Structure of a tRNA-specific deaminase with compromised deamination activity

Huijuan Liu¹, Saibin Wu¹, Dewei Ran¹ and Wei Xie¹,*

¹MOE Key Laboratory of Gene Function and Regulation, State Key Laboratory for Biocontrol, School of Life Sciences, The Sun Yat-Sen University, 135 W. Xiangang Rd., Guangzhou, Guangdong, 510006, People's Republic of China

*Correspondence: Wei Xie (xiewei6@mail.sysu.edu.cn)
Abstract

Nucleotide 34 in tRNA is extensively modified to ensure translational fidelity and efficacy in cells. The deamination of adenosine at this site catalyzed by the enzyme TadA gives rise to inosine (I), which serves as a typical example of the wobble hypothesis due to its diverse basepairing capability. However, recent studies have shown that tRNA^{Arg}_{ACG} in Mycoplasma capricolum contains unmodified adenosine, in order to decode the CGG codon. The structural basis behind the poorly performing enzyme M. capricolum TadA (named McTadA) is largely unclear. Here we present the structures of the WT and a mutant form of McTadA determined at high resolutions. Through structural comparison between McTadA and other active TadA enzymes as well as modeling efforts, we found that McTadA presents multiple structural conflicts with RNA substrates and thus offered support to previous studies from a structural perspective. These clashes would potentially lead to reduced substrate binding affinity of McTadA, consistent with our *in-vitro* deamination activity and binding assays. To rescue the deamination activity of McTadA, we carried out two rounds of protein engineering through structure-guided design. The unsuccessful attempts of activity restoration could be attributed to altered dimer interface and stereo hindrance from the non-catalytic subunit of McTadA, which could be the inevitable outcome of the natural evolution. Our study provides structural insight into an alternative decoding and evolutionary strategy by a compromised TadA enzyme at a molecular level.
Introduction

Protein synthesis relies primarily on the continuous interactions between the three major RNAs, i.e. rRNA, tRNA and mRNA respectively in ribosomes. The production of the polypeptide chain requires the recognition and binding of the codon in mRNA by the anticodon of tRNA \[1\]. tRNA acts as an adaptor molecule in translation and links the genetic information and the corresponding amino acid. Its maturation process generates a great variety of modified post-transcriptional nucleotides, and more than a hundred modifications have been found in tRNA to date (reviewed in \[2, 3\]). These modifications are extremely important for the structural stability and the decoding function of tRNA \[2, 4\], as defects in corresponding modification enzymes have often been linked to various human diseases. For example, the knockout of the zebrafish *mtu1* gene results in the abolition of the 2-thiouridine modification of U34 in mitochondrial tRNA\textsubscript{Lys}, tRNA\textsubscript{Glu} and tRNA\textsubscript{Gln}. This modification defect inhibits the synthesis of respiratory chain-related proteins, and eventually leads to diseases associated with mitochondrial dysfunction and pathogenesis in the auditory organs \[5, 6\].

The modifications of the nucleotide at position 34 in tRNA are particularly important to all species. Correctly modified tRNA at this site efficiently recognizes and binds the codon triplet on mRNA with enhanced efficiencies \[7, 8\]. tRNA-adenosine deaminase from bacteria (Tad) or adenosine deaminase acting on tRNA (TAD) from eukaryotes specifically catalyzes the deamination of A34, and this essential modification is conserved in both prokaryotes and eukaryotes \[9\]. An inosine (I) is subsequently produced and reads the last base of the codon, which could be A, C, U. Therefore, Tad plays an irreplaceable role in the decoding process.

Tads possess the characteristic H(C)XE and PCXXC zinc-binding hallmark sequences, and the zinc ion is critical for the deamination activity \[10-13\]. Three protein residues (two cysteines and a histidine) and a water molecule are involved in the coordination of Zn\(^{2+}\) \[10, 11\]. The zinc-coordinated water molecule is stabilized by the metal and acts as a nucleophile to attack C6 of the target adenosine in a later
step [10]. Although employing identical chemistry for catalysis, there are clear differences between the Tad enzymes from eukarya and bacteria, in terms of the deamination process. The bacterial enzyme named TadA usually functions a homodimer. In contrast, the eukaryotic Tad consists of the heterodimeric Tad2/Tad3 complex [12, 13]. Tad2 is the catalytic subunit and Tad3 serves as the regulatory subunit because it lacks a conserved glutamate in the conserved H(C)XE motif, which is replaced by a valine residue. Accordingly, the modification subjects (tRNA species) in the two kingdoms are quite different as well. Sequence alignment indicates that eukaryotic Tads share only limited homology with that of bacteria.

