Analysis of *Xenopus* dsRNA adenosine deaminase cDNAs reveals similarities to DNA methyltransferases

RONALD F. HOUGH and BRENDA L. BASS
Department of Biochemistry and Howard Hughes Medical Institute, University of Utah, Salt Lake City, Utah 84112, USA

ABSTRACT

We isolated two similar, but distinct, cDNA classes that encode *Xenopus* double-stranded RNA (dsRNA) adenosine deaminase. The longest, full-length open reading frame (ORF) predicts a 1,270-amino acid protein of 138,754 Da that is similar in size and about 50% identical to proteins encoded by mammalian cDNAs, yet larger than the 120-kDa protein purified from *Xenopus* eggs. Alignments of the *Xenopus* and mammalian ORFs show N-terminal heterogeneity, three conserved dsRNA binding motifs (dsRBMs), and strongly conserved carboxyl termini. Consistent with the observation of two cDNA classes, northern analyses of *Xenopus* oocyte poly A⁺ RNA show at least three mRNA species. Multiple nuclear polyadenylation hexamers and putative cytoplasmic polyadenylation elements were found in the 3' UTRs of cDNAs corresponding to the largest mRNA. In vitro translation experiments show that the cDNAs encode active deaminases and that the entire N-terminus and first dsRBM are dispensable for deaminase activity. Importantly, an analysis of the C-terminus of five known dsRNA adenosine deaminases, and two putative deaminases, reveals motifs that are strikingly similar to the conserved motifs of the DNA-(adenine-N6)-aminomethyltransferases and the DNA-(cytosine-5)-methyltransferases.

Keywords: ADAR; base-flipping; double-stranded RNA; dsRAD; ESL; inosine; multiple polyadenylation sites

INTRODUCTION

We previously purified a 129-kDa protein from *Xenopus laevis* eggs and named it “dsRAD,” for double-stranded RNA Adenosine Deaminase (Hough & Bass, 1994). The dsRAD enzyme, also called DRADA (Kim et al., 1994a), catalyzes deamination of adenosines to inosines within dsRNA and was first characterized in *Xenopus* embryos (Bass & Weintraub, 1987; Rebagliati & Melton, 1987; reviewed in Bass, 1993, 1997). Although dsRAD will deaminate adenosines within dsRNA of any sequence, it exhibits a 5' nearest neighbor preference, and modifies adenosines poorly near 3' termini (Polson & Bass, 1994).

Recent studies suggest one biologic function of dsRAD is to act as an RNA editing enzyme. Deamination of adenosines within certain mRNAs is proposed to alter codons so that multiple proteins can be synthesized from a single encoded sequence. Nucleotide changes consistent with adenosine deamination have been identified in the pre-mRNAs of certain gluteamate receptors (Sommer et al., 1991; Lomeli et al., 1994) and in the antigenome of hepatitis delta virus (Casey & Gerin, 1995). Some of these editing sites can be edited efficiently in vitro using purified dsRAD (Melcher et al., 1996; Polson et al., 1996). Although this suggests that dsRAD is the enzyme responsible for editing these sites in vivo, recent work shows dsRAD is a member of a family of enzymes with overlapping but distinct specificities (Melcher et al., 1996). Thus, definitive identification of the particular deaminase responsible for a given editing event awaits further experimentation.

cDNAs for human (Kim et al., 1994a; O'Connell et al., 1995; Patterson & Samuel, 1995) and rat (O'Connell et al., 1995) dsRAD have been isolated and found to code for proteins of ~136 kDa and ~130 kDa, respectively. In contrast, fractionation of extracts prepared from various mammalian tissues typically leads to the purification of smaller proteins, closer in size to the 120-kDa protein we purified from *Xenopus* (Hough & Bass, 1994). For example, smaller, active species have been purified from bovine liver nuclei (93–83 kDa; Kim et al., 1994a), bovine thymus (116 kDa; O’Connell & Keller, 1994), and bovine brain nuclei (120–70 kDa; Chen et al., 1995). Discrepancies between the pre-
dicted polypeptide sizes and purified dsRAD polypeptides has been attributed to nonspecific proteolysis (Kim et al., 1994b; O’Connell et al., 1995). Indeed, two polypeptides obtained from bovine liver nuclei yielded N-terminal sequences found subsequently within the ORF of the human cDNA (Kim et al., 1994b).

Northern analyses show that human dsRAD mRNAs migrate as a single ~7.0-kb band in all tissues examined, with higher levels of RNA in brain and lung, and detectable levels in heart, placenta, liver, skeletal muscle, kidney, and pancreas (Kim et al., 1994b; O’Connell et al., 1995; Patterson & Samuel, 1995). These mRNAs are transcribed from a single-copy gene on the long arm of Chromosome 1 (Wang et al., 1995; Weier et al., 1995). If all mammals, like humans, express only one dsRAD mRNA, and the smaller forms of dsRAD are not artifacts of purification, then they must be generated posttranscriptionally. The latter possibility has not been ruled out and, in particular, the 120-kDa Xenopus protein (Hough & Bass, 1994), the 116-kDa bovine thyrmus protein (O’Connell et al., 1995), and the 120-kDa bovine brain protein (Chen et al., 1995), may be true intracellular forms (see Discussion).

Here we describe the cloning of two distinct cDNAs that encode X. laevis dsRAD. Like the mammalian cDNAs, the longest Xenopus cDNA predicts a full-length protein (~139 kDa) that is longer than the protein purified previously from cells. Also similar to mammalian cDNAs, both cDNAs show ORFs containing three deORMs and a strongly conserved carboxyl terminus. In contrast to the single mRNA detected during northern analyses of mammalian RNA, northern analyses of Xenopus RNA show multiple dsRAD transcripts. Importantly, we also show that the conserved C-termini of five known and two putative dsRNA adenosine deaminases contain sequences that are remarkably similar to the signature motifs of the DNA-(adenine-N6a)-aminomethyltransferases (Malone et al., 1995) and the DNA-(cytosine-5)-methyltransferases (Kumar et al., 1994).

RESULTS

Cloning of Xenopus dsRAD cDNAs

Our initial attempts to microsequence intact Xenopus dsRAD implied the N-terminus was blocked (Hough & Bass, 1994). Thus, we re-purified the 120-kDa dsRAD protein from Xenopus S100 egg extracts and obtained sequence for internal peptides. Degenerate deoxyligonucleotides corresponding to these peptides were used in PCR to amplify dsRAD sequence from oligo-dT primed oocyte cDNA.

