The EMBO Journal Vol.17 No.4 pp.1120–1127, 1998

Inosine exists in mRNA at tissue-specific levels and is most abundant in brain mRNA

Michael S.Paul and Brenda L.Bass[1]

Department of Biochemistry and Howard Hughes Medical Institute, 50 North Medical Drive, University of Utah, Salt Lake City, UT 84132, USA

[1]Corresponding author
e-mail: bass@tulip.med.utah.edu

The general view that mRNA does not contain inosine has been challenged by the discovery of adenosine deaminases that act on RNA (ADARs). Although inosine monophosphate (IMP) cannot be detected in crude preparations of nucleotides derived from poly(A)$^+$ RNA, here we show it is readily detectable and quantifiable once it is purified away from the Watson–Crick nucleotides. We report that IMP is present in mRNA at tissue-specific levels that correlate with the levels of ADAR mRNA expression. The amount of IMP present in poly(A)$^+$ RNA isolated from various mammalian tissues suggests adenosine deamination may play an important role in regulating gene expression, particularly in brain, where we estimate one IMP is present for every 17 000 ribonucleotides.

Keywords: ADAR/deaminase/inosine/RNA editing

Introduction

During protein synthesis, amino acids are specified by mRNA codons, which consist of triplet combinations of the Watson–Crick nucleotides. Covalent modifications of the Watson–Crick nucleotides, like those that occur within tRNA (Limbach *et al.*, 1994), could potentially increase the number of ways in which a particular amino acid is specified, but are not thought to be prevalent in mRNA. In fact, although eukaryotic mRNAs and many viral mRNAs are 'capped' at their 5′ terminus with a 7-methylguanosine and proximal ribose methylations, only a single modified nucleoside has been detected within coding sequences (reviewed in Narayan and Rottman, 1992). However, this modified nucleoside, 6-methyladenosine, which is created by the methylation of adenosine (A), does not alter codon meaning since it also pairs with uridine (Engel and von Hippel, 1978).

The discovery of adenosine deaminases that act on RNA (ADARs; Bass and Weintraub, 1988; Melcher *et al.*, 1996; Bass *et al.*, 1997) has led to speculations that inosine (I), the product of adenosine deamination, may exist within mRNA (reviewed in Simpson and Emeson, 1996; Bass, 1997). Inosine prefers to base-pair with cytidine, and accordingly, is read by the translational machinery as guanosine (Basilio *et al.*, 1962). Except at certain wobble positions, an A to I modification within a codon would result in the specification of an alternate amino acid, and the synthesis of a protein that was not genomically encoded. Such post-transcriptional changes to the genomically encoded sequence of an mRNA are not unprecedented, and fall into a group of reactions known as RNA editing reactions (reviewed in Bass, 1997; Kable *et al.*, 1997; Smith *et al.*, 1997). However, all previously described RNA editing reactions that occur on mRNAs involve the conversion of one Watson–Crick nucleotide into another, or the insertion or deletion of Watson–Crick nucleotides, rather than the creation of a non-Watson–Crick nucleotide such as inosine. Since the first observation of RNA editing in 1986 (Benne *et al.*, 1986), many examples and types of RNA editing have been discovered, and the process is now recognized as an important way to regulate gene expression in eukaryotes.

Based on discrepancies between genomic and cDNA sequences that are consistent with adenosine deamination (e.g. A to G transitions), several mammalian glutamate receptor (gluR) mRNAs (Sommer *et al.*, 1991; Lomeli *et al.*, 1994), mammalian $5HT_{2C}$ serotonin receptor mRNA (Burns *et al.*, 1997), hepatitis delta virus (HDV) antigenomic RNA (Polson *et al.*, 1996) and mammalian α2,6-sialyltransferase mRNA (Ma *et al.*, 1997) have been proposed to be edited by ADARs *in vivo*. For each of these RNAs, adenosine deamination is proposed to cause a functionally important codon change. For example, an A to I change within gluR-B pre-mRNA is thought to be responsible for changing a glutamine codon to an arginine codon (Q/R editing site), as well as an arginine codon to a glycine codon (R/G editing site). Both of these editing events have been correlated with functionally important changes in the ion channels formed from the altered gluR subunits (reviewed in Seeburg, 1996). In support of the idea that the editing sites predicted by A to G transitions correspond to *in vivo* deamination sites, incubation of *in vitro* transcribed gluR-B, $5HT_{2C}$ and HDV antigenomic RNAs with various ADARs leads to deamination of adenosines at the proposed editing sites (Melcher *et al.*, 1995; Rueter *et al.*, 1995; Yang *et al.*, 1995; Polson *et al.*, 1996; Burns *et al.*, 1997). Also, a recently developed method for specifically cleaving RNA at inosine residues shows cleavage of endogenous rat brain gluR-B mRNA at both the Q/R and the R/G editing sites (Morse and Bass, 1997).

Despite the existence of a great deal of circumstantial evidence for the presence of inosine within mRNA, inosine monophosphate (IMP) has never been detected among ribonucleotides derived from mRNA. In fact, although inosine has long been known to exist in tRNA (reviewed in Grosjean *et al.*, 1996), the general view is that mRNA does not contain inosine (Limbach *et al.*, 1994). Indeed, analyses of nucleotides derived from ribonuclease-digested mRNA in our own laboratory have repeatedly emphasized this general view, and led to the detection of only the four Watson–Crick nucleotides.

© Oxford University Press

We reasoned that our inability to detect inosine within cellular mRNA might be analogous to the unfeasibility of detecting certain proteins within a crude cellular extract. Obviously, in such cases, the specific activity of the protein must be increased by purification in order for the protein to be detected above the background established by other, more abundant proteins. In this light, we designed a scheme by which IMP could be purified from the ribonucleotides of cellular poly(A)$^+$ RNA, using successive thin-layer chromatography (TLC) steps. Here we report that IMP is indeed a component of poly(A)$^+$ RNA, and is present at tissue-specific levels that correlate with the tissue-specific expression of ADAR mRNAs. The amount of IMP present in poly(A)$^+$ RNA isolated from various tissues suggests adenosine deamination may play an important role in regulating gene expression, particularly in brain, where we estimate one IMP is present for every 17 000 nucleotides. From our data, we estimate that mammalian brain contains ~1800-fold more inosine than can be accounted for by the editing sites within mammalian gluR-B mRNA, suggesting that there are many additional inosine-containing RNAs yet to be identified.

Results

Purification of IMP from rat brain poly(A)$^+$ RNA

Our general scheme for the purification of IMP from cellular RNA is outlined in Figure 1. We chose to analyze the RNA in the form of radiolabeled 5′ nucleoside monophosphates (5′ NMPs; pNs) since existing TLC systems, for the most part, were developed for the analysis of 5′ NMPs. RNA samples were first digested to 3′ NMPs (Nps) using RNase T2, to generate a substrate for T4 polynucleotide kinase (PNK), which requires a 3′ phosphate. The resulting 3′ NMPs were 5′ end-labeled with T4 PNK and [γ-^{32}P]ATP. The [^{32}P]Nps were then digested with nuclease P1 to yield 5′ [^{32}P]NMPs. At this point, after extracting the samples to remove all protein, the samples were spiked with non-radioactive 5′ IMP. This spike was an essential and key feature of our protocol since it allowed us to monitor IMP by UV absorbance, during initial purification steps, when radiolabeled IMP was undetectable.

