Functionally Distinct Double-stranded RNA-binding Domains Associated with Alternative Splice Site Variants of the Interferon-inducible Double-stranded RNA-specific Adenosine Deaminase*

(Received for publication, August 26, 1996, and in revised form, October 23, 1996)

Yong Liu‡, Cyril X. George‡, John B. Patterson§§, and Charles E. Samuel‡‡

From the 3Department of Molecular, Cellular and Developmental Biology, and 4Interdepartmental Graduate Program of Biochemistry and Molecular Biology, University of California, Santa Barbara, California 93106

The double-stranded RNA-specific adenosine deaminase (ADAR) is an interferon-inducible RNA-editing enzyme implicated in the site-selective deamination of adenosine to inosine in viral RNAs and cellular pre-mRNAs. We have isolated and characterized human genomic clones of the ADAR gene and cDNA clones encoding splice site variants of the ADAR protein. Southern blot and sequence analyses revealed that the gene spans about 30 kilobase pairs and consists of 15 exons. The codon phasing of the splice junctions of exons 3, 5, and 7 that encode the three copies of the highly conserved RNA-binding R-motif (R₁, R₂, and R₃) was exactly conserved and identical to those R-motif exons of the interferon-inducible RNA-dependent protein kinase. Alternative splice site variants of the 1226-amino acid ADAR-a protein, designated b and c, were identified that differed in exons 6 and 7. ADAR-b was a 5'-splice site variant that possessed a 26-amino acid deletion within exon 7; ADAR-c was a 3'-splice site variant that possessed an additional 19-amino acid deletion within exon 6. The wild-type ADAR-a, -b, and -c proteins all possessed comparable double-stranded RNA-specific adenosine deaminase activity. However, mutational analysis of the R-motifs revealed that the exon 6 and 7 deletions of ADAR-b and -c variants altered the functional importance of each of the three R-motifs.

Double-stranded RNA-specific adenosine deaminase (ADAR)* is an interferon-inducible, dsRNA-binding protein (1–6). ADAR, also known as dsRAD and DRADA, catalyzes the covalent modification of double-stranded RNA substrates by hydrolytic C-6 deamination of adenosine to yield inosine (7, 8). ADAR is implicated in two types of RNA editing processes. First, A-to-I modifications are found at multiple sites in viral RNAs, as exemplified by the biased hypermutations observed in negative-stranded RNA virus genomes during lytic and persistent infections, as in the case of measles virus (9, 10). Second, the C-6 adenosine deamination can be highly site-specific, occurring at one or a few sites in certain viral and cellular mRNAs as exemplified by hepatitis delta virus (HDV) RNA (11) and the GluR receptor channel pre-mRNAs (12), respectively. RNA editing plays an essential role in production of two HDV proteins from one ORF, proteins which have different functions in the life cycle of the closed circular single-stranded RNA HDV with the editing site located in a self-complementary dsRNA structure (11). Recent re-evaluation of HDV RNA editing demonstrated that the process occurs on antigenic RNA and involves a conversion of A to G (13, 14), raising the possible role of ADAR in the regulation of HDV replication. ADAR is also presumed to be responsible for the specific RNA editing of pre-mRNA transcripts in the brain encoding the a-amino-3-hydroxy-5-methyl-4-isoxazole propionate (AMPA) and kainate glutamate-gated ion channel GluR receptor subunits (15, 16). Editing of the GluR pre-mRNAs results in altered channel properties of the encoded proteins (12, 17). The highly selective Q to R, and R to G, amino acid changes in GluR subunits are dependent upon formation of dsRNA hairpin structures involving sequences derived from exon 11 and intron 11 for the Q/R site (18, 19) and exon 13 and intron 13 for the R/G site (17). Thus, the A to I deaminations observed both in viral RNA genomes and cellular pre-mRNA transcripts are dependent upon double-stranded regions within the substrate RNA.

The molecular cDNA cloning of ADAR from human, rat, and bovine cells has been described (6, 20, 21). The ADAR cDNA hybridizes to an interferon (IFN)-inducible 7-kb mRNA and possesses a single long open reading frame predicted to encode a 1226-amino acid protein (6, 20, 21). We isolated the 6.5-kb human cDNA encoding ADAR during a screen for cDNA clones of proteins regulated by IFN (5, 6). Genomic clones corresponding to the ADAR cDNA have likewise been isolated and the gene mapped by fluorescence in situ hybridization to human chromosome 1q21.1–21.2 (22).

The presence of two immunologically related forms of the human ADAR deaminase was demonstrated in a variety of human cell lines using antisera prepared against three non-overlapping regions of the human cDNA expressed in Escherichia coli (6); an interferon-inducible 150-kDa protein present in both the cytoplasm and nucleus; and a constitutively expressed 110-kDa protein present predominantly if not exclusively in the nucleus. A second cDNA for the rat dsRNA adenosine deaminase, designated RED1, likewise hybridizes to a 7-kb RNA but predicts a 771-amino acid protein (23), substan-
Splice Variants of the RNA-specific Adenosine Deaminase

The C domain. The other variant, designated as ADAR-c, has an amino acid residue deletion of 26 amino acids at the exon 7-intron VII junction. The RII domain subdomain and the catalytic C domain of ADAR was essential for both dsRNA deaminase activity and dsRNA-binding activity of the human PKR and vaccinia virus gene products (1–4) are similar to those elucidated for other known double-stranded RNA-binding proteins (28). The core lysine residue of the R core (32). This lysine residue was previously established as one of the R core residues essential for dsRNA-binding activity of the human PKR and vaccinia virus E3L proteins (29–31). Examination in vitro of the various R-motif mutants of ADAR revealed that the RII subdomain was essential for both dsRNA deaminase activity and dsRNA-binding activity, whereas RII was dispensable for both activities (32). Similar conclusions were derived from the analysis of deletion mutants of ADAR (34). These results together indicated that the three R-motifs of ADAR are functionally distinct from each other in the context of binding dsRNA substrates in a functional manner recognized by the enzyme catalytic center for deamination. The catalytic C domain of ADAR was mapped by mutagenesis to the C-terminal region of the protein (16, 32, 34).