Amplification between degenerate deoxyligonucleotide primers P85 and P95 (Table 1) produced a 251-bp PCR product with a continuous ORF that encoded a nearly complete deORBM (St Johnston et al., 1992; Bass et al., 1994; Kharrat et al., 1995), as well as the amino acid sequence MVNQGPGHDPK, identical to peptide-3 obtained by microsequencing (Fig. 1B). [32P]-dATP-labeled DNA probes were prepared by random priming of the 251-bp PCR product, then used to screen Xenopus head and ovary cDNA libraries. Partial cDNAs corresponding to ~2.6 kb of the 3’ ends of two distinct cDNAs were selected, and the coding regions were subcloned into plasmids for sequence analysis.

### TABLE 1. Deoxyligonucleotides.

<table>
<thead>
<tr>
<th>Name</th>
<th>Orientation</th>
<th>5’ Sequence-3’</th>
<th>Peptide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Degenerate Xenopus sequences</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P85</td>
<td>Sense</td>
<td>TTYGAYCCNWSHTAYAAAYGC</td>
<td>FDPSYNAP</td>
</tr>
<tr>
<td>P95</td>
<td>Antisense</td>
<td>SWNAACNCCGCAACCA</td>
<td>WFPAWS</td>
</tr>
<tr>
<td>P95&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Antisense</td>
<td>AACCACNCKNCNCCNAC</td>
<td>VGGRWF</td>
</tr>
<tr>
<td><strong>Degenerate mammalian sequences</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DF5</td>
<td>Sense</td>
<td>GGNRARGGNAARGCNACANAC</td>
<td>GECKATT</td>
</tr>
<tr>
<td>DE7</td>
<td>Sense</td>
<td>GTNGARAAYGNGCARLGC</td>
<td>VENGQEP</td>
</tr>
<tr>
<td><strong>Specific Xenopus sequences</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R5</td>
<td>Antisense</td>
<td>TCAITTAACCTCTGGAACTTGC</td>
<td>QVFEVNE</td>
</tr>
<tr>
<td>R9</td>
<td>Antisense</td>
<td>GTCCTGGTTCGCAACATCTTACAGT</td>
<td>TVKIGNQ</td>
</tr>
<tr>
<td>R2-Adapt&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Antisense</td>
<td>GACATGGTTCATGCGCCAGCG</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>GCCGCCGCCGCAACCTGGCTTG</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AAAACG</td>
<td></td>
</tr>
</tbody>
</table>

**Northern analysis probes**

<table>
<thead>
<tr>
<th>Name</th>
<th>Orientation</th>
<th>5’ Sequence-3’</th>
<th>Peptide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Antisense</td>
<td>GTTITTGCAAGAAGAAGTT</td>
<td>NFFLGNT</td>
</tr>
<tr>
<td>b</td>
<td>Antisense</td>
<td>AGTGTATGCACATTT</td>
<td>EIIGTF</td>
</tr>
<tr>
<td>c</td>
<td>Antisense</td>
<td>TCTGGGTTCTCCCTTG</td>
<td>QGDTQT</td>
</tr>
<tr>
<td>d</td>
<td>Antisense</td>
<td>GCCAAGTCTTGGCTAACA</td>
<td>VFOAKVG</td>
</tr>
<tr>
<td>e</td>
<td>Antisense</td>
<td>CACGTGTGGTTTCCTAGTGTA</td>
<td>3’ UTR</td>
</tr>
<tr>
<td>f</td>
<td>Antisense</td>
<td>TGGTTTGGATCTGCGCCTC</td>
<td>3’ UTR</td>
</tr>
</tbody>
</table>

<sup>a</sup>International Union of Biochemistry, Nomenclature Committee (1985).
cDNA classes (dsRAD 1 and 2, Fig. 1A) were isolated during the initial screenings. At this point, the mammalian dsRAD cDNA sequences were published (Kim et al., 1994b; O'Connell et al., 1995). This provided us with upstream sequence information that we used to prepare degenerate oligonucleotides (Table 1) corresponding to peptides conserved between the human and rat sequences (O'Connell et al., 1995). The oligonucleotides corresponding to the mammalian sequences, coupled with specific downstream Xenopus primers (Table 1), allowed the PCR amplification of an additional 1,328 bp of Xenopus cDNA. This was used to synthesize [32P]-random-primed deoxyoligonucleotides, which were used to rescreen the phage libraries.

After screening ~1.25 × 10^6 plaques from each library, a single full-length clone and nine other partial cDNAs encoding dsRAD-1 were isolated; 27 overlapping partial clones defined a contiguous dsRAD-2 cDNA consensus (Fig. 1). Excluding the highly variable 5' ends, which conceivably could reflect alterna-
tive splicing, the nucleotide sequences of dsRAD-1 (735-4195) and dsRAD-2 (262-375) were 90% identical. The nucleotides that differed between the two cDNAs were distributed throughout the sequence. The latter observation, coupled with the frequency of nucleotide substitutions (~10%), suggested dsRAD-1 and dsRAD-2 cDNAs were not alleles, but derived from two distinct genes that arose by duplication (Kobel & Du Pasquier, 1986). In contrast to the situation in mammals, duplicate active genes are relatively common in the pseudotetraploid X. laevis, whereas processed pseudogenes appear to be rare (Li & Graur, 1991; Larhammar & Riesinger, 1994).

Analysis of the two distinct classes of dsRAD cDNAs

The complete nucleotide sequence of dsRAD-1 showed an ORF of 3,810 nt encoding 1,270 amino acids (Fig. 1). Near the 5' end of the full-length dsRAD-1 cDNA, and within seven other partial dsRAD-1 cDNAs, a 435-nt repetitive region was found to encode 14 directly repeats of 11 amino acids rich in prolyl, seryl, arginyl, and glycyrl residues described by the consensus: P(R/H)(L/F)/(G/S) N(Q/R)/G(P/S)(P/S)AS. In addition, the full-length clone and three partial clones contained 33 nt of the 5' UTR with an in-frame UGA termination codon upstream of the presumed translation initiation site, GUAAUGA. This site conforms to commonly found eukaryotic translation initiation sites that contain purines in the +3 and +4 positions (Kozak, 1991).

The overlapping dsRAD-2 cDNAs defined an ORF of 1,124 amino acids that were 85% identical and 91% similar to the dsRAD-1 amino acid sequence (Fig. 1). This ORF lacked the repetitive region found in the N-terminus of dsRAD-1 and instead contained a single copy of the repeat consensus, as confirmed by eight cDNAs that overlapped this region. An initiation codon and unambiguous 5' UTR for dsRAD-2 were not identified, so we do not know if this ORF is complete. However, we note that, after accounting for differences due to the repetitive region, the dsRAD-1 and dsRAD-2 ORFs differed in length by only one amino acid. In addition, the length of the dsRAD-2 cDNA consensus was about the same length as the ~3.8 kb mRNA identified by northern analyses (see below).