Since several putative ADAR substrates were identified in mammalian brain tissue (Sommer *et al.*, 1991; Lomeli *et al.*, 1994; Burns *et al.*, 1997), we first attempted to purify IMP from rat brain poly(A)$^+$ RNA. We wanted to be able to quantify any observed radioactive IMP derived from the rat brain RNA, in a manner that accounted for the recovery and labeling efficiencies of each step of the protocol shown in Figure 1. Thus, we ran a number of control samples in parallel. These control samples contained RNA that was synthesized *in vitro* using T3 RNA polymerase and were spiked with various amounts of non-radioactive 3′ IMP. Since 3′ NMPs become labeled during our protocol (Np→*pNp, Figure 1), we anticipated that we would purify radioactive IMP in an amount proportional to the amount added, and that the radioactivity values we determined would take into account the losses and efficiencies intrinsic to our protocol. As a control for background radioactivity that might co-migrate with IMP, we also processed a T3 transcript that was not spiked with 3′ IMP.

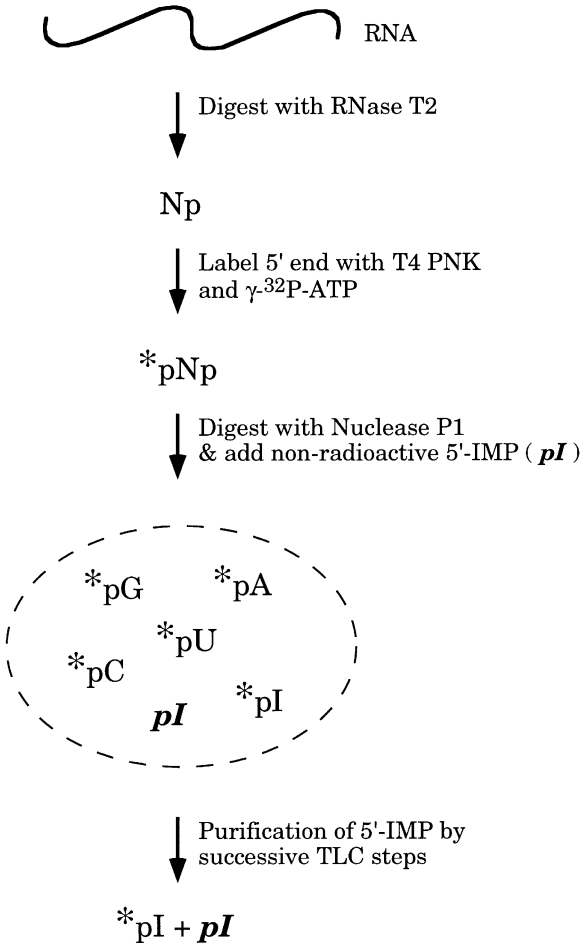

Fig. 1. Overall scheme for the purification of IMP. Asterisks indicate ^{32}P-labeled phosphates. Each step was optimized in control experiments, and RNase T2 digestion was determined to be >95% efficient. By monitoring the non-radioactive 5′ IMP spike (***pI***) with UV light, 5′ [^{32}P]IMP (*pI) was purified away from the crude nucleotide mixture (encircled with a dotted line) during successive TLC separations.

Ten μg of rat brain poly(A)$^+$ RNA, and 10 μg of each of the control synthetic RNAs, were processed as in Figure 1 and subjected to the first purification step, a two-dimensional (2D) TLC separation. The autoradiogram for each sample looked essentially identical, and that for the rat brain poly(A)$^+$ RNA is shown in Figure 2A. As anticipated, most of the radioactivity was distributed between radiolabeled 5′ AMP, 5′ CMP, 5′ GMP and 5′ UMP, and the [γ-^{32}P]ATP remaining from the PNK reaction. Although radioactive IMP was not detected, exposure of the TLC plate to short-wave UV light allowed the detection of the non-radioactive IMP spike, which we outlined with a dotted line (see Figure 2A). Reasoning that 5′ [^{32}P]IMP was present in the outlined spot, which overlapped with the 5′ [^{32}P]GMP, it was excised from each TLC plate with a pair of scissors. Utilizing a direct transfer method (Randerath and Randerath, 1980), the matrix side of each cut-out was brought into direct contact with a fresh TLC plate, and subjected to the second purification step.

The second purification step was a one-dimensional (1D) TLC system, and thus, multiple samples could be processed on a single plate (Figure 2B). Autoradiography

