The properties of a tRNA-specific adenosine deaminase from Drosophila melanogaster support an evolutionary link between pre-mRNA editing and tRNA modification

Citation for published version:

Digital Object Identifier (DOI):
10.1128/MCB.20.3.825-833.2000

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Molecular and Cellular Biology

Publisher Rights Statement:
Copyright © 2000, American Society for Microbiology

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
The Properties of a tRNA-Specific Adenosine Deaminase from *Drosophila melanogaster* Support an Evolutionary Link between Pre-mRNA Editing and tRNA Modification

LIAM P. KEEGAN,1 ANDRÉ P. GERBER,2 JIM BRINDLE,1 RONNY LEEHANS,3 ANGELA GALLO,1 WALTER KELLER,2 AND MARY A. O’CONNELL1*

MRC Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, United Kingdom, and Department of Cell Biology, Biozentrum of the University of Basel, CH-4056 Basel,2 and Zoological Institute, University of Basel, CH-4051 Basel,3 Switzerland

Received 16 August 1999/Returned for modification 30 September 1999/Accepted 1 November 1999

Adenosine and cytosine have exocyclic amino groups that participate in Watson-Crick base pairing during transcription and translation. RNA editing enzymes have been discovered that deaminate specific adenosine residues to inosine (4, 32, 38, 40) or specific cytosine residues to uridine in RNA molecules (11, 39). This can result in the incorporation of different amino acids at edited positions or the formation of a smaller protein due to the generation of a stop codon. Two closely related adenosine deaminases acting on RNA (ADAR1 and ADAR2) have been identified in vertebrates (3) that catalyze the deamination of specific adenosine residues to inosine in pre-mRNAs (for a review, see reference 22). These enzymes have homologous adenosine deaminase domains (32) and also contain multiple double-stranded RNA(dsRNA)-binding domains (47). ADARs recognize and deaminate specific adenosines within exons that form duplexes with flanking intronic sequences in pre-mRNAs (20, 32). ADAR activity is ubiquitous and has been found in all metazoans tested and in most tissues (52). The abundance of inosine in polyA+ RNA has been estimated to be 1 in 17,000 nucleotides in brain and less in other mammalian tissues, correlating with ADAR expression levels (35). Inosine in edited transcripts directs the incorporation of cytosine during first-strand synthesis of cDNA (5), and RNA editing events have usually been identified in cDNA sequences in which guanosine replaces a genomically encoded adenosine (8, 45). Pre-mRNAs encoding subunits of the glutamate-gated ion channels (for a review, see reference 42) and the G protein-coupled serotonin 2C receptor (8) undergo RNA editing of their sequences by this mechanism. Proteins encoded by edited mRNAs often have functional properties that differ from the genomically encoded versions.

Inosine was first observed as a noncanonical base occurring in a number of tRNAs (21). Inosine in tRNAs is generated by deamination of genomically encoded adenosine (2). Inosine occurs at position 34 in the anticodon in a number of different tRNAs (for a review, see reference 17). Eukaryotic alanine tRNA (tRNAAla), containing the anticodon ICyG, is the only class of tRNA that undergoes deamination of adenosine to inosine at position 37 adjacent to the anticodon (17). This inosine is subsequently methylated by an as-yet-uncharacterized enzyme (2). Recently, it has been found that inosine at position 34 and methylinosine at position 37 are major epitopes for anti-PL-12 myositis autoantibodies (6). An adenosine deaminase acting on tRNA (ADAT1) catalyzing the site-specific deamination of adenosine at position 37 in yeast tRNAAla has been cloned and characterized (15). This ADAT1 from *Saccharomyces cerevisiae*, named scTad1p (or scADAT1) and encoded by the *TADI* gene, is homologous to the ADARs throughout its sequence and has three characteristic adenosine deaminase motifs containing residues thought to chelate zinc and contribute to catalysis in the active site of the enzymes (9). Gerber et al. (15) have proposed that ADARs involved in pre-mRNA editing have a common evolutionary origin with ADATs involved in the modification of tRNAAla at position 37. These adenosine deaminase domains are related to a cytosine deaminase involved in mRNA editing (APOBEC) and to cytidine deaminases involved in the deamination of free nucleosides (9).