Here we report the identification of three naturally occurring variant forms of ADAR, including two alternative splice variants that are differentially expressed in different tissues. In comparison to the prototype ADAR (6, 20, 21), we now designated herein as ADAR-a, one isoform designated as ADAR-b contains a deletion of 26 amino acids at the exon 7-intron VII junction which lies between the RII motif subdomain and the catalytic C domain. The other variant, designated as ADAR-c, has an additional deletion of 19 amino acids at the intron V-exon 6 junction which lies between the RI and RII motif subdomains. These splice site variants retain the R-motif subdomains functionally intact, and all three of the ADAR variants encode active enzymes that possess comparable deaminase activity measured with a synthetic dsRNA substrate. Site-directed mutagenesis of the R-motifs revealed that the presence of one or both of the deletions in the spacer regions between the R-motifs and the catalytic domain of the ADAR-b and -c variants altered the functional importance of each of the individual R-motifs relative to that observed for ADAR-a.

**EXPERIMENTAL PROCEDURES**

**Oligonucleotides**—Oligonucleotides designed from the human ADAR-a cDNA sequence (6) were used for site-directed mutagenesis of the three dsRNA-binding R-motif subdomains (RI, RII, and RIII), for screening cDNA libraries by PCR, and for the construction of plasmid expression vectors. The following oligonucleotides were synthesized.

1. NosI (nt 31/RI* (nt 1900–1917); RII* (nt 1090–1107); RIII* (nt 2298)); 5'-AATGCCTCGCCGCGCGCCGAG-3' (nt 31/RII* (nt 1900–1917); RII* (nt 1090–1107); RIII* (nt 2298)); 5'-CAGTGGATGATGGTTCTGGAC-3'

These oligonucleotides were used respectively for the in vitro PCR and sequence analyses, as described by Sambrook et al.'s instructions (U.S. Biochemical Corp.). For those regions of ADAR DNA polymerase according to the manufacturer’s instructions (U.S. Biochemical Corp.). For those regions of ADAR where differences were identified from the originally described ADAR cDNA (6), a ZAP human placenta cDNA expression vector (Stratagene) and a mouse pY2 cDNA library derived from J774.1 mouse macrophase-cell line (kindly provided by Drs. F. Ferrier and C. Vandenberg, University of California, Santa Barbara) were also analyzed by PCR with multiple primer pairs.

**Oligonucleotide-directed Mutagenesis of the R-motifs and Construction of Human ADAR cDNA Expression Vectors**—The PCR-based method of site-directed mutagenesis of the prototype human ADAR cDNA (6), described previously in detail (32), was utilized for the newly isolated splice site variants. The methods for construction of the transcription vector plasmids were essentially as described by Sambrook et al. (37). Chemicals were reagent grade; enzymes were obtained from New England Biolabs unless otherwise noted.

**Starting with plasmid expression vector pDNANeoK88 (amino acids 296–1226) engineered to lack the 5'-GC-rich portion of the ORF and to initiate at Met-296 (6) as the wild-type parent, constructions possessing mutations in one or more of the three R-motifs had previously been described in the form of M296 ADAR (32). These included constructs with the K554E, R665E, and K776E substitutions, the three single mutants each with one of the R-motifs altered, and with R554E/R665E/R776E, the triple mutant in which all three of the R-motifs were altered. Combined PCR and sequence analyses, as described by the "Results," revealed the presence of two naturally occurring splice site variants of the prototype deaminase cDNA (6, 20), designated herein as ADAR-a. One of the variants (termed ADAR-b) possessed a 78-nt in-frame deletion between the RII motif and the catalytic C domain; the other (termed ADAR-c) possessed a 26-nt in-frame deletion between the RII and RIII motifs. The K88 3.7-kb starting parent previously constructed from overlapping λ cDNA clones (6), utilized because of the presence of convenient restriction sites that facilitated preparation of expression vector constructs, was found to correspond to ADAR-b. Using the previously reported mutagenesis strategy (32), PCR fragments corresponding to ADAR-a and -c generated with primer pairs Sall(+/KpnI-) and (+/KpnI-) from the

**GAG(-3' (nt 2023–2042); RII* (nt 1090–1107); RIII* (nt 2298)); 5'-CAGTGGATGATGGTTCTGGAC-3'

These oligonucleotides were used respectively for the in vitro PCR and sequence analyses, as described by Sambrook et al.'s instructions (U.S. Biochemical Corp.). For those regions of ADAR DNA polymerase according to the manufacturer’s instructions (U.S. Biochemical Corp.). For those regions of ADAR where differences were identified from the originally described ADAR cDNA (6), a ZAP human placenta cDNA expression vector (Stratagene) and a mouse pY2 cDNA library derived from J774.1 mouse macrophase-cell line (kindly provided by Drs. F. Ferrier and C. Vandenberg, University of California, Santa Barbara) were also analyzed by PCR with multiple primer pairs.

**Oligonucleotide-directed Mutagenesis of the R-motifs and Construction of Human ADAR cDNA Expression Vectors**—The PCR-based method of site-directed mutagenesis of the prototype human ADAR cDNA (6), described previously in detail (32), was utilized for the newly isolated splice site variants. The methods for construction of the transcription vector plasmids were essentially as described by Sambrook et al. (37). Chemicals were reagent grade; enzymes were obtained from New England Biolabs unless otherwise noted.