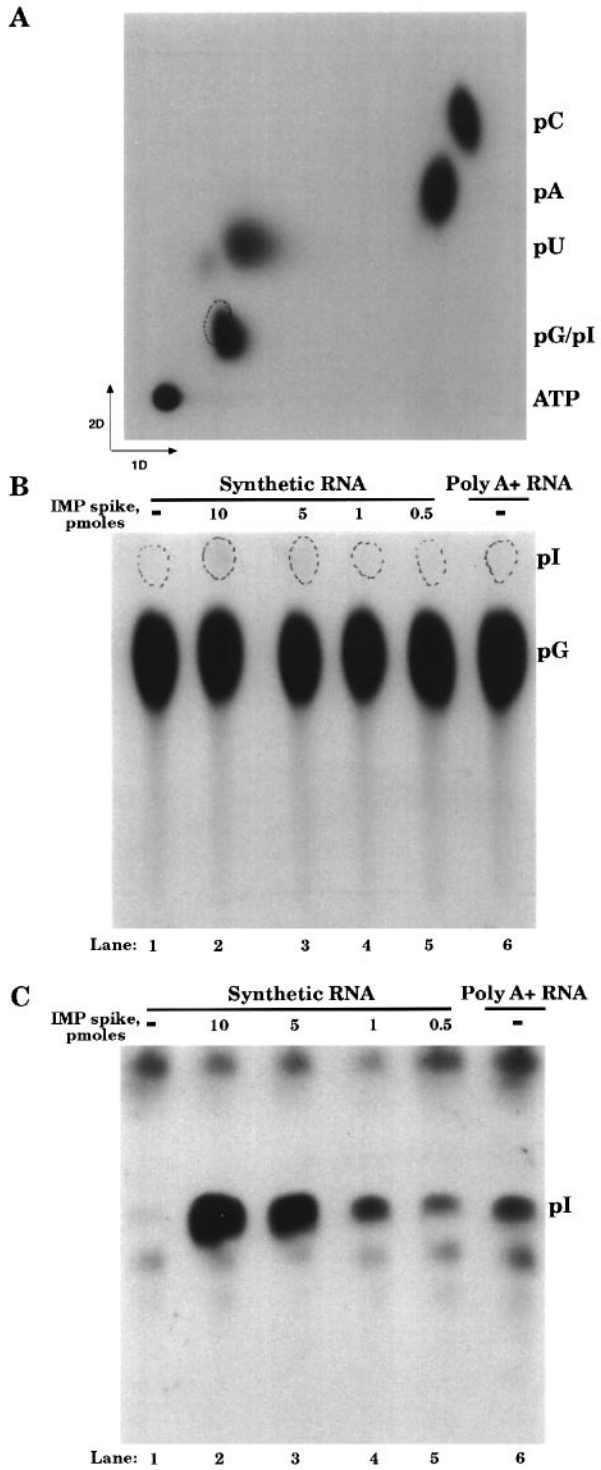

Fig. 2. Autoradiograms of the three successive TLC separations used to purify [^{32}P]IMP. Nucleotide migrations are shown to the right of each plate, and that of the non-radioactive 5′ IMP spike, with a dotted outline. For 1D separations, chromatography was bottom to top as shown. (**A**) Purification step 1: the crude nucleotide mixtures for each sample were separated by 2D TLC; only the autoradiogram of the rat brain poly(A)$^+$ sample is shown. (**B**) Purification step 2: the nucleotide composition is shown for samples derived from rat brain poly(A)$^+$ RNA (lane 6), and synthetic RNAs spiked with 0 (–), 10, 5, 1 and 0.5 pmol of IMP (lanes 1–5). (**C**) Purification step 3: the nucleotide composition is shown for samples derived from synthetic RNA or rat brain poly(A)$^+$ RNA; lanes are as in (B). The identity of radioactive spots that are not labeled to the right of the plate is unknown.

combined with the UV detection of the non-radioactive IMP indicated that purification step 2 effectively separated 5′ GMP from 5′ IMP. However, again, the level of background radioactivity was too high to allow detection of radioactive IMP in any of the samples; longer exposure showed substantial background radioactivity that co-migrated throughout the region of 5′ IMP migration (data not shown). As in purification step 1, non-radioactive 5′ IMP was visualized with UV light, outlined with a pencil, cut out, and taken to purification step 3 by the direct transfer method.

Purification step 3 was also a 1D TLC system, and the autoradiogram of the TLC plate is shown in Figure 2C. In contrast to the analyses performed after purification steps 1 and 2, after step 3 we observed radioactive IMP in the rat brain poly(A)$^+$ RNA sample, as well as each of the synthetic RNAs spiked with IMP. We verified that all material that co-migrated with IMP actually corresponded to IMP by excising each radioactive IMP spot and re-chromatographing it on a fourth TLC plate with a different solvent; no additional spots were observed, and the calculated amounts of IMP remained the same (data not shown).

A visual comparison of the rat brain poly(A)$^+$ sample in Figure 2C with the various spiked synthetic RNAs indicated there was ~1 pmol of IMP in 10 μg of rat brain poly(A)$^+$ RNA (compare lane 4 with lane 6). In order to obtain a more quantitative value, and to minimize inaccuracies due to loading differences, we normalized the radioactivity in the IMP spots using values determined for the amount of radioactivity in the AMP and CMP spots of our starting material (purification step 1; see Materials and methods). Although we had no convenient way to correct for errors that occurred during transfer to successive purification steps, control experiments showed that the direct transfer method gave quantitative recovery (data not shown). After normalization, we determined that the 10 μg of rat brain poly(A)$^+$ RNA analyzed in Figure 2 contained 1.4 pmol of IMP. Since 10 μg of poly(A)$^+$ RNA corresponds to ~30 nmol of NMP (assuming the molecular weight of each nucleotide is ~330 g/mol), this initial experiment indicated that, within rat brain poly(A)$^+$ RNA, there was approximately one IMP for every 21 000 nucleotides.

Importantly, the *in vitro* synthesized control RNA that was not spiked with IMP showed no detectable radioactivity that co-migrated with the UV-visualized IMP (Figure 2C, lane 1). This low background in the absence of exogenously added IMP was not by chance but took a great deal of trouble-shooting. In particular, we found that many commercially available reagents were contaminated with material, presumably 3′ NMPs, that became labeled during our protocol and resulted in substantial amounts of radioactivity that co-migrated with IMP. In our final protocol, we were able to exclude all but one of these reagents (see Materials and methods). The reagent we could not exclude was RNase T2, which was necessary for the first step of our labeling protocol (see Figure 1) and which appears to bind many nucleotides tenaciously at its active site, including 3′ IMP. The contaminating 3′ IMP becomes labeled during the protocol of Figure 1 and can be visualized in control experiments after very long exposure times (data not shown). Thus, the intrinsic contaminating IMP from this reagent defined our back-

ground (~0.5 fmol). Nevertheless, even in the presence of this background, our protocol was extremely sensitive and allowed the detection of 0.03 pmol of IMP in the background of 30 nmol of NMPs (1 part per million; data not shown).

IMP detected in mRNA was not due to contaminating tRNA

The oligo(dT) selection method we used to isolate the rat brain poly(A)$^+$ RNA is considered adequate for removing tRNA and, accordingly, we could not detect tRNA in our samples using various protocols to stain electrophoretically separated RNA (data not shown; see Materials and methods). However, we were worried that even an undetectable amount of contamination could yield the amounts of IMP we observed. Thus, we conducted additional experiments designed to rule out the possibility that the IMP observed in rat brain poly(A)$^+$ RNA was due to contaminating tRNA.

We first assayed whether the inosine-containing fraction of the poly(A)$^+$ RNA could be precipitated by high salt. The LiCl protocols we used were designed to precipitate large RNAs, such as mRNA and rRNA, while maintaining the solubility of small RNAs, such as tRNA and 5S RNA (Wallace, 1987). We reasoned that, if the IMP in our poly(A)$^+$ RNA derived from tRNA, it would not be recovered in the precipitated RNA, and we would not observe IMP in our final analysis. We tested the efficacy of our LiCl protocol by its ability to remove tRNA, and thus IMP, from the poly(A)$^-$ RNA fraction, i.e. the RNA that did not bind the oligo(dT) column. Figure 3A shows the amount of IMP purified from 10 μg of the poly(A)$^-$ RNA, with (lane 2) or without (lane 1) the LiCl precipitation steps. Clearly, the LiCl precipitation steps were very effective at excluding tRNA from the poly(A)$^-$ RNA. In contrast to its effect on the amount of IMP observed in the poly(A)$^-$ sample, essentially no change in the amount of IMP was observed in the poly(A)$^+$ sample after LiCl precipitation (compare lane 3 with lane 4; note that the small difference apparent by eye became insignificant when IMP amounts were normalized). Since mRNA, but not tRNA, is precipitated by high salt, this result was consistent with the idea that the IMP observed in the poly(A)$^+$ RNA derived from mRNA.