Embryonic nuclear extracts from *Drosophila melanogaster* contain an ADAR-like activity that converts adenosine to inosine in the antigenome RNA of hepatitis D virus (10). Editing of mRNAs has been proposed to occur in *Drosophila*, based on the detection of variant cDNAs having adenosine-to-guanosine substitutions in mRNAs encoded by the 4f-mp gene (37), the *cacophony* (cac) gene (36, 44), and the *paraletic* (para) gene (18). We describe here the cloning and characterization of a gene encoding an adenosine deaminase from *D. melanogaster*.

* Corresponding author. Mailing address: MRC Human Genetics Unit, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, United Kingdom. Phone: 44-131-467 8417. Fax: 44-131-343 2620. E-mail: Mary.O’Connell@hgu.mrc.ac.uk.
that is similar to vertebrate ADARs but lacks dsRNA-binding domains. Characterization of a purified recombinant protein shows that this protein is a tRNA-specific adenosine deaminase. This protein, dADAT1, has a higher degree of amino acid sequence homology to the ADARs than the previously characterized yeast ADAT1. These findings support an evolutionary relationship between pre-mRNA editing and tRNA modification.

**MATERIALS AND METHODS**

**Oligonucleotides used in this study.** The oligonucleotides used in this study are as follows: DR5 (5'-GGATCCCGAGAATGGTACGATC-3'), DR3 (5'-AAGCTTAAATGCTTCTAACTCAG-3'), dADAT81R (5'-GCTGTCATCCGCAATTTGACCA-3'), dADAT585R (5'-TCTGAAATTTACTTAAGCTTGGAC-3'), CCAC-3'),gt10d (5'-AGCAAGTACGCGTGGTATAGGT-3'), gt10I (5'-CGGCTGTCATCTGTGACGTC-3'),gt10A (5'-CCGCAATCTCCTAACAG-3'), dADAT585R (5'-CCGGCATTACGATAGTGA-3'), dADAT81 (5'-CTTI GTTCCGCCATCAGGCG-3'), and (d)13 adapter (5'-GAGCTGACCTGACATTCCCGGCCCCC-3').

**Isolation of cDNA clones and sequencing.** The P1 clone DS00941, sequenced by the Berkeley Drosophila Genome Project (BDGP), carries an open reading frame (ORF) of 1,940 bp that encodes the deaminase domain of mammalian ADAR1 and ADAR2. A 600-nucleotide fragment encoding this homologous domain was amplified from genomic DNA by PCR. The PCR product was subsequently subcloned in the T/A cloning vector pGEM-Teasy (Promega) and sequenced with a Dye Terminator Cycle Sequencing Ready Reaction kit on an Applied Biosystems 373A sequencer.

**Restriction enzyme mapping of dAdat1.** The P1 clone DS00941, sequenced by the Berkeley Drosophila Genome Project (BDGP), carries an open reading frame (ORF) of 1,940 bp that encodes the deaminase domain of mammalian ADAR1 and ADAR2. A 600-nucleotide fragment encoding this homologous domain was amplified from genomic DNA by PCR. The PCR product was subsequently subcloned in the T/A cloning vector pGEM-Teasy (Promega) and sequenced with a Dye Terminator Cycle Sequencing Ready Reaction kit on an Applied Biosystems 373A sequencer.

**A human homologue to the dAdat1 gene was identified in the Washington University expressed sequence tag (EST) database in a BLAST search with S. cerevisiae ADAT1. This clone (emb/AA16179/SAA16179), isolated from a StrataClone human lymphoblast library, was obtained from the IMTEK consortium (26). Further sequencing of this 1.9-kb cDNA clone indicated that the sequence in the database was from the 5' end of the clone and that the region encoding the amino terminus was missing. In order to obtain additional 5' coding sequence, PCR primers were designed from sequencing data at the 3' end and used to amplify the cDNA library (Clontech) with gt10d and an hADAT reverse primer, hADAT585R (5'-CTTGGATGAGTGA-3'). A 682-bp PCR fragment was subsequently subcloned in the T/A cloning vector pGEM-Teasy (Promega) and sequenced with a Dye Terminator Cycle Sequencing Ready Reaction kit on an Applied Biosystems 373A sequencer. The P1 clone DS00941, sequenced by the Berkeley Drosophila Genome Project (BDGP), carries an open reading frame (ORF) of 1,940 bp that encodes the deaminase domain of mammalian ADAR1 and ADAR2. A 600-nucleotide fragment encoding this homologous domain was amplified from genomic DNA by PCR. The PCR product was subsequently subcloned in the T/A cloning vector pGEM-Teasy (Promega) and sequenced with a Dye Terminator Cycle Sequencing Ready Reaction kit on an Applied Biosystems 373A sequencer.