Following the structure determination of numerous bacterial TadA enzymes such as *Escherichia coli* TadA (EcTadA, PDB 1Z3A), *Aquifex aeolicus* TadA (AaTadA, PDB 1WWR) [11, 14] etc., Losey *et al.* successfully obtained the structure of *Staphylococcus aureus* TadA (SaTadA, PDB 2B3J) complexed with the anticodon stem-loop of tRNA^Arg^ (Samini-RNA) [10]. All these TadAs invariably form homodimers in solution. The determination of the SaTadA-Samini-RNA complex structure provides insight into the molecular mechanism by which TadA specifically recognizes and catalyzes RNA deamination. The asymmetric unit contains two dimers of SaTadA [10]. Both dimers make extensive contacts with the Samini-RNA substrate (especially the five internal bases), with the C32-A38 forming an unusual base pair to cap the end of the loop. In addition, in the active site, the RNA substrate undergoes remarkable structural rearrangements within the loop region (nucleotides 33-37) [10].

Using the bioinformatics approach, Yokobori *et al.* found that more than 30 species from Mycoplasmas and other Mollicutes were missing the tRNA^Arg^-encoding gene that specifically pairs with the arginine codon CGG [15]. Additionally, 25 such species completely lack the *tadA* gene, suggesting non-classical decoding strategies in these species. In several species including *Mycoplasma capricolum*, only tRNA^Arg^_{ACG}, tRNA^Arg^_{TCT} and a copy of *tadA* gene were found to be present in the genome, and the former was presumably to decode the CGN quartet codons. Because inosine (resulted from A34 deamination) is considered as a poor basepairing partner of guanosine due to the unfavorable purine-purine base-pairing pattern (except for the A-I pairing) in
Crick’s theory, it raises the question as to how the CGG codon in these species was decoded. Further investigation of the *M. capricolum* tRNA<sub>Arg</sub><sup>ACG</sup> sequence by reverse transcriptase–PCR on extracted tRNA<sub>Arg</sub><sup>ACG</sup> indicated that some of these tRNA molecules contain un-deaminated A34, as adenosine at the wobble position were still found in 5 out of the 81 clones, while it was zero for *B. subtilis* [15], implying reduced deamination efficiency of the TadA enzyme in the former species. Moreover, sequence alignment and structural modeling suggested that this observation might be the result of the reduction in deamination efficiency due to selective mutations in TadA during evolution, which interferes with the binding and modification of the tRNA substrate. Consequently, the unmodified tRNA<sub>Arg</sub><sup>ACG</sup> is responsible for reading the CGG arginine codon [15]. However, their hypothesis was not confirmed due to a lack of structural information on *M. capricolum* TadA (McTadA).

In the current study, we report the crystal structures of McTadA and its mutants. Based on the structural alignment of McTadA and SaTadA, we designed a series of point mutations to help McTadA to regain the deaminase activity. In addition to multiple possible structural clashes between the enzyme and the RNA substrate, the tRNA-binding cleft of McTadA is found to be wider than that of other bacterial TadA enzymes, which may also contribute to the inability of the enzyme to efficiently interact with and deaminate tRNAs. In summary, our study on McTadA helps to explore the evolutionary scenario of the decoding process in Mycoplasmas and other Mollicutes, and aids in our understanding of the adenosine deaminase and the CDA superfamily from the structural and evolutionary perspectives.