Another effective method for removing tRNA from an RNA sample is to centrifuge the RNA through high molarity CsCl (Sambrook *et al.*, 1989). Thus, we performed another analysis in which the total RNA sample was subjected to a CsCl centrifugation step prior to oligo(dT) selection. Figure 3B shows the amount of IMP observed in 10 μg of the total RNA starting material (lane 1), 10 μg of the RNA fraction that pelleted through a 5.7 M CsCl cushion (lane 2) and 10 μg of the RNA fraction that was too small to centrifuge through the CsCl cushion and remained in the supernatant (lane 3). Clearly, the CsCl cushion effectively excluded the IMP-containing RNA from the pellet (lane 2), which contained the rRNA and mRNA. Consistent with the idea that the IMP observed in lane 1 derived from tRNA, the amount of IMP per 10 μg was greater in the supernatant fraction (lane 3), where tRNA had been concentrated.

To determine if the CsCl centrifugation affected the amount of IMP observed in the poly(A)$^+$ RNA, poly(A)$^+$

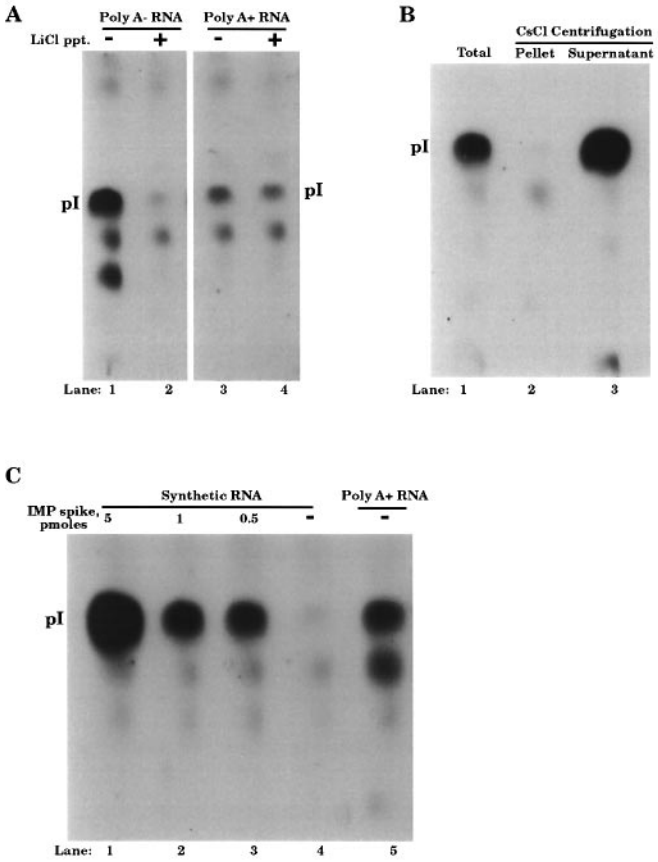

Fig. 3. Effects of high salt precipitation and CsCl centrifugation on the amount of IMP observed in poly(A)$^+$ RNA. RNAs were treated as outlined in Figure 1 and subjected to the three TLC purification steps of Figure 2. Only the purification step 3 analyses are shown; migration of 5′ IMP (pI) is indicated. (**A**) The autoradiogram shows nucleotides derived from 10 μg of rat brain poly(A)$^-$ or poly(A)$^+$ RNA, before (–) and after (+) precipitation with LiCl. The identity of radioactive spots that are not labeled to the right of the plate is unknown, but the spot of slowest migration in lane 1 is removed by LiCl precipitation, suggesting that it may derive from tRNA. (**B**) Instead of our normal procedure, which involves selecting poly(A)$^+$ RNA directly from total RNA, total RNA was first subjected to CsCl centrifugation. The autoradiogram shows IMP derived from 10 μg of the total rat brain RNA starting material (lane 1), from 10 μg of the CsCl pellet (lane 2) and from 10 μg of the CsCl supernatant (lane 3). (**C**) Oligo(dT) was used to purify poly(A)$^+$ RNA from the CsCl-pelleted RNA analyzed above (B, lane 2). The autoradiogram shows IMP derived from 10 μg of this poly(A)$^+$ RNA (lane 5), as well as that from 10 μg of synthetic RNA spiked with IMP as indicated (lanes 1–4).

RNA was isolated from the CsCl pellet by oligo(dT) selection, and subjected to the labeling protocol of Figure 1 and the three TLC purification steps used for the experiment of Figure 2. Figure 3C shows an autoradiogram of purification step 3. Ten μg of the poly(A)$^+$ fraction that had been purified by CsCl centrifugation and oligo(dT) selection contained 1.0 pmol of IMP (normalized value of lane 5, Figure 3C), a value that was almost identical to that found in previous samples not subjected to the CsCl purification step (e.g. Figure 2C, lane 6). Thus, beginning with two different RNA populations, that contained very different amounts of inosine-containing tRNA (compare lanes 1 and 2 of Figure 3B), our oligo(dT) method selected the same amount of inosine-containing poly(A)$^+$ RNA. Again, this meant that the IMP we observed in the poly(A)$^+$ RNA was not due to the non-specific contamina-

tion by tRNA, and that our normal oligo(dT) selection protocol was very effective in removing tRNA.

Although IMP was not readily visible in the RNA found in the CsCl pellet (lane 2, Figure 3B), a faint IMP spot became apparent after long exposure. Comparison of this faint spot with the spiked control RNAs run in parallel (data not shown) indicated that this RNA fraction (10 μg) contained ~0.1 pmol of IMP. At present we do not know if this IMP derives entirely from the mRNA in the CsCl pellet, or if some of it derives from the rRNA in the pellet; quantification of multiple experiments will be required to obtain a number accurate enough to make this determination. Nevertheless, the low amount of IMP observed in the CsCl pellet means that it is essentially impossible to attribute the IMP observed in the poly(A)$^+$ RNA to non-specific contamination by poly(A)$^-$ RNA, which after CsCl purification is largely rRNA. That is, contamination by 100 μg of the CsCl-purified poly(A)$^-$ RNA would be necessary to account for the ~1 pmol of IMP observed in only 10 μg of the poly(A)$^+$ RNA.

In summary, the results of Figure 3 demonstrated that the inosine-containing RNA within the poly(A)$^+$ RNA was precipitable by high salt and, in addition, could sediment through a 5.7 M CsCl cushion. Taken together, the results of Figure 3 indicated that the inosine-containing RNA in the poly(A)$^+$ fraction was mRNA, not tRNA.

IMP is found in mRNA at tissue-specific levels

For several reasons, we wanted to extend our analyses to additional tissues. First, we reasoned that differences between tissues would provide another line of evidence that observed IMP levels were intrinsic to mRNA, since it seemed unlikely that IMP derived from tRNA contamination would be tissue specific. Second, Northern analyses of human ADAR1 (Kim *et al.*, 1994; O'Connell *et al.*, 1995) and rat ADAR2 (Melcher *et al.*, 1996) indicate that the mRNAs for these enzymes are expressed in a tissue-specific manner, and we wondered if this tissue specificity might be reflected in overall IMP levels.

