Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 11457–11461, November 1994
Biochemistry

Molecular cloning of cDNA for double-stranded RNA adenosine deaminase, a candidate enzyme for nuclear RNA editing

UNKYU KIM, YI WANG, TAMARA SANFORD, YONG ZENG, AND KAZUKO NISHIKURA*

The Wistar Institute, 3601 Spruce Street, Philadelphia, PA 19104

Communicated by David Botstein, August 1, 1994

ABSTRACT **We have cloned human cDNA encoding double-stranded RNA adenosine deaminase (DRADA). DRADA is a ubiquitous nuclear enzyme that converts multiple adenosines to inosines in double-helical RNA substrates without apparent sequence specificity. The A → I conversion activity of the protein encoded by the cloned cDNA was confirmed by recombinant expression in insect cells. Use of the cloned DNA as a molecular probe documented sequence conservation across mammals and detected a single transcript of 7 kb in RNA of all human tissues analyzed. The deduced primary structure of human DRADA revealed a bipartite nuclear localization signal, three repeats of a double-stranded RNA binding motif, and the presence of sequences conserved in the catalytic center of other deaminases, including a cytidine deaminase involved in the RNA editing of apolipoprotein B. These structural properties are consistent with the enzymatic signature of DRADA, and strengthen the hypothesis that DRADA carries out the RNA editing of transcripts encoding glutamate-gated ion channels in brain.**

Double-stranded adenosine deaminase (DRADA) is an adenosine deaminase specific for double-stranded RNA (dsRNA) (1, 2). DRADA deaminates multiple adenosines to inosines by a hydrolytic deamination reaction (3). Although a relatively long double-helical RNA structure is required for efficient modification, the minimum length of the double-stranded region for substrate recognition may be as short as 15 bp (4). The precise biological function(s) of DRADA and the full range of its substrate RNAs are currently not known (5). However, several examples of *in vivo* interaction of the enzyme with cellular and viral gene transcripts have been reported in recent years (5). For instance, DRADA seems responsible for genesis of defective measles virus with biased hypermutation, which results in the lethal central nervous system disease measles inclusion body encephalitis (6). Furthermore, the coding of the DRADA-treated mRNAs is expected to be significantly altered, and therefore, the involvement of DRADA in the RNA editing process has been suggested (5, 7, 8). Indeed, DRADA now seems to be responsible at least for the RNA editing of glutamate-gated ion channel subunits (glutamate receptor, GluR) in mammalian brains (9).

Toward our ultimate goal of better understanding the physiological functions of DRADA, we have isolated and characterized cloned cDNA encoding human DRADA.† The deduced primary structure of DRADA revealed the presence of sequence motifs that underlie the functional and catalytic properties of this intriguing enzyme.

MATERIALS AND METHODS

DRADA Assay. The DRADA enzymatic activity was assayed *in vitro* as described (10). The radioactivity of the adenosine and inosine spots on TLC plates was quantified by a phosphor imaging system (Molecular Dynamics).

Determination of N-Terminal Sequence by Microsequencing. Approximately 100 pmol of purified bovine DRADA proteins (10) was fractionated by SDS/PAGE and transferred onto a poly(vinylidine difluoride) membrane. The 93- and 88-kDa bands were cut out and sequenced by Edman degradation (11).

Reverse Transcription–Polymerase Chain Reaction (PCR). Degenerate sets of oligonucleotides that represented the codons for N-terminal peptide sequences were synthesized. For the 93-kDa protein, the sense primer was CGGAATTC-CNGGNAAA/GGTNGA, and the antisense primer was CGGGATCCNGCT/CTCCTT/CTGGT/CTTNA, which correspond to amino acid residues PGKVE and AEQKL, respectively. For the 88-kDa protein, the sense primer was CGGAATTCAAA/GACNGGNTAC/TGTNGA, and the antisense primer was CGGGATCCG/ATCG/ATCNGGG/T/AATG/ATCG/ATC, which correspond to residues KTGYVD and DDPIDD, respectively. Restriction sites (underlined) for *Eco*RI for the sense primer and *Bam*HI for the antisense primer were included at the 5′ end. In addition, internal probes representing the residues flanked by the sense and antisense primers were synthesized. The sequences of the internal probe were for the 93-kDa protein C/TTTG/CAGC/CAGC/T/AGGCTCCTG and for the 88-kDa protein CGGGATCCAT/CTGNCCA/GTTC/TTCT/GTT. The first-strand cDNA synthesis was carried out (12) using total RNA prepared from the cultured bovine endothelial cell line, BFA-1C BPT (13). A portion of the PCR-amplified product was analyzed by Southern blot hybridization using the internal probe. The 75-bp cDNA that hybridized to the internal probe was purified from the agarose gel, digested with *Eco*RI and *Bam*HI, and ligated with pBluescript KS+ plasmid. Selected subclones, BUC1 and BUC2, were sequenced to confirm the identity.

