# Biochemical and Structural Studies of A-to-I Editing by tRNA:A34 Deaminases at the Wobble Position of Transfer RNA[†,‡]

Youssef Elias and Raven H. Huang*

*Department of Biochemistry, School of Molecular and Cellular Biology, University of Illinois, 600 South Mathews Avenue, Urbana, Illinois 61801*

*Received March 18, 2005; Revised Manuscript Received July 7, 2005*

ABSTRACT: Initial RNA transcription produces several tRNAs (one in prokaryotes and plant chloroplasts and seven or eight in eukaryotes) that contain an adenosine (A) at the wobble position (position 34). However, in all cases, adenosine at position 34 is post-transcriptionally converted to inosine (I), producing mature tRNAs without adenosine at the wobble position. The enzymes responsible for this A-to-I conversion in tRNA are tadA (acting as a homodimer) in prokaryotes and the heterodimeric ADAT2−ADAT3 complex in eukaryotes. The genes encoding these proteins are essential for cell viability, illustrating the biological importance of A-to-I editing at the wobble position of tRNA. In this study, recombinant tadA proteins from *Escherichia coli*, *Agrobacterium tumefaciens*, and *Aquifex aeolicus*, as well as the ADAT2−ADAT3 proteins from *Saccharomyces cerevisiae*, were overexpressed in *E. coli* and purified to homogeneity by chromatography. Crystallization of a proteolytically cleaved *A. tumefaciens* tadA (missing the last eight amino acids at the C-terminus) produced high-quality crystals, and the structure was determined at 1.6 Å resolution. In addition, enzymatic assays of the wild-type proteins as well as several mutants were carried out using both the full-length *E. coli* $tRNA^{arg2}$ and the truncated anticodon stem−loop motif as substrates. Our biochemical and structural studies, in combination with sequence and structural comparisons with other deaminases, allow us to propose a model of tadA−tRNA interaction that explains the molecular basis of tRNA recognition by tadA. In particular, a conserved FFxxxR motif at the C-terminus, which is unique to tadA, has been identified, and its critical role in tRNA substrate recognition is proposed. Furthermore, the structural study of prokaryotic tadA presented here also sheds light on tRNA substrate recognition and the possible evolutionary origin of the eukaryotic ADAT2−ADAT3 heterodimer.

Deaminases are a class of enzymes involved in catalyzing the hydrolysis of amino groups in biological molecules. Deamination reactions related to nucleic acids can be carried out at three levels depending on the type of substrate: the base level, the nucleotide/nucleoside level, and the nucleic acid (RNA and DNA) level. The former two substrate reactions allow for the disposal of toxic bases or for the biosynthesis of the building blocks of nucleic acids (*1*−*3*). The latter substrate reaction results in alteration of genetic information encoded in RNA and DNA. Hence, this type of modification was termed editing (RNA or DNA editing) because the deamination reaction alters the amino acid specificity of the codon, resulting in a translated protein whose sequence is different from the one predicted from its genome.

To date, biochemical studies and protein sequence analysis have suggested that all deaminases related to nucleic acids contain an active site that consists of a zinc ion, three amino acids of various combinations of Cys and His coordinated to the zinc ion, and an acidic residue (Glu or Asp) to mediate proton transfer. A water molecule is the fourth ligand of the zinc ion and is usually seen in the active sites of most available deaminase structures acting on a base or nucleotide/nucleoside (*4*−*6*). Therefore, the mechanism of the deamination reaction has been proposed to involve a nucleophilic attack by the zinc-activated water molecule at the carbon atom of the base to which the targeted amino group attaches, resulting in the elimination of ammonia.

While the mechanism of the deamination reaction has for the most part been resolved, the question of how a particular deaminase recognizes its specific substrate requires further study. In particular, structural information about the enzyme alone or in complex with a substrate and/or inhibitor will be required to reveal the molecular basis of substrate specificity. Deaminases that act on RNA are of particular interest because these enzymes (unlike deaminases that act on a base or nucleotide/nucleoside) must be able to recognize the targeted nucleotide (local structure) in the context of a properly folded RNA substrate (global structure). This includes the sequence and structure of nucleotides surrounding the targeted base, thus providing a system for studying RNA−protein interaction and recognition. In addition, RNA editing is of significant biological importance because the lack of RNA editing usually has severe biological consequences.

Depending on the type of RNA substrate, RNA editing enzymes can be generally divided into two classes: those involved in mRNA editing and those involved in tRNA editing. The mRNA deaminases can be further subdivided into two families based on the identity of the deaminated nucleotide (*7*, *8*). The ADAR[1] (adenosine deaminase acting on RNA) family of enzymes deaminates adenosine (*9*, *10*), while the APOBEC (apolipoprotein B editing catalytic

[†] This research was supported by NIH Grant CA90954.

[‡] The coordinates for the structure have been deposited with the Protein Data Bank as entry 2A8N.

* To whom correspondence should be addressed. Phone: (217) 333-3967. Fax: (217) 244-5858. E-mail: huang@uiuc.edu.

10.1021/bi050499f CCC: $30.25 
Published on Web 08/16/2005

subunit) family of enzymes deaminates cytidine (*11*). In addition, some members of the APOBEC family and AID (activation-induced cytidine deaminase) have been implicated in DNA deamination and play essential roles in host defense against infectious agents (*12, 13*).

In tRNA, only two positions are known to be edited by ADATs (adenosine deaminase acting on tRNA) (*14*). One is A-to-I editing at position 37, which occurs only in eukaryotic $tRNA^{ala}$ and is carried out by ADAT1 (*15*). The other is A-to-I editing at the wobble position (position 34), which occurs in both prokaryotic and eukaryotic tRNAs and is carried out by a tadA homodimer in prokaryotes and a ADAT2−ADAT3 heterodimer in eukaryotes (*16*−*18*). I34 is believed to play a critical role in protein synthesis due to its alternative base pairing with U, C, or A in the third position of codons (*19*−*21*). This is consistent with findings that the genes encoding tadA and the ADAT2−ADAT3 heterodimer are essential for cell viability. In addition, C-to-U editing at the wobble position has also been reported, but the enzyme involved in this C-to-U editing has not been identified (*22*).

To date, various structural studies have been conducted on deaminases that act on a base or nucleotide/nucleoside. Despite this, it was only recently that the first structure of a bona fide deaminase acting on nucleic acid (RNA or DNA) has been reported. Yokoyama and co-workers reported the crystal structure of *Aquifex aeolicus* tadA (*23*). The study reported here extends our understanding of deaminases by examining the crystal structure of tadA from *Agrobacterium tumefaciens* (39% identical to *Aq. aeolicus* tadA). Furthermore, proteolytic cleavage of the last eight C-terminal residues of *A. tumefaciens* tadA was required to generate crystals that diffracted to high resolution. Enzymatic assays of the wild-type enzymes as well as several mutants showed that the last eight residues of the C-terminus are important for enzymatic activity and potentially represent a flexible region necessary for RNA substrate binding and catalysis. Systematic sequence alignment of tadA family members, in combination with a docking model of the RNA−tadA complex, allowed us to propose a role for various conserved tadA residues in RNA substrate recognition. Finally, the collective information of the structure of tadA, the sequence alignment of tadA with the ADAT2−ADAT3 heterodimer, and the enzymatic assay results reported here provides insight into tRNA substrate recognition in general and the possible evolutionary origin of the eukaryotic ADAT2−ADAT3 heterodimers.