In addition to rat brain, we chose to analyze rat skeletal muscle, heart, lung and thymus. For all five tissues, multiple samples were analyzed from multiple animals using the protocol of Figure 1, and the three TLC purification steps of Figure 2A–C. A representative autoradiogram corresponding to a TLC plate for purification step 3 is shown in Figure 4A, and the average amounts of IMP observed, with standard deviations, are shown in Table I. For each experiment, IMP values were normalized as described previously.

Inspection of the data presented in Figure 4A and Table I revealed that the amount of IMP found in rat mRNA varied in a tissue-specific manner. Interestingly, the relative levels of IMP in different tissues correlated well with the levels of ADAR1 and ADAR2 mRNA found in various tissues. For example, Northern analyses of human ADAR1 (Kim *et al.*, 1994; O'Connell *et al.*, 1995) and rat ADAR2 (Melcher *et al.*, 1996) always show high level mRNA expression in brain, and very low level expression in skeletal muscle. Since our IMP analyses were performed with rat tissues, the rat ADAR2 Northern analyses are perhaps most relevant. In this case, the amount of IMP we observed in various tissues agreed well with published results that show tissue-specific expression of ADAR2

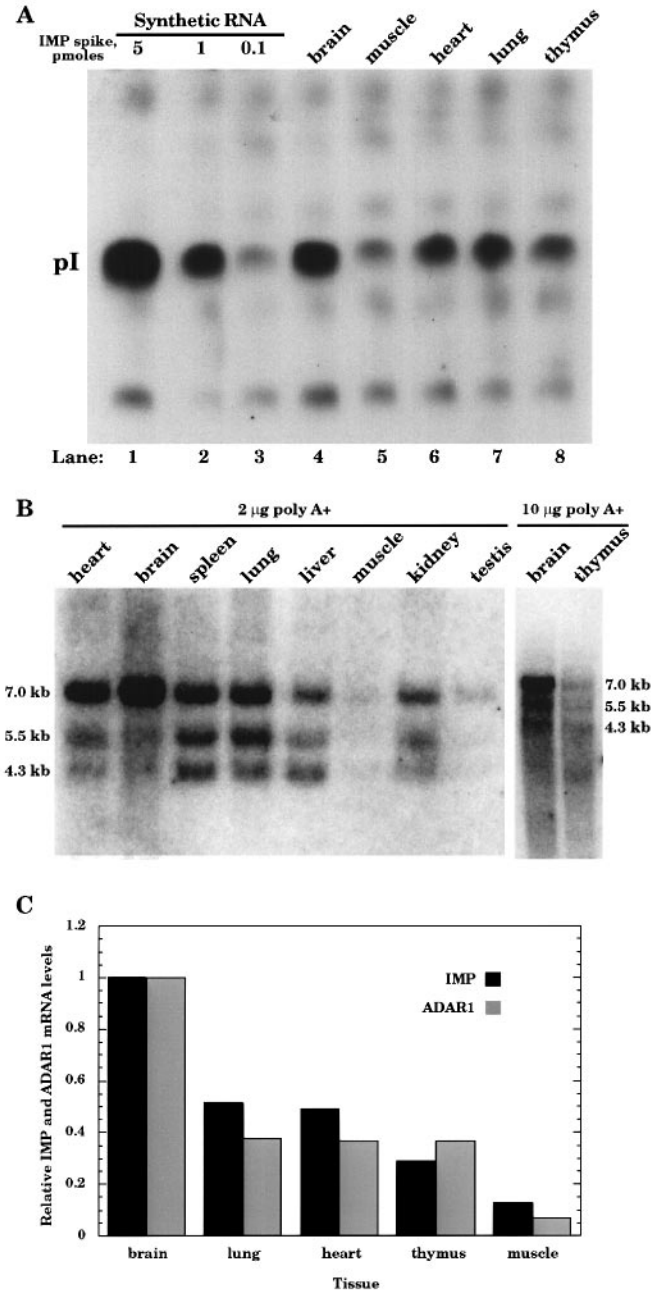

Fig. 4. Analyses of IMP and ADAR1 mRNA in rat tissues. (**A**) A purification step 3 analysis of IMP derived from 10 μg of poly(A)$^+$ RNA of various tissues, compared with that from 10 μg of synthetic RNAs spiked with various amounts of IMP. (**B**) Northern blot analyses of the amounts of ADAR1 in various rat tissues. The rat ADAR1 probe hybridized to three bands, and average lengths (kb; $n = 3$) were estimated from RNA markers. We do not know the mechanism that leads to multiple transcripts of rat ADAR1, but multiple ADAR1 transcripts have been observed in other organisms (as cited in Bass *et al.*, 1997). The probe was specific for ADAR1 and did not hybridize to ADAR2 (data not shown). (**C**) Amounts of IMP (black bars) or ADAR1 mRNA (gray bars) in various tissues were normalized to brain amounts. IMP values derive from data shown in Table I, and ADAR1 mRNA values were determined from Northern analyses (e.g. B). The height of each ADAR1 mRNA column represents an average of at least three determinations, made by quantifying the band of largest molecular weight (see B) on multiple Northern blots; estimates made by summing all three bands gave similar values. Relative values determined on different Northern blots varied, on average, by 22%.

Table I. Amounts of IMP in different rat tissues

Tissue	IMP, pmol[a] [in 10 μg of poly(A)$^+$]	No. of exp.	One IMP for every:
Brain	1.8 ± 0.9	6	17 000 nucleotides
Lung	0.9 ± 0.7	3	33 000 nucleotides
Heart	0.9 ± 0.6	3	33 000 nucleotides
Thymus	0.5 ± 0.3	3	60 000 nucleotides
Muscle	0.2 ± 0.1	4	150 000 nucleotides

[a]Each value is followed by a standard deviation (±).

following the pattern brain>lung>heart>skeletal muscle (Melcher *et al.*, 1996). To complement these published data, we performed several Northern analyses using a rat ADAR1 probe (Figure 4B), and quantified the relative amounts of ADAR1 in the poly(A)$^+$ RNA of various rat tissues. Again, the relative amount of ADAR1 expression correlated remarkably well with the relative amounts of IMP in various tissues (Figure 4B and C). Note that we also included several other tissues in our Northern analyses. Based on the strong correlation of ADAR expression and IMP levels, we would predict that spleen, liver and kidney mRNA would also contain significant levels of IMP (see Figure 4B).

Discussion

Here we show that inosine is a normal constituent of poly(A)$^+$ RNA, that can only be detected after it is purified away from the more abundant Watson–Crick nucleotides. We demonstrate that mRNA isolated from different tissues has different amounts of IMP, and that the relative amounts correlate with the tissue-specific expression of ADAR1 and ADAR2, enzymes known to convert A to I within RNA. Importantly, we quantify the amount of IMP contained within mRNA derived from different tissues, thus providing the first estimates for the number of inosine-containing RNAs yet to be discovered.