The amino acid sequences predicted from the Xenopus dsRAD cDNAs are ~50% identical to the mammalian polypeptides. The three dsRBM and the C-termini are strongly conserved, whereas the spacer regions and the N-termini are much less conserved (Fig. 1B). A putative bipartite nuclear localization signal (NLS), as well as an additional conserved patch of basic residues that could serve as an NLS (reviewed in Görlich & Mattaj, 1996), are indicated by asterisks in Figure 1B. The regions surrounding the putative bipartite NLS (dsRAD-1 residues 272-318) and adjacent to the putative down-stream NLS (dsRAD-1 residues 388-424) are both ~50% similar to the N-terminus of Vaccinia E3L (residues 20-66) as described previously for the human cDNA (Patterson & Samuel, 1995).

The human dsRAD cDNA predicts a 48-amino acid direct repeat immediately upstream of the conserved aspartyl residue (dsRAD-1 D363, Fig. 1B) that is not found in the rat (O'Connell et al., 1995) or Xenopus cDNA sequences. In addition, the mammalian cDNAs lack sequences corresponding to the 145-amino acid repetitive region found in Xenopus dsRAD-1, but do have nucleotide sequences that code for 22-25 amino acids, rich in Arg and Gly (O'Connell et al., 1995), that may be distantly related to the Xenopus repeats. The significance of these differences in the dsRAD N-termini is presently unknown.

The 1,117-nt 3' UTR of dsRAD-1 contained an AAUAAA polyadenylation signal 14-nt upstream of the polyA tail. In contrast, the similar, but distinct, 378-nt 3' UTR of dsRAD-2 contained a less-favored polyadenylation hexamer, AAUAAA (Wickens, 1990a), 12-nt upstream of the polyA tail (Fig. 2). Three additional nuclear polyadenylation hexamers were found in the 3' UTR of dsRAD-1, including the hexamer found in the 3' UTR of dsRAD-2. Both 3' UTRs also contained U-rich elements similar to the cytoplasmic polyadenylation elements (CPEs) (UUUUUAU; Wickens, 1990b; Bilger et al., 1994) that are present in some Xenopus mRNAs that undergo cytoplasmic polyadenylation during oocyte development (Fig. 2).

Multiple Xenopus dsRAD mRNAs are expressed in oocytes

The observation of two distinct classes of Xenopus dsRAD cDNAs, as well as the existence of multiple polyadenylation sites within the 3' UTR of dsRAD-1, suggested that Xenopus might express multiple dsRAD mRNAs. To explore this possibility, we performed northern analyses of Xenopus oocyte polyA+ RNA. Indeed, prominent bands at ~5.3 kb and ~3.8 kb were observed, as well as a minor band at ~4.3 kb. All three bands were identified with several different probes, including antisense RNA, and random-primed probes (data not shown), as well as oligonucleotide probes specific for widely separated regions of identity between dsRAD-1 and dsRAD-2 (Fig. 3, lanes a and d). We also probed similar northern blots with oligonucleotide probes that were designed to be specific for dsRAD-1 or dsRAD-2 sequences. Although most of these short oligonucleotide probes gave a high background hybridization, they clearly showed that dsRAD-1 sequences could be found among the 5.3-kb and 4.3-kb transcripts, but not the smallest 3.8-kb transcript. In contrast, dsRAD-2 specific probes hybridized to the 5.3-kb and 3.8-kb bands, but not to the 4.3-kb band.
As mentioned before, duplicate genes are fairly common in *Xenopus* and the sequence divergence between dsRAD-1 and dsRAD-2 cDNAs suggests that they derive from different genes. Because duplicate genes in *Xenopus* sometimes show different expression patterns (Kobel & Du Pasquier, 1986), the simplest explanation for our northern results is that the nascent transcripts from the putative dsRAD-1 and dsRAD-2 genes are each processed to give a long and a short form of the dsRAD mRNA. Certainly further studies will be required to prove this scenario, and to correlate definitively the isolated cDNAs with the various transcripts observed by northern analyses. However, we note that the length of the dsRAD-1 cDNA (4,958 nt), which contains 33 nt of the 5' UTR, most likely derives from the longest ~5.3-kb mRNA, whereas the length of the dsRAD-2 consensus (3,750 nt, excluding the polyA tail) correlates with a nearly full-length copy of the ~3.8-kb mRNA.

At present, we do not know the identity of the lower molecular weight bands that show strong hybridization with probes c and e, but note that they are specific to these probes (Fig. 3). Database searches suggested that one of these oligonucleotide sequences has a spurious similarity to certain repetitive elements, and thus, these bands may represent hybridization that is unrelated to dsRAD. Regardless, these transcripts are too small to encode dsRAD and further experimentation will be required to determine their origin.

dsRAD-1 and dsRAD-2 cDNAs encode active deaminases
To verify that the dsRAD-1 and dsRAD-2 cDNAs encoded active dsRNA adenosine deaminases, we performed activity assays on proteins prepared by in vitro translation. RNAs transcribed from the full-length dsRAD-1 cDNA (~5.0 kb), and a partial dsRAD-2 cDNA (~2.2 kb) were translated in rabbit reticulocyte lysate or wheat germ extract. The dsRAD-2 cDNA encoded an N-terminal truncated protein that began slightly before the second dsRBM (C-terminal amino acids 511–1124, Fig. 1). In the absence of sequence information for the full-length dsRAD-2 protein, we chose to assay this truncated protein, reasoning that such an assay could potentially determine whether dsRAD-2 sequences encode an active deaminase, as well as provide information on whether N-terminal sequences were dispensable. The latter was of particular interest because the C-terminus is thought to encode the dsRAD catalytic domain (see Discussion).

Equal volumes of the translation mixtures were incubated with a 794-bp dsRNA internally labeled with [32P]-α-ATP. After the reaction, the nucleotides of the RNA were assayed for the presence of inosine by thin-layer chromatography. Figure 4 shows that proteins translated from full-length dsRAD-1 (F), as well as the carboxy-terminal half of dsRAD-2 (P), could convert adenosines to inosines in vitro. Figure 4 shows results for proteins synthesized in the wheat germ extract, but
the reticulocyte system gave similar results (data not shown). In both translation systems, the molar amount of protein synthesized from the longer dsRAD-1 mRNA was typically threefold lower than that made from the partial dsRAD-2 mRNA (see Fig. 4 legend). Because we always incubated equal amounts of the in vitro translation mixtures to normalize for any nonspecific effects, the dsRAD-2 reactions presumably contained more active enzyme. This seems to be the most likely reason for the difference between the amount of inosine produced by the full-length (F) and truncated protein (P) samples, but kinetic analyses with purified proteins will be required to determine if the two proteins differ in their intrinsic activities. We conclude