**Starting with plasmid expression vector pDNANeoK88 (amino acids 296–1226) engineered to lack the 5'-GC-rich portion of the ORF and to initiate at Met-296 (6) as the wild-type parent, constructions possessing mutations in one or more of the three R-motifs had previously been described in the form of M296 ADAR (32). These included constructs with the K554E, R665E, and K776E substitutions, the three single mutants each with one of the R-motifs altered, and with R554E/R665E/R776E, the triple mutant in which all three of the R-motifs were altered. Combined PCR and sequence analyses, as described by the "Results," revealed the presence of two naturally occurring splice site variants of the prototype deaminase cDNA (6, 20), designated herein as ADAR-a. One of the variants (termed ADAR-b) possessed a 78-nt in-frame deletion between the RII motif and the catalytic C domain; the other (termed ADAR-c) possessed a 26-nt in-frame deletion between the RII and RIII motifs. The K88 3.7-kb starting parent previously constructed from overlapping λ cDNA clones (6), utilized because of the presence of convenient restriction sites that facilitated preparation of expression vector constructs, was found to correspond to ADAR-b. Using the previously reported mutagenesis strategy (32), PCR fragments corresponding to ADAR-a and -c generated with primer pairs Sall(+/KpnI-) and (+/KpnI-) from the
human kidney λ CDNA library, respectively, were used to produce the wild-type, the three single R mutants, and the triple R mutant for M296 ADAR-a and -c. The wild-type construct encoding ADAR-R which contained only the 57-nt deletion between the R1 and R2H motifs was also engineered likewise. To obtain constructs encoding the full-length M1 ADAR protein, a PCR product corresponding to the N-terminal region of ADAR was amplified from the human kidney cDNA library using primers BamHI (B) EcoRI (E), HindIII (H), and XhoI (X). C, five of the overlapping λ phage genomic clones (6, 7, 143, 151, and 176) are shown to scale, along with the three P1 genomic clones (249, 652, and 959) that span the entire length of the ADAR gene and continue into flanking sequences.

**Splice Variants of the RNA-specific Adenosine Deaminase**

**Double-stranded RNA-specific Adenosine Deaminase Assay—**Measurement of dsRNA-specific adenosine deaminase activity was essentially as described previously (3, 6, 32). The [32P]-labeled synthetic dsRNA substrate was prepared by annealing opposing transcripts in 20 mM Tris-HCl, pH 7.9, containing 0.15 M NaCl by first heating at 75 °C for 5 min and then slowly cooling to room temperature. The standard deaminase reaction mixture contained approximately 10 fmol of [32P]-labeled dsRNA and varying amounts of ADAR protein; following incubation at 30 °C for 2 h, the reaction mixture was extracted with phenol: chloroform and then chloroform. RNA was recovered by ethanol precipitation in the presence of 0.5 µg of poly(rI) (Sigma) added as carrier, washed with 70% ethanol, dried, and suspended in 10 µl of nuclease PI buffer (30 mM potassium acetate, pH 5.3, containing 10 mM zinc sulfate) before digestion with 1.5 U of nuclease PI (Pharmacia Biotech Inc.) for 2 h at 50 °C (3). IMP and AMP were resolved from each other by thin layer chromatography (TLC) on cellulose NM 300 glass plates (Macherey and Nagel) in a solvent consisting of saturated (NH4)2SO4, 100 mM sodium acetate, pH 6.0, and isopropl alcohol (79:19:2). Autoradiography was usually performed for 8 h at ~80 °C with a screen; radioactivity associated with the excised TLC spots was quantified using a Beckman LS1801 liquid scintillation system.

**Isolation and Characterization of Genomic Clones—**A human genomic library in the λ phage vector EMBL-3 SP6/T7 prepared from human placenta DNA (Clontech) was screened initially by filter hybridization using random primed [32P]-labeled cDNA fragments of the human ADAR-a cDNA as the probes (6, 37). Subsequently, probes included random-primed [32P]-labeled genomic fragments. Lambda phage DNA was prepared from twice-rescreened plaques, and genomic inserts were characterized by restriction mapping and Southern blot analysis (6, 39). A human genomic P1 library in the pAD10SacBII vector (40) was screened by PCR using synthetic oligonucleotide primers based on sequences determined from previously isolated λ genomic clones. The library screening by PCR followed the protocol described previously (22). Restriction fragments of positive genomic clones were subcloned into the pBluescript plasmid (Stratagene) for detailed restriction mapping and DNA sequencing by the dideoxynucleotide procedure (36).

**Materials—**Unless otherwise specified, materials and reagents were as described previously (6, 22, 32).

**RESULTS**

**Isolation of Genomic Clones and Determination of the Structural Organization of the Human ADAR Gene—**A human genomic library in the λ phage vector EMBL3 was screened using as probes restriction fragments of the human ADAR cDNA, previously designated as the K88 or DSRAD cDNA (6). Several overlapping λ phage clones were isolated (Fig. 1C). These genomic clones were characterized by restriction map-
ping and Southern blot analysis. However, because the typical insert size of the λ phage genomic clones was about 15 kb and the overlap between λ clones was limited, a P1 phage genomic library was subsequently screened. Three overlapping P1 clones with inserts of about 85 kb, designated as clones 249, 652, and 959, were isolated that covered the entire ADAR gene.

A composite map of the human ADAR gene was determined (Fig. 1B). The precise exon-intron organization was established by sequencing plasmid subclones and by comparison of the genomic sequences to the previously determined cDNA sequences (6, 20). The 5′-region of the published ADAR cDNA (6, 20) was verified and extended by the 5′-rapid amplification of cDNA ends procedure. The human ADAR gene contains 15 exons and spans about 30 kb (Fig. 1). The AUG translation initiation site for the 1226-amino-acid ADAR protein is located in exon 1, the A of which corresponds to position 1 of the cDNA sequence. The 3′-terminal exon number 15, the largest exon, includes the UAG translation termination site and the 3′-untranslated region. The complete exon-intron boundaries of the human ADAR gene are summarized in Table I. Exons range in size from 94 to 2984 base pairs. Introns range from 127 base pairs to about 6.5 kb pairs. All splice site junctions conform to the GT—AG rule (41).

