A phylogenetic analysis reveals an unusual sequence conservation within introns involved in RNA editing

P. JOE ARUSCAVAGE and BRENDA L. BASS

Department of Biochemistry/Howard Hughes Medical Institute, University of Utah, Salt Lake City, Utah 84132, USA

ABSTRACT

Adenosine deaminases that act on RNA (ADARs) are RNA editing enzymes that convert adenosines to inosines within cellular and viral RNAs. Certain glutamate receptor (gluR) pre-mRNAs are substrates for the enzymes *in vivo*. For example, at the R/G editing site of gluR-B, -C, and -D RNAs, ADARs change an arginine codon (AGA) to a glycine codon (IGA) so that two protein isoforms can be synthesized from a single encoded mRNA; the highly related gluR-A sequence is not edited at this site. To gain insight into what features of an RNA substrate are important for accurate and efficient editing by an ADAR, we performed a phylogenetic analysis of sequences required for editing at the R/G site. We observed highly conserved sequences that were shared by gluR-B, -C, and -D, but absent from gluR-A. Surprisingly, in contrast to results obtained in phylogenetic analyses of tRNA and rRNA, it was the bases in paired, helical regions whose identity was conserved, whereas bases in nonhelical regions varied, but maintained their nonhelical state. We speculate this pattern in part reflects constraints imposed by ADAR's unique specificity and gained support for our hypotheses with mutagenesis studies. Unexpectedly, we observed that some of the gluR introns were conserved beyond the sequences required for editing. The ~600-nt intron 13 of gluR-C was particularly remarkable, showing >94% nucleotide identity between human and chicken, organisms estimated to have diverged 310 million years ago.

Keywords: ADAR; adenosine deaminase; evolution; glutamate receptor; inosine

INTRODUCTION

Adenosine deaminases that act on RNA (ADARs) are RNA editing enzymes that change adenosines to inosines within double-stranded RNA (dsRNA; reviewed in Bass, 1997; O'Connell, 1997). The enzymes are found in all metazoa, and at least two distinct ADARs, ADAR1 and ADAR2, have been characterized (as cited in Bass et al., 1997). One function of ADARs is to deaminate adenosines in codons of mRNAs so that multiple protein isoforms can be synthesized from a single encoded sequence. ADARs are known to catalyze functionally important codon changes in the antigenomic RNA of hepatitis delta virus (HDV; Polson et al., 1996), and the mRNAs for glutamate receptors (gluRs; Egebjerg & Heinemann, 1993; Higuchi et al., 1993; Lomeli et al., 1994) and serotonin receptors (Burns et al., 1997). For the latter two cellular transcripts, editing occurs on the pre-mRNA, and the requisite double-stranded structure is

formed by base pairing of exon sequences with nearby introns. At present these substrates are the most well-characterized, but the amount of inosine detected within poly A⁺ RNA suggests many substrates are yet to be discovered (Paul & Bass, 1998).

When ADARs act within open reading frames (ORFs) they must target specific adenosines to produce a protein isoform of precise function. How this specificity is achieved is not well understood, but is thought to depend on slight sequence preferences (Polson & Bass, 1994), as well as the structure of the RNA substrate (Lehmann & Bass, 1999; reviewed in Bass, 1997). Comparisons of homologous RNA molecules from different organisms, or phylogenetic analyses, are one of the best methods for elucidating functionally important sequences and structures in an RNA molecule (reviewed in Pace et al., 1999). Sequences or structures that are conserved between distantly related organisms are likely to be important or essential for the function of the molecule. In hopes of increasing our understanding of what sequences and structural features are important for ADAR specificity, we performed a phylogenetic analysis of an ADAR substrate.

Reprint requests to: Brenda L. Bass, Department of Biochemistry/HHMI, University of Utah, Salt Lake City, Utah 84132, USA; e-mail: bbass@howard.genetics.utah.edu.

We chose to analyze sequences required for editing at the R/G site of gluR pre-mRNAs. As its name implies, this editing site changes an arginine codon (AGA) to a glycine codon (IGA) and results in protein isoforms that alter the kinetic properties of ion channels assembled from the gluR subunits (Lomeli et al., 1994). Ionotropic glutamate receptors are classified according to their preferred agonists, and those responsive to α -amino-3-hydroxy-5-methyl-4-isoxazole propionate, or AMPA, are gluR-A, gluR-B, gluR-C, and gluR-D (reviewed in Hollmann & Heinemann, 1994). Three of the four AMPA receptors, gluR-B, -C, and -D, are edited at the R/G site, whereas the fourth AMPA receptor, gluR-A, is not (reviewed in Seeburg et al., 1998). As a group, the AMPA receptors share ~70% amino acid and nucleotide identity and can be easily aligned for sequence comparisons. We reasoned that a phylogenetic analysis of the AMPA receptors might reveal sequences important for editing, and that these sequences would be shared by gluR-B, -C, and -D, but absent from gluR-A.

Here we report that, indeed, we observed highly conserved sequences that were unique to gluR-B, -C, and -D. Notably, and in contrast to results obtained previously for other RNAs, it was the bases in paired, helical regions whose identity was conserved, whereas bases in nonhelical regions varied, but maintained their nonhelical state. This pattern most likely reflects constraints imposed by ADAR's unique specificity, and we speculate it may also exist in other RNAs, especially in the substrates of other dsRNA-binding proteins (dsRBPs). Unexpectedly, and quite remarkably, we observed that some of the gluR introns were conserved beyond the sequences required for editing. The ~600-nt intron 13 of gluR-C was particularly remarkable, showing >94% nucleotide identity between human and chicken, organisms estimated to have diverged 310 million years ago (Kumar & Hedges, 1998).

RESULTS

Editing at the R/G site of the AMPA gluRs is known to require base pairing between exon 13 and the downstream intron (intron 13; Lomeli et al., 1994), so we began our study by attempting to clone genomic DNA spanning exon 13 to exon 14. The PCR was used to amplify genomic DNA, and primers within exon sequences were designed based on existing cDNA sequences (see Materials and Methods). AMPA gluRs undergo alternative splicing in this region to produce flop (exon 13–exon 14) and flip (exon 13–exon 15) transcripts (Sommer et al., 1990). When a flop version of the cDNA existed in the database, primers were designed to amplify genomic sequence from exon 13 to exon 14. If only a flip version of the cDNA existed, primers were designed based on exon 13 and exon 15 sequences, and genomic DNA extending from exon 13 all the way to exon 15 was amplified.

Figure 1 lists the genomic sequences we successfully cloned, and their general features. We obtained PCR products that included all of intron 13 for each gluR sequence except gluR-A of rat and mouse. The gluR-A intron 13 appears to be quite long, and we were never able to amplify the entire intron. However, based on published sequences of primers used to clone regions internal to intron 13 (Lomeli et al., 1994), we were able to amplify the sequences relevant to our study. Extreme care was taken to avoid contamination during the PCR, and despite long stretches of sequence identity among clones, all clones contained unique exonic sequences corresponding to the appropriate cDNA.