## MATERIALS AND METHODS

*Cloning, Overexpression, and Purification of tadA and ADAT2−ADAT3 Proteins.* The open reading frames (ORFs) of putative tadA enzymes from *A. tumefaciens* and *Aq. aeolicus* bacteria were identified from a BLAST search of the NCBI database based on sequence homology to the published *Escherichia coli* tadA gene. Each putative tadA gene was amplified from the respective genomic DNA using PCR and cloned into a PLM-1 vector (*24*) between SacI and HindIII restriction sites. The *E. coli* tadA gene was amplified from *E. coli* genomic DNA using PCR. The amplified gene was cloned into a pGEX-2TK (N-terminal GST fusion vector, Amersham Bioscience) vector between BamHI and EcoRI restriction sites. The *Saccharomyces cerevisiae* ADAT2 and ADAT3 genes were amplified from *S. cerevisiae* cDNA using PCR. ADAT2 and ADAT3 genes were sequentially cloned into the PLM-1H1 vector, a derivative of the PLM-1 vector that allows two proteins to be expressed simultaneously (*25*). *A. tumefaciens* mutants were generated using a QuickChange site-directed mutagenesis kit (Stratagene), and the introduction of mutations was confirmed by DNA sequencing.

Recombinant tadA-PLM-1 and ADAT2−ADAT3-PLM-1H1 vectors were transformed into BL-21(DE3) *E. coli* expression cells. A 10 mL overnight preculture was used to inoculate 3 L of LB supplemented with 50 $\mu$g/mL ampicillin. Cells were cultured at 37 °C (20 and 30 °C for *E. coli* tadA and *S. cerevisiae* ADAT2−ADAT3, respectively) until an $OD_{600}$ of 0.6 and then induced with 1 mM isopropyl $\beta$-D-glactopyranoside (IPTG) for an additional 4 h. *Aq. aeolicus*, *A. tumefaciens*, and *S. cerevisiae* ADAT2−ADAT3 heterodimers were purified similarly, except the buffer for the *S. cerevisiae* ADAT2−ADAT3 heterodimer had a higher pH (pH 8.3). Purification using an FPLC system was conducted as follows. Cell pellets were suspended in DEAE buffer A [20 mM Tris-HCl (pH 7.5), 5 mM NaCl, and 1 mM DTT], and cells were lysed using a French press. Cell lysates were centrifuged, and supernantants were passed through a preparatory DEAE anion exchange column. The protein was eluted with DEAE buffer B (same as DEAE buffer A except 1 M NaCl). Fractions containing tadA were pooled and then injected into a HiLoad hydrophobic column equilibrated with HiLoad buffer A containing 20 mM Tris-HCl (pH 7.5), 1.25 M $(NH_4)_2SO_4$, and 1 mM DTT. Proteins were eluted with HiLoad buffer B [same as HiLoad A except no $(NH_4)_2SO_4$]. Fractions containing tadA were pooled, concentrated, and washed with DEAE buffer A. The sample was then applied to the MonoQ anion exchange column, and DEAE buffers were used to elute proteins. Finally, fractions containing tadA from MonoQ were purified by Superdex 200 size exclusion using a buffer containing 20 mM HEPES (pH 7.0), 200 mM NaCl, and 1 mM DTT. On the basis of the retention volume of elution, the tadA protein acted as a dimer. *E. coli* tadA was prepared similarly except that single-step chromatography using a GST column was used during purification. When required, protein was concentrated using an Amicon centrifugal filter device from Millipore. Selenomethionine-labeled protein was prepared using the following protocol: 5 mL of overnight culture was used to inoculate 3 L of minimal medium supplemented with ampicillin and thiamine at final concentrations of 100 and 50 $\mu$g/L, respectively. The cell culture was grown at 37 °C to $OD_{600}$ of 0.6, and then 40 mL of methionine suppression buffer (comprised of 200 mg of L-Lys-HCl, L-threonine, and L-phenylalanine, and 100 mg of L-leucine, L-isolutine, and L-valine in water) was added. After incubation for an additional 30 min at 37 °C, the cell culture was induced with 0.5 mM IPTG and addition of 50 mg of selenomethionine/L. Incubation at 37 °C was continued for an additional 18 h, and the cells were then harvested.

[1] Abbreviations: tadA, tRNA-specific adenosine deaminase; ADAT, adenosine deaminase acting on tRNA; ADAR, adenosine deaminase acting on RNA; AID, activation-induced cytidine deaminase; APOBEC, apolipoprotein B editing catalytic subunit; GD, guanine deaminase; CD, cytosine deaminase; CDA, cytidine deaminase; ADA, adenosine deaminase; rmsd, root-mean-square deviation.

Table 1: Statistics of Data Collection and Refinement[a]

| | inflection | peak | hard remote | native |
|---|---|---|---|---|
| space group | | | | $P2_12_12_1$ |
| unit cell | | | | $a = 48.05$ Å, $b = 56.34$ Å, $c = 90.67$ Å $\alpha = 90.00°$, $\beta = 90.00°$, $\gamma = 90.00°$ |
| resolution (Å) | 100−1.65 | 100−1.65 | 100−1.65 | 50−1.60 |
| wavelength (Å) | 0.9794 | 0.9792 | 0.9686 | 0.9792 |
| no. of unique reflections | 29782 | 29771 | 30744 | 39316 |
| completeness (%) | 97.9 | 97.9 | 97.9 | 97.7 |
| average $I/\sigma(I)$[b] | 16.3 | 16.2 | 16.0 | 15.7 |
| redundancy | 9.3 | 9.2 | 9.3 | 7.2 |
| $R_{sym}$[c] (%) | 5.0 | 6.0 | 4.6 | 6.9 |
| refinement | | | | |
| resolution (Å) | | | | 50−1.6 |
| no. of reflections (free) | | | | 31216 (2548) |
| $R_{crystal}$[d] ($R_{free}$[e]) (%) | | | | 22.5 (25.8) |
| rmsd for bonds (Å) | | | | 0.0049 |
| rmsd for angles (deg) | | | | 1.21 |

[a] The mean figure of merit (FOM) for phasing is 0.75 from SOLVE and 0.83 after RESOLVE. [b] $I/\sigma(I)$ is the mean reflection intensity/estimated error. [c] $R_{sym} = \sum|I - \langle I\rangle|/\sum I$, where $I$ is the intensity of an individual reflection and $\langle I\rangle$ is the average intensity over symmetry equivalents. [d] $R_{crystal} = \sum||F_o| - |F_c||/\sum|F_o|$, where $F_o$ and $F_c$ are the observed and calculated structure factor amplitudes, respectively. [e] $R_{free}$ is equivalent to $R_{crystal}$ but calculated for a randomly chosen set of reflections that were omitted from the refinement process.

*Crystallization, Data Collection, Structural Determination, and Refinement.* Crystals of proteolytically cleaved *A. tumefaciens* were obtained at room temperature using the hanging drop vapor diffusion method by mixing tadA (10 mg/mL) with an equal volume of a well solution containing 16% PEG 6000 (w/v), 0.1 M NaCl, 0.1 M MES (pH 6.0), and 50 mM spermine. Crystals of selenomethionine-labeled protein were obtained under similar conditions. Crystals grew as a diamond shape and were harvested after growth for 1 week. To carry out data collection at a low temperature, the crystals were first soaked in a cryo-protecting solution containing all the components of the well solution supplemented with 20% glycerol, mounted in a nylon loop, and then flash-frozen in liquid nitrogen. Both the native and three-wavelength anomalous dispersion (MAD) data were collected at the 14-BMC beamline at the Advanced Photon Source (APS). Data were reduced with Denzo and Scalepack (*26*). For phase determination, the resolution range from 30 to 1.8 Å was chosen. Five of the six expected sites (assuming tadA forms a dimer in an asymmetric unit) were located using Solve (*27*). Phases were improved with solvent flattening, and ~80% of the model was automatically built using Resolve (*28*). Several rounds of manual building using O (*29*), followed by refinement with CNS (*30*), resulted in a final model with an *R*-factor of 22.5% ($R_{free}$ = 25.8%). Final refinement statistics are given in Table 1.