### Table I
Exon-intron sizes and junction sequences of human ADAR gene

| Exon No. | Size (bp) | Junction (cDNA Position) | Intron No. | Size (kb) | Junction (cDNA position) | Exon no. |
|----------|-----------|--------------------------|------------|-----------|--------------------------|----------|
| 1        | 202       | CCGCCAgtaagcgggcct (15)  | I          | 3.4       | cttatttctgcagGGGTAT (16) | 2        |
| 2        | 1586      | ACCTCgtaagagaccac (1601) | II         | 2.5       | ctttccgtaagATTTAA (1602) | 3        |
| 3        | 184       | GAGAGgtaagtgctcct (1785) | III        | 0.4       | cattttctctagACTGCA (1786) | 4        |
| 4        | 149       | ACACAagtatgctcctacg (1934) | IV         | 0.6       | atctccgtccagGTCCCA (1935) | 5        |
| 5        | 145       | AACCAGgtaaggcgggtttt (2079) | V          | 0.1       | atctccctctttagCTGAA (2080) | 6α       |
|          |           |                          | VI         | 6.5       | ctcatccccaaagGTTCGT (2137) | 6β       |
| 6α       | 191       | GCCCAAggtgatgtctcta (2270) | IX         | 0.7       | tttttttttacagCTCCCT (2271) | 7α       |
| 6β       | 134       | GCCCAAggtgatgtctcta (2270) | X          | 0.3       | cctttttttcalagCTCCCT (2271) | 7β       |
| 7α       | 226       | AAGACAgtaagagcctct (2496) | VII        | 0.3       | tttttttcccacagCTCCCT (2497) | 8        |
| 7β       | 148       | ACAGAggttaaccccaagtg (2418) |            |           |                           |          |
| 8        | 172       | GACACggtgatgtgagct (2668) | VIII       | 0.3       | acacctctctagGGAATC (2669) | 9        |
| 9        | 94        | CATCAGgtgacgaggtct (2762) | IX         | 0.7       | ttttttttttacagCTCCCT (2763) | 10       |
| 10       | 123       | TATCAGgtctgtacagtt (2885) | X          | 0.3       | ccttttttttacagCTCCCT (2886) | 11       |
| 11       | 134       | AGAACGgtaggtgatcaca (3019) | XI         | 1.8       | tctccacacagGGAAG (3020) | 12       |
| 12       | 183       | CATGGgtaggggccctg (3202) | XII        | 0.3       | ttttttacagGTTCACC (3203) | 13       |
| 13       | 113       | CCCAAGgtagctcatataaccc (3315) | XIII       | 0.4       | tgggattctctagGTTCGCC (3316) | 14       |
| 14       | 128       | GGATGGgtaggagacag (3443) | XIV        | 0.15      | tgggattctctagGCGACAG (3444) | 15       |
| 15       | 2984      | AATAAAAAAAAAAACAGAATCTG (6427) |            |           |                           |          |

The human ADAR gene revealed a non-random distribution. Intron phase zero before the first base of the codon occurred in 4 out of 14 introns (29%) in the protein coding region of the ADAR gene; 3 out of 14 (21%) were intron phase 1 after the first base, and 7 out of 14 (50%) were intron phase 2 after the second base of the junction codon. Comparison of cDNA and genomic sequences revealed that the three RNA-binding subdomain motifs RI, RII, and RIII of the ADAR deaminase (6, 32) are located in exons 3, 5, and 7, respectively (Fig. 1A). The RNA-dependent protein kinase PKR possesses two copies of the dsRNA-binding subdomain R-motif (24, 26). Interestingly, the codon phasing was exactly conserved between the exons that specify the R subdomains in the ADAR and PKR genes in the human, and the pkr gene in the mouse (Fig. 2). This conserved phasing would readily accommodate exon skipping involving one of the R-motif exons, while still retaining the open reading frame of the mRNA encoding the downstream enzyme catalytic region.

### Multiple Forms of ADAR Due to Alternative Splicing—Previous studies revealed that two forms of the human dsRNA-specific adenosine deaminase are likely to exist, as demonstrated by Western immunoblot analysis of nuclear extracts from cultured cells (6). This earlier finding, together with the present observation that the intron phases prior to the R-motif exons of ADAR were all phase two (Fig. 2), raised the possibil-
Splice Variants of the RNA-specific Adenosine Deaminase

Four ADAR variants are theoretically possible if two exons, exons 6 and 7, are present in two (α, β) forms (Table I). There-
 Jordi Barceló-Gonzalez: Dependency graph

Splice Variants of the RNA-specific Adenosine Deaminase

FIG. 4. Schematic structure of the three ADAR splice variants. ADAR-a indicates the 1226-amino acid prototype version of ADAR. ADAR-b is the exon 7b version lacking 26 amino acids of ADAR-a in the region between the R_{III} motif and the C domain as a result of alternative splicing between a cryptic 5'-splicing site in exon 7 and intron VII; ADAR-c is the exon 6j variant lacking 19 amino acids from exon 6 of ADAR-b in the region between the R_{III} and R_{II} motifs as a result of alternative splicing between intron V and a cryptic 3'-splicing site in exon 6.

mouse library gave only two (lane c) upon PCR analysis with the primer pair PP1 that flanked both deletions. As compared with the human kidney library, the human placenta library did not include a detectable ADAR message possessing the 57-nt deleted exon 6j (Fig. 3B, lane f). As shown in Fig. 3B, PCR products obtained with two primer pairs, PP3 (lanes g and i) and PP4 (lanes h and j), indicated that human kidney (lanes g and h) expressed three ADAR species, one with full-length exons 6a and 7a, designated as ADAR-a, one with full-length exon 6a and the 78-nt deleted exon 7b, designated as ADAR-b, and one with both the 78- and 57-nt deleted exons 6j and 7b, respectively, designated as ADAR-c. By contrast, the human placenta library only yielded the ADAR-a and -b forms (lanes i and j). The fourth theoretical variant combination, containing only the 57-nt deletion exon 6j and the full-length exon 7a (denoted ADAR-d), was not detected in the cDNA libraries examined (lanes h and j). These results demonstrated the utilization of two alternative splice sites for exon 6 and exon 7, yielding three variant forms of ADAR as summarized by the schematic diagram of Fig. 4. The variant forms of ADAR are differentially expressed in placenta and kidney tissues and macrophage cells, suggesting the possibility of a tissue-specific function for variant ADAR isoforms.