Figure 2A shows an alignment of sequences spanning the exon 13/intron 13 junction for all of the genomic sequences we obtained; the general features of this region continued throughout the rest of the intron. As expected, the exonic sequences (boxed) were highly conserved (blue nucleotides, identity) with most of the variability (red) occurring at codon wobble positions. In contrast, but consistent with commonly accepted dogma, intronic sequences were not conserved between all members of the AMPA family. However, when we excluded the gluR-A sequences, which are not edited and thus should not have sequences required for editing, a region of strong conservation was observed in the 5' end of intron 13 (Fig. 2B).

Organism	gluR	5' ss	intron length (total)	3' ss	accession #
Rat	A	CCGCCCTGAG	GTAAG---668+?-----	AAATCCAGTA	AF201341
Mouse	A	CCGCCCTGAG	GTAAG---685+?-----		AF201342
Human	B	CCTCATTAAG	GTGGG---868---TTCAAG	AAATGCGGTT	AF201343
Rat	B	CCTCATTAAG	GTGGG---849---TTCAAG	AAATGCGGTT	AF201344
Mouse	B	CCTCATTAAG	GTGGG---863---TTCAAG	AAATGCGGTT	AF201345
Chicken	B	CCTCATTAAG	GTGGG---900---TTCAAG	AAATGCGGTT	AF201346
Pigeon	B	CCTCATTAAG	GTGGG---919---TTCAAG	AAATGCGGTT	AF201347
Tilapia	B	CCTCATTAAG	GTGGG---1322---TTCAAG	AAATGCGGTT	AF201348
Human	C	CAGCATTAAG	GTGGG---561---TTCAAG	AAATGCTGTT	AF201349
Rat	C	CAGCATTAAG	GTGGG---556---TTCAAG	AAATGCTGTT	AF201350
Mouse	C	CAGCATTAAG	GTGGG---556---TTCAAG	AAATGCTGTT	AF201351
Chicken	C	CAGCATTAAG	GTGGG---564---TTCAAG	AAATGCTGTT	AF201352
Pigeon	C	CAGCATTAAG	GTGGG---559---TTCAAG	AAATGCTGTT	AF201353
Tilapia	D	CTGCATTAAG	GTGGG---734---TTCAAG	AAATGCTGTT	AF201354

FIGURE 1. Overall features of genomic sequences from human (*Homo sapiens*), rat (*Rattus norvegicus*), mouse (*Mus musculus*), chicken (*Gallus gallus*), pigeon (*Columba livia*), and tilapia (*Oreochromis niloticus*, a fish). Except for the tilapia gluR-D sequence, both strands of three or more clones from each PCR product were sequenced, and a consensus sequence corresponding to the sequence found in the majority of the clones was compiled (see accession numbers). The tilapia gluR-D consensus derived from only two clones, but these were identical. The figure shows consensus sequence for the exon-13/intron-13 junction (5' ss) and the intron-13/exon-14 junction (3' ss). If the entire intron was cloned, its complete length is indicated. Except for the gluR-A sequences, the exon sequences shown derive from our own genomic clones and match existing cDNA sequences. For rat gluR-A, exonic sequence at the 3' splice site derives from the existing cDNA sequence; the corresponding sequence is not shown for mouse gluR-A, as only the flip sequence has been reported, and as discussed, we were unable to clone the complete intron for either of the gluR-A sequences (+ ?). Note that cDNA sequences were not available for mouse gluR-C, and this sequence was amplified using primers based on the rat sequence. cDNA sequence was also unavailable for tilapia gluR-D, but this sequence was cloned serendipitously during PCR amplification of the tilapia gluR-B gene.

Editing at the R/G site is thought to occur within a short RNA hairpin, and mutagenesis studies of the *gluR-B* sequence indicate the hairpin exists and is important for editing (Lomeli et al., 1994; Yang et al., 1997). The regions of strong conservation observed in our phylogenetic analysis were confined to this hairpin, suggesting sequences most important for editing are confined to this region. When we superimposed our phylogenetic data on the hairpin structure, an intriguing pattern was observed (Fig. 2C). In complete contrast to what is usually observed in phylogenetic studies of RNA molecules (e.g., see Rich & RajBhandary, 1976; Noller & Woese, 1981; Gutell et al., 1994), it was the bases predicted to be involved in canonical base pairing that were most conserved, whereas those shown as unpaired in the secondary structure varied between organisms (Fig. 2C, red dots). In most cases the nucleotides that varied between organisms did not appear to vary randomly, but so as to maintain the structural features of the hairpin. For example, the loop varied so as to maintain a loop, and the unpaired nucleotides at the base of the stem varied so as to maintain their unpaired state (Fig. 3).

The identity of bases at mismatch/pair 1 and 2 (see labels on Fig. 2C) varied between different organisms (Table 1). Interestingly, although we expected both of these pairs to be mismatches, as is found in the *gluR-B* sequence, in some cases AU and GU wobble pairs were observed. Mismatch/pair 1 was an AC in all *gluR-B* sequences, an AU in all *gluR-C* sequences, and a GU in *tilapia gluR-D*; mismatch/pair 2 was always an R•R pair (GG, AG, AA) except for the *tilapia gluR-D* sequence, which had a UG pair at this site.

Do biochemical and mutagenesis studies support the phylogenetic data?

Usually phylogenetic analyses provide support for functionally important helices by identifying bases that covary. For example, the observation that base *x* is an A when base *y* is a U, but when base *x* is a G, base *y* is a C, supports the existence of a canonical base pair between *x* and *y*. Using such conventional criteria, our phylogenetic analysis did not provide support for the R/G hairpin structure because the identity of all paired bases was absolutely conserved. For this reason we felt it was important to confirm the hairpin structure. We used structure-specific ribonucleases to probe the structure of a 79-nt RNA that spanned the predicted hairpin (Fig. 4); this sequence is known to be accurately and efficiently edited *in vitro* (see below, Fig. 5). As ribonucleases are rather large molecules, they cannot provide high-resolution details about a structure, but our analysis conclusively verified that the sequence forms an RNA hairpin. Sequences shown as base paired in Figure 2C were resistant to cleavage with the single-strand specific ribonucleases T2 and T1, but were

cleaved with nuclease V1, which preferentially cleaves helical sequences. The T2 and T1 cleavages that did occur were within the hairpin loop, or at nucleotides immediately adjacent to the loop, which may be more accessible to the nucleases because of their proximity to the loop.