Screening of the Recombinant Library. A human natural killer cell cDNA library was screened using BUC2 as a specific probe. The positive cDNA plasmid, HUC1, contained ≈4 kb of insert DNA, which hybridized to both BUC1 and BUC2 by Southern blot analysis. The insert of HUC1 was then used to rescreen the original cDNA library, from which additional overlapping cDNA clones, HUC2, HUC3, and HUC4, were obtained.

Southern Blot Analysis. Southern blot hybridization was carried out in buffer containing 2× SSC, 1× Denhardt's solution, 40% (vol/vol) formamide, 10% (wt/vol) dextran sulfate, 1% SDS, denatured salmon sperm DNA (0.05 mg/ml), and heat-denatured ^{32}P-labeled probe (10^6 cpm/ml) at 37°C for 18 hr. The membrane was washed with 2× SSC/1% SDS at 37°C for 30 min.

Expression of the Recombinant Baculovirus DRADA Protein. For a full-length construct, pVLDRADA140, *Xba* I–*Kpn*

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

Abbreviations: dsRNA, double-stranded RNA; DRADA, dsRNA adenosine deaminase; DRBM, dsRNA binding motif; GluR, glutamate receptor; RNP, ribonucleoprotein.

*To whom reprint requests should be addressed.

†The sequence reported in this paper has been deposited in the GenBank data base (accession no. U10439).

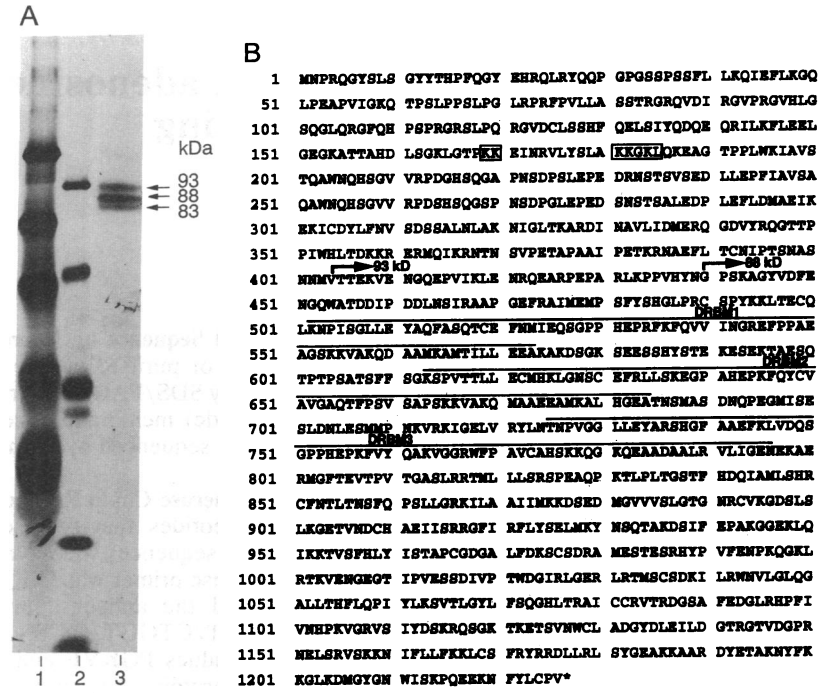

FIG. 1. Cloning of cDNA encoding human DRADA. (*A*) The final purified fraction was analyzed by SDS/PAGE (7% gel) and three protein bands were visualized by silver staining. Lanes: 1 and 2, molecular mass standards; 3, bovine DRADA. (*B*) The deduced amino acid sequence of the 136-kDa form of human DRADA is shown. The putative bipartite nuclear localization signal is boxed. The N-terminal sequences of 93- and 88-kDa proteins are indicated by arrows; three DRBM repeats are indicated by lines.