A human homologue to the dAdat1 gene was identified in the Washington University expressed sequence tag (EST) database in a BLAST search with S. cerevisiae ADAT1. This clone (emb/AA16179/SAA16179), isolated from a StrataClone human lymphoblast library, was obtained from the IMTEK consortium (26). Further sequencing of this 1.9-kb cDNA clone indicated that the sequence in the database was from the 5' end of the clone and that the region encoding the amino terminus was missing. In order to obtain additional 5' coding sequence, PCR primers were designed from sequencing data at the 3' end and used to amplify the cDNA library (Clontech) with gt10d and an hADAT reverse primer, hADAT585R (5'-CTTGGATGAGTGA-3'). A 682-bp PCR fragment was subsequently subcloned in the T/A cloning vector pGEM-Teasy (Promega) and sequenced with a Dye Terminator Cycle Sequencing Ready Reaction kit on an Applied Biosystems 373A sequencer. The P1 clone DS00941, sequenced by the Berkeley Drosophila Genome Project (BDGP), carries an open reading frame (ORF) of 1,940 bp that encodes the deaminase domain of mammalian ADAR1 and ADAR2. A 600-nucleotide fragment encoding this homologous domain was amplified from genomic DNA by PCR. The PCR product was subsequently subcloned in the T/A cloning vector pGEM-Teasy (Promega) and sequenced with a Dye Terminator Cycle Sequencing Ready Reaction kit on an Applied Biosystems 373A sequencer. The P1 clone DS00941, sequenced by the Berkeley Drosophila Genome Project (BDGP), carries an open reading frame (ORF) of 1,940 bp that encodes the deaminase domain of mammalian ADAR1 and ADAR2. A 600-nucleotide fragment encoding this homologous domain was amplified from genomic DNA by PCR. The PCR product was subsequently subcloned in the T/A cloning vector pGEM-Teasy (Promega) and sequenced with a Dye Terminator Cycle Sequencing Ready Reaction kit on an Applied Biosystems 373A sequencer.
overnight following standard protocols (49). Hybridized transcripts were detected with an alkaline phosphatase-conjugated anti-digoxigenin Fab fragment (Boehringer Mannheim), with nitroblue tetrazolium (Sigma) and 5-bromo-4-chloro-3-indolylphosphate (BCIP) (Sigma) as chromogenic substrates. Embryos were whole mounted in Canada balsam (Serva) and photographed with an Olympus 35-mm camera on an Olympus Provus microscope with differential interference contrast optics.

**RESULTS**

Characterization of a *Drosophila* gene encoding an adenosine deaminase domain. Homology searches with hADAR1 identified an ORF encoding an ADAR-like adenosine deaminase domain in a 3-kb subclone of *D. melanogaster* P1 genomic clone DS00941 that had been sequenced by the BDGP (Fig. 1). The sequence conservation in the predicted adenosine deaminase domain included the three motifs containing a histidine and two cysteine residues, proposed to be involved in zinc coordination at the active site, as well as a conserved glutamate residue in motif I which is also believed to participate in the deamination reaction (23, 34). The coding sequence in DS00941 did not contain any potential dsRNA-binding domain (47).

The genomic sequence from DS00941 was used to design PCR primers to amplify a 600-bp DNA fragment encoding this putative adenosine deaminase sequence from genomic DNA (Fig. 1), which was then used to screen embryonic cDNA libraries. Low-stringency Southern blots with this fragment indicated that the gene is single copy (results not shown). One long cDNA clone (clone 12) and another shorter clone (clone H) were isolated, confirming that this ORF is part of a transcribed sequence.

The P1 clone DS00941 is now fully sequenced by the BDGP and is part of a 2.9-Mb stretch of contiguous sequence in the *Adh* region on the left arm of *Drosophila* chromosome II. The gene encoding the adenosine deaminase domain is the second predicted gene product from the centromere distal end of this 82-kb P1 clone and is referred to as BG:DS00941.2 (1). The relationship between the isolated cDNAs and the structure of the BG:DS00941.2 transcript as predicted from the genomic sequence by *Drosophila* GRAIL and Genefinder programs are indicated in Fig. 1, together with the 600-nucleotide fragment that was used as a probe to isolate the cDNAs. Comparison of cDNA sequences with the genomic DNA sequence revealed that the cDNA clones are interrupted by a short intron of 57 bp within the protein coding sequence and by a second short intron of 54 bp within the 3′ untranslated region of the longer cDNA clone. Clone H begins at base pair 574 of clone 12 and terminates at a more proximal polyadenylation site. Two different polyadenylation sites are used, one located 30 bp after the ORF and a second site located 1.8 kb more 3′, which produces a transcript with a long 3′ region in which no additional ORFs can be identified. Both polyadenylation sites are preceded by a polyadenylation consensus sequence.