We suspected our phylogenetic data did support the existence of the hairpin, but according to rules based on the substrate requirements of double-stranded RNA binding proteins (dsRBPs), such as ADARs, rather than the more conventional rules derived from studies of rRNAs and tRNAs. All evidence to date suggests that dsRBPs lack sequence specificity, but have a strong preference for A-form helical RNA. In our favorite model, ADARs do not discriminate between substrates at the level of sequence, but rather, according to the unique pattern of paired and unpaired sequences that characterize the substrate, that is, the position of the loops and mismatches within the RNA helix (reviewed in Bass, 1997). The pattern of sequence conservation we observed in our phylogenetic analysis is consistent with this hypothesis, as for the most part, the position of the loops and mismatches were conserved between all organisms. Although the strict conservation of base-pair identity we observed in our analysis may seem surprising, it is entirely consistent with the idea that the precise number and position of mismatches is essential for accurate and efficient editing *in vivo*. For a compensatory change to occur *in vivo*, an RNA molecule must pass through a mismatched state. If additional mismatches within the hairpin altered editing efficiency so as to alter viability, base-pair identity would be strictly conserved.

Although substrate discrimination by ADARs is probably more complex than this (see Discussion), the latter idea predicted that compensatory mutations might actually have little effect on R/G site editing, if the mutations were made artificially so the molecule did not have to pass through a mismatched state. On the other hand, the model predicted changing base pairs to mismatches would have a more dramatic effect on R/G site editing. To test this idea we synthesized a number of mutant R/G hairpins, and using an *in vitro* assay, compared the amount of R/G site editing between the mutant and wild-type sequences (Fig. 5). As predicted, most of the mutations that changed one base pair to another (i.e., compensatory mutations; Fig. 5A) had little effect on the efficiency of editing at the R/G site. A few compensatory mutations did alter editing at the R/G site by a small but statistically significant amount; importantly, such slight changes could have significant consequences *in vivo* where even minor changes in editing efficiency have been correlated with lethality (Brusa et al., 1995). In contrast, but as predicted, all but two of the mutations that changed a base pair to a mismatch altered the amount of editing at the R/G site; notably, one of the mutations that did not alter editing

was a GU-to-CC change, whereas the other created a new mismatch immediately adjacent to the hairpin loop. For the mutations that decreased editing, two caused a dramatic decrease, whereas the others decreased editing only slightly (Fig. 5B). As part of our mutagenesis studies we also made a number of changes at the positions of mismatch/pair 1 and 2 (Fig. 5C). Consistent with the phylogenetic data, results with these mutants suggested ADARs might have specific requirements at these nucleotide positions (see Discussion).

It is important to emphasize that these mutagenesis studies are only the first step in understanding the im-

portance of the R/G hairpin sequence to RNA editing. Our assay only monitored editing at the R/G site, so we do not know which mutations created new deamination sites that might be deleterious in vivo. Further, our assays monitored R/G site editing at complete reaction, and thus we cannot evaluate differences in rate that may be significant in vivo. We note that the experiments shown in Figure 5 were performed using the native *Xenopus* ADAR1 enzyme with the human R/G hairpin sequence. However, when evaluated using the hepatitis delta virus antigenomic RNA, the *Xenopus* enzyme shows a specificity identical to that observed

FIGURE 2. (Legend on facing page.).

in vivo in human liver cells (Polson et al., 1996). Regardless, we have recently purified recombinant human ADAR1 (H. Ley & B.L. Bass, unpubl. results) and used this enzyme to assay four of our mutants (data not shown). We found that three mutants affect R/G site editing similarly, whether assayed by the human or the frog enzyme. For example, using human ADAR1 and conditions identical to those used for the assays of Figure 5, mutant S2#14 showed wild-type levels of deamination ($99 \pm 9\%$), whereas mutant H23m ($57 \pm 4\%$) and 50m ($74 \pm 3\%$) showed significant decreases in editing at the R/G site. Using the human enzyme the mutations found

in S2#15 also decreased R/G site editing ($82 \pm 6\%$), but to a lesser degree than observed with the frog enzyme ($20 \pm 7\%$). The latter result is consistent with the idea that mismatch/pair 1 and 2 are used to fine-tune editing efficiency and vary according to the requirements of different organisms and gluRs (see Discussion).

Intron 13 of gluR-C is highly conserved throughout its length

For both gluR-B and gluR-C, we were able to clone sequences from multiple organisms. Thus, we also

FIGURE 2. Sequence identities among AMPA-type glutamate-receptor genes. Consensus sequences were aligned using GCG software (Wisconsin Package Version 10.0, Genetics Computer Group (GCG), Madison, Wisconsin). PILEUP was used to create the alignment (gapweight = 5; gaplength weights = 1), and the program PRETTY was used to create the figure shown. Nucleotides found to be identical among all organisms analyzed are shown in capital, blue letters, while those that varied in one or more organisms are shown in small, red letters. The general features of the sequence shown were observed throughout the remaining intron and exon sequences. A position where a consensus was not reached among four clones of rat gluR-A is indicated with a k (nucleotide number 57; two clones were G and two T). Codon triplets for a portion of the exon are indicated. **A:** Alignments that included all AMPA-type receptors (gluR-A, -B, -C, -D) reveal strong conservation of coding sequences (boxed) and exon/intron junctions, but little conservation of intronic sequences. **B:** In contrast, gluRs known to be edited at the R/G site (gluR-B, -C, -D) share a strongly conserved region that extends into the intron. **C:** The region of strong conservation is shown in the context of the R/G hairpin with positions conserved in gluR-B, -C, and -D shown in blue and positions that vary in at least one organism shown in red. The 5' splice site of intron 13, the R/G site, and two positions that vary between mismatches and pairs are indicated. Asterisks mark nucleotides missing in the gluR-C sequence.

FIGURE 3. Comparison of the hairpin loop and unpaired sequences that flank the stem of the R/G hairpin. Sequence identities from Figure 2 are shown in dark blue, and subgroups that share additional identities are indicated. **A:** The loop sequence shared by human, chicken, and pigeon gluR-B is shown in black, with sequences that differ in other subgroups indicated with various colors. The length of the nonconserved sequences in the hairpin loop suggest gluR-B and gluR-D have a 5-nt loop, whereas those of gluR-C have a 4-nt loop, or possibly, a dinucleotide loop (?). Note that previous studies of the R/G hairpin focused on the gluR-B sequence and structures of the gluR-C and -D hairpins are inferred. **B:** Nonconserved, unpaired sequences flanking the terminal CG pair are shown for various subgroups that share the same sequence. The data suggest gluR-C and tilapia gluR-D hairpins may have an additional base pair at the end of the stem.

aligned each type of AMPA receptor individually to discern if there were features unique to each group. For gluR-B, in addition to the expected similarity of the mammalian sequences with each other, we observed a region of similarity beginning ~500 nt downstream of the 5' splice site and continuing for 217 nt (5'-AATGCA to CCCAAAG-3'; see accession numbers, Fig. 1). Pairwise comparisons showed that these sequences were

TABLE 1. Variation of mismatch/pairs.