**Materials and methods**

**Cloning, overexpression, protein purification**

The gene encoding *M. capricolum* tadA (accession ABC01127) was chemically synthesized, which was digested by NdeI and XhoI and ligated into the vector pET-21b (+) (designated as TadA-pET21b). The construct contains the full-length gene fused with a C-terminal 6 × His tag. The vector was then transformed into the *Transetta* (DE3) strain (Transgen Biotech). A 20-mL overnight culture was used to
inoculate 2 L of LB supplemented with 30 µg/mL kanamycin, 64 µg/mL chloramphenicol and 100 mM ZnCl₂. Cells were grown to an OD₆₀₀ of 0.6-0.8 at 37 ℃, at which point the temperature was lowered to 16 ℃, and induced with 0.5 mM IPTG. Overnight cells were harvested and sonicated in the lysis buffer (20 mM Tris-HCl pH 8.0, 250 mM NaCl, 20 mM imidazole, 10 mM β-mercaptoethanol (β-ME), 15 µM ZnCl₂ and 15 % glycerol). The lysate was clarified by centrifugation at 23500 g for 60 min at 4 ℃. The resulting supernatant was incubated with 1-2 mL of Ni-NTA agarose (Qiagen) equilibrated with buffer A (20 mM Tris-HCl pH 8.0, 250 mM NaCl, 20 mM imidazole, 10 mM β-ME, 15 µM ZnCl₂ and 15 % glycerol) for 60 min. Gradual removal of indigenous proteins was carried out using an imidazole gradient. Buffer A supplemented with 250 mM imidazole was used to elute the TadA protein. The purity of the target protein was more than 90 % as judged by SDS-PAGE, but degrTadion was observed at room temperature. TadA-containing fractions were pooled and diluted two-fold with buffer A. Next the protein was loaded onto a 5-mL HiTrap Heparin column (GE Healthcare) and eluted with a NaCl gradient supplemented with 10% glycerol. The eluate was dialyzed to a buffer containing 20 mM Tris-HCl pH 8.0, 250 mM NaCl, 2 mM PMSF, 10 mM β-ME and 15 µM ZnCl₂. Fractions were pooled and eventually concentrated to 2-3 mg/mL, with a yield of ~0.5 mg/L medium.

The genes for TadA from *Bacillus subtilis* (BsTadA), *Staphylococcus aureus* (SaTadA) and *Escherichia coli* (EcTadA) were chemically synthesized, which were digested by *Nde*I and *Xho*I and ligated into the pET-21b (+) or pET-28a (+) vectors (Merck). The vectors were then transformed into the *Transetta* (DE3) strain (Transgen Biotech). The yeast Tad2 and Tad3 genes were amplified by PCR from *S. cerevisiae* cDNA using the primers Tad2-F1/R1 and Tad3-F1/R1 respectively. After double digestion by corresponding restriction enzymes, the two genes were sequentially ligated into the multiple cloning site (MCS) of the pETDuet-1 vector (Merck). After Tad3 was inserted into MCS1 using the *Eco*RI and *Hind*III restriction sites, Tad2 was inserted into MCS2 of the same vector using the *Bgl*II and *Xho*I sites. Then the entire DNA fragment carrying both genes was further amplified from the vector using the
primers Tad3-F2 and Tad2-R1, and subcloned into pET-28a (+) vector (Merck). All the primers used in this study were listed in Table 1. The over-expression, purification and follow-up treatment of the target proteins followed the same protocol of that of the WT McTadA protein.

**Point mutation and protein engineering**

Site-directed mutagenesis was performed using the Fast Mutagenesis System (Transgen Biotech) according to the manufacturer’s instructions. Primers for mutation of specific amino acids were shown in Table 1. The WT plasmids were used as the templates to obtain corresponding mutants for overexpression in Transetta (DE3). The purification method of the mutant protein follows the protocol of the WT proteins.