Clone 12 encodes an ORF (from nucleotides 23 to 1207) encoding a protein of 394 amino acids (Fig. 2) that contains a putative ADAR-type adenosine deaminase domain but lacks a dsRNA-binding domain. The predicted adenosine deaminase domain shows 38% identity and 47% similarity to human ADAR2, whereas the *S. cerevisiae* Tad1p has 25% identity and 37% similarity to human ADAR2 (Wisconsin Package; best fit, gap weight 8 and length weight 2) (13). The predicted *Drosophila* adenosine deaminase domain shows 29% identity and 40% similarity to the *S. cerevisiae* Tad1 protein.

This was the first gene of the ADAR family identified in *Drosophila* and we considered it a candidate enzyme for catalyzing the proposed editing events in the *cac* and *para* pre-mRNAs (18, 36, 44). Since the sequence of the encoded deaminase domain resembles vertebrate ADAR2 more than yeast Tad1p, we initially considered the possibility that the ORF was the 3′ end of a longer mRNA that would also encode dsRNA-binding domains. We therefore directed considerable effort to search for cDNAs that could encode a longer protein. Other embryonic cDNA libraries were screened and, despite the isolation of additional positive clones, none contained ORFs longer than that of clone 12. Neither amplification of 5′ ends from embryonic cDNA libraries nor 5′ rapid amplification of DNA ends with poly(A) + RNA from embryos, larvae, or adults gave any indication of an RNA significantly longer than clone 12 that could encode additional amino acids on the 5′ end. Northern blot analysis of poly(A) + RNA from embryos with a fragment encoding the deaminase domain used as a probe revealed two bands, one of approximately 4.3 kb and a shorter one of 3 kb, which is the same size as clone 12 (data not shown). We therefore conclude that the 4.3-kb band must encode an mRNA with an extended 3′ untranslated region.

The genomic sequence upstream of the start of clone 12 contains no good consensus for a TATA box, but there is a consensus for an initiator (Inr) element (27) 15 nucleotides from the predicted start codon.

A rabbit polyclonal antiserum was raised against the deaminase domain. The BamHI-HindIII fragment encoding part of the deaminase domain was subcloned into the *E. coli* histidine tag expression vector pTREHisB to obtain a fusion protein with six histidines at the N terminus. This protein was purified and used for polyclonal antibody production in rabbits. Antiserum raised against the adenosine deaminase domain recognized a protein of 43.5 kDa, which is the size predicted from the coding sequence in clone 12, on immunoblots of *Drosophila* embryonic nuclear extracts (data not shown).

BG:DS00941.2 encodes a *Drosophila* ADAT that deaminates adenosine 3′ in a silk moth tRNA-alanine substrate. The
FIG. 2. Multiple sequence alignment showing the Drosophila adenosine deaminase encoded by clone 12, compared to vertebrate ADARs and yeast scTad1p and spTad1p. The adenosine deaminase motifs containing cysteine and histidine residues that are proposed zinc ligands (#) and a glutamate residue (Δ) that is thought
methylo-trophic yeast *P. pastoris* was used to overexpress the protein encoded by clone 12 so as to determine its function. Previous work has shown this yeast to be excellent for the production of active ADARs and ADATs (15, 16, 33). The coding sequence of clone 12 was subcloned into the *P. pastoris* expression vector pSK-FLIS6 to express a fusion protein with a FLAG epitope at the amino terminus and six histidine residues at the carboxy terminus. This construct was transformed into *P. pastoris* by standard methods. Extracts from transformed cells contain a protein with a molecular mass of 48 kDa detectable on immunoblots with α-FLAG M2 antibody. This molecular mass is consistent with the length of the coding sequence of clone 12 and the N- and C-terminal epitope tags (results not shown).