Splice Variant Versions of ADAR Are Active Enzymes—Four versions of the dsRNA-specific adenosine deaminase protein, the three naturally occurring variants ADAR-a, -b, and -c, and an engineered variant not yet detected naturally (ADAR-d), were expressed both in vitro and in vivo. The rabbit reticulocyte lysate cell-free protein synthesizing system, programmed with ADAR mRNA transcribed in vitro from the ADAR cDNA templates, was utilized initially to confirm the protein coding capacity of the constructs. The ADAR protein products synthesized in vitro showed the predicted sizes as analyzed by SDS-PAGE (data not shown). Because the full-length ADAR constructs possessing the natural 5'-untranslated region from the cDNA were expressed poorly (6, 32), a truncated 5'-terminating untranslated region was generated by PCR using the BamHI (+)44 primer; this engineered construct also included the AUG1 translation start in optimal context. ADAR-a mRNA transcribed from this construct was efficiently translated in vitro (data not shown). This ADAR construction in the pcDNA1/neo vector, was used for expression of the M1 full-length 1226-amino acid ADAR proteins in COS cells. Expression of ADAR proteins in vivo was detected by Western immunoblot analysis, using antibodies 2 and 3 generated against recombinant ADAR proteins produced in E. coli (6). Antibody 3 generated against the N-terminal region of the cDNA ORF was previously shown to specifically recognize an interferon-inducible p150 ADAR protein found in the cytoplasm and nucleus but not the p110 nuclear protein (6). This is illustrated in Fig. 5A by analysis of extracts prepared from human SY5Y cells, which were included as a positive control and reference marker; p150 was greatly enhanced in both nuclear (lanes a and b) and cytoplasmic (lanes c and d) extracts prepared from IFN-treated (Fig. 5A, lanes b and d) as compared with untreated (lanes a and c) cells. Only the p150 protein was detected with antibody 3 (lanes c and d), whereas both p150 and p110 were detected by antibody 2 (lanes a and b) as previously reported (6).

The three cDNA constructs encoding the wild-type M1 full-length deaminase variants, ADAR-a, -b and -c, were all efficiently expressed in COS-1 cells (Fig. 5A, lanes f–h). The ADAR-a, -b, and -c proteins were present in both the cytoplasmic (Fig. 5A, lanes f–h) and nuclear fractions (data not shown) of transfected COS cells. Although antibodies prepared against recombinant human ADAR react poorly with monkey COS ADAR (6), a COS cell p150-like protein was detectable in both the cytoplasmic and nuclear fractions in cells transfected with vector alone (lanes c and i). The IFN-inducible p150 protein of human SY5Y cells (lanes c and d) possessed an electrophoretic mobility comparable to the full-length p150 ADAR-a variant expressed in COS cells (lane f). Curiously, both vector-coded and endogenous chromosome-coded ADAR proteins were detected as two bands on Western immunoblots. The lower signal is believed to represent a proteolytic cleavage product (2, 6).

All four variants of the deaminase cDNA, ADAR-a, -b, -c, and -d, were also efficiently expressed in transfected COS cells as the M296 N-terminally truncated form of ADAR (Fig. 5B, lanes k–n). This is illustrated by the Western blot of the cytoplasmic fractions (Fig. 5B, lanes k–n), using antibody 2 which was previously shown to recognize both the p150 and the p110 human proteins (6) as shown also in Fig. 5A (lanes a and b). Similar to previously reported findings for human cells (6), COS cells also expressed high levels of the nuclear p110-like protein (compare lane o with j), which was comparably sized to the M296 ADAR-a variant (lane k) as shown in Fig. 5B.

Because the nuclear fractions prepared from COS cells exhibited high deaminase activity due to the presence of the
endogenous M296-like p110 protein (3, 6, 42), the cytoplasmic fractions prepared from transfected cells were employed to examine the enzymic activity of the cDNA-encoded ADAR splice variants. All three of the full-length wild-type versions of the recombinant ADAR variants, ADAR-a, -b and -c, and all four N-terminally truncated M-296 wild-type versions of the ADAR variants, ADAR-a, -b, -c and -d, possessed dsRNA adenosine deaminase activity (Fig. 6, b–h). By contrast, the cytoplasmic fraction prepared from control COS cells transfected with the expression vector alone without an ADAR cDNA insert showed much lower endogenous deaminase activity (Fig. 6, lane a). Relative specific enzyme activities were calculated for the ADAR variants by quantifying the extent of A-to-I conversion relative to the amount of expressed ADAR protein (Fig. 7). The wild-type ADAR-a, -b, and -c variants of the full-length form (Fig. 7A), as well as the wild-type ADAR-a, -b, -c, and -d variants of the M296 form (Fig. 7B), displayed comparable specific deaminase activity measured with a synthetic dsRNA substrate. These results indicate that the deletions in exons 6 and 7 resulting from alternative splicing caused no discernible effect on the enzyme activity of the wild-type ADAR.

Functionally Distinct R-motif Subdomains Associated with Exon 6 and 7 Variants Resulting from Alternative Splicing—Previously we introduced an equivalent mutation into each of the three repeated R-motif subdomains of ADAR in the M296 form by substituting a highly conserved R-core lysine residue that had been shown by mutagenesis studies of PKR and E3L to be essential for RNA binding activity (29–33). Analysis of R-motif mutants of ADAR revealed that the three R-motif subdomains are functionally distinct (32, 34); RIII was the most important for enzyme function, whereas the single substitution mutants at either the R1 or R2 subdomain retained significant enzyme activity.