Purification of a nucleotide from cellular RNA

A feature of essential and key importance to our protocol was the addition of non-radioactive IMP, the 'spike', to the crude mixture of radiolabeled cellular nucleotides. This allowed us to monitor the IMP fraction by UV absorbance, during the initial steps of purification when the specific activity of radiolabeled cellular IMP was too low for detection. The methods we describe are straightforward and reproducible and should be amenable to the detection and quantification of IMP derived from any isolatable RNA population, of any organism. In addition, our purification methods should be applicable to the detection of other modified nucleotides that may be suspected to exist in cellular RNA. In this case, the only prerequisite is that the non-radioactive 5′ nucleotide corresponding to the particular nucleotide that is to be purified should be available in quantities sufficient to allow detection by UV.

How many inosine-containing RNAs are there?

Probably the most significant data of our study are those of Table I, where we have tabulated the amounts of IMP contained in mRNA of various tissues. These numbers allow predictions with regard to the number of inosine-containing RNAs yet to be discovered, and provide the first clues in regard to the importance of adenosine deamination to the regulation of gene expression. To date, very few RNAs have been proposed to be ADAR substrates (reviewed in Bass, 1997), and the amount of IMP we observed in mRNA was far in excess of what we expected. Although most of our analyses were performed on rat tissues, we also purified significant amounts of IMP from *Caenorhabditis elegans* poly(A)$^+$ RNA (data not shown). Thus, IMP is not limited to the mRNA of rat, but probably exists wherever ADARs are found, namely, metazoa.

As shown in Table I, among the tissues we analyzed, mRNA isolated from brain contained the highest amount of IMP [1.8 pmol in 10 μg of poly(A)$^+$ RNA], while skeletal muscle contained the least [0.2 pmol of IMP in 10 μg of poly(A)$^+$ RNA]. Since 10 μg of RNA corresponds to ~30 nmol of nucleotide, there was an average of one molecule of IMP for every 17 000 nucleotides of rat brain mRNA, or every 150 000 nucleotides of rat skeletal muscle mRNA. If all inosine-containing mRNAs contained, on average, a single inosine, and an average mRNA length of 1500 nucleotides was assumed, our numbers indicate that, on average, one out of every 11 rat brain mRNAs, or one out of every 100 skeletal muscle mRNAs, would contain inosine. If each of these inosine residues correlated with a biologically relevant codon change, certainly adenosine deamination would play a major role in the regulation of gene expression.

Of course, at present, we do not know what fraction of the observed IMP correlates with a biologically relevant codon change. In some cases, adenosine deamination could have a neutral effect, for example, by creating a synonymous codon. In addition, although the functionally important editing sites within gluR, serotonin receptor, HDV and α2,6-sialyltransferase RNAs occur in coding regions, any fraction of the IMP we observed could derive from non-coding sequences; in this case, adenosine deamination could function to alter the RNA structure or the binding of regulatory proteins. At present, we also do not know whether the purified IMP arises from the presence of one inosine residue in each of many different messages, or many inosine residues in a minimal number of mRNAs. Furthermore, our calculations say nothing about the complexity of the inosine-containing population of mRNA, which could be composed of a few very abundant RNAs, or many RNAs of very low abundance. Finally, we cannot rule out the possibility that many RNAs form fortuitous double-stranded structures at a low frequency, which are deaminated by ADARs without a biological consequence. However, given the large amounts of IMP we observe, even assuming an extreme scenario in which 90% of the observed IMP derives from such random events leads to the conclusion that there are many biologically relevant inosines yet to be discovered (see discussion below).

With regard to how many inosine-containing RNAs are yet to be identified, it is interesting to consider what proportion of the IMP we observed in rat brain poly(A)$^+$ RNA can be accounted for by the editing sites within rat brain gluR-B mRNA. Discounting a few sites within intronic sequences, which should not be present in our steady-state RNA population, gluR-B mRNA contains

three deamination sites (the Q/R, +4 and R/G editing sites). By sequencing cDNAs, it has been determined that 99% of the rat gluR-B mRNA population is edited at the Q/R site, and 38% at the nearby +4 site (Rueter *et al.*, 1995); editing at the R/G site varies with splice variant and developmental stage but reaches a maximum of ~85% (Lomeli *et al.*, 1994). Using quantitative Northern analyses, we have determined that 10 μg of rat brain poly(A)$^+$ RNA contains ~0.5 fmol of gluR-B mRNA (M.Paul and B.Bass, unpublished data). Thus, in 10 μg of poly(A)$^+$ RNA, only 1 fmol of IMP should derive from inosine within gluR-B mRNA [(0.5 fmol)×(0.99 + 0.38 + 0.85)]. Since this represents only 0.06% of the IMP we observed in rat brain, if gluR-B mRNA typified all inosine-containing mRNAs, there would be ~1800 different inosine-containing RNAs yet to be identified in mammalian brain.

Of course, it is possible that the gluR-B transcript is not representative of all inosine-containing RNAs. In fact, based on the number of A to G transitions mapped in cDNA sequences of putative ADAR substrates, it appears that there are two types of substrates (Bass, 1997): those whose cDNAs show many A to G transitions (hypermutation or low selectivity substrates) or very few (high selectivity substrates). So far, all of the RNAs for which ADARs are proposed to edit specific codons (e.g. gluR, serotonin receptor, HDV RNAs) fall into the high selectivity category; in a given base-paired region, <10% of the adenosines appear as guanosines in cDNA sequences. In contrast, in the hypermutation or low selectivity type of sequence, ~50% of the adenosines in a given base-paired region appear as guanosines in cDNAs. (It has been proposed that hypermutations occur within RNA substrates that are completely double-stranded, while the more selective type of deamination occurs within RNAs whose base pairs are interrupted by mismatches, bulges and loops; Bass, 1997). Although a number of viral RNAs appear to show the hypermutation type of deamination, only one cellular mRNA, the *Drosophila* 4f-rnp mRNA, shows the hypermutation type of deamination (Petschek *et al.*, 1996). Unfortunately, there are no estimates for the abundance of the 4f-rnp mRNA, and no similar hypermutation type cDNAs have been identified for mammals. However, if we assume one-half of the adenosines in a 1500 nucleotide RNA were deaminated (assuming equal parts of the Watson–Crick nucleotides), so that each inosine-containing mRNA contained ~188 inosines, the amount of IMP we observed in rat brain poly(A)$^+$ RNA could be achieved if only one out of 2130 transcripts was an inosine-containing RNA. Or taking another view, if we assume a single mRNA of the same abundance as the gluR-B mRNA (~0.5 fmol/10 μg), and of average length (1500 nucleotides), deamination of 50% of its adenosines could account for 5% of the IMP we observed in rat brain (94 fmol/1.8 pmol).