Organism	Mismatch/pair ^a	
	1	2
Human-B	AC	GG
Rat-B	AC	GG
Mouse-B	AC	GG
Chicken-B	AC	GG
Pigeon-B	AC	GG
Tilapia-B	AC	GG
Human-C	AU	AG
Rat-C	AU	AA
Mouse-C	AU	AG
Chicken-C	AU	AG
Pigeon-C	AU	AG
Tilapia-D	GU	UG
Consensus	RY	(R,U)R

^aPairs are shown as interacting within a helix, left to right, 5' base, 3' base.

≥96% identical between mammals and birds. Although tilapia gluR-B contained a portion of this region of similarity, it was much more degenerate.

By far the most remarkable sequence conservation was observed when the gluR-C sequences were compared as a group. These introns were conserved throughout their length (556–564 nt), showing >94% identity even between the most distantly related organisms. Table 2 compares the evolutionary distances, expressed as substitutions per 100 nt, for the gluR-C intron-13 sequences and the gluR-C exon sequences. When substitutions at all codon positions are considered, the intron-13 sequences are more conserved than the exon sequences. When substitutions at synonymous and nonsynonymous positions are considered separately, intron 13 is much more conserved than nucleotides at synonymous positions, and contains only slightly more substitutions than nucleotides at nonsynonymous positions.

DISCUSSION

We performed a phylogenetic analysis of the AMPA-type gluR genes to determine what features of the gluR pre-mRNA are important for editing at the R/G site. Sequences within a short hairpin previously shown to be important for editing are strongly conserved among gluRs edited at the R/G site, but not within gluR-A, which is not edited at this site. Interestingly, in contrast to what is usually observed in a phylogenetic analysis of an RNA sequence, the identity of bases in double-stranded regions of the R/G hairpin were most conserved. We did not observe sequence similarities between the introns of gluR-B, -C, and -D beyond the short hairpin, suggesting that if additional sequences are required for editing in vivo, they have a function specific to a particular gluR (B, C, or D).

FIGURE 4. Structure mapping with ribonucleases confirms the R/G hairpin structure. A 79-nt RNA corresponding to a region spanning the human gluR-B R/G editing site (see Fig. 5) was labeled at its 5' end and incubated with various structure-specific reagents. Digestion of the RNA under native conditions with single-strand specific ribonucleases (T2, T1) or reagents specific for helical sequences (V1) confirmed the overall hairpin structure. Control reactions included a sample of RNA incubated in buffer only (--) and for sequence orientation, RNA treated with various hydrolysis conditions (H) or digested with T1 under denaturing conditions (65°C; Δ). According to convention, nucleotides are numbered with respect to their distance from the R/G site (zero), positive referring to the 3' direction and negative to the 5' direction. Similar experiments on two of the mutants (H15m, H23m) and the tilapia gluR-D wild-type sequence showed no significant changes in secondary structure.

A different type of sequence conservation

Phylogenetic analyses previously performed for other RNA molecules show that, in general, the identity of bases within helical regions are less conserved than

those in nonhelical regions (e.g., see Rich & Raj-Bhandary, 1976; Noller & Woese, 1981; Gutell et al., 1994). For the most part this is because it is easier to make sequence-specific interactions with nonhelical regions of RNA. In a double helix, most of the functional groups that distinguish one base from another are within the major groove (Seeman et al., 1976), which for A-form RNA is narrow and difficult to access. In the absence of sequence-specific interactions, bases within an RNA helix are free to vary, as long as the double-helical structure is maintained. Such covariation (e.g., AU base pair to a GC base pair), is often used as proof of the existence of a base-paired region of RNA.

The pattern of sequence conservation we observed in the R/G site hairpin is opposite from that described above, as it is the bases in helices that are most conserved, whereas bases in nonhelical regions covary so as to maintain their nonhelical state. This pattern is not so surprising given what is known about ADARs. The number of deaminations that occur within an RNA helix, or the *selectivity* of the ADAR reaction, depends on the length and stability of the RNA helix (reviewed in Bass, 1997; Lehmann & Bass, 1999). Analyses of endogenous RNA suggest that the R/G hairpin is edited at a single, specific adenosine (the R/G site; Lomeli et al., 1994). To maintain this selectivity, mutations that lengthen the stem of the R/G hairpin by pairing bases in the loop, or at the base of the stem, would be selected against. Instead, these sequences would covary so as to maintain their unpaired state. Similarly, mismatches would be conserved so as to maintain the particular stability required for the desired selectivity.

The striking conservation of the bases in the paired regions of the R/G hairpin may also be related to the effect of mismatches on ADAR selectivity. As mentioned, the path to a compensatory mutation requires that the base pair first mutate to a mismatch. Additional mismatches within the R/G hairpin would be predicted to alter the efficiency of editing at the R/G site, as emphasized by our mutagenesis data (Fig. 5). Although it may be fairly easy for a molecule such as rRNA to accommodate additional mismatches, especially because it exists in multiple copies, the transient mismatched state could be lethal to an organism dependent on precise and accurate editing of its gluR mRNAs. Thus, it would be difficult to accumulate compensatory mutations. Interestingly, our mutagenesis data suggest that GU can be replaced with a canonical base pair as well as certain mismatches with little effect on RNA editing (see H52Nm and 52m, Fig. 5). This suggests that changes of canonical pairs to GU pairs, which would not require an intermediate mismatched state, should be tolerated. In fact, among the gluR-B sequences, a base pair near the top of the stem adjacent to the hairpin loop does vary between an AU and GU pair (see Fig. 3A).

Importantly, because phylogenetic analyses are usually performed on organisms much more distant than

FIGURE 5. (Legend on facing page.)

TABLE 2. Sequence divergence between gluR-C exons and intron 13^a.

	Human				Rat				Chicken				Pigeon			
	Intron	Exon			Intron	Exon			Intron	Exon			Intron	Exon		
		All ^b	s ^c	ns ^d		All	s	ns		All	s	ns		All	s	ns
Human	0.0	0.0	0.0	0.0	0.36	4.35	16.46	0	5.98	14.51	75.58	0.68	4.67	15.39	88.57	0.37
Rat					0	0	0	0	6.06	14.30	80.63	0.68	4.94	15.23	91.82	0.37
Chicken									0	0	0	0	2.58	7.81	36.65	0.75
Pigeon													0	0	0	0

^aOnly partial cDNAs, encoding 270 amino acids, were available for pigeon and chicken, so exon divergence numbers were based on this 810-nt sequence from human, rat, chicken, and pigeon. Substitutions per 100 bases were estimated using software from the GCG package (Wisconsin Package Version 10.0, Genetics Computer Group (GCG), Madison, Wisconsin). Sequences were aligned with PILEUP and used as input for the programs DISTANCES and DIVERGE. Kimura's two-parameter method was used to correct for multiple hits and to account for the difference in substitution rates for transitions and transversions (Kimura, 1980). For accession numbers, see Table 3 (cDNAs) and Figure 1 (introns).