**Crystallization, data collection and structure determination**

2.5 mg/mL TadA and 10 mM β-ME, 15 µM ZnCl₂ were mixed and incubated on ice for 40 min, and the sample was then mixed with crystallization well solutions at a ratio of 1:1 (v/v). The crystals of the WT McTadA was grown at 4 °C using the hanging drop vapor diffusion method and the best crystallization condition was 0.49 M sodium phosphate monobasic monohydrate, 0.91 M potassium phosphate dibasic, pH 6.9. The crystals of McTadA-TM were obtained under the condition of 1.4 M ammonium sulfate, 0.1 M sodium citrate pH 6.0. Crystals grew to full size over 3 days at 4 °C. Before flash-cooling, crystals were transferred into the cryo-protective solution composed of the components from the well solution and glycerol of increasing concentrations, to a final concentration of 20 % (v/v). Crystals were then mounted on a nylon loop and flash-frozen in liquid nitrogen. Data collection of the WT crystals was carried out on an in-house XtaLab Synergy single crystal diffractometer (Oxford Diffraction), operating at 40 kV and 30 mA with a rotating copper-anode generator of a wavelength of 1.54 Å. The diffraction dataset was recorded with a 51.5-mm Onyx CCD detector. Data reduction was performed using the CrysAlisPro software (Oxford Diffraction). The McTadA-TM data were collected at Beamline 19U (BL19U1) at the Shanghai Synchrotron Radiation Facility (SSRF) with a wavelength of 0.979 Å, which was subsequently processed with the program HKL3000 [16]. All data were collected at 100 K with a nitrogen cryo stream.
The WT crystals belonged to the P4_3212 space group with 2.40-Å resolution and the asymmetric unit was predicted to contain one monomer. Molecular replacement (MR) was first performed with Phenix using the AfTadA (PDB 2A8N) as the search model [17]. After a plausible solution was obtained, the model was manually rebuilt by COOT according to the electron density map [18, 19]. The rebuilt model was fed to the phenix.refine and multiple cycles of refinement were conducted, followed by model rebuilding [20, 21]. The R_free/R_work factors were 21.9- and 19.8 % respectively and the final model was validated by Molprobity [22]. The structure of the TM mutant was solved by MR using the WT structure as the model. The R_free/R_work factors for the mutant were 25.1- and 19.8 % respectively. The structural presentation of McTadA was prepared by PyMOL (http://www.pymol.org/). All data collection and refinement statistics are presented in Table 2.

**In-vitro transcription of RNA**

The MctRNA^Arg was prepared by *in-vitro* transcription [23]. The transcribed RNA was purified by 12 % denaturing Urea-PAGE gel, extracted, and was then ethanol-precipitated. The pellet was washed and redissolved in the TE buffer (10 mM Tris-HCl pH 8.0 and 1 mM EDTA) to a concentration of 5 mg/mL and annealed by heating to 65 °C and allowed to cool to room temperature. The annealed RNA was stored at −80 °C for further use. The Samini-RNA based on the anticodon stem-loop sequence of SatRNA^Arg for the ITC assay was chemically synthesized.

**Coupled deaminase activity assays**

2.5 μM MctRNA^Arg or Samini-RNA was incubated with 2.5 μM McTadA (WT or McTadA-TM) in a reaction buffer of 10 mM Tris-HCl pH 8.0, 25 mM KCl, 2.5 mM MgCl_2 and 1 mM DTT at 30 °C for 60 min. 50 mM CHES pH 9.5, and 2 μM EcEndoV [24-27] were added to this mixture, and the reaction was incubated at 37 °C for another 50 min. The reaction was stopped by adding 2 × Urea loading buffer. The product was analyzed by 18 % urea-PAGE gel electrophoresis followed by ethidium bromide staining and UV-light exposure.

**Circular dichroism (CD) analysis**

The WT and mutant proteins were first diluted to a concentration of 0.3-0.5 mg/mL in
the buffer containing 20 mM Tris-H$_2$SO$_4$, 250 mM Na$_2$SO$_4$, and 1 mM DTT. Far-UV CD measurements were performed on J-810 spectropolarimeter (Jasco, Japan) using a Teflon-sealed quartz glass cuvette (1-mm path length, Starna, UK). A wavelength increment of 1 nm, a response time of 4 s, and a scan speed of 20 nm/min were employed during the measurements. All the resulting spectra were buffer-corrected.

**Isothermal titration calorimetry (ITC)**

The ITC experiments were conducted at 25 °C using a PEAQ ITC titration calorimeter (Malvern instruments). To exactly match their buffer compositions, the Samini-RNA and TadA proteins (WT McTadA or BsTadA/E55Q) were dialyzed against the same buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl. The affinities were measured by titrating Samini-RNA (200-300 μM) into WT McTadA or BsTadA/E55Q (30-35 μM). The first injection of 0.4 μL was followed by 25 injections of 1.5 μL drops. The MICROCAL ORIGIN software was used to determine the site-binding models that produced good fits. Individual peaks from titrations were integrated and displayed on a Wiseman plot. The first reading was removed from the analysis. The binding affinity ($K_p$) and change in enthalpy ($\Delta H$) associated with the binding events were calculated after fitting the dataset.