I (the 5′-end 3.7 kb of HUC1) and *Kpn* I–*Xba* I (the 3′-end 1 kb of HUC2) fragments were ligated with pVL1393 (Invitrogen). For pVLDRADAΔ, a mutant lacking the C-terminal 346 amino acids, a new termination codon was created at residue 880 by filling in the overhang of *Xma* III located downstream of the dsRNA binding motifs (DRBMs). Infected *Spodoptera frugiperdera* (Sf9) cells were labeled with 50 μCi of [^{35}S]methionine (1 Ci = 37 GBq) for 1 hr (14), and the labeled protein was analyzed by SDS/PAGE and fluorography.

Determination of the cDNA Sequence. DNA sequencing of sets of nested deletion mutants of the cDNA clones was carried out with the model 373A DNA sequencing system (Applied Biosystems).

Sequence Analysis. The overlapping sequences were aligned and combined by the FRAGMENT ASSEMBLY program (15). Alignments of DRBM and deaminases were determined by PILEUP, BESTFIT, and GAP programs, and identification of various protein sequence motifs was by the MOTIFS program (15). A homology search of the National Center for Biotechnology Information peptide sequence database (Release 82.0) was performed by Internet electronic mail.

RESULTS

Purification of DRADA Proteins and Isolation of DRADA cDNAs. By following the A → I modification activity, we have purified (10) DRADA to homogeneity from bovine liver nuclear extracts. The final fraction contained three DRADA polypeptides of 93, 88, and 83 kDa (Fig. 1*A*). The N-terminal amino acid sequences of the 93-kDa and 88-kDa proteins were determined by microsequencing, and two short bovine cDNA clones encoding the N termini were obtained. Using

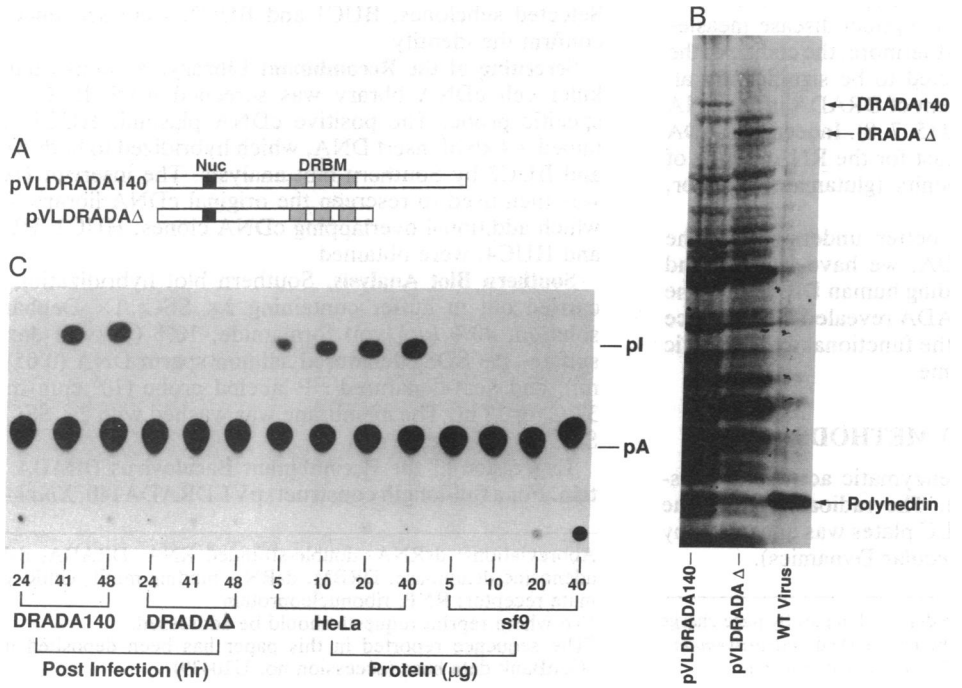

FIG. 2. Characterization of the recombinant DRADA proteins. (*A*) Full-length DRADA140 and C-terminally truncated DRADAΔ constructs are shown. (*B*) Forty-eight hours after infection with recombinant or wild-type baculoviruses, Sf9 cells were labeled with [^{35}S]methionine and proteins were separated by SDS/PAGE. (*C*) Crude extracts made after infection at various time points were analyzed by a modified-base assay using 20 μg of extract as described (10). Increasing amounts of extracts made from Sf9 cells or HeLa cells were also assayed. pI, inosine 5′-monophosphate; pA, adenosine 5′-monophosphate.