*Enzymatic Assays and HPLC Analysis.* Full-length tRNA was prepared by in vitro transcription, and stem−loop RNA was chemically synthesized and purchased from Dharmacon Research (Boulder, CO). RNAs were purified using denaturing polyacrylamide gel electrophoresis. For a reaction involving wild-type tadA, 4 $\mu$M purified tadA was incubated with 40 $\mu$M stem−loop RNA in a 25 $\mu$L reaction mixture containing 100 mM MES (pH 6.0), 100 mM NaCl, and 1 mM DTT. The mixture was incubated at 37 °C for 1 h, and reactions were stopped by adding 0.2 unit of P1 nuclease and incubation for an additional 1 h at 37 °C. The mononucleotides generated by P1 nuclease were dephosphorylated by adding 1 unit of shrimp alkaline phosphatase (SAP) and incubating at 37 °C for 15 min. The resulting mixture of nucleosides was directly loaded onto a C18 reverse-phase HPLC column (18-LC-S, Supelco), and nucleosides were separated with a gradient of solvent A and B. Solvent A was 20 mM ammonium acetate (pH 6.0), and solvent B was 40% acetonitrile (v/v) in water: gradient, 0% B at 0.0 min, 0% B at 3.0 min, 5% B at 10 min, 20% B at 21 min, 50% B at 22 min, 50% B at 25 min, 0% B at 26 min, and 0% B at 35 min with a constant flow rate of 2 mL/min. Similar reaction conditions were used when the full-length *E. coli* $tRNA^{arg2}$ was used as a substrate except the concentration of tRNA was elevated to 60 $\mu$M. For reactions involving the proteolytically digested *A. tumefaciens* tadA and other mutants, the concentration of the enzyme was increased to 40 $\mu$M (a 10-fold increase from the wild-type value), and the incubation time was increased to 10 h (a 10-fold increase). The peaks corresponding to I and A were integrated, and the relative enzymatic activity was calculated, with the wild-type enzyme as 100%.

*Construction of a Docking Model of RNA−tadA Complex.* The structure of the anticodon stem−loop motif of $tRNA^{phe}$ (*31*) was employed to build a docking model. The nucleotide sequence of anticodon was mutated from GAA of $tRNA^{phe}$ to ACG of $tRNA^{arg2}$ using O (*29*). A34 was then flipped out from the anticodon loop. The modified stem−loop RNA was then manually docked into the cleft formed by the dimerization of the tadA using PyMOL (*32*). It was necessary to adjust the side chains of R70 and R94 in the *trans* monomer to accommodate the RNA molecule. Finally, the structure of additional C-terminal amino acids (residues 131−145 for A monomer and 141−145 for B monomer) that was missing in our structure was built into our docking model. The structure of the missing C-terminal residues was based on the structure of *Aq. aeolicus* tadA recently reported by Yokoyama and co-workers (*23*). The C-terminus and most loops that are close to the docked RNA are not involved in crystal packing. The only exception is L8 that contains H124, which does not affect the docking.

## RESULTS AND DISCUSSION

*Crystallization of tadA.* The genes encoding tadA from *E. coli*, *A. tumefaciens*, and *Aq. aeolicus*, as well as *S. cerevisiae* ADAT2−ADAT3 heterodimer, were cloned into PLM1 vectors (*24*). For the *E. coli* tadA, it was necessary to have a GST tag at the N-terminus to increase its solubility during overexpression and purification. The recombinant proteins were overexpressed in *E. coli* and purified to homogeneity. Crystallization trials were carried out with tadA from *A. tumefaciens* and *Aq. aeolicus*, and the *S. cerevisiae* ADAT2−ADAT3 heterodimer (i.e., those recombinant proteins without a tag). Both *Aq. aeolicus* and *A. tumefaciens* tadA yielded small needle crystals from initial crystallization trials. However, the crystals from *Aq. aeolicus* tadA diffracted to only 4 Å resolution, and the crystals from *A. tumefaciens* tadA were too small to be tested. Attempts to improve the quality of the crystals were not successful. The structure of the same *Aq. aeolicus* tadA was recently reported by Yokoyama and co-workers (PDB entry 1WWR) (*23*). Their recombinant protein contained a six-His tag at the N-terminus, and the tag was involved in crystal packing. Thus,

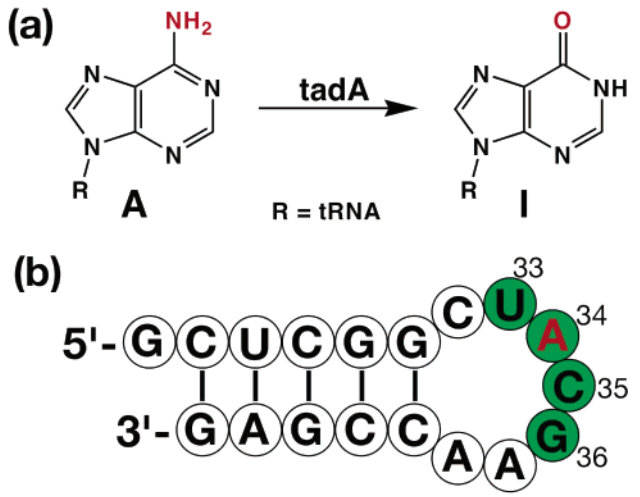


FIGURE 1: (a) Hydrolytic deamination reaction carried out by a prokaryotic tadA and eukaryotic ADAT2−ADAT3 heterodimer. The change in chemical structure that resulted from the reaction is colored red. (b) Nucleotide sequence of the stem−loop RNA employed in this study. The conserved nucleotides required to be the substrate of tadA are colored green, with the letter of the targeted nucleotide in red.

contrary to many other cases, the presence of the His tag may have facilitated the production of high-quality crystals.

SDS gel analysis of *A. tumefaciens* tadA stored at 4 °C showed proteolytic cleavage, and MS analysis of the cleaved tadA indicated that eight amino acids at the C-terminus were cleaved. We reasoned that the cleavage might result from the relative flexibility of the C-terminus, which might have prevented us from obtaining high-quality crystals. Indeed, crystallization trials of the proteolytically cleaved *A. tumefaciens* tadA yielded crystals that diffracted to 1.6 Å resolution. To obtain a large quantity of the purified protein for structural and biochemical studies, a new construct that encodes the amino acids corresponding to the proteolytically cleaved *A. tumefaciens* tadA was constructed. MAD data were obtained on the basis of crystals produced from the purified protein of this construct.

*Structure, Active Site, and RNA Binding Cleft of A. tumefaciens tadA.* The crystal structure of the proteolytically cleaved *A. tumefaciens* tadA (residues 1−144, eight residues shorter at the C-terminus than the full-length protein) was determined using the MAD phasing, and the structure was refined with several rounds of model building and simulated annealing. The asymmetric unit contains two molecules of tadA that form a dimer with a buried surface area of 2400 Å$^2$ (out of ∼6850 Å$^2$ surface area per monomer). The dimer of tadA seen in our structure is consistent with the form of the purified protein observed by gel filtration chromatography (data not shown). The electron density map of the final round of refinement lacks electron density of the last 13 amino acids at the C-terminus of the A monomer. Therefore, the structure of the A monomer contains only residues 2−131. On the other hand, the observed electron density of the B monomer extends to residue 141. Hence, the structure of the B monomer contains residues 4−141, three residues short of the proteolytic cleavage site.