**Electrophoretic mobility shift assay (EMSA)**

2 μM tRNA was incubated with McTadA (WT or McTadA-TM) at different molar ratios in a buffer of 20 mM Tris-HCl pH 8.0 and 1 mM DTT on ice for 60 min. EcTadA, BsTadA, SaTadA and ScTad2/3 were used as controls. The reaction mixture was then mixed with glycerol (10% v/v) before being loaded and analyzed by 6% native gel with a running buffer containing 40 mM Tris-HCl pH 9.2. The gels were electrophoresed at 100 V for 60 min and stained either with ethidium bromide.

**Results**

**The in-vitro deamination activity of McTadA**

To test the deamination activity toward the ACG codon-containing RNA substrates, we first expressed the full-length recombinant McTadA. The protein was prone to precipitation upon isolation, and had to be stabilized by 10-15% glycerol. The
deamination assays were then performed to check McTadA’s activity, using the Samini-RNA fragment or the full-length tRNAArg of *M. capricolum* (MctRNAArg) (Fig. 1A). This was a coupled assay utilizing the combined action of both McTadA for deamination and *E. coli* Endonuclease V (EcEndoV) for RNA cleavage. The latter specifically recognizes inosine (I) and cleaves the second phosphodiester bond after I produced by the A34 deamination. The results showed that the enzyme was inactive on both two RNA molecules, while the positive control EcTadA in conjunction with EcEndoV gave rise to cleaved bands from both substrates (Fig. 1B). Additionally, the incubation of McTadA or EcTadA with the RNA substrates alone did not generate any cleaved products, meaning that there was no unwanted nuclease contamination. We also supplemented Ca^{2+}, Mg^{2+}, Zn^{2+}, or Mn^{2+} ions at different concentrations separately to the assays, but obtained the same results (Fig. 1C). ICP-MS spectrometry demonstrated that the protein contained the Zn^{2+} ion (data not shown), which ruled out the mis-folding possibility due to its tendency to precipitation. Of note, the addition of the Mn^{2+} ion resulted in protein precipitation immediately and a smear was seen in the gel. On the other hand, the EcTadA result was interesting, because tRNAArg of *E. coli* (EctRNAArg) differs in anticodon loop sequence from MctRNAArg or the tRNAArg of *Staphylococcus aureus* (SatRNAArg), which suggested its broad substrate specificity.

**The overall structure of *M. capricolum* TadA.**

To explore the underlying structural basis behind McTadA’s poor binding affinity, we crystalized the protein and solved the structure at a resolution of 2.4 Å. The crystal was in *P*4_32_2 crystallographic symmetry and there was only one molecule in the asymmetric unit (Fig. 2A). The final structure displayed a good geometry with the R_work and R_free values being refined to 19.8 % and 21.9 % respectively, and no residues fell in disallowed region in the Ramachandran plot. The protein displayed a typical tRNA adenosine deaminase fold, with two long helices flanking each side of the central beta-sheet. The sheet is composed of five strands, with b1 antiparallel to all the others. The N- (Asn5-His21) and C-terminus (Gln129-Arg147) each forms a long helix, and they run in parallel direction, but do not form any inter-helical contacts.
The density from Ser115 to Asp118 was rather poor, and was considered a disorder and not modeled. Otherwise, residues from Asp2 to Leu148 were completely resolved. In accordance to the ICP-MS result, there is a zinc ion bound to the catalytic center. The zinc ion is fixed by four ligands, comprising His54, Cys84, Cys87, and a water molecule. The four ligands form a tetrahedral geometry, with the water molecule held in position by the general base Glu56, whose OE2 atom is only 2.9 Å away (Fig. 2B). Additionally, Phe108 from the β4-β5 loop forms possible hydrophobic interactions with Phe143, Leu139, Ile136, all of which are located on the same side of the final helix (Fig. 2C). These interactions suggest that there might be crosstalks between these two regions. Crystallographic symmetry operations could generate the enzyme dimer, i.e. the functional catalytic unit, with the 2-fold axis coinciding with the dimer interface (Fig. 2A).

**Comparison to other TadAs and generation of a McTadA-RNA complex model**

A Dali search showed that the closest homologs to McTadA are SaTadA in complex with Samini-RNA (PDB 2B3J) [10], AaTadA (PDB 1WWR) [11], and TadA from *Salmonella enterica* (SeTadA, PDB 3OCQ), respectively. SaTadA shares moderate sequence homology (26 %) and structural similarity with McTadA (an RMSD of 1.6 Å over 135 residues), indicating considerable structural deviation from the latter (Fig. 3A, B).