one of these bovine cDNAs as a specific probe, we screened a recombinant cDNA library made from human natural killer cells, which contain a high level of DRADA activity (16). The nucleotide sequence of 6671 bp derived from the overlapping cDNA clones contained a short 5′ untranslated region (154 bp) and a long 3′ untranslated region (2839 bp), including a poly(A)$^+$ tract (GenBank accession no. U10439). We found a single open reading frame encoding a 1226-amino acid protein with a calculated molecular mass of 136 kDa (Fig. 1*B*). The proposed initiation codon is in reasonable agreement with the mammalian translation initiation consensus sequence (17) and is preceded by an in-frame stop codon. This open reading frame contained the N-terminal sequences of both 93- and 88-kDa proteins, which appear to be truncated forms, lacking, respectively, 403 and 439 amino acid residues of the N terminus of the full-length 136-kDa DRADA protein.

Analysis of the DRADA Activity of the Recombinant Baculovirus Protein. We analyzed the protein encoded by the isolated cDNA clone by recombinant expression in Sf9 cells (Fig. 2). A unique protein band of ≈140 kDa was detected in cells infected with the recombinant virus containing the entire coding sequence (DRADA140), indicating that a full-length protein was expressed. The cells infected with a recombinant virus containing a C-terminally truncated DRADA sequence (DRADAΔ) produced a protein of ≈110 kDa. Only the extracts of the cells expressing DRADA140 showed the A→I conversion activity at a high level, whereas cells expressing DRADAΔ and uninfected Sf9 cells displayed very little activity (Fig. 2*C*). These results strongly suggest that the cloned cDNA indeed codes for a functional DRADA enzyme, although we cannot exclude the very unlikely possibility that the cloned cDNA activated a latent form of DRADA protein or the DRADA gene itself. Interestingly, the N-terminally truncated forms of DRADA, 93 and 88 kDa, were not detected in Sf9 cells infected with recombinant virus carrying the entire open reading frame.

Conservation of the DRADA Gene and Ubiquitous Expression of DRADA. A Southern blot analysis indicated that DRADA is well conserved in mammalian cells as genomic DNA prepared from human, bovine, and mouse cells hybridized strongly with the human cDNA probe (Fig. 3*A*). However, this probe did not detect sequences in amphibian or insect genomes. Since the enzymatic activity of DRADA has been reported in these two species (1, 2, 16), we tested additional cDNA probes, including a DNA fragment encoding the C-terminal region, predicted to contain the conserved catalytic domain (see *Discussion*). All probes gave negative results. Thus, the DRADA sequence may not have been well conserved during evolution, except in certain short stretches of amino acid sequences possibly involved in the catalysis (see below).

As shown in Fig. 3*B*, all tissues analyzed contained DRADA transcripts. The size of the DRADA mRNA (7 kb) indicates that our overlapping cDNAs (6671 nt) contain nearly the entire DRADA mRNA. The DRADA gene appeared to be expressed ubiquitously. Brain tissue contains relatively high levels of the DRADA transcript, consistent with proposed involvement of DRADA in the RNA editing of GluR subunits (7–9).

Deduced Structural Features of DRADA. Computer-assisted inspection of the predicted primary structure revealed several features that illuminate the known and proposed functional properties of DRADA. Consistent with the documented localization of DRADA in the cell nucleus (16), a bipartite nuclear localization signal (18) is found at residues 169–185 (Fig. 1*B*). The central region of the protein harbors three repeats of DRBM (Fig. 4). The presence of these motifs explains the selectivity of DRADA for dsRNA (1, 2, 4, 10) and identifies DRADA as a member of a growing family of DRBM-containing proteins (19, 20) (Fig. 4).