FIGURE 4. Deamination of dsRNA by in vitro-translated dsRAD protein. Full-length dsRAD-1 and partial dsRAD-2 transcripts were translated in wheat germ extract. Aliquots (15 μL) of the translation mixtures were incubated with 3 fmol of a 294-bp dsRNA (internally labeled with [32P]-α-ATP) for 3 h at 25°C. The F1 nuclease-digested reaction products were resolved by thin-layer chromatography on cellulose plates in saturated ammonium sulfate, 0.1 M sodium acetate, pH 6.0, isopropanol (79:19:1). Left to right: (−) no template control; (F) full-length dsRAD-1 translation mixture; (P) partial dsRAD-2 translation mixture; (+) no template control supplemented with 3 μL purified Xenopus dsRAD (Hurst et al., 1995) to rule out the presence of inhibitors; migration positions of radiolabeled (p) 5’ IMP and (pA) 5’ AMP markers relative to the origin (o) are designated. Full-length dsRAD-1 translation mixtures contained three bands with apparent molecular mass of −146 kDa, −133 kDa, and −117 kDa (data not shown), similar to the predicted native mass of −139 kDa. A main band of −79 kDa, and two smaller bands, −71 kDa and −64 kDa, were seen in the partial dsRAD-2 mixtures, close to the predicted native size of −68 kDa. Approximately 0.7 fmol of in vitro translated full-length dsRAD-1 deaminated 1% of the dsRNA adenosines (lanes F) and −20 fmol of translated N-terminal truncated dsRAD-2 showed 9% deamination (lane P).

Comparison of Xenopus dsRAD with other deaminases

Recently, a number of cDNAs encoding proteins showing sequence similarities to dsRAD have been identified. The schematic alignment shown in Figure 5A compares the full-length Xenopus dsRAD-1 with this potential family of deaminases. Family members show completely different N-termini and one, two, or three dsRBMs. Spacing between the dsRBMs, and between the dsRBMs and the conserved carboxyl termini varies, even among the dsRAD sequences. For example, the spacer between dsRBM-3 and the carboxyl domain of Xenopus dsRADs is shortened by 24 amino acids compared with the mammalian dsRADs. The C-termini of all family members are strongly con-
FIGURE 5. Comparison of Xenopus dsRAD with the putative deaminase family. A: Schematic representation of the predicted ORFs of several known dsRNA adenosine deaminases and two putative deaminases is shown. Identity and similarity scores (relative to dsRAD-1) within the dsRBM (filled boxes) and highly conserved C-termini (speckled boxes) were determined with GCG software (Bestfit). dsRAD-1 ORF was compared successively with human dsRAD (Kim et al., 1994b, U10439), rat dsRAD (O'Connor et al., 1995, U10544), red-1 (Melcher et al., 1996, U43354), mouse Tenr (Shumacher et al., 1995, X84693), and C. elegans T20H4.4 (U00357; R.F. Hough, A.T. Lingam, & B.L. Bass, in press). B: C-terminal consensus sequence from 5A were confirmed by realignment of only the C-terminal regions (A-E) of the seven sequences using the clustal V method (Higgins & Sharp, 1989) with the PAM 250 residue weight table (Lasergene/Megaalign from DNASTAR Inc.; default parameters). This alignment is shown with the four dsRADs represented by their consensus sequence (dsRAD Cons). Identical residues among dsRAD sequences are in upper-case letters in the dsRAD consensus and are highlighted in black if conserved in all aligned sequences. Strong similarities (three of four positions) within the dsRAD consensus and the family consensus (Consensus) are indicated by lower-case letters. For discussion purposes, conserved amino acids are grouped into five regions, indicated A-E.
erved (Fig. 5A), containing 58 identical amino acids (Fig. 5B). In addition, a portion of the highly conserved region C (1063-1086 of the dsRAD consensus, Fig. 5B) showed ~50% similarity to Vaccinia E3L (residues 58-84) in a BLASTP database search (Altschul et al., 1990) performed with the dsRAD-1 sequence. Note that Saccharomyces cerevisiae HRA400 fits the alignment reasonably well, but was omitted from the alignment because it lacks a dsRBM (see Patterson & Samuel, 1995).

Conserved C-terminus has similarities to DNA-methyltransferases

Amino acid sequence alignments of the DNA-(cytosine-C5)-methyltransferases (5mC Mtase; Kumar et al., 1994) and the DNA-(adenine-N6/cytosine-N4)-aminomethyltransferases (N6mA and N4mC Mtaases; Malone et al., 1995; Timinskas et al., 1995) have identified up to 10 conserved motifs within each of the four different type II Mtaase groups (5mC Mtaases and the α, β, γ amino Mtaases). Motifs I-III and X contain residues that bind the S-adenosylmethionine (ado-Met) cofactor, and motifs IV-VIII contain residues of the catalytic sites that capture the base to be modified and activate it for methylation. Each Mtaase also contains a variable DNA target recognition domain (TRD). Figure 6A shows that the ado-Met and catalytic domains and TRDs appear in different linear orders within the Mtaase primary sequences. Because the Mtaases are an extremely diverse family of proteins (Wilson, 1992), it was only by using the structure-based local alignments of Malone et al. (1995) and Kumar et al. (1994) that we were able to recognize motifs within the C-terminal consensus of the deaminase family (Fig. 5B) that are remarkably similar to those of the 6Nαα and 5mC Mtaases. We show an alignment that compares the putative deaminase ado-Met domains (XN6, IN6-IIIα) and catalytic domains (IVVIII) with members of the N6αα Mtaases, and the putative TRD (IVCS, VIIC5, and VIIIIC5) with 5mC Mtaases (Fig. 6B).

Among the many conserved residues within the catalytic domain of the Mtaases, the function of three residues, two in motif IV, and one in motif VIII (see asterisks, Fig. 6B) have been well characterized. As shown in Figure 6B, the conserved aromatic residue in motif VIIIIV appears in every deaminase considered. The motif IV consensus (F/Y)x(D/N/S)P(Y/F), derived from all DNA amino Mtaases (α, β, γ; Malone et al., 1995), compares favorably with the deaminase consensus Fx(E/Q/D)P(V/I)Y, provided that substitution of E/Q/D for D/N/S is reasonable. It is noteworthy that Caenorhabditis elegans T2014.4 sequence contains an ASP at this position (Fig. 6B). In addition, among the amino Mtaases, one PRo is always present in this motif, but substitutions can occur at either position (Malone et al., 1995).

The ado-Met binding motif fN6 appeared to be the least conserved, which may not be surprising because the deaminases are not thought to bind ado-Met. However, the deaminase consensus in this motif, DfIAeisiRrgf (954-DChAeVSRrgf-965 in dsRAD-1, Fig. 1), is remarkably similar to the sequence 138-DVRKRRG-145 in Vaccinia VP39 cap-dependent 2'-O-methyltransferase that is critical for ado-Met binding (Shi et al., 1996). The carboxyloxdide chain of D39 interacts with the α-amino group of the ado-Met methionine moiety in the co-crystal structure (Hodel et al., 1996). Perhaps CHAE was inserted into the dsRAD sequence, or deleted from the viral sequence. In the Mtaases, residues of motif I are involved in making a tight turn to form a classical Rossmann fold, which binds the ado-Met cofactor. If the deaminases adopt a structure similar to that of the Mtaases, motif I would be in close contact with the active site. Thus, the well-conserved deaminase consensus residues may be active site residues that are uniquely suited for deamination.