Extracts were made from 300-ml cultures grown in methanol medium. Soluble extract was mixed with Ni²⁺-NTA, poured into a column, and eluted with 250 mM imidazole. This eluate was applied to an affinity column bearing the α-FLAG M2 monoclonal antibody, and proteins were eluted with FLAG peptide. Figure 3A shows SDS-polyacrylamide gel electrophoresis of fractions from a FLAG affinity column in which the proteins have been stained with Coomassie blue. A protein of approximately 40 kDa is present in the eluate fractions, and this protein is recognized by the antibody raised against the deaminase domain (Fig. 3B). Approximately 40 μg of pure dADAT1 was obtained per 300 ml of starting culture.

Peak fractions containing the 48-kDa protein from the FLAG affinity column were tested for adenosine deaminase activity on in vitro-synthesized [α-³²P]ATP-labelled tRNA*Ala* from the insect *B. mori* (2, 46, 51) or from *S. cerevisiae*. scTad1p has been shown to specifically deaminate the adenosine position at 37 in yeast and *B. mori* tRNA*Ala* (15). The recombinant 48-kDa *Drosophila* protein converts adenosine to inosine in tRNA*Ala* from *B. mori* with a specific activity of 8.4 U/mg (Fig. 4A, lane 5). Maturated *B. mori* tRNA*Ala* in which the adenosine at position 34 has been changed to guanosine has the same level of adenosine-to-inosine conversion as the wild-type substrate (lane 6). However, a change from adenosine to guanosine at position 37 (lane 7) or a double change of cytosine 36 to uridine and adenosine 37 to guanosine (lane 8) eliminates all adenosine-to-inosine conversion in the *Bombyx* tRNA substrate, indicating that it is adenosine 37 and not adenosine 34 that is deaminated. In contrast, yeast tRNA*Ala* is not an efficient substrate for the *Drosophila* enzyme that has a specific activity of 0.08 U/mg with this substrate (Fig. 4A, lane 3). The location of the inosine formed in the tRNA*Ala* substrates was confirmed by performing reverse transcription PCR on the reaction products and sequencing the resulting cDNAs (data not shown). Therefore, the *Drosophila* clone encodes a protein that is clearly a functional homolog of scTad1p, and we have named it *Drosophila* ADAT1 (dADAT1).

The *Drosophila* enzyme cannot bind efficiently to the yeast tRNA*Ala*. UV cross-linking of dADAT1 to [α-³²P]ATP-labelled tRNA*Ala* substrates from *B. mori* or yeast confirm that dADAT1 binds much more efficiently to the tRNA from *B. mori* (Fig. 4B, lanes 1 and 3). Surprisingly, dADAT1 also binds to a *Bombyx* tRNA substrate in which positions 36 and 37 are mutated (Fig. 4B, lane 2) although less tightly than to the wild-type tRNA*Ala* from *B. mori* (Fig. 4A) and the *B. mori* tRNA efficiently (15), and UV cross-linking experiments confirm that scTad1p binds efficiently to both tRNA*Ala* from *B. mori* (Fig. 4B, lane 7) and from yeast (lane 9).

We also tested whether purified recombinant dADAT1 could deaminate adenosine to inosine on extended dsRNA, but no conversion was observed (results not shown), suggesting that dADAT1 has no role in pre-mRNA editing.

Expression of dADAT1 transcripts in *Drosophila* embryos. Digoxigenin-labelled sense or antisense *dADAT1* RNA probes derived from clone 12 were hybridized to *Drosophila* embryos. As expected, a signal could be visualized only with the antisense *dADAT1* probe. The expression patterns observed with antisense probes from clone H and clone 12 were identical, and those obtained with clone 12 are shown in Fig. 5. A strong overall staining was detected at the blastoderm stage (Fig. 5A), suggesting that *dADAT1* mRNA is maternally provided, since no zygotic transcripts are expressed at this early stage. Zygotic expression was very strong during germ band extension, especially in the mesoderm and neuroectoderm (Fig. 5B). From germ band retraction onwards, high levels of transcripts were confined to the central nervous system and transcript levels remained high in the entire brain and ventral nerve cord throughout late embryogenesis (Fig. 5C).

Isolation of a human ADAT gene and comparison with *Drosophila* *Adat1* reveals metazoan ADATs with adenosine deaminase domains surprisingly similar to ADARs involved in pre-mRNA editing. Database searches with *S. cerevisiae* ADAT1 identified a partial sequence of a potential human homolog in the Washington University EST database. The clone from which this sequence derived was obtained, and additional sequence information for the amino-terminal half of this protein was determined (see Materials and Methods). The resulting sequence revealed an ORF with homology over its entire coding sequence to *Drosophila* *Adat1*, which is not surprising considering that tRNA*Ala* species from *B. mori* and humans are very similar (6). While this paper was in preparation, Maas et al. (28) reported the characterization of the human ADAT1 and showed that this protein deaminates adenosine 37 of the human tRNA*Ala*. The amino acid sequence of this human ADAT is included in the multiple sequence alignment shown in Fig. 2. Both human and *Drosophila* ADATs are more similar to adenosine deaminase domains of vertebrate ADARs than to the yeast ADATs.