The structure of each monomer is composed of a central five-stranded $\beta$-sheet, flanked by four α-helices (three in the A monomer because of the disordered C-terminus, Figure 2a,b). The overall fold is shared by other families of deaminases. Therefore, detailed discussion of the fold has been described extensively elsewhere (*33*, *34*). As expected from the sequence, the active site contains a zinc ion, which is coordinated by three strictly conserved amino acids, His53, Cys83, and Cys86 (Figure 2c). The fourth strictly conserved amino acid, Glu55, is also in the active site. In addition, a water molecule can be clearly seen in our structure, serving as the fourth ligand for the zinc ion (Figure 2c). Thus, the active site is very similar to those of other deaminases.

Dimerization of tadA with *C*2 symmetry results in formation of two clefts on opposite faces of the dimer. We presume that these clefts are where the RNA substrate binds. At the base of the cleft resides the active site, where the zinc ion is located in the structure. The active site is readily accessible as can be clearly seen from the top view of the dimer (Figure 2b). The base of the active site is surrounded by three loops (L4, L6, and L8) from the *trans* monomer (the one without the active site), and also by three loops (L3, L5, and L7) with the C-terminal helix (α4) from the *cis* monomer (the one contributing the active site) (Figure 2a). As discussed later, these loops and the C-terminal helix are involved in RNA substrate binding based on our docking model.

Structural comparison and a database search using Dali (*35*) revealed various structures that have similar folding (most of them deaminases). Three of these structures, *Aq. aeolicus* tadA (PDB entry 1WWR, *Z* score = 23.4, rmsd = 1.2 Å), *Bacillus subtilis* guanine deaminase (PDB entry 1WKQ, *Z* score = 17.6, rmsd = 2.0 Å), and *S. cerevisiae* cytosine deaminase (PDB entry 1UAQ, *Z* score = 16.5, rmsd = 2.2 Å), are particularly useful for our understanding of RNA substrate recognition by tadA. The recently published structure of *Aq. aeolicus* tadA provides structural information about the missing C-terminal tail in our structure. Unlike tadA, both *B. subtilis* guanine deaminase and *S. cerevisiae* cytosine deaminase act on small molecule substrates, yet both of them share a high degree of sequence and structural homology with tadA (Figures 3 and 4). Therefore, detailed analysis of the similarities and differences, in terms of both sequence and structure, between tadA and *B. subtilis* GD or *S. cerevisiae* CD may reveal information about the key features of tadA that are responsible for RNA recognition.

For the first 100 amino acids, a high level of sequence similarity exists between tadA and *B. subtilis* GD or *S. cerevisiae* CD (Figure 3), indicating that the core of the deaminase domain is within the first 100 amino acids. Despite considerably less sequence homology beyond the first 100 amino acids, the overall structure of tadA is very similar to those of *B. subtilis* GD and *S. cerevisiae* CD (Figure 4).

Within the tadA family, a high level of sequence homology extends beyond the first 100 amino acids. Notable conservations beyond the first 100 amino acids are K106 (R in a few cases), H124, and the FFxxxR motif at the end of the C-terminus (Figure 3, and Figure S2 of the Supporting Information). We were particularly interested in the FFxxxR motif at the end of the C-terminus because the proteolytically cleaved *A. tumefaciens* tadA employed in our structural study, which lacks the C-terminal residues after the first F of the FFxxxR motif, only exhibited residual activity. To further our understanding of the role of the conserved FFxxxR motif, we have generated several mutants in which the conserved residues (F144, F145, and R149) were individually mutated to an alanine. The enzymatic activity

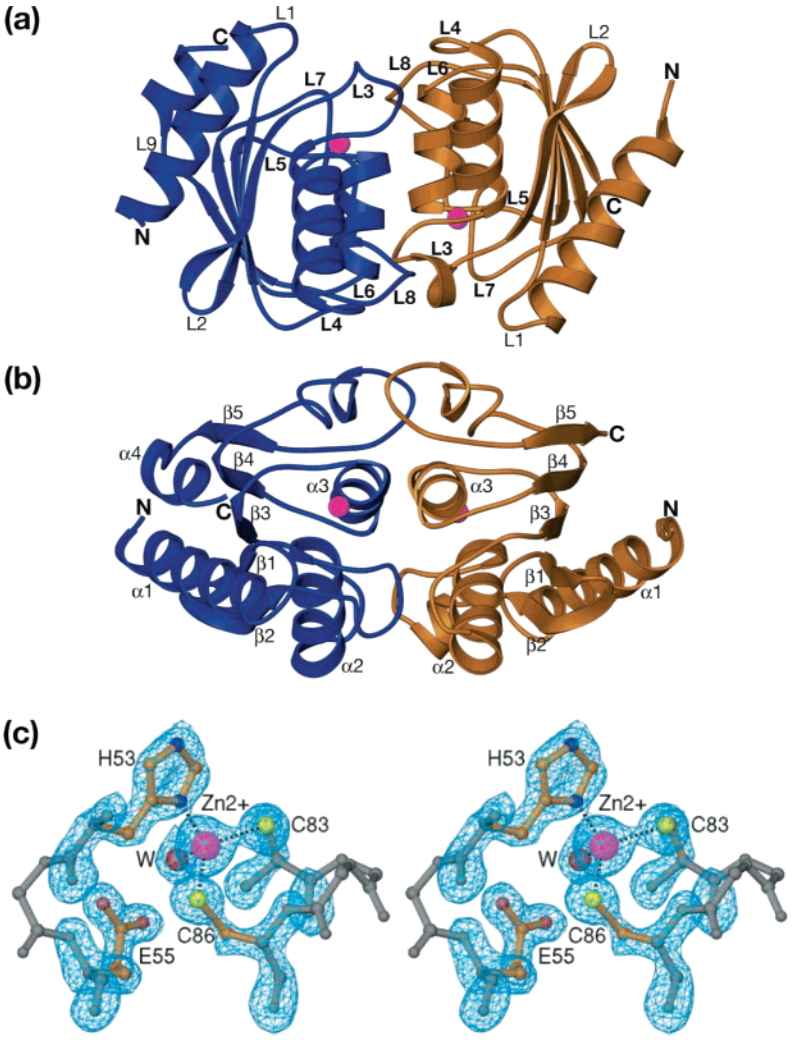


FIGURE 2: Structure of the proteolytically cleaved *A. tumefaciens* tadA homodimer. Ribbon representation of the overall structure in side (a) and top (b) views. One monomer of the homodimer is colored blue, and the other is colored orange. The zinc ions in the active sites are represented as magenta spheres. Only loops are labeled in panel a for clarity, and those likely to be involved in RNA substrate binding are in bold letters. Only α-helices and *β*-sheets are labeled in panel b for clarity. (c) Stereoview of the active site with the simulated annealing $F_o - F_c$ omit map (within 4 Å of the zinc ion). The main chains are colored gray, and the side chains are colored orange. The heteroatoms are individually colored, with nitrogen in blue, oxygen in red, sulfur in yellow, and zinc in magenta. The map is contoured at 4.0$\sigma$.