We noticed that the Xenopus dsRAD sequence elements PC...ENV...GERLR (dsRAD-1 amino acids 1010-1077, Fig. 1) were remarkably similar to the 5mC Mtaase catalytic motif sequences, PC...ENV...QxRxxR (Cheng & Blumenthal, 1996). Among these strongly conserved residues are the Glu and Arg that bind cysteine in the catalytic pocket and the Cys that provides the catalytic thiol (Wu & Santi, 1987; Klimasauskas et al., 1994; Kumar et al., 1994). However, because the deaminase sequences diverge significantly in this region (IVCS-VIIIIC), we feel this domain of the deaminases may be more analogous to a TRD than a 5mC Mtaase catalytic domain (see Discussion).

DISCUSSION

Here we describe two classes of cDNAs (dsRAD-1 and dsRAD-2) that code for X. laevis dsRNA adenosine deaminase. The dsRAD-1 cDNA contains a complete ORF that encodes a 1,270-amino acid protein of 138,754 Da. This predicted molecular weight is very close to that of proteins encoded within mammalian cDNAs, but larger than the 120-kDa protein we purified previously from Xenopus eggs. A consensus derived from many partial dsRAD-2 cDNAs shows an 1,124-amino acid ORF that is 85% identical to that of dsRAD-1, but lacks the 145-amino acid N-terminal repetitive region (Fig. 1). The dsRAD-2 consensus does not contain an obvious 5' UTR, and we do not know how much of this ORF remains to be defined. The ORFs encoded by dsRAD-1 and dsRAD-2 cDNAs show about 50% amino acid identity with those encoded by mammalian cDNAs and, like their mammalian counterparts, contain three dsRBM and a strongly conserved C-terminus.

Northern analyses of Xenopus oocyte polyA + RNA show three dsRAD transcripts of ~5.3 kb, ~4.3 kb, and ~3.8 kb (Fig. 3). Preliminary experiments using deoxynucleotidylate probes specific for dsRAD-1 or dsRAD-2 indicate that dsRAD-1 sequences are present...
in mRNAs migrating with the major ~5.3-kb and minor ~4.3-kb bands, whereas dsRAD-2 sequences can be found in mRNAs migrating with the ~5.3-kb and ~3.8-kb bands. Although we have not yet performed genomic analyses, the 10% nucleotide sequence difference between dsRAD-1 and dsRAD-2 suggests the cDNAs derive from different genes, and our northern analyses suggest the nascent transcripts of these putative genes are alternatively processed. We have not yet investigated the mechanisms that lead to the multiple transcripts. However, the presence of multiple polyadenylation signals in the 3’ UTR of the dsRAD-1 cDNA suggests that alternate polyadenylation sites could be utilized to produce messages of different sizes. Interestingly, the putative polyadenylation sites (Fig. 2) contain upstream polyU elements similar to those found in messages selected for cytoplasmic polyadenylation during oocyte maturation (Wickens, 1990b; Bilger et al., 1994), suggesting that dsRAD mRNAs could be further modified by this process.

Similar to the studies of mammalian dsRAD cDNAs (see Introduction), our studies of Xenopus cDNAs raise questions in regard to how smaller dsRAD proteins purified from cells relate to longer proteins encoded by isolated cDNAs. In contrast to human cells, which contain a single dsRAD mRNA, Xenopus cells have multiple mRNAs. Potentially, one of the smaller mRNAs could code for the 120-kDa protein we purified originally. In fact, if the ORF of the dsRAD-2 cDNA proves to be nearly complete, the native molecular mass of the encoded protein (124 kDa) would be very close to that of the protein we purified. However, the internal peptides we sequenced from the purified 120-kDa protein indicate at least some of this protein derived from dsRAD-1 specific sequence. Of the six peptides sequenced from the 120-kDa Xenopus protein, four occur in regions of complete identity between the dsRAD-1 and dsRAD-2 proteins (Fig. 1B, peptides 3–6), and thus could derive from either sequence. The sequence of peptide-1 could have potentially distinguished between the proteins. Unfortunately, this peptide was a minor species that co-eluted with peptide-4, and, although it clearly derived from the region indicated in Figure 1B, it gave a sequence that

---

**Figure 6.** Similarities between deaminase family C-termini and DNA Mases. A: Linear order of motifs in dsRAD and the Mases: open boxes, ado-Met motifs [I–III, X]; filled boxes, catalytic motifs [IV–VIII]; speckled box, motif IX specifically found in 5MC Mases; TRD, target recognition domain. Motif size and distances between motifs are not drawn to scale. B: Alignment of C-terminal deaminase sequences with the ado-Met domains (XN6A, I6A–IIIb), and catalytic domains (IVV–VIIIb) of N6mA Mases, and with motifs IVf, Vf, VIIc of the 5MC Mases. Abbreviations and accession numbers for the Mases are given in Malone et al. (1995) and Kumar et al. (1994) and conserved amino acids were grouped as in these previous studies: (E, D, Q, N), (V, L, I, M), (F, Y, W), (G, P, A), (K, R), and (S, T). Black and gray shading shows amino acids represented by the majority (10 or more); hydrophobic residues (L, I, V, M) were distinguished in gray to aid visibility. Residues found in all N6mA Mases are underlined in the consensus. Asterisks denote the conserved D and Y (IVf) and Y (VIIc) inferred to interact with adenine in the catalytic site by analogy to N/Y and F residues in the crystal structure of M.TaqI (Malone et al., 1995). A–F refer to conserved regions shown in Figure 5B. VP39 designates region of similarity to the ado-Met binding region of Vaccinia 2’-O-methyltransferase, and E3L refers to a region of similarity to Vaccinia E3L protein residues 58–94 (see Results). A global alignment of the seven deaminase C-termini and N6mA Mases was used to establish the initial orientation, using the Clustal V method with the PAM 250 matrix (Lasergene/Megalign, see Fig. 5 legend). The Mase sequences in this alignment were then adjusted manually and with the Clustal V method, to match previous alignments using the conserved motifs as anchors for local alignments (Malone et al., 1995). Local alignments of the deaminases to the aligned Mases were performed similarly using conserved deaminase residues as anchor points, e.g., dsRAD-1 TG (XN6A), PCG (IVf); ENV (Vf): RLR (VIIc), PVEPVY (IVf); SRV (Vf); and SY (VIIc). The most significant sequence difference was the distance between motifs IVf and Vf in the deaminases (52–58 amino acids) compared to the N6mA Mases (0–8 amino acids). The distance found in the deaminases is more consistent with the N6Aβ (6–25 amino acids) and N6Aγ (19–37 amino acids) Mases (Malone et al., 1995). (Figure continues on facing page.)
FIGURE 6. Continued.
was too ambiguous to distinguish between dsRAD-1 and dsRAD-2. However, the unambiguous peptide-2 sequence clearly derived from dsRAD-1. Thus, we are certain our purified preparation contained a protein derived from dsRAD-1 sequences, but have not ruled out the possibility that the 120-kDa protein band was a mixture of dsRAD-1 and dsRAD-2 polypeptides.