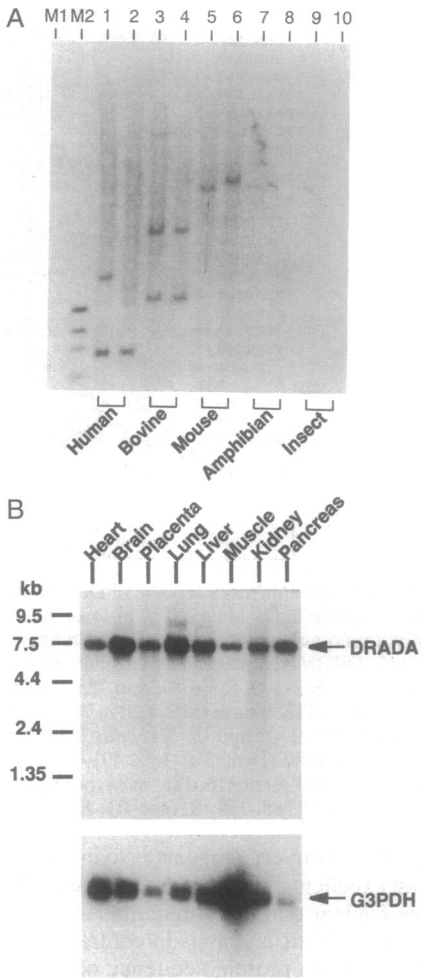

FIG. 3. Conservation of the DRADA gene structure and ubiquitous expression of DRADA. (*A*) Twenty micrograms of chromosomal DNA obtained from HeLa (human), BFA-1C BPT (bovine), MOPC11 (mouse), XTC-2 (*Xenopus laevis*), and Sf9 (insect) cells was digested with either *Eco*RI and *Bam*HI (odd numbered lanes) or *Eco*RI and *Hin*dIII (even numbered lanes) and analyzed by Southern blot hybridization with a human DRADA cDNA probe containing the DRBM region. Lanes M1 and M2 contain molecular mass standards. (*B*) (*Upper*) A Northern blot containing 2 μg of human poly(A)$^+$ RNA was hybridized with the human DRADA cDNA probe. (*Lower*) The blot was rehybridized with a glyceraldehyde-3-phosphate dehydrogenase (G3PDH) cDNA control probe.

Finally, as shown in Fig. 5, the C-terminal region of DRADA contains the tripeptide sequences HAE and PCG, which are conserved in several cytidine and dCMP deaminases, including REPR, a recently cloned enzyme involved in the RNA editing of apolipoprotein B (21). These tripeptides are likely to be involved in the coordination of a zinc atom and formation of the catalytic center of DRADA (see below). Curiously, a database search revealed a nematode gene (T20H4.4) of unknown function (22) with a considerable degree of sequence conservation to the C-terminal region of DRADA, particularly around and including the tripeptide sequences HAE and PCG (Fig. 6). This nematode gene may encode a prototype of the vertebrate version of DRADA.

DISCUSSION

In this study, we have reported isolation and characterization of cloned cDNA encoding human DRADA. In spite of its ubiquitous occurrence in all animal tissues, the *in vivo* role of DRADA has remained obscure (5). On the strength of its