^bAll: all codon positions.

^cs: synonymous positions.

^dns: nonsynonymous positions.

those we were constrained to use (by the availability of cDNA sequences for primer design), we considered the possibility that the evolutionary distance between the organisms of our study was too small for compensatory changes to occur. To explore this idea we wanted to compare other RNA sequences among the organisms used in our analysis, in particular, sequences for which an RNA structure had been determined. We found that such comparisons were greatly restricted by the limited availability of accurate sequences for the non-mammalian species. However, partial sequence for the chicken large subunit rRNA was available (Michot et al., 1990) and comparison of this sequence to mammalian sequences clearly showed compensatory mutations. For example, three compensatory changes occur within the 14-nt stem of domain 2 (H14, see Michot et al., 1990), between rat (accession number V01270) and chicken (accession number X51919).

In general our mutagenesis results are consistent with the idea that base pairs within the R/G helix are conserved because of the deleterious effects of additional mismatches. Mutations that changed one base pair to another were less likely to affect editing at the R/G site than those that changed a base pair to a mismatch. However, in many cases changing a base pair to a mismatch affected R/G site editing only slightly,

and further, similar small effects were observed for some mutations that changed one base pair to another. Even though these mutations have only small effects, they could constrain the sequence in vivo. Importantly, transgenic mice exhibiting only a 25% drop in editing at the Q/R site, another editing site within AMPA gluRs, suffer epilepsy and die about three weeks after birth (Brusa et al., 1995). Although similar experiments have not been done for the R/G site, these experiments support the idea that the sequence of the R/G hairpin may be under strict constraints.

Another entirely plausible reason for the conservation of base pairs that do not seem absolutely essential for editing is that these base pairs are constrained because they are important for a process other than RNA editing, like RNA splicing. Future experiments will be required to determine this. Regardless, the fact that the sequence conservation in intron 13 of gluR-B, -C, and -D is limited to the short hairpin known to be required for RNA editing, and is not found in intron 13 of the unedited gluR-A, makes it likely that the conservation is in part dictated by ADARs. In addition to depending on the length and stability of the RNA helix (selectivity), ADAR specificity is dictated by slight sequence *preferences*. For example, ADAR1 preferentially targets adenosines that do not have a G as their 5' nearest neighbor,

FIGURE 5. Analysis of the effect of mutations on editing at the R/G site. A primer extension assay was used to determine the efficiency of editing at the R/G site for a wild-type sequence and various mutant molecules (see Materials and Methods). Data from primer extension assays were quantified, and the bar graphs indicate averages from multiple experiments (≥ 3), with standard deviations. (All mutations named with an H prefix derived from ≥ 9 determinations). Mutations listed bottom to top on the hairpin correlate with bars on the graphs from left to right, and are also color-coded. Overall, mutations that change the identity of a base pair (**A**) have the least effect, whereas those that create a mismatch (**B**) or change the identity of a mismatch (**C**) have more dramatic effects. For the data shown, deamination was catalyzed by *Xenopus* ADAR1; see text for results using human ADAR1. Asterisks indicate sequences that were altered from the human sequence to facilitate cloning.

and those that are ≥ 8 nt from a 3' end (Polson & Bass, 1994). When ADARs are required to target a specific adenosine, such preferences, together with selectivity, may result in tight sequence constraints. Adenosines not targeted for deamination should, if possible, be put in the context of a 5' G or very close to a 3' terminus. These sequence choices would need to be balanced by considerations of the thermodynamic stability of the molecule to maintain selectivity. The sequence would be constrained to promote accurate and efficient editing at the R/G site, but also to avoid creating secondary deamination sites. If a new deamination site were created, perhaps one that was more rapidly edited than the R/G site, editing at the R/G site might be altered because it would now occur in a molecule with an additional mismatch (AU $\rightarrow$ IU). For the R/G hairpin, secondary sites might cause undesired changes in the ORF or in RNA splicing. In this regard we favor the idea that if other processes or factors contribute to the sequence conservation of the hairpin, it is because they are affected by changes in ADAR specificity.

The phylogenetic and mutagenesis data suggest certain noncanonical features at mismatch/pair 1 and 2 of the R/G hairpin may be important, although it is difficult to predict a particular structural feature that fits all data. Among the organisms analyzed in our study, mismatch/pair 1 was either an AC, AU, or GU pair (see Table 1). Mutation at this site to a CA or CC mismatch caused little effect on editing at the R/G site, suggesting other mismatches are tolerated at this site (see Fig. 5). Consistent with the idea that mismatch/pair 1 must either be a mismatch or less stable base pair, a mutation to a GC dramatically decreased editing at the R/G site. However, a mutation to a CG at this site had little effect; of course, this pair may create an additional editing site that is selected against *in vivo*. At mismatch/pair two, consistent with the idea that an RR base pair exists, canonical CG and GC pairs are tolerated but not a pyrimidine–pyrimidine, CC pair. Importantly, as shown in Table 1, the phylogenetic variation at mismatch/pair 1 and 2 occurs between different types of gluR, raising the possibility that different gluRs require slightly different amounts of R/G site editing and that these bases are used to fine-tune editing efficiency. Finally, the conserved location of mismatch/pair 1 and 2 suggest the distance of the mismatches from the R/G editing site may be important, perhaps for positioning the target adenosine in the active site.

Implications for future studies

Because the substrate specificity of ADARs is likely mediated at least in part by its dsRNA binding motifs (dsRBMs), we speculate the pattern of sequence conservation found in our phylogenetic analysis will be observed in substrates of other proteins that contain

dsRBMs. Further, we suspect the pattern of conserved helices, with covariation in adjacent nonhelical regions, may be more common than thought, even in RNAs that are not recognized by dsRBPs, such as rRNA. Of course, this type of sequence conservation would generally be overlooked because covariation is typically a prerequisite for the proof of an RNA helix.

During the course of our study, work by others showed that editing at the R/G site occurs in chicken gluR-B (Lee et al., 1998), as would be predicted by the phylogenetic conservation of the sequence and structure of the R/G hairpin. Further, a cDNA corresponding to a partially processed mRNA of chicken gluR-D has been noted to show a high degree of sequence conservation in the region of the R/G hairpin (Paperna et al., 1996). Although many of the gluRs from the different organisms used in our analysis have not been assayed for R/G site editing, the strong sequence conservation revealed by our study suggests all of these receptors will be edited. In this regard, our analysis suggests that when ADAR substrates are identified in the future, their homologs in other organisms may be isolated simply by sequence similarity.