Conceivably, the dsRAD-1-specific ~4.3-kb mRNA could represent an alternatively processed form of dsRAD-1 whose ORF encoded a protein closer in size to the protein we purified. However, because all dsRAD-1 specific probes we have used so far hybridize to the major 5.3-kb as well as the minor 4.3-kb transcripts, we cannot yet rule out the possibility that these RNAs code for the same size protein. Thus, similar to researchers studying the mammalian enzymes, we must consider the possibility that the smaller protein somehow derives from an mRNA encoding a longer protein.

As mentioned in the Introduction, smaller dsRAD proteins could be due to artificial proteolysis during purification. Indeed, the initial identification of a ~40-kDa Z-DNA binding polypeptide from chicken lungs (Herbert et al., 1993), which turned out to be a portion of the full-length chicken dsRAD purified subsequently (Herbert et al., 1995), is consistent with proteolysis. However, more recent observations support the idea that the large and small forms of dsRAD are true intracellular forms and derive from a single message. For example, western analyses of human cell nuclear extracts clearly show a large p150 form of the protein, which is observed most readily after interferon induction, as well as a smaller protein, p110, which is proposed to be the more abundant, constitutively expressed form of dsRAD (Patterson et al., 1995; Patterson & Samuel, 1995). Further, using antibodies to different regions of the full-length protein, these authors convincingly show p110 is an N-terminal truncated form of p150. Possibly, a large protein precursor is processed proteolytically to yield the final, constitutively expressed form, or both proteins are synthesized from the same messenger RNA. With regard to the latter, we note that alternative initiation at an in-frame methionine (EDMA in Fig. 1) that is conserved in all mammalian and Xenopus dsRADs would yield proteins of a size consistent with the smaller proteins from human (Patterson & Samuel, 1995), bovine (O’Connell & Keller, 1994; Chen et al., 1995), and Xenopus (Hough & Bass, 1994) sources. In fact, the recombinant human dsRAD expressed from this methionine (Met<sup>299</sup>) migrates almost identically to the endogenous human p110 protein during SDS-PAGE (Patterson et al., 1995; Patterson & Samuel, 1995). Certainly future experiments will help distinguish among these interesting possibilities.

Our activity assays showed that an in vitro-translated dsRAD-2 ORF that lacked the N-terminal half of the protein had deaminase activity. This observation indicates catalytic residues reside in the C-terminus of dsRAD (Fig. 4), consistent with studies of the mammalian enzymes. For example, analyses of the human enzyme show that even short deletions from the C-terminal end, and point mutations of four positions, result in loss of deaminase activity (Lai et al., 1995). These four positions are the histidyl and glutamyl residues in the element, DCHAE, and two downstream cysteinyl residues in the elements, PCG and MSCSDK found in regions A, B, and C, respectively (Fig. 5B). Of course, the fact that RED-1, a smaller, but active deaminase (Melcher et al., 1996), does not have sequences corresponding to the N-terminal half of dsRAD (Fig. 5A), is also consistent with the active site residing in the C-terminal domain.

Our alignment of the C-termini of the five known and two putative dsRNA adenosine deaminases revealed 58 identical residues, including 15 aromatics or leucines, 12 basic residues, 9 glycines, 6 serines, and 5 acidic residues (Fig. 5B). Of these, a single histidyl and glutamyl, and four aspartyl residues, are potential ligands for coordinating zinc (Vallee & Auld, 1990). Our alignment predicts that both mouse Tenr and C. elegans T20f4.4 will prove to be active deaminases, although the C-terminus of testis-specific Tenr is the most divergent sequence. Thus, unambiguous identification of the function of Tenr, or clear demonstration that it has deaminase activity, would greatly advance our understanding of the C-terminal residues required for activity.

Both murine adenosine deaminase (ADA; Wilson et al., 1991) and Escherichia coli cytidine deaminase (CDA; Bettis et al., 1994) have been crystallized with transition-state analogues and shown to contain zinc at their catalytic centers. Their tertiary structures are completely different, however, and they are thought to have converged during evolution (Bettis et al., 1994). Many of the amino acids shown by crystallographic studies to be in the active site, or coordinated with zinc atoms, were predicted previously to be important for catalysis based on their conservation among adenine and AMP deaminases (Chang et al., 1991), or cytidine, CMP, dCMP, and deoxyuridine deaminases (Bhattacharya et al., 1994; Reizer et al., 1994).

Possibly, dsRAD also requires zinc for catalysis. Conserved sequence elements similar to regions that coordinate zinc ligands in both the cytidine and adenosine deaminase families are found in the dsRAD C-termini. The DCHAE consensus sequence (region A, Fig. 5B) fits the consensus for cytidine deaminases (I'H/C'JA; Bhattacharya et al., 1994; Reizer et al., 1994) as well as adenosine deaminases (HAX; Chang et al., 1991). Although the downstream elements—PCxxC in the CDA family consensus and SL(S/N)TDDP in the ADA family consensus—are not found in the dsRNA deaminases, residues similar to both of these elements, such as PCG (region B), SCSDKIL (region C), and SVN
WCLAD (region D) are present in the C-terminal sequences (Fig. 5B). These similarities between the dsRNA adenosine deaminases and the nucleoside and nucleotide deaminases are intriguing, but certainly the dsRNA deaminases could employ a mechanism that does not require zinc. In fact, other catalytic schemes for hydrolytic deamination exist, as emphasized by studies of cytosine deaminase, an iron-dependent enzyme reversibly inhibited by zinc (Porter & Austin, 1993), which does not contain amino acid sequences found in the cytidine deaminase consensus.

A model for catalysis by the C-terminus of dsRAD

Catalysis by dsRAD involves a nucleophile attack at C6 of adenine. The C6 atom is buried deep within the major groove of the A-form helix, and how dsRAD gains access to its target is unclear. We previously proposed that dsRAD might use a mechanism similar to that of the type II DNA Mtases and related base-flipping enzymes (Polson & Bass, 1994). These enzymes sequeseter target nucleotides into their active sites by extracting them from the DNA helix, leaving the helical structure relatively unperturbed (reviewed in Roberts, 1995; Cheng & Blumenthal, 1996).