Although the function of the hypermutation type of deaminations is unclear, at least in some cases it is very clear that the more selective type of deamination functions to alter codon meaning, thus allowing multiple proteins to be synthesized from a single encoded sequence. Certainly the identification of additional inosine-containing sequences will allow more insight into the function of adenosine deamination. We hope that our data will guide such investigations by directing studies towards tissues whose mRNA contains an abundance of IMP. Although we have only analyzed five tissues in this study, the strong correlation of IMP levels with ADAR levels suggests that any tissue where ADARs are expressed at high levels will have a correspondingly high level of IMP within its mRNA.

Materials and methods

RNA isolation

Poly(A)$^+$ RNA was isolated directly from tissue (Pel-Freez) using a FastTrack 2.0 mRNA isolation kit (InVitrogen). Alternatively, as indicated, tissues were homogenized in 4 M guanidinium isothiocyanate, 20 mM NaOAc (pH 5.5), 0.1 mM dithiothreitol (DTT) and 0.5% *n*-lauryl sarcosine, loaded onto a 5.7 M CsCl, 0.01 M EDTA cushion (~200 μg RNA/cushion) and centrifuged (SWTi 40 rotor, 30 000 r.p.m., 22 h). Pellets were resuspended in H_2O, ethanol precipitated, then poly(A)$^+$ RNA was selected as above (FastTrack 2.0). The supernatant was extracted successively with H_2O-saturated phenol, phenol:chloroform (1:1, v/v) and chloroform, then ethanol precipitated. When indicated, high salt precipitation was performed on poly(A)$^+$ and poly(A)$^-$ RNA as described (Wallace, 1987); the entire procedure of two LiCl precipitations was then repeated. When poly(A)$^-$ RNA was required, the oligo(dT) flow-through was extracted as for the CsCl supernatant. Aliquots of poly(A)$^+$ RNA samples were electrophoresed prior to use, and rRNA quantified with cyber green II (Molecular Probes); the average amount of rRNA contamination was 9.3 ± 0.6%. Transfer RNA was not detected in the poly(A)$^+$ RNA after staining with either cyber green II or ethidium bromide. Synthetic RNA was transcribed using T3 polymerase MAXIScript (Ambion). To validate the quality of commercially obtained tissues, in some cases, RNA was isolated from rat brains dissected in our own laboratory; freshly dissected brains were frozen in liquid nitrogen moments after the rats were sacrificed. Resultant IMP values were identical to those determined with commercially obtained samples.

Preparation of 5′ ^{32}P-NMPs

Ten μg of rat poly(A)$^+$ or synthetic RNA was incubated (65°C, 5 min), snap-cooled on dry ice, and digested to 3′ NMPs with 5 U of RNase T2 (Gibco-BRL) in 50 mM NaOAc (pH 4.5) for 30 min at 37°C (25 μl final). Reactions were again heated (65°C, 5 min), and the process repeated. Samples were extracted with phenol:chloroform, chloroform, then ether, followed by lyophilization.

Lyophilized RNase T2 digestions were resuspended in 4 μl of H_2O, and reacted (37°C, 30 min) with 1 μM [γ-^{32}P]ATP (Amersham, 6 000 Ci/mmol, 150 mCi/ml) and 6 U of T4 PNK (Pharmacia) in 2× One-Phor-All Buffer Plus (Pharmacia) supplemented with 5 mM DTT (final volume, 10 μl), followed by heat-inactivation of PNK (75°C, 10 min). The volume was adjusted by adding 36.4 μl of H_2O, and [^{32}P]Nps were digested to 5′ [^{32}P]NMPs with 1.5 U of nuclease P1 (BMB; 37°C, 30 min). Reactions were phenol:chloroform-, chloroform- and ether-extracted, then dried. PNK reactions were optimized by varying ATP and 3′ IMP concentrations.

Control experiments showed that spermidine, molecular biology grade glycogen and RNase T2 were contaminated with 3′ IMP and other 3′ NMPs, that became labeled during our protocol (Figure 1). Spermidine and glycogen could be excluded from all protocols, but RNase T2 was essential. Attempts to remove contaminating nucleotides by size-exclusion chromatography failed, presumably because the contaminating nucleotides are tightly bound by RNase T2. Thus, the contaminating IMP from the RNase T2 defined our background (~0.5 fmol/reaction).

Purification by thin-layer chromatography

For purification step 1, samples were resuspended in 10 μl of H_2O plus 1 μl of 100 mM non-radioactive 5′ IMP, and spotted in 2.5 μl aliquots onto a PEI thin layer plate, 2.5 cm from the plate bottom, and 2.5 cm from its right side. After loading, plates were soaked in anhydrous methanol (10 min), dried, then developed in two dimensions as described (Gupta *et al.*, 1976), except that plates were not neutralized between dimensions. Thin layers were exposed on film (30 s) and non-radiolabeled 5′ IMP was visualized with UV light, outlined with a pencil, and cut out with scissors. The matrix side of the cut-outs was brought into contact with a fresh PEI thin layer plate pre-soaked in anhydrous methanol (Randerath and Randerath, 1980), in preparation for purification

step 2. For purification step 2, samples were re-chromatographed with H_2O to the origin, then 1 M sodium formate buffer (pH 3.4) to 4 cm onto a Whatmann 3MM paper wick stapled to the top of the plate. Plates were dried, soaked in anhydrous methanol (10 min), exposed to film (15–30 min) and non-radioactive 5′ IMP excised as before. For the third purification step, the matrix side of each cut-out was brought into contact with a cellulose TLC plate (Kodak) and re-chromatographed in saturated NH_4SO_4:isopropanol:0.1 M NaOAc, pH 6.0 (79:2:19). Plates were dried, non-radiolabeled 5′ IMP was outlined, and autoradiography performed.

Radioactivity was quantified with a Molecular Dynamics PhosphorImager using ImageQuant software. The radioactivity in IMP spots (purification step 3) was normalized using values determined for the amount of radioactivity in the AMP and CMP spots of the starting material (purification step 1, see Figure 2A). Analysis of the excised AMP and CMP spots on additional TLC systems showed no contaminating material, indicating that the radioactivity in these spots was directly proportional to the amount of AMP and CMP in the sample. Comparison of the AMP and CMP spots between samples required identical autoradiography or PhosphorImager exposure times. Because the first purification step was a 2D TLC system, and each sample was run on a different TLC plate, the size of available film and PhosphorImaging screens was too small to expose the TLC plates for all samples simultaneously. Thus, for each sample, the region of the TLC plate containing the AMP and CMP spots was excised, and the excised pieces for all samples were placed adjacent to one another and re-exposed. IMP amounts were calculated by linear regression analysis, using the normalized amounts of 5′ [^{32}P]IMP observed in the spiked RNAs to construct a standard curve.