Highly conserved introns

Finally, in regard to the strong conservation of intron 13 of gluR-C, we can only conclude that the sequence of this intron is functionally important. There are short ORFs throughout intron 13, so conceivably, it could serve a coding function, either as part of the gluR-C mRNA, or an opposing transcript. Although we have not tested this idea experimentally, we have searched available databases, including those containing EST sequences, and can find no significant similarities. We favor the idea that the intron is conserved because it is important for a noncoding function, for example, for regulating the alternative splicing required to generate flip and flop transcripts.

It is typically assumed that sequences of introns, unlike those of exons, are not conserved. Although evidence to support this view exists, in reality, it is very hard to make general conclusions because few introns have been sequenced from multiple organisms. In fact, comparisons of large regions of human and rodent genomes indicate noncoding sequences are likely to vary in their conservation (reviewed in Koop, 1995; Hardison et al., 1997). Some show very little conservation, some a great deal, and some show a mosaic pattern of conservation with divergent sequences intermittent with highly conserved sequences. The latter studies involved comparisons of organisms that diverged 40–80 million years ago (human and rodents), and our studies suggest highly conserved introns can be observed even between more distantly related organisms.

MATERIALS AND METHODS

DNA isolation

Genomic DNA from human, mouse, and chicken was purchased (Promega), and that of rat and tilapia prepared from commercially obtained rat brains (Pel-Freez) or whole tilapia, using standard procedures (Ausubel et al., 1994–1997). Tilapia were kindly provided by Dr. Tom Kocher (University of New Hampshire). Pigeon blood was obtained from Avian Resources, The Bowman Gray School of Medicine, and pigeon genomic DNA prepared using a technique of Dr. Peter T. Boag (Queen's University, pers. comm.).

Cloning and sequencing

Initial upstream and downstream primers were designed according to cDNA sequence, near the previously determined splice sites (e.g., see Lomeli et al., 1994; see Table 3). For each primer set, optimal amplification conditions were determined by performing a gradient annealing temperature profile using a constant magnesium concentration of 2.5 mM. Conditions that produced a single DNA product of the expected size were considered optimal, and fragments were

excised from an agarose gel and cloned. When none of the temperatures appeared optimal, additional parameters were varied; these included elongation times and type of polymerase (see Table 4). In some cases, particular primer sets proved problematic and alternate primer sets were designed. Except where noted in Table 4, PCR amplification was performed in 50 μ L containing 1 $\times$ Promega PCR buffer, 2.5 U of polymerase (Taq or Pfu), 2.5 mM MgCl₂, 250 μ M of each dNTP, and 500 nM of each primer.

PCR fragments generated with Taq polymerase were ligated into pGEM-T (Promega), those generated with Pfu polymerase into pCR-Script Amp SK(+) (Stratagene) and both were transformed into DH5 alpha cells. Purified plasmid DNA was sequenced using vector primers (T3, T7, and/or SP6), and if necessary, internal primers. Internal primers were needed to sequence the entire clone for human, rat, mouse, chicken, pigeon, and tilapia gluR-B, and for tilapia gluR-D, and were based on intron sequence obtained during sequencing with vector primers. Optimal primer choice was facilitated by analysis with OLIGO 4.0, 5.0 software (National Biosciences, Inc., Plymouth, Minnesota). Sequences were confirmed visually by chromatogram (EditView, Perkin-Elmer), and aligned with Megalign (DNASTar package) using the Clustal method.

TABLE 3. Primers used for amplifying gluR genomic sequence.

Gene	Primer binding site	Primer sequence	cDNA accession number
Rat GluR-A	Exon 13	5'-GCTATGGCATTGCGACACCC	M36418
Rat GluR-A	Exon 14	5'-GTAGCAAATTCTAGAATTAAGTC	M36418
Mouse GluR-A	Exon 13	5'-GCTATGGCATTGCGACACCC	X57497
Mouse GluR-A	Exon 14	5'-GTAGCAAATTCTAGAATTAAGTC	X57497
Human GluR-B	Exon 13	5'-TGGAGTCCACGATGAACGAG	L20814
Human GluR-B	Exon 15	5'-CTCCACATTCACCTTTATCG	L20814
Rat GluR-B	Exon 13	5'-CCTTGTGACACCATGAAAGT	M85035
Rat GluR-B	Exon 14	5'-TTTCAATTTGTCCAACAGGC	M85035
Mouse GluR-B	Exon 13	5'-CCTTGTGACACCATGAAAGT	*
Mouse GluR-B	Exon 14	5'-TTTCAATTTGTCCAACAGGC	X57498
Chicken GluR-B	Exon 13	5'-AGGAAACCCTGTGATACCATGAA	X89508
Chicken GluR-B	Exon 15	5'-TTGGCTCCACATTCACCTTTATC	X89508
Pigeon GluR-B	Exon 13	5'-CCAAAGGCTACGGCATC	Z29713
Pigeon GluR-B	Exon 14	5'-CAGGCCTTGTTTCATTTCAGTT	Z29713
Tilapia GluR-B	Exon 13	5'-CGAAAACCCTGCGACACCATGAAGGTG	L37836
Tilapia GluR-B	Exon 14	5'-CACCTCCCCCGCTGCCGCACTCTCCTT	L37836
Human GluR-C	Exon 13	5'-AAAGTCCAAGGGAAAGTTCG	U10302
Human GluR-C	Exon 14	5'-GGCCTTGCTCATTTCAGTTTT	U10302
Rat GluR-C	Exon 13	5'-CGTGCGATACGATGAAAGTT	M85036
Rat GluR-C	Exon 14	5'-TTTCAATTTGTCCAAGAGGC	M85036
Mouse GluR-C	Exon 13	5'-CGTGCGATACGATGAAAGTT	*
Mouse GluR-C	Exon 14	5'-TTTCAATTTGTCCAAGAGGC	*
Chicken GluR-C	Exon 13	5'-CCTGCGACACCATGAAAGTT	X89509
Chicken GluR-C	Exon 14	5'-GTCTTGCTCTTGGAGTCACCTC	X89509
Pigeon GluR-C	Exon 13	5'-AAGGGCTATGGTGTAGCAACCC	Z29714
Pigeon GluR-C	Exon 14	5'-TCGTCTTGCTCTTGGAGTCACC	Z29714
Tilapia GluR-D	Exon 13	5'-CGAAAACCCTGCGACACCATGAAGGTG	**
Tilapia GluR-D	Exon 14	5'-CACCTCCCCCGCTGCCGCACTCTCCTT	**

*Mouse sequences amplified using primers based on rat sequence.