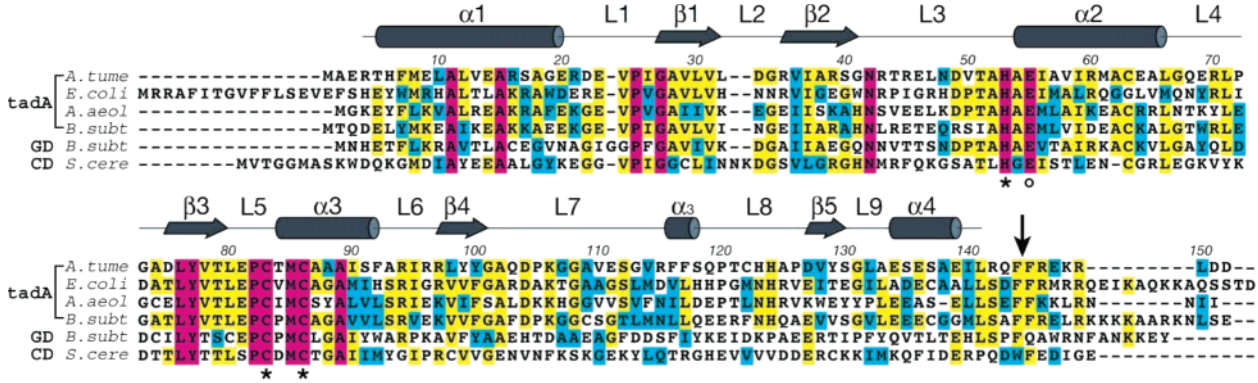


FIGURE 3: Alignment of tadA sequences from *A. tumefaciens*, *E. coli*, *Aq. aeolicus*, and *B. subtilis* together with the sequences of *B. subtilis* GD and *S. cerevisiae* CD. Residue numbers over the alignment correspond to the *A. tumefaciens* tadA sequence. The conserved residues are boxed in color, with completely conserved residues in magenta, identical residues in yellow, and similar residues in cyan. The secondary structure of the proteolytically cleaved *A. tumefaciens* tadA is depicted schematically above the primary sequence, with α-helices highlighted as cylinders and *β*-sheets as arrows. The site of proteolytic cleavage is denoted with an arrow. Asterisks denote residues acting as ligands for a zinc ions, and the degree denotes the conserved glutamate in the active site.

of these mutants, together with the wild-type enzyme and the proteolytically cleaved one, was assayed using an HPLC system.

*Enzymatic Assays of tadA.* To analyze enzymatic activity, a stem−loop RNA corresponding to the anticodon region of *E. coli* $tRNA^{arg2}$ was chemically synthesized and purified

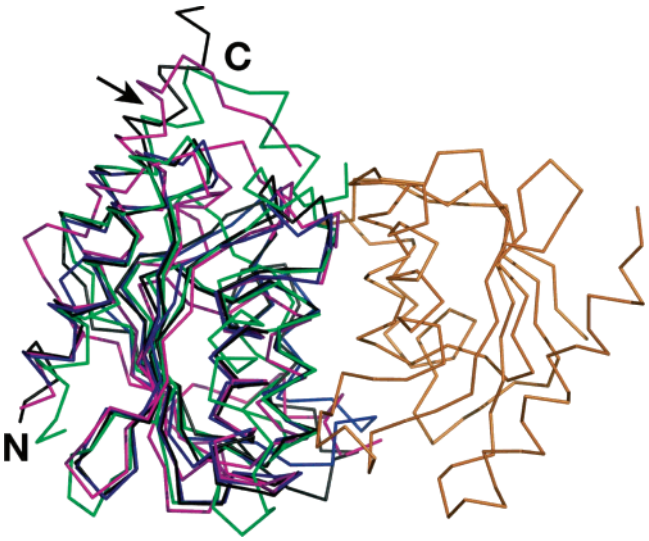


FIGURE 4: Cα superposition of the structure of the proteolytically cleaved *A. tumefaciens* tadA (blue and orange) with those of *Aq. aeolicus* tadA (black), *B. subtilis* GD (magenta), and *S. cerevisiae* CD (green). The second monomer from *Aq. aeolicus* tadA, *B. subtilis* GD, and *S. cerevisiae* CD homodimers is omitted for clarity. The equivalent position of the proteolytic cleavage site in *A. tumefaciens* tadA is denoted with an arrow.

(Figure 1b). Such a stem−loop RNA was shown to be an effective substrate for *E. coli* tadA by Keller and co-workers (*17*). The RNA substrate was incubated with various enzyme preparations, and the resulting RNA products were enzymatically digested and dephosphorylated into mononucleosides. Nucleosides were separated in a reverse-phase C18 column (Figure 5). Incubation of the stem−loop RNA with full-length tadA of *A. tumefaciens* resulted in the appearance of a new peak with a retention time slightly ahead of the G peak (Figure 5b). The identity of this new peak as I was confirmed by its coelution in the C-18 column with the commercial nucleoside (data not shown). As expected, a single-site mutant (E55A) at the strictly conserved glutamic acid within the active site resulted in a loss of activity (Figure 5c). The *E. coli* tadA was also shown to be active using the stem−loop RNA as a substrate (Figure 5d). The stem−loop RNA, however, is not an efficient substrate for the *S. cerevisiae* ADAT2−ADAT3 heterodimer, as I was not produced (Figure 5e). This is consistent with the results reported by Keller and co-workers (*16*). However, the *S. cerevisiae* ADAT2−ADAT3 heterodimer was enzymatically active as it carried out the A-to-I conversion when the full-length tRNA was used as the substrate (Figure S1 of the Supporting Information).

When the stem−loop RNA was treated with the proteolytically cleaved *A. tumefaciens* tadA with a 10-fold increase in enzyme concentration and a 10-fold increase in reaction time, a very small I peak was observed (Figure 5f). Although the peak was very small, it could be accurately integrated, and the result was reproducible. On the basis of peak integration, truncation of the last eight amino acids of *A. tumefaciens* tadA (starting at the second F of the FFxxxR motif) resulted in a reduction of enzymatic activity by 2000-fold (compare panels b and f of Figure 5), indicating the importance of the C-terminus of *A. tumefaciens* tadA. Enzymatic assays of the individual mutations of the conserved FFxxxR motif were also consistent with this conclusion. The enzymatic activity of F144A and R149A mutants was slightly better than the activity of the proteolytically cleaved one (panels g and i of Figure 5, and compare to panel f). Still, these mutations resulted in a reduction in enzymatic activity of ~800-fold. Interestingly, the F145A mutation has no enzymatic activity detected by our HPLC assay, and therefore, it is predicted to be a lethal mutation (Figure 5h). Mutation of another strictly conserved amino acid R94 to an alanine also substantially reduced the enzymatic activity (220-fold reduction, Figure 5j), although, when compared to mutations in the FFxxxR motif, the reduction was less pronounced. Reactions using the full-length tRNA instead of stem−loop RNA gave similar results (Figure S1 of the Supporting Information).

*Molecular Recognition of the RNA Substrate by tadA.* With the assumption that the cleft formed by dimerization of tadA is the binding site of the RNA substrate, a docking model of the RNA−protein complex was generated (Figure 6). The structure of the anticodon stem−loop portion of $tRNA^{phe}$ was used to build the model (*31*). It was necessary to flip out the targeted A34 nucleotide during docking because it was not possible to model A34 into the active site without a severe steric clash between other parts of the RNA and tadA if this base flipping did not occur. Generally, base flipping is regarded as a required mechanism for enzymes to catalyze chemical reactions on bases in DNA or RNA. This was first observed in DNA methyltransferase (*36*, *37*), later in many DNA glycosylases (*38*−*40*), and more recently in RNA modification enzymes (*41*, *42*). We believe that the tadA-catalyzed deamination reaction will be similar in this regard.