To examine phylogenetic relationships among adenosine deaminase domains, ADAR and ADAT sequences were recovered from databases and multiple sequence alignments were performed with the Wisconsin Genetics Computer Group (GCG) Pileup program (13). The most consistent alignments involve residues close to the histidine and cysteine residues of deaminase motifs I, II, and III (Fig. 2). All ADAT proteins are approximately the same size and there is homology among the proteins over the full length of the sequences that is observed when lower gap insertion and extension penalties are used in the Pileup program. Aligned sequences were truncated to re-
move dsRNA-binding domain sequences from the ADAR genes, and multiple sequence alignments of adenosine deaminase domains were used to generate phylogenetic trees with the PAUP (phylogenetic analysis using parsimony) program in the Wisconsin GCG program package (13).

Many adenosine deaminase domain trees generated from Pileup alignments made with different gap insertion and extension penalties have topologies similar to that shown in Fig. 6. Human ADAT and Drosophila ADAT1 are closely related and their relationship to S. cerevisiae Tad1p and Schizosaccharomyces pombe Tad1p pair is distant. The ADAR genes are always found to branch from the ADAT line. Among the ADARs, human ADAR2 and the putative Drosophila ADAR N35H14 (accession no. AL035207) are closely related, while the branching order of the other two ADARs is less consistent. In order to obtain the best possible alignment with fewer gaps and with aligned residues representing true homologies, a region around the conserved deaminase motifs was selected and aligned more stringently to produce the tree shown in Fig. 6.

**DISCUSSION**

An enzymatic activity capable of converting adenosine to inosine in dsRNA is present in D. melanogaster (10). To find a gene encoding this activity, we performed a homology search with hADAR1. An ORF was found that had high sequence homology to the deaminase domain of hADAR1. We isolated this gene and showed that it encoded an adenosine deaminase that specifically converts adenosine to inosine at position 37, adjacent to the anticodon in B. mori tRNAAla(IGC). We have named this gene dAdat1 because of its homology to the yeast scADAT1.

The dAdat1 gene is located 2 to 7 kb centromere distally to Son of sevenless (Sos) (7, 43) in the Adh region on the left arm of chromosome II. It is within a 2.9-Mb stretch of contiguous sequence that has been correlated with extensive sets of deletions and point mutations (1). dAdat1 is flanked centromere distally by a predicted carbonate dehydratase gene (BG: DS00941.1) and centromere proximally by a predicted proteasome subunit gene (BG:DS00941.3). The predicted carbonate dehydratase coding sequence is fully contained within the 3′ untranslated region of the longer dAdat1 cDNA clone 12. dAdat1 and the predicted proteasome subunit gene are both present on genomic fragments that contained Sos and 10 kb of upstream DNA that was transformed into Drosophila to identify the correct mRNA for Sos by genetic rescue (7, 43).

Drosophila ADAT1 binds and deaminates tRNAAla from B. mori 100-fold more efficiently than the tRNAAla from S. cerevisiae (data not shown). The yeast Tad1p deaminates A37 more effectively in yeast tRNAAla than in the tRNAAla from B. mori (15). Each ADAT shows preference for tRNAAla from the same or more closely related organisms.

The tRNAAla from B. mori (Fig. 3C) that was used as a substrate for dADAT1 in this study is a tRNA expressed specifically in the silk moth salivary gland, where alanine is incorporated into the silk fibroin protein (46, 51). Wild-type and mutant forms of this tRNA substrate have been used in earlier studies of adenosine-to-inosine conversion and other modifications in tRNAAla (2, 17). This salivary gland-specific tRNAAla differs by one base at position 40 in the anticodon arm from the constitutive silk moth tRNAAla, which also has an inosine at position 37 (Fig. 4C). One gene encoding a tRNAAla has been characterized in Drosophila (12), but it is not known whether some Drosophila tRNAs have tissue-specific expression similar to that found for the silk moth tRNAAla genes (46). It is not known if the conversion of adenosine to inosine at position 37 in tRNAAla is always complete, and the degree of conversion could vary in different tissues.