**Amplified using tilapia gluR-B primers (see text).

TABLE 4. Summary of PCR conditions used to amplify genomic sequence.

Gene	Enzyme	Anneal. (°C)	Cycles	Elong. (min)	Product length
Rat GluR-A	Pfu	55	35	2.5	705 bp
Mouse GluR-A	Pfu	55	35	2.5	722 bp
Human GluR-B	Pfu	51	35	2.5	1609 bp*
Rat GluR-B	Pfu	56	31	3.0	990 bp
Mouse GluR-B	Pfu	56	31	3.0	1004 bp
Chicken GluR-B	Pfu	53	35	2.5	1573 bp*
Pigeon GluR-B	Taq	65	33	3.5	1008 bp
Tilapia GluR-B	Taq	60	35	3.0	1515 bp
Human GluR-C	Taq	68	35	2.5	752 bp
Rat GluR-C	Taq	68	35	2.5	696 bp
Mouse GluR-C	Pfu	56	31	3.0	696 bp
Chicken GluR-C	Pfu	65	35	2.5	766 bp
Pigeon GluR-C	Pfu	65	35	2.5	725 bp
Tilapia GluR-D	Taq	60	35	3.0	927 bp

*Amplified from exon 13 to 15. Anneal. = annealing. Elong. = elongation.

Synthesis of wild-type and mutant R/G hairpins

Complementary oligonucleotides were synthesized and annealed to generate a double-stranded DNA fragment containing an *Xba*I site at one end, followed by a T7 promoter, a portion of the human *gluR-B* genomic DNA that spanned the R/G hairpin, and finally, a *Sa*I site at the downstream terminus; this fragment was ligated into pKL36S (Lehmann & Bass, 1999) that had been digested with *Xba*I and *Sa*I. Plasmid templates for the mutant molecules were made similarly. *Sa*I was used to linearize all plasmids for transcription, producing the 79-nt RNAs shown in Figure 5. RNA was transcribed with T7 polymerase using standard protocols, treated with DNase, passed through a chromospin-10 column (Clontech) to remove unincorporated nucleotides, and extracted with phenol and chloroform and ethanol precipitated. RNA was dried, resuspended in water and gel purified after electrophoresis on an 8 M urea, 12% polyacrylamide, 1× TBE gel. RNA concentration was determined by optical density at A260, using the relationship 40 O.D. A260 = 1 μ g/mL.

RNA structure mapping

RNA was treated with calf intestinal phosphatase (CIP), and then 5'-end labeled using γ -³²P-ATP and T4 polynucleotide kinase. RNA was again put through a Chromaspin-10 column, and gel purified a second time after electrophoresis on an 8 M urea, 10% polyacrylamide, 1× TBE gel. RNA was treated with proteinase K, and extracted with phenol, phenol-chloroform, and chloroform, then ethanol precipitated, washed with 70% ethanol, dried, and resuspended in 100 μ L ddH₂O. For digestion with T2 or V1, 10,000 cpm of RNA was mixed with 4 μ L of 20 mM Tris-HCl (pH 7.5), 20 mM MgCl₂, 200 mM KCl, 1 μ L of 10 μ g/ μ L torula RNA, and 0.1 U T2 ribonuclease or V1, diluted to 10 μ L with water, and incubated for 20 min at 37°C. For T1 reactions 10,000 cpm of RNA were mixed

with 2 μ L 100 mM sodium acetate (pH 4.6), 1 μ L of 10 μ g/ μ L torula RNA, and 0.05 U T1 ribonuclease, and diluted to 10 μ L with water and incubated for 15 min at 37°C (native) or 60°C (denaturing). All reactions were stopped by adding 10 μ L containing 0.05% dyes in 8 M urea, and immediately putting samples on dry ice. Samples were electrophoresed on an 8 M urea, 12% polyacrylamide, 1× TBE gel at 50 W for 1 h, 45 min, dried, and then put on a PhosphorImager screen overnight.

In vitro editing and primer extension assays

RNAs were incubated with ADAR1 in 100 μ L containing 40 fmol of R/G RNA hairpin, 40 mM Tris (pH 8.0), 5% glycerol, 25 mM KCl, 10 mM NaCl, 1.1 mM MgCl₂, 1 mM DTT, 50 μ M EDTA, 20 μ M ATP, 0.16 U/ μ L RNasin. The amounts of ADAR1 were as indicated in figure legends. Two preparations of ADAR1 were used, one purified from *Xenopus laevis* eggs (Hough & Bass, 1994) and the other from *Saccharomyces cerevisiae* overexpressing a recombinant form of the human ADAR1 (H. Ley and B. Bass, unpubl. data). Reactions were incubated for 30 min at 30°C, then stopped by extracting with an equal volume of phenol:chloroform:isoamylalcohol (25:24:1), followed by a second extraction with chloroform. The sample was then ethanol precipitated, washed with 70% ethanol, dried, and resuspended in 5 μ L ddH₂O in preparation for the primer extension assay.

A primer extension assay was used to quantify the amount of editing at the R/G site. In the presence of ddT, reverse transcription terminates at the R/G site for unedited molecules, but continues to the next adenosine for molecules that are edited. A small amount of editing was observed at the adenosine 5' of the R/G site, as indicated by a small fraction of molecules that were extended to the next adenosine. Forty femtomoles of RNA in 5 μ L of ddH₂O were mixed with 1 μ L 10× dNTP/ddTTP mix, 1 μ L radiolabeled R/G RT primer, 2 μ L 5× AMV RT buffer (Boehringer Mannheim) and heated at 70°C for 10 min, then at 55°C for 10 min; 1 μ L containing 4 U of AMV RT (Boehringer Mannheim) was then added, and the reaction was incubated at 42°C for 15 min (10× ddNTP/ddTTP contained 0.1 mM of dATP, dCTP, and dGTP and 2.5 mM ddTTP). Thirty picomoles of the primer were labeled with T4 polynucleotide kinase and γ -³²P-ATP, and after inactivating the kinase (65°C for 10 min), passed successively through two chromaspin-10 columns). The primer was designed based on complementarity to the wild-type sequence (5'-ATTGTTATACTATTCCACCC), but in preliminary tests was found to work well for all mutants except H15m. For the latter mutant an alternate primer was used (5'-ATTGTTATACTATTCCACCCAT). Reverse transcription reactions were stopped by adding 7 μ L of formamide loading buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol), and 10 μ L of the sample were electrophoresed on a 20% polyacrylamide (19:1), 8 M urea, 1× TBE sequencing gel, and run at ~15,000 V/h. The wet gel was wrapped with saran wrap and exposed for 2 h in a PhosphorImager cassette and quantified using ImageQuant software (Molecular Dynamics). Percent editing was determined by summing the intensity of all editing dependent bands and then dividing by the sum of the intensities of all bands in the lane.