On the basis of our docking model, the contribution of RNA substrate recognition from the *trans* monomer is primarily from three loops, L4, L6, and L8 (Figure 2a). Specifically, our model shows that three amino acids, R70 from L4, R94 from L6, and H124 from L8, are close to C35 and G36 of the RNA. R94 and H124 are strictly conserved (Figure S2 of the Supporting Information), and they are likely to be involved in the recognition of nucleotides C35 and G36 that are required for recognition of RNA as a substrate by tadA (Figure 1b). The enzymatic assay of the R94A mutant was consistent with this assessment. Most members of the tadA family contain an Arg at position 70. For those enzymes that are missing R70, there is always a Lys or His at the preceding residue (Figure S2 of the Supporting Information). Therefore, because it is less conserved compared to R94 and H124, R70 may interact with RNA in a less specific manner (e.g., interaction with the phosphate backbone).

The contribution of RNA binding from the *cis* monomer also contains three loops, L3, L5, and L7, as well as the C-terminal helix, α4 (Figure 2a). Amino acids constituting the active site are primarily located in L5, interacting with targeted nucleotide A34. L3 contains residue H53, which is also part of the active site. Amino acids from L7 interact with the 5′ end of the RNA. In particular, the conserved K106 (R in a few other tadA members, Figure S2 of the Supporting Information) is found to interact with the phosphate backbone of nucleotide U32. In addition to these loops, our analysis indicates that the amino acids at the very end of the C-terminus also play a critical role in RNA substrate binding, as discussed below.

Sequence comparison of 20 tadA family members from different organisms shows strict conservation of the C-terminus, with a consensus motif of FFxxxR (Figure S2 of

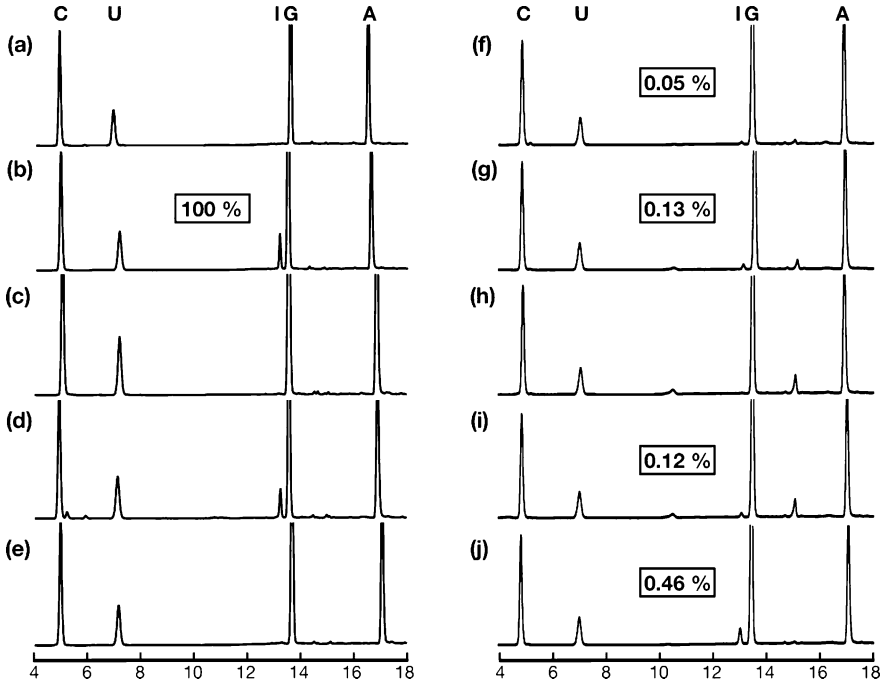


FIGURE 5: HPLC analysis of reaction products of a stem−loop RNA incubated with various tadAs. The stem−loop RNA was incubated with different enzymes, and the resulting RNA was enzymatically digested into nucleosides and analyzed by HPLC with a reverse-phase C18 column: (a) RNA alone, (b) RNA and *A. tumefaciens* tadA, (c) RNA and the *A. tumefaciens* tadA E55A mutant, (d) RNA and *E. coli* tadA, (e) RNA and the *S. cerevisiae* ADAT2−ADAT3 heterodimer, (f) RNA and proteolytically cleaved *A. tumefaciens* tadA, (g) RNA and the *A. tumefaciens* tadA F144A mutant, (h) RNA and the *A. tumefaciens* tadA F145A mutant, (i) RNA and the *A. tumefaciens* tadA R149A mutant, and (j) RNA and the *A. tumefaciens* tadA R94A mutant. Numbers in insets are enzymatic activity relative to the wild-type enzyme. The ruler at the bottom shows the retention time in minutes.

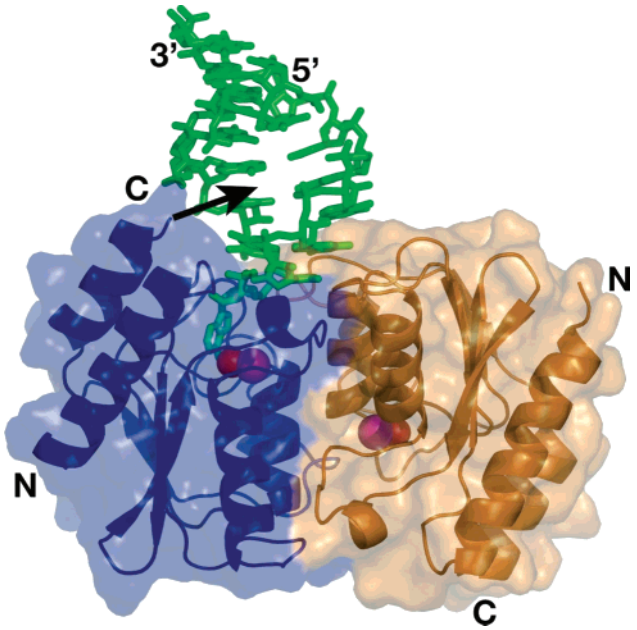


FIGURE 6: Docking model of *A. tumefaciens* tadA bound to a stem−loop RNA. RNA is shown as a stick and is colored green. The protein is depicted as a ribbon representation embedded in a semitransparent surface representation and colored as described in the legend of Figure 2. The zinc ions and their coordinated water molecules in the active sites are represented as magenta and red spheres, respectively. The structure of the C-termini of both monomers in this model ends with FF of the FFxxxR motif. A black arrow points to the major groove of stem−loop RNA and is likely to be the direction of the missing C-terminal residues upon its association with RNA substrate.

the Supporting Information). This motif is unique to tadA because neither *B. subtilis* GD nor *S. cerevisiae* CD contains such a motif (Figure 3). Furthermore, among the three intervening residues within the strictly conserved FF and R, at least one and in most cases two are either K or R (Figure S2 of the Supporting Information). This sequence analysis, together with our enzymatic assay as discussed above, and structural analysis as discussed below, points to the role of the conserved C-terminal FFxxxR motif for RNA substrate binding through the major groove of RNA.