In order to investigate the biochemical properties of the two newly identified alternative splice variants ADAR-b and -c relative to the originally identified 1226 ADAR-a protein (6, 20, 21), the same Lys to Glu substitution (see Fig. 2 schematic) was introduced into each of the R-motifs of the three versions of ADAR, utilizing the PCR-based mutagenesis strategy we previously described (32). All three of the single R mutants, and the triple R mutant for ADAR-a, B, comparison of enzyme activities of ADAR proteins expressed in the truncated M296 form, including ADAR-d, whose expression was not detectable at the transcription level. Normalization was done with ADAR-a as the standard.

FIG. 6. Functional analysis of enzyme activity of ADAR proteins expressed in COS cells. Autoradiogram shows enzyme activity of ADAR proteins expressed in vivo using the cytoplasmic fractions from COS-1 cells transfected with expression constructions encoding ADAR in the full-length form (including versions ADAR-a, b, and c) and in the N-terminally truncated M296 form (including versions ADAR-a, b, c, and d). 32P-Labeled dsRNA substrate was incubated with equivalent amounts of cytoplasmic extracts under the standard assay conditions as described under “Experimental Procedures.” Following subsequent P1 nuclease digestion, the labeled nucleotides were analyzed by thin layer chromatography. The positions to which the adenosine (AMP) and inosine (IMP) 5′-nucleoside monophosphate standards migrated, as well as the origin, are indicated.

FIG. 7. Comparison of enzyme activities of expressed ADAR proteins. Relative specific enzyme activities were calculated based on the percentage of A-to-I conversion catalyzed by expressed ADAR proteins that were quantitated from Western immunoblot analyses. A, comparison of enzyme activities of ADAR proteins expressed in the full-length form, including the wild-type (WT) for the three isoforms (ADAR-a, -b, and -c), and the three single R mutants and the triple R mutant for ADAR-a. All the relative specific activities were normalized to that of the wild-type ADAR-a. B, comparison of enzyme activities of wild-type ADAR proteins expressed in the truncated M296 form, including ADAR-d, whose expression was not detectable at the transcription level. Normalization was done with ADAR-a as the standard.
comparison of R-motif subdomain mutants of the ADAR-a, -b, and -c variants in the M296 form revealed distinctly different profiles of enzyme activity. No substantial difference was observed for ADAR-a in the M296 form among the three single R-motif mutants and the triple R mutant, in the N-terminally truncated M296 form. The activities of R mutants for each of the three ADAR isoforms (ADAR-a, -b, and -c) were respectively normalized to that of the wild type, since the wild-type proteins possessed comparable specific enzyme activities as exhibited in Fig. 7B.

We have isolated and characterized genomic clones of human ADAR, the interferon-inducible dsRNA-specific adenosine deaminase also known as DRADA and dsRAD (6). Three important points emerge from the results reported herein. First, and most important, is that previously unknown splice site variants of the ADAR editing enzyme were identified that are expressed in human and mouse tissues and cell lines; the two ADAR variants result from alternative splicing of exons 6 and 7. Second, that the variants encode active editing enzymes in which the three RNA-binding R-motif subdomains are functionally distinct. Third, that the ADAR gene structure involving 15 coding exons possesses a precisely maintained codon phasing for the three R-motif exons of human ADAR that is exactly conserved in the human and mouse genes of the dsRNA-dependent protein kinase PKR that possesses two R-motif exons.

We earlier reported the isolation of genomic clones for the ADAR deaminase, \( \lambda_{6c} \) and \( \lambda_{151} \), which were mapped to a single locus on human chromosome 1q21.1–21.2 by fluorescence in situ hybridization (22). We now have isolated additional overlapping \( \lambda \) phage clones as well as P1 phage genomic clones for human ADAR. The exon-intron organization of the 30-kb human ADAR gene was defined by Southern blot and direct sequence analyses. The exon structure for the ADAR gene that we determined is in general agreement with that described by Nishikura and co-workers (43), although small apparent differences exist between some intron sizes. Comparison of the genomic sequence with the cDNA sequence revealed that the coding region of the human ADAR gene is specified by 15 exons. The intron positions within the codonts of the human ADAR gene were not evenly distributed; 50% were intron phase 2 after the second base of the junction codon (Table 1). By contrast, results obtained for a large set of animal genes revealed that phase two was represented by only 22% of the total (44). Curiously, the intron phases prior to the three R-motif-containing exons of the human ADAR gene, exons 3, 5, and 7, were all phase two. Even more striking is the finding that this codon phasing observed for the human ADAR gene R-motif exons was conserved exactly for the two R-motif exons of the mouse \( \text{pkr} \) (39) and human (45) PKR genes (Fig. 2). This conserved phasing is consistent with a common ancestor for the R-motif coding exons of the ADAR and PKR genes and the occurrence of early exon shuffling events. The conserved phasing would readily accommodate exon skipping involving R-motif exons, while still retaining the ORF of the mRNA encoding the downstream catalytic domain of the deaminase in the case of ADAR, or the kinase in the case of PKR. However, PCR analysis did not detect the existence of PKR splice variants lacking one or both R-motif exons (45). Although our PCR analyses of human placenta and kidney and mouse macrophage cDNA libraries likewise did not reveal the presence of ADAR splice variants lacking one or more of the R-motif exons as predicted for such an exon-skipping event, we did identify two other different kinds of ADAR splice variants.