ACKNOWLEDGMENTS

We thank Tanya Baker for a helpful discussion, Herbert Ley and Mark Macbeth for critically reading this manuscript, and members of the Bass laboratory for advice and helpful suggestions. We are grateful to Ron Hough and Herbert Ley for purified ADAR1, Katrina Lehmann for technical advice, Tom Kocher for tilapia, and Peter Boag for protocols for isolating DNA from pigeon blood. Oligonucleotides were synthesized by the Howard Hughes Medical Institute (HHMI) oligonucleotide synthesis facility at the University of Utah supported by the Department of Energy (Grant #DE-FG03-94ER61817) and HHMI; DNA was sequenced by the University of Utah Health Sciences DNA Sequencing Facility supported in part by the National Cancer Institute (Grant #5P30CA42014). This work was supported by funds to B.L.B. from the National Institute of General Medical Sciences (GM44073). B.L.B. is an HHMI Associate Investigator.

Received August 24, 1999; returned for revision, September 21, 1999; revised manuscript received November 10, 1999

REFERENCES

- Ausubel F, Brent R, Kingston R, Moore D, Seidman J, Smith J, Struhl K, eds. 1994–1997. *Current protocols in molecular biology*. John Wiley & Sons.
- Bass BL. 1997. RNA editing and hypermutation by adenosine deamination. *Trends Biochem Sci* 22:157–162.
- Bass BL, Nishikura K, Keller W, Seeburg PH, Emeson RB, O'Connell MA, Samuel CE, Herbert A. 1997. A standardized nomenclature for adenosine deaminases that act on RNA. *RNA* 3:947–949.
- Brusa R, Zimmermann F, Koh DS, Feldmeyer D, Gass P, Seeburg PH, Sprengel R. 1995. Early-onset epilepsy and postnatal lethality associated with an editing-deficient GluR-B allele in mice. *Science* 270:1677–1680.
- Burns CM, Chu H, Rueter SM, Hutchinson LK, Canton H, Sanders-Bush E, Emeson RB. 1997. Regulation of serotonin 2C receptor G-protein coupling by RNA editing. *Nature* 387:303–308.
- Egebjerg J, Heinemann SF. 1993. Ca^{2+} permeability of unedited and edited versions of the kainate selective glutamate receptor GluR6. *Proc Natl Acad Sci USA* 90:755–759.
- Gutell RR, Larsen N, Woese CR. 1994. Lessons from an evolving rRNA: 16S and 23S rRNA structures from a comparative perspective. *Microbiol Rev* 58:10–26.
- Hardison RC, Oeltjen J, Miller W. 1997. Long human-mouse sequence alignments reveal novel regulatory elements: A reason to sequence the mouse genome. *Genome Res* 7:959–966.
- Higuchi M, Single FN, Kohler M, Sommer B, Sprengel R, Seeburg PH. 1993. RNA editing of AMPA receptor subunit GluR-B: A base-paired intron-exon structure determines position and efficiency. *Cell* 75:1361–1370.
- Hollmann M, Heinemann S. 1994. Cloned glutamate receptors. *Annu Rev Neurosci* 17:31–108.
- Hough RF, Bass BL. 1994. Purification of the *Xenopus laevis* double-stranded RNA adenosine deaminase. *J Biol Chem* 269:9933–9939.
- Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16:111–120.
- Koop BF. 1995. Human and rodent DNA sequence comparisons: A mosaic model of genomic evolution. *Trends Genet* 11:367–371.
- Kumar S, Hedges SB. 1998. A molecular timescale for vertebrate evolution. *Nature* 392:917–920.
- Lee JC, Greig A, Ravindranathan A, Parks TN, Rao MS. 1998. Molecular analysis of AMPA-specific receptors: Subunit composition, editing, and calcium influx determination in small amounts of tissue. *Brain Res Brain Res Protoc* 3:142–154.
- Lehmann K, Bass B. 1999. The importance of internal loops within RNA substrates of ADAR1. *J Mol Biol* 291:1–13.
- Lomeli H, Mosbacher J, Melcher T, Hoyer T, Geiger JR, Kuner T, Monyer H, Higuchi M, Bach A, Seeburg PH. 1994. Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. *Science* 266:1709–1713.
- Michot B, Qu LH, Bachelier JP. 1990. Evolution of large-subunit rRNA structure. The diversification of divergent D3 domain among major phylogenetic groups. *Eur J Biochem* 188:219–229.
- Noller HF, Woese CR. 1981. Secondary structure of 16S ribosomal RNA. *Science* 212:403–411.
- O'Connell MA. 1997. RNA editing: Rewriting receptors. *Curr Biol* 7:R437–R439.
- Pace N, Thomas B, Woese C. 1999. Probing RNA structure, function, and history by comparative analysis. In: Gesteland RF, Cech TR, Atkins JF, eds. *The RNA world*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 113–141.
- Paperna T, Lamed Y, Teichberg VI. 1996. cDNA cloning of chick brain alpha-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptors reveals conservation of structure, function and post-transcriptional processes with mammalian receptors. *Brain Res Mol Brain Res* 36:101–113.
- Paul MS, Bass BL. 1998. Inosine exists in mRNA at tissue-specific levels and is most abundant in brain mRNA. *EMBO J* 17:1120–1127.
- Polson AG, Bass BL. 1994. Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *EMBO J* 13:5701–5711.
- Polson AG, Bass BL, Casey JL. 1996. RNA editing of hepatitis delta virus antigenome by dsRNA-adenosine deaminase. *Nature* 380:454–456.
- Rich A, RajBhandary UL. 1976. Transfer RNA: Molecular structure, sequence, and properties. *Annu Rev Biochem* 45:805–860.
- Seeburg PH, Higuchi M, Sprengel R. 1998. RNA editing of brain glutamate receptor channels: Mechanism and physiology. *Brain Res Rev* 26:217–229.
- Seeman NC, Rosenberg JM, Rich A. 1976. Sequence-specific recognition of double helical nucleic acids by proteins. *Proc Natl Acad Sci USA* 73:804–808.
- Sommer B, Keinänen K, Verdoorn TA, Wisden W, Burnashev N, Herb A, Kohler M, Takagi T, Sakmann B, Seeburg PH. 1990. Flip and flop: A cell-specific functional switch in glutamate-operated channels of the CNS. *Science* 249:1580–1585.
- Yang JH, Sklar P, Axel R, Maniatis T. 1997. Purification and characterization of a human RNA adenosine deaminase for glutamate receptor B pre-mRNA editing. *Proc Natl Acad Sci USA* 94:4354–4359.