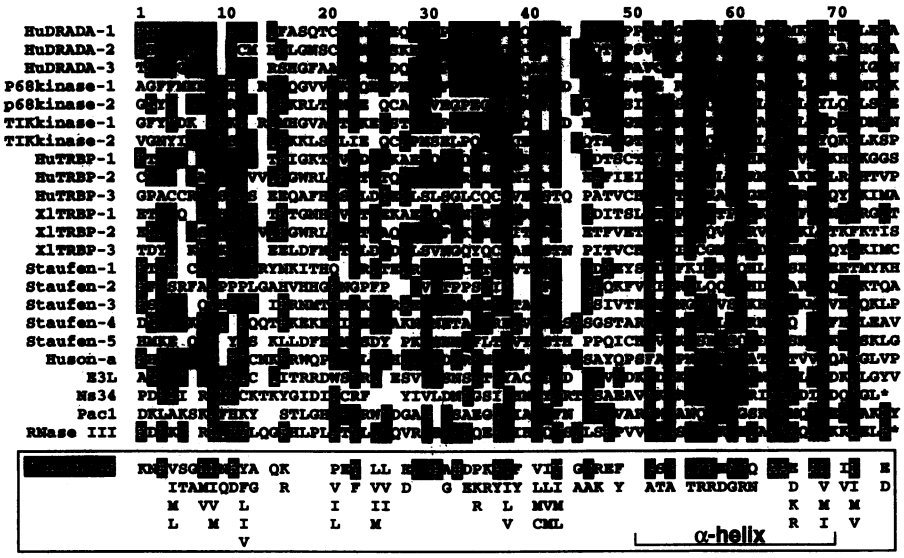

FIG. 4. DRADA contains three repeats of DRBM. The following groups of amino acids were weighted equivalent: V,L,I,M; K,R; S,T; E,D; Y,F; N,Q; A,G. Amino acids identical in at least six sequences or similar in at least eight sequences are boxed. For the deduced consensus sequence, only amino acids identical in at least nine sequences are boxed. The C-terminal region of the DRBM, predicted to form an α-helix (19), is indicated. DRBM sequences of the following genes (accession numbers indicated in parentheses) were aligned: HuDRADA-1, -2, and -3, human DRADA nt 502–573, 613–684, and 725–796; p68 kinase-1 and -2, human dsRNA-dependent kinase (M35663) nt 7–79 and 99–169; TIK kinase-1 and -2, mouse TIK kinase (M65029) nt 7–78 and 93–164; HuTRBP-1, -2, and -3, human TAR-binding protein (M60801) nt 8–78, 137–208, and 270–342; X1TRBP-1, -2, and -3, *Xenopus* TAR-binding protein (M96370) nt 19–89, 111–182, and 224–295; Staufen-1, -2, -3, -4, and -5, *Drosophila* Staufen (M69111) nt 310–380, 491–559, 577–647, 710–782, and 951–1020; Huson-a, human son-a (X63753) nt 1363–1436; E3L, vaccinia virus E3L (M36339) nt 116–186; ns34, porcine group C rotavirus ns34 protein (M69115) nt 334–402; PacI, *Schizosaccharomyces pombe* PacI protein (X54998) nt 288–358; RNase III, *Escherichia coli* RNase III (X02673) nt 154–227.

unique property to convert adenosines to inosines in dsRNA, DRADA is an excellent candidate enzyme responsible for the site-selective nuclear editing of mammalian transcripts such as those encoding GluR channels in central nervous system (5, 7–9). The deduced protein sequence of human DRADA reveals several interesting structural features that appear to underlie the functional facets of this enzyme.

DRADA and dsRNA Interaction. The region making up the DRBMs is likely to constitute the dsRNA binding domain. DRADA has uncommonly high affinity for dsRNA (10), which can now be explained in part by the three copies of DRBM present in DRADA. Since each motif is capable of binding independently to dsRNA (19, 20), having three such motifs may allow DRADA to make three contacts with dsRNA, which may increase the affinity for dsRNA in a cooperative manner (4, 10). Although an increasing number of proteins that interact with dsRNA have been discovered in recent years (19, 20), the exact nature of the interaction between these proteins and dsRNA is poorly understood. DRADA and other proteins containing DRBM bind to dsRNA without strict sequence specificity (1, 2, 19, 20). Thus, interaction between dsRNA and DRBM may be mediated by hydrogen bonding of protein side chains with the backbone phosphate, which is proposed to be involved in nonspecific protein–nucleic acid interactions (23). Binding of DRADA may also cause distortion in the dsRNA, including disruption of the limited base pairings around the sites of modification to render the adenosines accessible to the catalytic domain of DRADA. The fact that adenosines in an AU-rich region are modified by DRADA more frequently (9) could be explained by the relative ease of melting A·U base pairs compared to G·C pairs. Interestingly, there is a partial conservation of a ribonucleoprotein (RNP) core consensus sequence just 62 residues upstream of DRBM-1 (GYVDF). The RNP consensus found in many single-stranded RNA binding proteins consists of a 90-residue stretch of loosely conserved sequence within which reside highly conserved core sequences of eight (RNP-1) and five (RNP-2) residues (24). The short RNP-2-like stretch found in DRADA may participate in destabilizing A·U base pairs and in creating a local single-stranded RNA region (24).

Catalytic Mechanism of DRADA and Conservation of Residues Required for Deamination. A set of evolutionarily conserved amino acid residues arranged and spaced in a specific sequence context has been reported for adenosine and AMP deaminases (25) and also for cytidine deaminases

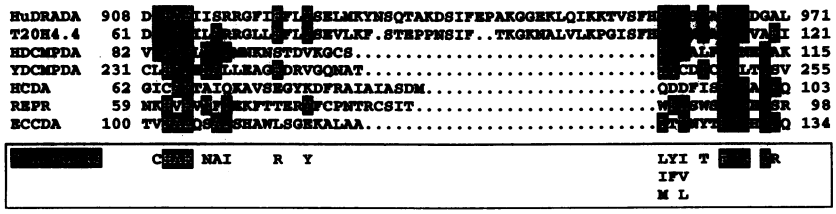

FIG. 5. Amino acid residues involved in the catalytic mechanism of DRADA. The regions previously known to contain highly conserved amino acid residues of the following deaminases (accession numbers indicated in parentheses) were aligned with those of the conserved regions found in human DRADA and nematode T20H4.4 (U00037): human dCMP deaminase (HDCMPDA; L12136), yeast dCMP deaminase (YDCMPDA; M13010), human cytidine deaminase (HCDA; S52873), REPR (L07114), and *Escherichia coli* cytidine deaminase (ECCDA; M60916). Amino acid residues identical in at least three sequences, or similar in five sequences, are boxed. For the deduced consensus sequence, only amino acids identical in at least five sequences are boxed.

