., Kaczmarek, R., Baraniak, J., Stec, W. J., and Brenner, C. (2004) *J. Biol. Sci.* **279,** 18711–18716
32. Yung, T. M., Sato, S., and Satoh, M. S. (2004) *J. Biol. Sci.* **279,** 39686–39696

Aprataxin Forms a Discrete Branch in the HIT (Histidine Triad) Superfamily of Proteins with Both DNA/RNA Binding and Nucleotide Hydrolase Activities

Amanda W. Kijas, Janelle L. Harris, Jonathan M. Harris and Martin F. Lavin

J. Biol. Chem. 2006, 281:13939-13948.
doi: 10.1074/jbc.M507946200 originally published online March 16, 2006