The crystal structure of *Aq. aeolicus* tadA recently reported by Yokoyama and co-workers shows that the structure of the C-terminal amino acids missing in our structure is the further extension of the C-terminal helix, α4 (Figure 4, colored black). Our docking model allowed us to incorporate part of the structure of the missing C-terminal extension seen by Yokoyama and co-workers up to residue F145 (the second F of the FFxxxR motif, one amino acid beyond the proteolytic cleavage site in our *A. tumefaciens* tadA, our structure ends at L141, four amino acids shorter than the model). Further incorporation of the missing C-terminal amino acids caused steric clash with the docked RNA (Figure 6). Analysis of molecules involved in crystal packing of the *Aq. aeolicus* tadA structure reported by Yokoyama and co-workers indicates that part of each C-terminal helix of one monomer (amino acids after FF of the FFxxxR motif) is inserted into the dimer interface (presumably RNA binding cleft) of two other molecules. Because of its involvement in crystal packing, it is possible that these C-terminal residues may adopt a different conformation(s) in solution. Our observation that this C-terminus is susceptible to proteolytic cleavage between the conserved FF pair is consistent with this possibility. In addition, the structures of the C-terminal tails of *B. subtilis* GD and *S. cerevisiae* CD are in conformations different from those of the *Aq. aeolicus* tadA

structure (Figure 4, colored magenta and green, respectively). Taken together, we believe that the additional C-terminal amino acids after the conserved FF are flexible and may occupy the major groove of the RNA substrate upon its association with tadA, as indicated by the arrow in Figure 6. There could be three driving forces for this re-direction of the C-terminal basic residues into the major groove: (i) favorable electrostatic interaction of the C-terminal basic residues in tadA (positively charged) with the phosphate backbone in RNA (negatively charged), (ii) unfavorable interaction of the side chains of conserved FF (hydrophobic) with the backbone of RNA (hydrophilic) upon RNA binding, and (iii) hydrophobic interaction of the FF side chains with the nearby conserved ExPxGAxxxx motif (residues 23−31, where each x is a hydrophobic residue). In that regard, the conserved FF residues in FFxxxR motif may act as a "switch", facilitated by the possible driving forces mentioned above, to re-direct the rest of the C-terminal amino acids into the major groove of the RNA.

Mutational studies of the conserved FFxxxR motif are consistent with this hypothesis. Both F144A and R149A mutations reduced the enzymatic activity by approximately 800-fold (Figure 5g,i). Truncation of the C-terminus starting with the FxxxR sequence reduced the enzymatic activity even further (by 2000-fold, Figure 5f). Interestingly, the F145A mutation had the most damaging effect on the enzymatic activity of tadA, even more so than the truncation of the eight C-terminal residues. In our assay, the enzymatic activity of the F145A mutant was not detectable (Figure 5h). One possible explanation for the most damaging effect of the F145A mutation is that this mutation alters the conformation of the C-terminus in such way that the RNA substrate is no longer able to bind in a productive manner. It is even possible that the C-terminus of the F145A mutant blocks the entrance of the RNA binding cleft. In that sense, the truncation of the eight C-terminal residues should be in theory less damaging because it may still allow RNA substrate to access the active site for the reaction. Ultimately, a structure of an RNA−tadA complex is necessary to reveal the molecular basis of RNA interaction by the conserved FFxxxR motif at the C-terminus of tadA.

*Implication of tRNA Substrate Recognition by the Eukaryotic ADAT2−ADAT3 Heterodimer*. Since at least seven tRNAs in eukaryotes require A-to-I editing at the wobble position, it must be the global features of tRNA (i.e., those shared by all substrate tRNAs), not the local nucleotide sequences of the anticodon region (that differ from one another), that are the determining factors for substrate recognition by eukaryotic enzymes. Biochemical studies carried out by Keller and co-workers, as well as by us, are consistent with this assumption (*16*). A-to-I editing at the wobble position catalyzed by the eukaryotic enzyme requires the full-length tRNA (Figure S1h of the Supporting Information). The stem−loop RNA corresponding to the anticodon region of tRNA, which is an effective substrate of bacterial tadA, is not a substrate for the *S. cerevisiae* ADAT2−ADAT3 heterodimer (Figure 5g). Protein sequence alignments of tadA with ADAT2 and ADAT3, together with the structure of *A. tumefaciens* tadA and the docking model presented here, shed some light on the molecular basis of tRNA substrate recognition by the eukaryotic ADAT2−ADAT3 heterodimer.

Pairwise sequence alignments indicate that the homology region of the prokaryotic tadA in ADAT2 is located in the N-terminal two-thirds of the protein. On the other hand, it is the C-terminal half of ADAT3 that shares sequence similarity with prokaryotic tadA (Figure S3 of the Supporting Information). If we assume that the tadA homology regions of the ADAT2−ADAT3 heterodimer possess a structure similar to that of the tadA homodimer presented here, then the location of additional amino acids in the ADAT2−ADAT3 heterodimer that are missing from the tadA homodimer may have implications in the possible roles of these amino acids in molecular recognition of the full-length tRNA. In this regard, the monomer in blue shown in Figure 6 would be the equivalent to ADAT2 because it contributes the catalytic active site, whereas, the monomer in orange would be the equivalent to ADAT3. Since both the C-terminus of the tadA monomer in blue and the N-terminus of the monomer in orange are on the same side as the bound RNA substrate, the additional residues of the C-terminus of ADAT2 (∼90 amino acids) and the N-terminus of ADAT3 (∼160 amino acids) that are missing in the tadA homodimer would be located on the same side to which the RNA substrate binds (Figure 6). Therefore, these additional amino acids from the ADAT2−ADAT3 heterodimer could represent domains that are responsible for further interaction with the full-length tRNA in regions other than the anticodon stem−loop motif.

Finally, from an evolutionary point of view, it is not difficult to imagine how deamination activity on one tRNA could have evolved to encompass eight tRNAs (seven in yeast), if we consider the different modes by which prokaryotic tadA and the eukaryotic ADAT2−ADAT3 heterodimer recognize their tRNA substrates. The eukaryotic ADAT2−ADAT3 heterodimer could have evolved from a tadA-like homodimer by simple addition to the N-terminus of one monomer and the C-terminus of the other. Further mutations of amino acids within the protein, together with these additions, could result in an enzyme that relies less on local structure and more on the global structure of the tRNA substrate. Since the same tRNA editing capability seen on the ADAT2 side could not be sustained on the ADAT3 side due to a lack of N- and C-terminal addition, the critical glutamate in the active site of ADAT3 was mutated during evolution (which abolished the enzymatic activity of ADAT3), possibly to prevent detrimental deamination reactions.

## ACKNOWLEDGMENT

We thank the staffs of beamline 14-BMC at APS (K. Brister, N. Lei, and R. Pahl) for their support during data collection. We also thank B. Budke for constructing the expression vector of *E. coli* tadA and D. Kranz, J. Gerlt, and J. Yu for helpful discussions and critical reading of the manuscript.