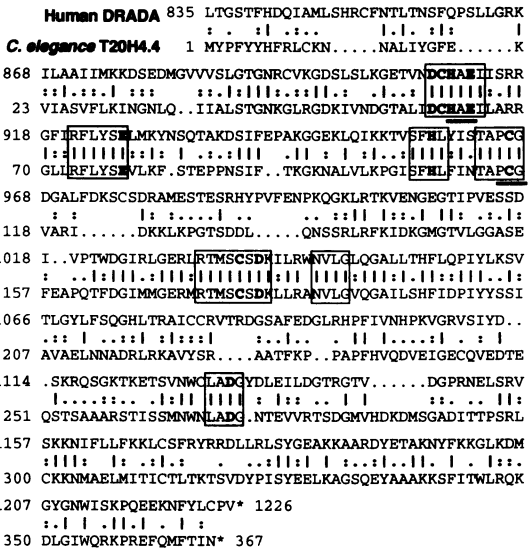

FIG. 6. Amino acid sequence comparisons of human DRADA and *Caenorhabditis elegans* T20H4.4. A comparison of the entire two sequences shows 37% identity. A conservative amino acid substitution is also indicated (· or :). Seven boxes within the C-terminal region indicate the highly conserved contiguous amino acid sequences. Boldface type highlights amino acid residues potentially capable of coordinating a zinc atom.

and dCMP deaminases (21). Among these residues, histidine, cysteine, glutamic acid, and aspartic acid, often found in highly conserved short amino acid stretches, are believed to be involved in the coordination of a zinc atom and to participate directly in the hydrolytic deamination mechanism commonly employed by these deaminases (21, 25, 26). We proposed earlier that the catalytic mechanism of DRADA might be different from adenosine or AMP deaminases (10). Indeed, DRADA does not contain sequences homologous to other adenosine deaminases (25) but, instead, contains in its C-terminal region the tripeptides HAE and PCG, conserved in cytidine and dCMP deaminases, including the RNA editing enzyme REPR (21). The three-dimensional structure of *Escherichia coli* cytidine deaminase complexed to a transition state analogue has been solved (26). The glutamic acid residue of HAE provides all of the necessary proton transfer functions of the *E. coli* enzyme. In addition, the histidine of HAE and the cysteine of PCG appear to participate in the zinc chelation (26). Our prediction that these tripeptide sequences are directly involved in the catalytic mechanism of DRADA can now be tested by site-directed mutagenesis.

DRADA as an RNA Editing Enzyme. Because of the enzyme's ability to introduce base changes in its substrate RNA sequence, the involvement of DRADA in the RNA editing process has been proposed (1, 2, 5). Inosine is treated as guanosine by the translation machinery (5). In addition, reverse transcriptase recognizes inosine as guanosine, and thus, the A → I change introduced by DRADA is revealed by an A → G change in the corresponding cDNA. To date, three cases of site-selective RNA editing in mammalian cellular RNA have been reported. One occurs in the rat kidney *WT1* (Wilms tumor suppressor oncogene) transcript, in which a single genomically specified uridine residue in RNA is changed to cytidine resulting in the Pro → Leu replacement (27). The second example is the tissue-specific editing of human intestinal apolipoprotein B mRNA in which a cytidine residue is deaminated to uridine, creating a new translation stop codon required for the synthesis of a shorter apolipoprotein B-48 (21). The third case concerns the editing of GluR transcripts. There, all documented instances result in the conversion of particular adenosines in the gene to guanosines in the cDNA but may involve deamination of the adenosines to inosines in the RNA by DRADA (7, 8). The ensuing amino acid changes lead to altered gating, ion conductance, and ion permeability properties of the edited GluR channels (7, 8).

Although circumstantial evidence suggested that the RNA editing of GluR transcripts is carried out by DRADA, its lack of sequence specificity was puzzling in light of the high degree of precision expected for the editing enzyme (5, 7, 8). Furthermore, any substrate RNA for DRADA must be in duplex form. Surprisingly, recent findings (9) have shown how the strict requirement of DRADA for dsRNA can be used to achieve sequence-specific editing of GluR transcripts. In one case, the editing replaces the gene-encoded glutamine (Q encoded by CAG) with an arginine (R encoded by CGG). For this Q/R site editing in GluR-B, formation of a 17-bp dsRNA structure between exon and intron sequences is a prerequisite (9). In addition, the editing further requires, for efficient modification, a 45-nt inverted repeat structure. Additional adenosines in the intron were found to be modified also. These two characteristics, requirement for dsRNA and multiple adenosine modifications, strongly suggest that DRADA may indeed be responsible for the editing of this particular transcript (9). In view of the ubiquitous expression of DRADA in different tissues (16), one may reasonably expect the involvement of DRADA in the editing of other genes. The cloned cDNA for DRADA should facilitate the testing of these hypotheses.