One of the splice variants that we identified, ADAR-b, possessed a 26-amino acid deletion in exon 7; the other splice variant that we identified, ADAR-c, contained an additional 19-amino acid deletion in exon 6. The 26-amino acid deletion of ADAR-b was located between the RI subdomain and the C-terminal catalytic domain; the 19-amino acid deletion of ADAR-c was located between the RI and RII subdomains. In both the ADAR-b and -c variants, the integrity of all three R-motifs was retained. What was altered was the spacing between the R-motifs themselves or between the R-motifs and the catalytic domain of the variant editing enzyme forms. The prototype ADAR, now designated as ADAR-a, corresponds to the form of the editing enzyme for which cDNA clones were isolated (6, 20, 21). The human ADAR-a cDNA predicts a 1226-amino acid protein that is inducible by IFN (5, 6). Three lines of evidence are consistent with the conclusion that the newly discovered ADAR variants are indeed splicing variants and not products of separate genes.
First, the positions of the deletions in exons 6 and 7 identified for the ADAR-b and ADAR-c variants coincide exactly with the previously determined exon-intron junctions at one terminus and conform at the other terminus to the GT—AG rule (41) for splice sites. Second, the deduced amino acid sequence of the ADAR-b and -c variant cDNA clones agrees exactly to the amino acid sequence previously determined for the prototype ADAR-a cDNA clones (6, 20). Third, PCR analyses of genomic DNA, either of the λ library, the P1 genomic clones, or high molecular weight chromosomal DNA isolated from human U cells, yields only a single PCR fragment of predicted size and does not yield smaller products that would be expected if the deletions found in cDNA clones also existed in the genome (data not shown).

Recent studies indicate that the dsRNA-specific adenosine deaminase deaminase ADAR plays a central role in the A-to-I RNA editing process (10, 46). However, how the site selectivity is achieved for the precise editing required in the cases of some of the natural cellular and viral RNA substrates is an important issue that remains unresolved. The highly conserved RNA-binding subdomain R present in triplicate in ADAR is found in duplicate in the interferon-inducible RNA-dependent protein kinase PKR (24–28, 47). The fact that no RNA-binding sequence specificity has been identified for such R-motif subdomains in the case of the PKR kinase (24, 48) raises the question how ADAR catalyzes the site-selective deamination of adenosine in natural substrates, for example in Glur pre-mRNA (12, 15, 16, 49) and HDV antigenic RNA (11, 14). The existence of multiple forms of the dsRNA adenosine deaminase would provide one possible mechanism by which the specificity problem could be resolved (6, 23). Purified ADAR (dsRAD) from Xenopus oocyte nuclei will catalyze the deamination of Glur-B RNA in vitro, although the deamination of the biologically relevant Q/R ed-nuclei will catalyze the deamination of GluR-B RNA resolved (6, 23). Purified ADAR (dsRAD) from HDV antigenic RNA (11, 14). The existence of multiple catalyzes site-selective deamination of adenine in natural case of the PKR kinase (24, 48) raises the question how ADAR catalyzes the site-selective deamination of adenosine in natural substrates, for example in Glur pre-mRNA (12, 15, 16, 49) and HDV antigenic RNA (11, 14). The existence of multiple forms of the dsRNA adenosine deaminase would provide one possible mechanism by which the specificity problem could be resolved (6, 23). Purified ADAR (dsRAD) from Xenopus oocyte nuclei will catalyze the deamination of Glur-B RNA in vitro, although the deamination of the biologically relevant Q/R editing site occurs much less frequently in vitro by the Xenopus enzyme than observed in vivo (15). One possibility is the absence of accessory factors that may play indirect roles in vivo; another possibility is the presence of multiple closely related variants of the editing enzyme in the purified enzyme preparation, only one of which is relevant for in vivo editing. Although dsRNA-specific adenosine deaminase enzymes have been purified extensively and characterized from bovine liver (2), calf thymus (3), and chicken lung (50) in addition to Xenopus oocytes (1), the apparent size of the enzymes varies considerably among the four sources, and the purified preparations often contain multiple bands as determined by gel electrophoresis.

ADAR has been shown to be present ubiquitously in primary tissues and cell lines by Northern analysis or direct enzyme assay (6, 20, 21, 42). However, our results establish that at least one alternative splice variant, ADAR-c, was expressed differentially in different tissues (Fig. 3). Both ADAR-a and -b were expressed in human kidney and placenta, as well as in a mouse macrophage-like cell line. By contrast, ADAR-c was detected only in the human kidney library from the three cDNA libraries examined. All of the ADAR variants were efficiently expressed in COS cells (Fig. 5). The various splice variants of ADAR were active enzymes, with comparable specific activity when examined as the wild-type version of the three R-motif subdomains (Figs. 6 and 7).

As an approach to examining the functional importance of each of the three copies of the R-motif subdomains, we constructed a family of ADAR point mutants in which each of the R-motifs was mutated in each of the splice variants. A highly optimized lysine residue (Fig. 2) established as essential for the R-motif mediated RNA-binding activity of the cellular PKR kinase and the viral E3L protein (29, 30, 33) was substituted with glutamic acid in either any one, or all three, of the R copies in the ADAR splice variants. For ADAR-a, either in its full-length p150 form or the truncated M296 form, mutation of any of the three R-motif subdomains reduced but did not abolish enzyme activity (Figs. 7A and 8). By contrast, the ADAR-a triple RRIIRIII mutant showed no detectable deaminase activity either as the full-length p150 protein or the M296 protein relative to the respective wild-type ADAR-a prototype parent. This suggests that each of the three R-motif is comparably important for enzyme function in the case of ADAR-a, and the N-terminal portion of ADAR-a upstream of M296 and the R-motif has little effect on the function of the dsRNA-binding domain.