By superimposing SaTadA complexed with Samini-RNA (PDB 2B3J) [10], we could generate the structure of the McTadA-RNA complex (Fig. 4A). While the general fold and the zinc ion position were conserved, inspection and analyses of the structures revealed the structural uniqueness for McTadA. Clashes with RNA at two regions are evident, which coincide with the unusual structural features of McTadA: 1) the Ser105-Ile110 fragment would severely clash with U33; 2) the Ile19-Asn22 fragment would bump into C35; and 3) K146 may be in close distance from C35 (3.1 Å, Fig. 4B, C). These findings were consistent with the prediction made by Yokobori *et al.* [15], based on the homology modeling of the SaTadA-Samini-RNA complex (PDB 2B3J) [10]. Therefore, the poor activity of McTadA toward RNA is likely the result of steric hindrance between the enzyme and substrates.
Attempts to restore the deamination activity and possible causes of failure

As an attempt to restore the activity of McTadA, we introduced structure-based mutations to the clashed regions: 1) replacement of the Ser105-Ile110 fragment by the Asp-Pro-Lys tri-peptide; 2) substitution of the Ile19-Asn22 fragment by the Ala-Gln-Leu-Gly tetra-peptide; 3) the K146A mutation. These mutations were either based on the SaTadA sequence, or simply reduced the side chain sizes of the clashing residues (in the K146A case). The mutant, which we name McTadA-TM, was even more unstable than wild type (WT). Nevertheless, we managed to obtain the pure protein for both biochemical characterization and follow-up crystallization trials. The McTadA-TM mutant was successfully crystallized based on the WT crystallization condition, and the structure was determined at 2.3 Å. The mutant structure resembles that of the WT, with only an RMSD value of 0.56 Å for 133 Cαs. There was still one monomer in the ASU, but the space group changed to P321. The structural superposition indicated that the key regions of the mutant were rearranged as we have designed, and would remarkably reduce its possible clashes with Samini-RNA, according to our model. Specifically, the previous Ile19-Asn22 fragment was originally located at the top portion of the α1-helix, and after the mutation the replacement fragment rotated to a certain degree to avoid the clashes with C35 (Fig. 5A). The K146A mutation considerably reduced the size of the side chain, and now poses little hindrance to C35 as well. Lastly, the substitution at the Ser105-Ile110 region produced local disorder between Lys107-Tyr111 (new numbering according to the TM mutant) and pulled the fragment away from U33. Additionally, the unresolvable density suggested flexibility of this region, which may help to relieve the rigidity and structural clashes (Fig. 5B). From our model, the current the structure of the mutant would not interfere with the binding of RNA at these two sites.

With these desired engineering purposes presumably accomplished, we next performed activity assays as described above. However, we still could not detect the deamination activity for the McTadA-TM. Circular dichroism studies showed that the TM mutant had quite similar profiles to that of WT, suggesting little structural perturbation by the mutations (Fig. S1). Thorough structural analyses revealed that the
possible reasons might be due to the fact that only the conflicts from one monomer of McTadA were considered when we design the McTadA-TM mutant, but we failed to take account for the possible influence of the other monomer. Moreover, the PDBePISA server [28] indicated that when WT McTadA forms a dimer, it forms quite a wide interfacial cleft (Fig. 6A). The buried surface area of McTadA is only 1890 Å², which is considerably smaller than that of EcTadA (2870 Å² based on PDB 1Z3A) [14], AaTadA (2800 Å² based on PDB 1WWR) [11], or Agrobacterium fabrum tRNA adenosine deaminase TadA (AfTadA, 2620 Å² based on PDB 2A8N) [17]. When RNA substrates bind, they would make contacts with both enzyme subunits. While the anticodon loop bases (especially A34) primarily interact with the catalytic subunit, other bases also require support from the non-catalytic subunit. A larger interfacial area suggests a weaker association between the enzyme and RNA substrates. This is probably caused by the translation of α4 (Ala55-Leu67), which moves away from the dimer interface (Fig. 6A). Additionally, close examination of the McTadA dimer further revealed that the R45Y46 fragment from the catalytic subunit and Ser69 from the non-catalytic subunit would block bases C35 and G37 respectively (Fig. 6B, C). Of note, the Ca of Ser69 is only 2.1 Å from N1 of G37, a backbone conflict that could not be resolved by simple point mutations. Furthermore, the quintuple mutant (double mutation R45P/Y46I on top of the triple mutant) failed to express in E. coli and we therefore stopped our engineering effort on this aspect. Taken together, the inactivity might still stem from the unresolved steric clashes due to unusual structural features of the McTadA enzyme, where the non-catalytic subunit further reduces the binding possibility of RNA substrates.