This work was supported by National Institutes of Health Grants GM40536, CA09171, and CA10815.

1. Bass, B. L. & Weintraub, H. (1988) *Cell* **55**, 1089–1098.
2. Wagner, R. W., Smith, J. E., Cooperman, B. S. & Nishikura, K. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2647–2651.
3. Polson, A. G., Crain, P. F., Pomerantz, S. C., McCloskey, J. A. & Bass, B. L. (1991) *Biochemistry* **30**, 11507–11514.
4. Nishikura, K., Yoo, C., Kim, U., Murray, J. M., Estes, P. A., Cash, F. E. & Liebhaber, S. A. (1991) *EMBO J.* **10**, 3523–3532.
5. Kim, U. & Nishikura, K. (1993) *Semin. Cell Biol.* **4**, 285–293.
6. Bass, B. L., Weintraub, H., Cattaneo, R. & Billeter, M. A. (1989) *Cell* **56**, 331.
7. Sommer, B., Köhler, M., Sprengel, R. & Seeburg, P. H. (1991) *Cell* **67**, 11–19.
8. Köhler, M., Burnashev, N., Sakmann, B. & Seeburg, P. H. (1993) *Neuron* **10**, 491–500.
9. Higuchi, M., Single, F. N., Köhler, M., Sommer, B., Sprengel, R. & Seeburg, P. H. (1993) *Cell* **75**, 1361–1370.
10. Kim, U., Garner, T. L., Sanford, T., Speicher, D., Murray, J. M. & Nishikura, K. (1994) *J. Biol. Chem.* **269**, 13480–13489.
11. Matsudaira, P. (1987) *J. Biol. Chem.* **261**, 10035–10038.
12. Lee, C. C., Wu, X., Gibbs, R. A., Cook, R. G., Muzny, D. M. & Caskey, C. T. (1988) *Science* **239**, 1288–1291.
13. Grinspan, J. B., Mueller, S. N. & Levine, E. M. (1983) *J. Cell. Physiol.* **114**, 328–338.
14. O'Reilly, D. R., Miller, L. K. & Lucklow, V. A. (1992) *Baculovirus Expression Vectors* (Freeman, Oxford).
15. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
16. Wagner, R. W., Yoo, C., Wrabetz, L., Kamholz, J., Buchhalter, J., Hassan, N. F., Kim, S. U., Perussia, B., McMorris, F. A. & Nishikura, K. (1990) *Mol. Cell. Biol.* **10**, 5586–5590.
17. Kozak, M. (1989) *J. Cell Biol.* **108**, 229–241.
18. Laskey, R. A. & Dingwall, G. (1993) *Cell* **74**, 585–586.
19. McCormack, S. J., Thomis, D. C. & Samuel, C. E. (1992) *Virology* **188**, 47–56.
20. St. Johnson, D., Brown, N. H., Gall, J. G. & Jantsch, M. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10979–10983.
21. Navaratnam, N., Morrison, J. R., Bhattacharya, S., Patel, D. P., Funahashi, T., Giannoni, F., Teng, B.-B., Davidson, N. O. & Scott, J. (1993) *J. Biol. Chem.* **268**, 20709–20712.
22. Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., *et al.* (1994) *Nature (London)* **368**, 32–38.
23. Anderson, J. E., Ptashne, M. & Harrison, S. C. (1987) *Nature (London)* **326**, 846–852.
24. Haynes, S. R. (1992) *New Biol.* **4**, 421–429.
25. Chang, Z., Nygaard, P., Chinault, A. C. & Kellem, R. K. (1991) *Biochemistry* **30**, 2273–2280.
26. Betts, L., Xiang, S., Short, S. A., Wolfenden, R. & Carter, C. W., Jr. (1994) *J. Mol. Biol.* **235**, 635–656.
27. Sharma, P. M., Bowman, M., Madden, S. L., Rauscher, F. J., III, & Sukumar, S. (1994) *Genes Dev.* **8**, 720–731.