The significance of adenosine-to-inosine conversion adjacent to the anticodon in tRNAAla(IGC) is not known. Yeast tad1 mutants are viable and contain unmodified adenosine 37 in their tRNAAla species (15). No deletion in dAdat1 is available that only removes dAdat1 and not neighboring genes. Almost all the lethal mutations that have been found near Sos have been assigned to genes that were previously known or that have been identified in the genomic sequence (1). None of the known lethal mutations are in dAdat1. The effect of loss of function in Drosophila Adat1 therefore cannot yet be assessed. Transcripts of dAdat1 are produced by nurse cells and loaded into the egg before fertilization (Fig. 5). Many mRNAs encoding proteins that will be required for transcription and translation during the early stages of embryonic development are provided in this way. Zygotic cells throughout the embryo contain dAdat1 mRNA in germ band-extended embryos, but most of the expression in later germ band-retracted embryos is in the central nervous system. Much of the mesoderm and ectoderm is larval tissue that will be histolyzed at metamorphosis; thus, it is possible that early dAdat1 expression is sufficient for the remaining lifetime of these cells. The high level of expression in the nervous system is intriguing. It is possible that dAdat1 activity is ubiquitous, even though the transcript expression pattern observed by in situ hybridization is concentrated in the nervous system. Considering the high degree of homology between dADAT1 and the ADAR enzymes, it is curious that the embryonic expression of dAdat1 is found in the central nervous system and is reminiscent of rRED2 expression, which is confined to the brain (31).

It is surprising that Drosophila ADAT1 shows a greater similarity to ADARs involved in pre-mRNA editing than to the yeast Tad1 proteins, considering that these proteins have very different substrates (Fig. 2). Based on functional groupings, dADAT1 might have been expected to resemble S. cerevisiae Tad1p more than vertebrate ADARs. A human ADAT gene also exhibits similarity to dAdat1 and to vertebrate
ADARs. It is possible that this similarity between dADAT1 and ADARs reflects a role for dADAT1 in editing pre-mRNA. However, such an activity is unlikely, since recombinant dADAT1 protein cannot convert adenosine to inosine in dsRNA substrates. An ADAR-like gene (N35H14) with two dsRNA-binding domains has been sequenced by the European Drosophila Genome Project and independently cloned from Drosophila (R. Reenan, personal communication). The protein encoded by this gene converts adenosine to inosine in dsRNA substrates (M. O’Connell and R. Reenan, unpublished results) and is a strong candidate for catalyzing editing in the cac and para pre-mRNAs (18, 36, 44).

One major question is how the ADAR family of enzymes recognize their target adenosines in pre-mRNA. The dsRNA-binding domains show little sequence specificity, recognizing continuous dsRNA structures (41). With the emergence of the homologous ADAT family of enzymes having a deaminase domain that binds to specific tRNA, it seems probable that the adenosine deaminase domain of ADARs contributes more than was previously anticipated to RNA binding and hence to substrate recognition. It has previously been shown that adenosine deaminase domains alone are not sufficient for editing by ADARs in vitro (25, 29), but it has not been determined if the deaminase domain alone can bind to RNA. A domain of ADAR corresponding to the region of ADAT homology can now be expressed in recombinant form. Detailed RNA-binding studies can then be performed in vitro to determine if indeed the deaminase domain alone can bind to pre-mRNA substrates and contribute to the target specificity of these enzymes.

ADATs are likely to have a higher affinity for RNA substrates than an isolated ADAR adenosine deaminase domain, since ADARs probably depend on their dsRNA-binding domains for RNA affinity. It is interesting therefore to compare ADAR and ADAT sequences around the putative zinc-chelating residues. The number of residues between deaminase motifs I and II is well conserved, whereas the number of residues between deaminase motifs II and III differs (22). Structural adjustments near deaminase motif III could be associated with changes in the target specificity of adenosine deaminase domains. A cluster of positively charged residues is well conserved between dADAT1, hADAT1, and the yeast ADATs in the region amino terminal to deaminase motif III (Fig. 2) (28). This cluster of positively charged residues could be a region of the protein that contacts RNA.