Our amino acid sequence alignments reveal that the C-terminal consensus sequence (Fig. 5B) can be described largely as a series of motifs that are in the same linear order as, and very similar in sequence to, the motifs found in group A N-6-adenine methyltransferases (Fig. 6). The putative catalytic motifs [IV$_{N_6}$-VII$_{N_6}$], which in the N6mA Mtases are involved in binding adenine, are especially well conserved in the C-termini of the five dsRNA deaminases, and two putative deaminases. Inserted between the putative cofactor site [X$_{N_6}$, I$_{N_6}$-III$_{N_6}$] and catalytic site [IV$_{N_6}$-VII$_{N_6}$] are sequences similar to the 5mC Mtase catalytic motifs [IV$_{C_5}$, VI$_{C_5}$, and VIII$_{C_5}$], which, in the Mtases, are involved in pyrimidine binding (Kumar et al., 1994). However, the similarity to the 5mC Mtases is more tenuous, and thus, these motifs may be part of a TRD rather than a functional catalytic site. Of course, it is possible that the deaminases have binding sites for multiple nucleic acid bases. This idea is consistent with the gene duplication model that proposes Mtases evolved from a primordial nucleotide binding protein (Lauster, 1989). Moreover, 5mC catalytic motifs and surrounding sequences are evident in the TRDs of some N6As and N6Ay Mtases (Lauster, 1989).

Conservation of the functional residues of the N6mA Mtase active site suggests that dsRAD functions in a manner analogous to these enzymes, and flips its target adenine out of the RNA helix into the active site (Roberts, 1995; Cheng & Blumenthal, 1996). In this model, upon binding dsRNA, dsRAD would stabilize the flipped out adenine using the conserved [YY] pair (asterisks, Fig. 6B) in a hydrophobic pocket formed by motifs IV$_{N_6}$-VII$_{N_6}$. Hydrolysis of the N6 amino group could be catalyzed by a zinc cofactor coordinated with any of the potential ligands conserved in the carboxyl terminus (Fig. 5B). In addition, the Asp, His, and Arg residues found in the putative ado-Met motif (I$_{N_6}$-subdomain A) could participate in catalysis because they would be close to the active site if dsRAD adopts a structure similar to the Mtase catalytic domain (Schluckebier et al., 1995). It is also possible that other amino acid residues of motif IV$_{N_6}$ (Fig. 6B), which in theaminomethyltransferases are thought to facilitate methylation by polarizing the N6 amino group through hydrogen bonding (Malone et al., 1995), could be used to polarize the N6-C6 bond for hydrolytic attack. In this case, alternative mechanisms that do not require zinc become more probable.

Finally, in support of our base-flipping model, we note that M.HpaII DNA-5mC-methyltransferase, in the absence of its ado-Met cofactor, catalyzes the deamination of cytosine with a rate enhancement on the order of $-10^4$ (Shen et al., 1992). This indicates that the basic architecture of the Mtases is compatible with catalysis of a deamination reaction. Although beyond the scope of this paper, our identification of Mtase motifs within dsRAD suggests an $\alpha/\beta$ topological model (Malone et al., 1995) for the structure of the dsRAD C-terminus that is directly amenable to experimental testing.

MATERIALS AND METHODS

Protein purification and peptide sequencing

Poly (G+C)-sepharose active fractions obtained from frog eggs as described previously (Hough & Bass, 1994) were pooled and concentrated by precipitation in 10% trichloroacetic acid, then resolved on a 1.5-mm, 10% SDS-polyacrylamide gel, followed by electroblotting to PVDF (Problot, ABI) as described by the manufacturer. The $-120$-kDa band stained with Ponceau S was excised and sent to the Harvard Microchemistry Facility for in situ digestion (Aebi et al., 1989) with endolysoprotease-C and sequencing. Solubilized peptides were resolved by reverse-phase HPLC and selected peaks analyzed in an externally calibrated Finnigan LaserMat mass spectrometer, followed by sequencing on an ABI model 477A, or HP G1000A, protein sequencer.

Cloning and sequencing of cDNAs

Degenerate deoxyoligonucleotides corresponding to Xenopus dsRAD peptide sequences were synthesized in both directions and used to amplify PCR products from oocyte cDNA prepared by reverse transcription of total RNA as described (Lee & Caskey, 1990). A product of $-250$ bp obtained by amplification with P85 and P99 primers (Table 1) was gel purified by electrophoresis in 8% TBE-polyacrylamide, then reamplified with P85 and a nested 3' degenerate deoxyoligonucleotide (p59). The reamplified PCR product was cloned
by ligation of the PCR generated A-overhangs (Taq DNA polymerase, Perkin-Elmer Cetus) with linearized pCRII (Invitrogen) containing T-overhangs. Nucleotide sequencing of the insert was by the dyeideoxy method using T7 and SP6 primers and the Sequenase 2.0 kit (USB).

The 251-bp insert was excised by EcoRI I digestion and used for random-priming (MBI kit) utilizing [32P]-α-dATP (DuPont NEN, >3,000 Ci/mmol). These random-primed probes (specific activity >10^6 cpm/μg) were used to screen two Xenopus cdNA libraries: a head library constructed in A-ZAP II, Net I-EcoRI I, from stage 28-30 tailbud tadpoles (Hemmati-Brivanlou et al., 1991) kindly provided by Richard Harland (UC, Berkeley); and a Uni-Zap XR, EcoRI I-Xho I, ovary library (Stratagene). Standard screening protocols were employed (Sambrook et al., 1989). Briefly, ~250,000 plaques from each library were bound to 0.45-μ nitrocellulose filters (Schleicher & Schuell). Filters were washed overnight at 65°C in 3X SSC/0.1% SDS, then incubated in 6X SSC, 1X Denhardt’s solution, 0.05% sodium pyrophosphate, 0.1% SDS, 50% formamide, 50 μg/mL denatured salmon sperm DNA for 2 h at 65°C, followed by overnight incubation at 42°C. Hybridization was for ~22 h at 42°C in fresh hybridization mixture containing 32P-labeled random-primed oligonucleotides (≥2 x 10^6 cpm/mL) and 50 μg/mL torula RNA substituted for salmon sperm DNA. The filters were washed 30 min at 42°C in 6X SSC/0.1% SDS, then washed three times for 15 min at room temperature in solutions containing 0.1% SDS and 3X, 1X, 0.5X, and 0.2X SSC, respectively. Blotted filters were exposed to Kodak XAR-5 film using intensifying screens for 17-72 h at -80°C. Positive plaques were purified through two additional plateings and recovered as pBluescript SK(−) phagemids with the ExAssist/SOLR system according to instructions from Stratagene.