## SUPPORTING INFORMATION AVAILABLE

HPLC analysis of reaction products of the full-length tRNA incubated with various tadA's (Figure S1), alignment of tadA sequences from 20 organisms (Figure S2), and sequence alignment of *A. tumefaciens* tadA versus *S. cerevisiae* ADAT2 and ADAT3 (Figure S3). This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES


1. Vogels, G. D., and Van der Drift, C. (1976) Degradation of purines and pyrimidines by microorganisms, *Bacteriol. Rev. 40*, 403−68.
2. Nygaard, P. (1983) Utilization of preformed purine bases and nucleotides, in *Metabolism of Nucleotides, Nucloesides, and Nucleobases in Microorganisms* (Munch-Petersen, A., Ed.) pp 27−93, Academic Press, New York.
3. Nygaard, P. (1993) Purine and pyrimidine salvage pathways, in *Bacillus subtilis and Other Gram-positive Bacteria* (Sonenshein, A. L., Hoch, J. A., and Losick, R., Eds.) pp 359−78, American Society for Microbiology, Washington, DC.
4. Xiang, S., Short, S. A., Wolfenden, R., and Carter, C. W., Jr. (1995) Transition-state selectivity for a single hydroxyl group during catalysis by cytidine deaminase, *Biochemistry 34*, 4516−23.
5. Ireton, G. C., Black, M. E., and Stoddard, B. L. (2003) The 1.14 Å crystal structure of yeast cytosine deaminase: Evolution of nucleotide salvage enzymes and implications for genetic chemotherapy, *Structure 11*, 961−72.
6. Xie, K., Sowden, M. P., Dance, G. S., Torelli, A. T., Smith, H. C., and Wedekind, J. E. (2004) The structure of a yeast RNA-editing deaminase provides insight into the fold and function of activation-induced deaminase and APOBEC-1, *Proc. Natl. Acad. Sci. U.S.A. 101*, 8114−9.
7. Gerber, A. P., and Keller, W. (2001) RNA editing by base deamination: More enzymes, more targets, new mysteries, *Trends Biochem. Sci. 26*, 376−84.
8. Schaub, M., and Keller, W. (2002) RNA editing by adenosine deaminases generates RNA and protein diversity, *Biochimie 84*, 791−803.
9. 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*, 947−9.
10. Bass, B. L. (2002) RNA editing by adenosine deaminases that act on RNA, *Annu. Rev. Biochem. 71*, 817−46.
11. Teng, B., Burant, C. F., and Davidson, N. O. (1993) Molecular cloning of an apolipoprotein B messenger RNA editing protein, *Science 260*, 1816−9.
12. Turelli, P., and Trono, D. (2005) Editing at the crossroad of innate and adaptive immunity, *Science 307*, 1061−5.
13. Sheehy, A. M., Gaddis, N. C., Choi, J. D., and Malim, M. H. (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein, *Nature 418*, 646−50.
14. Grosjean, H., Auxilien, S., Constantinesco, F., Simon, C., Corda, Y., Becker, H. F., Foiret, D., Morin, A., Jin, Y. X., Fournier, M., and Fourrey, J. L. (1996) Enzymatic conversion of adenosine to inosine and to N1-methylinosine in transfer RNAs: A review, *Biochimie 78*, 488−501.
15. Gerber, A., Grosjean, H., Melcher, T., and Keller, W. (1998) Tad1p, a yeast tRNA-specific adenosine deaminase, is related to the mammalian pre-mRNA editing enzymes ADAR1 and ADAR2, *EMBO J. 17*, 4780−9.
16. Gerber, A. P., and Keller, W. (1999) An adenosine deaminase that generates inosine at the wobble position of tRNAs, *Science 286*, 1146−9.
17. Wolf, J., Gerber, A. P., and Keller, W. (2002) tadA, an essential tRNA-specific adenosine deaminase from *Escherichia coli*, *EMBO J. 21*, 3841−51.
18. Auxilien, S., Crain, P. F., Trewyn, R. W., and Grosjean, H. (1996) Mechanism, specificity and general properties of the yeast enzyme catalysing the formation of inosine 34 in the anticodon of transfer RNA, *J. Mol. Biol. 262*, 437−58.
19. Crick, F. H. (1966) Codon−anticodon pairing: The wobble hypothesis, *J. Mol. Biol. 19*, 548−55.
20. Jukes, T. H. (1973) Possibilities for the evolution of the genetic code from a preceding form, *Nature 246*, 22−6.
21. Curran, J. F. (1998) Modified nucleosides in translation, in *Modification and Editing of RNA* (Grosjean, H., and Benne, B., Eds.) pp 493−516, American Society for Microbiology Press, Washington, DC.
22. Alfonzo, J. D., Blanc, V., Estevez, A. M., Rubio, M. A., and Simpson, L. (1999) C to U editing of the anticodon of imported mitochondrial tRNA(Trp) allows decoding of the UGA stop codon in *Leishmania tarentolae*, *EMBO J. 18*, 7056−62.
23. Kuratani, M., Ishii, R., Bessho, Y., Fukunaga, R., Sengoku, T., Sekine, S. I., and Yokoyama, S. (2005) Crystal structure of tRNA adenosine deaminase TadA from *Aquifex aeolicus*, *J. Biol. Chem. 280*, 16002−8.
24. MacFerrin, K. D., Chen, L., Terranova, M. P., Schreiber, S. L., and Verdine, G. L. (1993) Overproduction of proteins using the expression-cassette polymerase chain reaction, *Methods Enzymol. 217*, 79−102.
25. Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: Implications for drug resistance, *Science 282*, 1669−75.
26. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, in *Methods in Enzymology*, pp 307−26, Academic Press, San Diego.
27. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution, *Acta Crystallogr. D55* (Part 4), 849−61.
28. Terwilliger, T. C. (1999) Reciprocal-space solvent flattening, *Acta Crystallogr. D55*, 1863−71.
29. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−9.
30. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination, *Acta Crystallogr. D54*, 905−21.
31. Shi, H., and Moore, P. B. (2000) The crystal structure of yeast phenylalanine tRNA at 1.93 Å resolution: A classic structure revisited, *RNA 6*, 1091−105.
32. DeLano, W. L. (2002) *The PyMOL Molecular Graphic System*, DeLano Scientific, South San Francisco, CA.
33. Ko, T. P., Lin, J. J., Hu, C. Y., Hsu, Y. H., Wang, A. H., and Liaw, S. H. (2003) Crystal structure of yeast cytosine deaminase. Insights into enzyme mechanism and evolution, *J. Biol. Chem. 278*, 19111−7.
34. Liaw, S. H., Chang, Y. J., Lai, C. T., Chang, H. C., and Chang, G. G. (2004) Crystal structure of *Bacillus subtilis* guanine deaminase. The first domain-swapped structure in the cytidine deaminase superfamily, *J. Biol. Chem.*
35. Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices, *J. Mol. Biol. 233*, 123−38.
36. Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. (1994) HhaI methyltransferase flips its target base out of the DNA helix, *Cell 76*, 357−69.
37. Reinisch, K. M., Chen, L., Verdine, G. L., and Lipscomb, W. N. (1995) The crystal structure of HaeIII methyltransferase convalently complexed to DNA: An extrahelical cytosine and rearranged base pairing, *Cell 82*, 143−53.
38. Slupphaug, G., Mol, C. D., Kavli, B., Arvai, A. S., Krokan, H. E., and Tainer, J. A. (1996) A nucleotide-flipping mechanism from the structure of human uracil-DNA glycosylase bound to DNA, *Nature 384*, 87−92.
39. Lau, A. Y., Scharer, O. D., Samson, L., Verdine, G. L., and Ellenberger, T. (1998) Crystal structure of a human alkylbase-DNA repair enzyme complexed to DNA: Mechanisms for nucleotide flipping and base excision, *Cell 95*, 249−58.
40. Bruner, S. D., Norman, D. P., and Verdine, G. L. (2000) Structural basis for recognition and repair of the endogenous mutagen 8-oxoguanine in DNA, *Nature 403*, 859−66.
41. Hoang, C., and Ferre-D'Amare, A. R. (2001) Cocrystal structure of a tRNA Psi55 pseudouridine synthase: Nucleotide flipping by an RNA-modifying enzyme, *Cell 107*, 929−39.
42. Xie, W., Liu, X., and Huang, R. H. (2003) Chemical trapping and crystal structure of a catalytic tRNA guanine transglycosylase covalent intermediate, *Nat. Struct. Biol. 10*, 781−8.


BI050499F