This region of sequence similarity among ADATs has also been noted by Maas et al. (28) in an alignment in which all hADAT residues between deaminase motifs II and III were included. Due to the variability in the number of residues between motifs II and III, sequence alignments that include this region can have different gaps and therefore highlight different blocks of homology. In addition to the ADAT signature sequence near deaminase motif III, Maas et al. (28) observed an ADAR specific signature sequence 19 residues amino terminal to the ADAT signature sequence. However,
visual inspection of many different computer-generated alignments suggests that this ADAR-specific signature could be a variant of the ADAT signature sequence which has been moved 19 residues further from deaminase motif III. Residues 178 to 189 of hADAR1 (LRTKvenGEgTi) (Fig. 2), which are conserved among ADARs, resemble residues 176 to 187 of dADAT1 (LRTKgrGErT1), which are conserved among ADATs. It is possible that the evolution of ADATs to ADARs involved an insertion close to motif III that could have affected target sequence recognition.

It is possible that an ADAT signature sequence is only visible close to the active site of the enzyme, as other tRNA-protein contacts may involve multiple weakly conserved sites over the entire protein and therefore be difficult to recognize. One or more of the 19 differences between tRNAAla from \textit{S. cerevisiae} and \textit{B. mori} have a more significant effect on the activity of dADAT1 than on Tad1p, suggesting that these proteins may not make all RNA contacts in precisely the same manner.

Pre-mRNA editing may have evolved when an original ADAT acquired dsRNA-binding domains and a new set of targets in pre-mRNA. An interesting interpretation of sequence homologies and of the phylogenetic tree shown in Fig. 6 would be the following. (i) Tad1p does not encode an essential function in yeast, and the sequence conservation with metazoan ADATs is correspondingly low and focused on the zinc-chelating residues in the deaminase domain. (ii) ADARs evolved from ADATs after the divergence of fungal and metazoan lines. (iii) Double-stranded RNA-binding domains may have been added to an ADAT gene after a gene duplication event, and such a protein might have retained ADAT function for a period until it evolved to recognize a more ADAR-like range of RNA targets. Other evolutionary scenarios can be envisaged, however; \textit{Drosophila} and human ADAT genes appear to be more closely related than yeast TAD1 to a precursor gene that gave rise to the ADAR family. Studies on the recognition of tRNA substrates by \textit{Drosophila} or human ADAT1 should help to explain how pre-mRNA editing evolved and how target sequences are recognized in pre-mRNA editing.

FIG. 5. Expression of \textit{daDat1} transcripts in \textit{Drosophila} embryos. Embryos were hybridized with digoxigenin-labelled antisense probe to \textit{dADAT1} clone 12. (A) Embryo at cellular blastoderm formation, showing maternally supplied transcript. (B) Ventral view of an embryo during germ band retraction, showing widespread expression of \textit{daDat1} transcripts in mesoderm, ectoderm, ventral nerve cord, and supraesophageal ganglion. (C) Lateral view of a later-stage embryo, showing expression confined to the ventral nerve cord and supraesophageal ganglion. (D) Lateral view of late embryo hybridized with sense probe. All embryos have their anterior to the left, and photographs were taken at \times60 magnification.

FIG. 6. Phylogenetic tree of ADAR and ADAT adenosine deaminase domain sequences. The sequence region from the TGXKC block before deaminase motif I to the PIYL conservation after deaminase motif III (see Fig. 2) was taken for each sequence and used to generate a more-stringent Pileup alignment (gap insertion penalty, 8; extension penalty, 2). This alignment was used to generate the tree and bootstrap values shown with PAUP in the Wisconsin GCG software package (48). GenBank database accession numbers for sequences used here are as follows: \textit{Drosophila} ADAR sequence N3SH14, AL055207; and \textit{C. elegans} ADAR sequence T20H4.4, Q22618.

ACKNOWLEDGMENTS

We are grateful to Hyouta Himeno, Hirosaki University, for yeast tRNA-Ser; to H. Grosjean for tRNA substrates; to M. Affolter and A. Jarman for cDNA libraries; and to R. Reenan for sharing unpublished results.

This work was supported by the Medical Research Council, the University of Basel, the Swiss National Science Foundation, and the Louis-jeantet Foundation for Medicine. A.P.G. was the recipient of a predoctoral fellowship from the Boehringer Ingelheim Fonds, and R.L. was supported by a grant from the European Union (via the Bundesamt für Bildung und Wissenschaft, Bern, Switzerland).

REFERENCES


Inosine and N1-methylinosine within a synthetic oligomer mimicking the anticodon loop of human (RNAAla) are major epitopes for anti-PL-12 autoantibodies. RNA 5:465–875.