To generate additional 5′ probes, degenerate deoxyoligonucleotides corresponding to conserved N-terminal peptides in the mammalian sequences (O’Connell et al., 1995) were used with specific downstream Xenopus primers for PCR amplification. First-strand cdNA, enriched in dsRAD sequences by reverse transcription of 2.9 μg polyA⁺ oocyte RNA with the “R2-Adapt” primer (Table 1), was used in a 50-μL PCR: “R2-adapt”-primed oocyte cdNA template (1 μL); 4 μL either DF5 or DF7 degenerate mammalian upstream primers in pairwise combinations with 1 μM of Xenopus specific primers, R5 or R9 (Table 1); 2.5 units AmpliTaq DNA polymerase (Perkin-Elmer Cetus). The reactions also contained 67 mM Tris-HCl, pH 8.8, 6.7 mM MgCl2, 16.6 mM (NH4)2SO4, 170 μg/mL BSA, 1.5 mM dNTPs, 10% DMSO (Frohman, 1984). After a 5 min 94°C denaturation step, the reactions were initiated by addition of AmpliTaq polymerase at 72°C followed by thermocycling (94°C/30 s; 58°C/30 s; 72°C/2 min) × 30 cycles; 72°C/10 min extension.

Several degenerate deoxyoligonucleotides either failed or gave spurious products, but DF3 and DF7 degenerate primers, in combination with Xenopus primers R5 and R9, gave PCR products of approximately the sizes expected. The gel-purified products were reapplied, isolated by TA cloning (Invitrogen), and confirmed by sequencing. Although there were several mismatches between the mammalian degenerate deoxyoligonucleotides and the Xenopus templates, there was sufficient identity at the 3′ ends of the deoxyoligonucleotides for successful priming at the annealing temperature used. The repetitive nature of the nucleotide sequences at the 5′ end of the dsRAD ORFs, coupled with our use of specific downstream Xenopus primers and template cdNA enriched for dsRAD sequences, probably contributed to the success of our approach. The 1,328-bp [DF3-R5] PCR fragment was used to prepare random-primed deoxyoligonucleotides as described above to rescreen the cdNA libraries.

DNA sequencing was done either by the dyeideoxy method (Sequenase 2.0, USB) or by the Health Sciences Center Sequencing Facility, using T3 and T7 vector primers, and dsRAD-1 or dsRAD-2 specific primers to walk on each strand. Full-length dsRAD-1 was sequenced completely on both strands. Additional confirmation was provided by a contiguous consensus sequence prepared with nine other overlapping cdNAs. The dsRAD-2 sequence is a consensus derived from 27 cdNAs. At least 2, and as many as 10 representative sequences along each strand contributed to the consensus.

Northern analyses

X. laevis females were primed and their ovaries excised and treated with collagenase B as described previously (Bass et al., 1994). Poly A⁺ RNA was selected from total RNA prepared by the proteinase K method (Sambrook et al., 1989) using an Oligotex-dT kit (Qiagen), or direct poly A⁺ selection from proteolyzed oocyte homogenates using Fast Track 2.0 kit (Invitrogen). Standard protocols (Ausubel et al., 1987) were followed for electrophoresis and northern blot transfer to nylon membranes (Biotrans, ICN or Gene Screen, Dupont NEN), except that the agarose gel and electrophoresis buffers contained 0.66 M formaldehyde as described (Tsang et al., 1993). RNA markers 0.3-6.6 kb (Promega) and 0.2-9.4 kb (Life Technologies) were loaded in separate lanes and the blots were stained with 0.04% methylene blue in 0.5 M sodium acetate, pH 5.2, prior to hybridization to visualize the marker positions and to insure that the RNA was intact.

Deoxyoligonucleotide probes (Table 1) were labeled at their 5′ termini using T4 polynucleotide kinase and [32P]-γ-ATP, and passed through Chroma spin-10 columns (Clontech) to remove unincorporated radioactivity, followed by ethanol precipitation. Analyses were also done using 32P-labeled random-primed probes, or antisense RNA, prepared with the 251-bp DNA fragment. Prehybridization and hybridization were done as described above for cdNA clones except that formamide was omitted from the solutions containing the deoxyoligonucleotides. The short deoxyoligonucleotide probes were selected from nonredundant regions showing the greatest difference between the dsRAD-1 and dsRAD-2 cdNA sequences, and were designed to have similar melting temperatures (OLIGO primer analysis software). Membrane slices hybridized with different probes were subjected to the same washing protocol as described above, except that the final 0.2X SSC/0.1% SDS was done at 42°C for 1 h. Images were captured by phosphorimagery (Molecular Dynamics). Although all blots gave similar results, persistent differences in background were observed, for unknown reasons.

In vitro translation and dsRAD assays

Capped mRNAs were prepared by transcription of Xho I-linearized pBluescript SK(−) (Stratagene) containing full-length dsRAD-1 (~5.0 kb) and partial dsRAD-2 cdNA (~2.2
kb) inserts using T3 polymerase with a 2:1 ratio of GpppG to GTP (Bass et al., 1994). The transcripts contained the complete ORF and 3′ UTR of dsRAD-1, or nt 1525–3572 containing the C-terminal 614 amino acids (Met113 to Leu188) and 3′ UTR of dsRAD-2 (Fig. 1). In vitro translations were done in nuclelease-treated rabbit reticulocyte lysates or wheat germ extracts (Promega) as described by the manufacturer in the presence of [35S]-methionine. Incubations were for 2 h with 2 pmol mRNA templates. Total protein synthesis was monitored by TCA precipitation and protein band densities were resolved by SDS-PAGE followed by autoradiography. Activity assays were done as described previously for purified dsRAD (Hough & Bass, 1994). Aliquots of the reticulocyte or wheat germ reactions were incubated with various amounts of a 794-bp chloramonophenol acetyltransferase (CAT) duplex RNA (1–10 fmol) at 25 °C for 1–3 h.

ACKNOWLEDGMENTS

We thank R. Harland (University of California, Berkeley) for providing a superb X. laevis head library, R. Robinson and W. Lane (Harvard Microchemistry Facility) for performing protein digestion and sequencing of dsRAD peptides, and M. Robertson and E. Lawrence (Health Sciences Center Sequencing Facility) for their patience and extensive sequencing of numerous cDNA clones. We thank T. Formosa, D. Morse, and H. Ley III for reading the manuscript. We also appreciate helpful suggestions and discussions from the members of our research group. This work was supported by funds to B.L.B. from the National Institutes of Health (GM46703) and the David and Lucile Packard Foundation. Some of the deoxoyglycosidases were synthesized by the Utah Cancer Center Protein/DNA core facility, supported by the National Cancer Institute (#5 P30 CA42014). B.L.B. is a Howard Hughes Medical Institute assistant investigator.

Received November 8, 1996; revised for return December 16, 1996; revised manuscript received January 27, 1997

REFERENCES