Results obtained for the R2 mutant of variants ADAR-a, -b, and -c suggest that functionally distinct R-motif subdomains result from the alternative splicing of exon 6 and exon 7 (Fig. 8). When R2 was mutated, the 26-amino acid deletion (exon 7β form) between the RII1 motif and the C domain in ADAR-b, as well as the additional 19-amino acid deletion (exon 6β form) between the RII1 and RII1 motifs in ADAR-c, appeared to gradually reduce the functional significance of R2. By contrast, exons 6β and exon 7β seemed to gradually enhance the functional importance of the RII1 subdomain, as revealed by the relative decrease in enzyme activity for the RII1 mutants of the ADAR-b and -c variants relative to ADAR-a containing exons 6a and 7a. The most dramatic alteration was observed for the R1 mutants; the presence of the exon 7β rendered the R1 motif dispensable, as demonstrated by the higher enzyme activities for ADAR-b and -c variant R1 mutants relative to the corresponding wild-type parents, and also to ADAR-a. These results suggest that the presence of deletions flanking the RII1 motif due to alternative splicing, especially the one located between the RII1 motif and the C domain, leads to distinct functionality of each of the three R-motifs. This implies that shortening the distance between the dsRNA substrate binding domain and the catalytic domain, as well as the distance within the three R subdomains, can consequently change the manner by which the three repeated R-motifs and the C domain interact, thus altering the way by which dsRNA substrates are bound and correctly positioned at the catalytic center for deamination. These results are consistent with the notion that the splice variants of ADAR likely contribute to mediating the required site selectivity in editing of natural substrates.

Although preferential deamination of certain adenosine residues has been demonstrated in vitro using synthetic RNA substrates (4, 51), the basis of the site selectivity presumed to occur in vivo as illustrated by the postulated modification of hepatitis delta virus and human immunodeficiency virus RNAs (13, 14, 52) and also the cellular mRNAs encoding the gluta-mate-activated cation channel proteins where one or two adenosine residues are modified (12, 19) has not been resolved. In support of the speculation that the repeated nature of the R-motif subdomain found within ADAR may be important in determining the relative substrate selectivity of the enzyme, there exist multiple forms of ADAR with distinct positioning of the three R-motifs relative to the C domain that showed distinct involvement of each individual R-motif in the process of deamination. They are most likely to represent ADAR enzymes differing in substrate specificity, which are involved in recognizing distinct natural dsRNA substrates via interacting with dsRNA region in distinct structural environment. A similar situation has been described for the Wilms tumor susceptibility gene WT1, where the alternative splicing between the third and fourth zinc fingers results in an insertion of KTS that is of great physiological importance (53, 54). This insertion of three amino acids altered the binding properties of the WT1
isoforms for their DNA targets (55) as well as their subnuclear localization (56).

The demonstration of multiple splice variants of ADAR with functionally distinct R-motif subdomains, as shown herein, when taken together with the reported rat RED1 dsRNA adenosine deaminase presumably encoded by a different gene than the prototype rat ADAR deaminase (23), clearly establishes that there is indeed a family of dsRNA adenosine deaminase enzymes. These ADAR enzymes undoubtedly act to recognize a variety of natural viral and cellular RNA substrates with distinct structures. What is not yet clear is whether a single deaminase protein is sufficient to confer the necessary site selectivity required in the editing reaction, or whether the process is more complex at the level of structural organization. For example, whether an accessory protein interacts with either the substrate RNA or the deaminase to help confer the required editing specificity, or whether an editing complex formed by interactions among a family of ADAR deaminases and accessory proteins contributes to the specificity (49), remains to be established.

REFERENCES

1. Hough, R. F., and Bass, B. L. (1994) J. Biol. Chem. 269, 9903–9909
2. Kim, U., Garner, T. L., Sanford, T., Speicher, D., Murray, J. M., and Nishikura, K. (1994) J. Biol. Chem. 269, 13480–13489
3. O’Connell, M. A., and Keller, W. (1994) Proc. Natl. Acad. Sci. U.S.A. 91, 10596–10600
4. Nishikura, K., Yoo, C., Kim, U., Murray, J. M., Estes, P. A., Cash, F. E., and Seeburg, P. H. (1994) Cell 75, 1361–1370
5. Yang, J.-H., Sklar, P., Axel, R., and Maniatis, T. (1995) Mol. Cell. Biol. 15, 5701–5711
6. Higuchi, M., Single, F. N., Kohler, M., Sommer, B., Sprengel, R., and Seeburg, P. H. (1995) Cell 75, 1361–1370
7. Higuchi, M., Single, F. N., Kohler, M., Sommer, B., Sprengel, R., and Seeburg, P. H. (1995) Cell 75, 1361–1370
8. Wang, Y., Zeng, Y., Sanford, T., Zeng, Y., and Nishikura, K. (1995) Proc. Natl. Acad. Sci. U.S.A. 92, 11457–11461
9. O’Connell, M. A., Krause, S., Higuchi, M., Haas, J. J., Trotty, N. F., Jenny, A., and Keller, W. (1995) Mol. Cell. Biol. 15, 1389–1397
10. Weier, H.-U. G., George, C. X., Greulich, K. M., and Samuel, C. E. (1995) Genomics 30, 372–375
11. Melcher, T., Maan, S., Herb, A., Sprengel, R., Seeburg, P. H., and Higuchi, M. (1996) Nature 379, 460–464
12. Samuel, C. E. (1993) J. Biol. Chem. 268, 7603–7606
13. McMillan, N. A. J., Carpick, B. W., Hollis, B., Toone, W. M., Zamanian-Daryush, M., and Williams, B. R. G. (1995) J. Biol. Chem. 270, 2661–2666
14. St. Johnston, D., Brown, N. H., Gall, J. G., and Jantsch, M. (1992) Proc. Natl. Acad. Sci. U.S.A. 89, 10979–10983
15. Kharrat, A., Maestas, M. J., Gibson, T. J., Nilges, M., and Pastore, A. (1995) EMBO J. 14, 3572–3584
16. Drummond, I. A., Rupprecht, H. D., Rohwer-Nutter, P., Lopez-Guisa, J. M., and Keller, W. (1996) Mol. Cell. Biol. 15, 358–364
17. McCormack, S. J., Ortega, L. G., Doohan, J. P., and Samuel, C. E. (1994) Virology 198, 92–99
18. Hollmann, M., and Heinemann, S. (1994) Annu. Rev. Biochem. 63, 77–81
19. Polson, A. G., Bass, B. L., and Casey, J. L. (1996) Mol. Cell. Biol. 15, 5376–5388
20. Casey, J. L., and Gerin, J. L. (1995) Science 269, 5686–5590