Northern blot analyses

Rat poly(A)$^+$ RNA was electrophoresed, then transferred to a nylon membrane (Sambrook *et al*., 1989), or rat Multiple Tissue Northern Blots were used (Clontech). Blots were probed with a 543 bp PCR fragment (random-primed with [α-^{32}P]dCTP) corresponding to the 5′ most coding sequence of ADAR1 (forward primer, GGGCCC-ACTATGGCTCAGGGGTTCAGG; reverse primer, GTCACACGGCA-GCAAACAATATACGATTGA TGTCC). Radioactivity was quantified as above.

Acknowledgements

We thank K.Randerath, J.McCloskey, G.Carmichael and members of our laboratory for helpful advice and discussion. We thank D.Morse for critically reading this manuscript. This work was supported by funds to B.L.B. from the National Institute of General Medical Sciences, National Institutes of Health (GM 44073) and the David and Lucile Packard Foundation. B.L.B. is a Howard Hughes Medical Institute associate investigator.

References

Basilio,C., Wahba,A.J., Lengyel,P., Speyer,J.F. and Ochoa,S. (1962) Synthetic polynucleotides and the amino acid code, V. *Proc. Natl Acad. Sci. USA*, **48**, 613–616.

Bass,B.L. (1997) RNA editing and hypermutation by adenosine deamination. *Trends Biochem. Sci.*, **22**, 157–162.

Bass,B.L. and Weintraub,H. (1988) An unwinding activity that covalently modifies its double-stranded RNA substrate. *Cell*, **55**, 1089–1098.

Bass,B.L., Nishikura,K., Keller,W., Seeburg,P.H., Emeson,R.B., O'Connell,M.A., Samuel,C.E. and Herbert,A. (1997) A standardized nomenclature for adenosine deaminases that act on RNA. *RNA*, **3**, 947–949.

Benne,R., Van den Burg,J., Brakenhoff,J.P., Sloof,P., Van Boom,J.H. and Tromp,M.C. (1986) Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell*, **46**, 819–826.

Burns,C.M., Chu,H., Rueter,S.M., Hutchinson,L.K., Canton,H., Sanders-Bush,E. and Emeson,R.B. (1997) Regulation of serotonin 2C receptor G-protein coupling by RNA editing. *Nature*, **387**, 303–308.

Engel,J.D. and von Hippel,P.H. (1978) Effects of methylation on the stability of nucleic acid conformations. Studies at the polymer level. *J. Biol. Chem.*, **253**, 927–934.

Grosjean,H. *et al*. (1996) Enzymatic conversion of adenosine to inosine and to N1-methylinosine in transfer RNAs: a review. *Biochimie*, **78**, 488–501.

Gupta,R.C., Randerath,E. and Randerath,K. (1976) An improved separation procedure for nucleoside monophosphates on polyethyleneimine-(PEI)-cellulose thin layers. *Nucleic Acids Res.*, **3**, 2915–2921.

Kable,M.L., Heidmann,S. and Stuart,K.D. (1997) RNA editing: getting U into RNA. *Trends Biochem. Sci.*, **22**, 162–166.

Kim,U., Wang,Y., Sanford,T., Zeng,Y. and Nishikura,K. (1994) Molecular cloning of cDNA for double-stranded RNA adenosine deaminase, a candidate enzyme for nuclear RNA editing. *Proc. Natl Acad. Sci. USA*, **91**, 11457–11461.

Limbach,P., Crain,P. and McCloskey,J. (1994) Summary: the modified nucleosides of RNA. *Nucleic Acids Res.*, **22**, 2183–2196.

Lomeli,H. *et al*. (1994) Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. *Science*, **266**, 1709–1713.

Ma,J., Qian,R., Rausa,F.M.I. and Colley,K.J. (1997) Two naturally occurring alpha2,6-sialyltransferase forms with a single amino acid change in the catalytic domain differ in their catalytic activity and proteolytic processing. *J. Biol. Chem.*, **272**, 672–679.

Melcher,T., Maas,S., Higuchi,M., Keller,W. and Seeburg,P.H. (1995) Editing of alpha-amino-3-hydroxy-5-methylisoxazole-4-propionic acid receptor GluR-B pre-mRNA *in vitro* reveals site-selective adenosine to inosine conversion. *J. Biol. Chem.*, **270**, 8566–8570.

Melcher,T., Maas,S., Herb,A., Sprengel,R., Seeburg,P.H. and Higuchi,M. (1996) A mammalian RNA editing enzyme. *Nature*, **379**, 460–464.

Morse,D.P. and Bass,B.L. (1997) Detection of inosine in messenger RNA by inosine-specific cleavage. *Biochemistry*, **36**, 8429–8434.

Narayan,P. and Rottman,F.M. (1992) Methylation of mRNA. *Adv. Enzymol.*, **65**, 255–285.

O'Connell,M.A., Krause,S., Higuchi,M., Hsuan,J.J., Totty,N.F., Jenny,A. and Keller,W. (1995) Cloning of cDNAs encoding mammalian double-stranded RNA-specific adenosine deaminase. *Mol. Cell. Biol.*, **15**, 1389–1397.

Petschek,J.P., Mermer,M.J., Scheckelhoff,M.R., Simone,A.A. and Vaughn,J.C. (1996) RNA editing in *Drosophila* 4f-rnp gene nuclear transcripts by multiple A-to-G conversions. *J. Mol. Biol.*, **259**, 885–890.

Polson,A.G., Bass,B.L. and Casey,J.L. (1996) RNA editing of hepatitis delta virus antigenome by dsRNA-adenosine deaminase. *Nature*, **380**, 454–456.

Randerath,K. and Randerath,E. (1980) Thin-layer separation methods for nucleic acid derivatives. *Methods Enzymol.*, **12A**, 323–334.

Rueter,S.M., Burns,C.M., Coode,S.A., Mookherjee,P. and Emeson,R.B. (1995) Glutamate receptor RNA editing *in vitro* by enzymatic conversion of adenosine to inosine. *Science*, **267**, 1491–1494.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Seeburg,P.H. (1996) The role of RNA editing in controlling glutamate receptor channel properties. *J. Neurochem.*, **66**, 1–5.

Simpson,L. and Emeson,R. (1996) RNA editing. *Annu. Rev. Neurosci.*, **19**, 27–52.

Smith,H.C., Gott,J.M. and Hanson,M.R. (1997) A guide to RNA editing. *RNA*, **3**, 1105–1123.

Sommer,B., Köhler,M., Sprengel,R. and Seeburg,P.H. (1991) RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell*, **67**, 11–19.

Wallace,D.M. (1987) Precipitation of nucleic acids. *Methods Enzymol.*, **152**, 41–48.

Yang,J.H., Sklar,P., Axel,R. and Maniatis,T. (1995) Editing of glutamate receptor subunit B pre-mRNA *in vitro* by site-specific deamination of adenosine. *Nature*, **374**, 77–81.

Received October 7, 1997; revised November 25, 1997; accepted December 3, 1997