Lastly, to check our hypothesis on the weak binding capacity of the enzyme, we performed binding studies using the ITC method. Along with it, TadA from EcTadA, BsTadA, SaTadA and Tad2/3 heterodimer from S. cerevisiae (ScTad2/3) were used as controls. To eliminate the interference of possible deamination during the binding process, we also mutated the key catalytic glutamate residue (Glu59 in EcTadA, Glu55 in BsTadA, Glu55 in SaTadA respectively, and Glu56 in ScTad2). Most of these mutants except for BsTadA/E55Q turned out to be quite unstable and the proteins were
also heavily bound by co-purified nucleic acids that were difficult to remove. As a result, they constantly precipitated from solution immediately upon elution from the affinity column (data not shown). Nevertheless, we managed to perform ITC study using BsTadA/E55Q and WT McTadA. We found that while BsTadA showed quite strong affinity toward the Samini-RNA (2.8 ± 0.4 μM), McTadA did not show signs of a binding event (as suggested by the little amount of heat generated during the titration) even if we raised the protein and RNA concentrations to 35 μM and 300 μM respectively (Fig. 7). This result supported our theory that structural conflicts would arise upon the binding of tRNA and the inert activity was due to reduced binding affinity (kD) rather than catalysis (kcat). Additionally, EMSA was also carried out using full-length tRNA^Arg from *M. capricolum* or tRNA^Ala* from *S. cerevisiae*. ScTad2/3 showed reasonable binding toward SctRNA^Ala* as well as MctRNA^Arg* (although with reduced affinity). By contrast, McTadA was not found to bind MctRNA^Arg* at all (Fig. S2).

**Discussion**

The post-transcriptional modification scenarios present in tRNAs differ greatly from one organism to another, and represent the largest modification repertoire found in all RNA types. Nucleotide 34 in tRNA plays an essential role in determining the degenerate codons during the translation process by pairing with the last base of the corresponding codons. First proposed by Francis Crick more than a half century ago, the wobble hypothesis states that the third position of the codon on mRNA and the first position of the anticodon on tRNA could pair with each other in a less stringent manner than that of the canonical base pairs, and therefore expands the codon usages. In accordance with this theory, base 34, or the “so-called” wobble base is usually extensively modified to adapt to the different basepairing needs.

One frequently occurring modification is that the wobble A34 in some precursor tRNAs is enzymatically converted to inosine by specific tRNA:A34 deaminases during tRNA maturation. The deaminated inosine is capable of reading the C, U and A-ending codons, but the basepairing with a G-ending codon is not allowed [9].
the A-to-I deamination provides more possibilities and diversities of pairing combinations, and is adopted by many organisms including eukaryotes. For example, in most bacteria, the four arginine codons (CGN) are decoded by two tRNAs. One tRNA harboring a wobble inosine (tRNA\textsuperscript{Arg}_{ICG}) reads the CGU, CGC and CGA codons, while a second tRNA harboring a wobble cytidine (tRNA\textsuperscript{Arg}_{CCG}) reads the remaining CGG codon. However, \textit{M. capricolum} and some closely related Mollicutes present exceptions to the rule, in that the gene encoding tRNA\textsuperscript{Arg}_{CCG} or even \textit{tadA} is missing. \textit{M. capricolum} harbors only one copy of \textit{tadA} and tRNA\textsuperscript{Arg}_{ACG} genes. As A34 deamination is considered an early event in tRNA maturation process, the 6 % non-deaminated A34-tRNA molecules are not to be easily overlooked, comparing to the 100 % A-to-I conversion observed in \textit{B. sublitis}. On the other hand, decoding of the CGG codon still poses a problem, as I34-tRNA\textsuperscript{Arg}_{CCG} would not be a good decoder. After studying the tRNA\textsuperscript{Arg} sequences of \textit{M. capricolum}, Yokobori \textit{et al}. therefore deduced that tRNA\textsuperscript{Arg}_{ACG} molecules with a non-deaminated wobble adenosine in this organism is used to decode the CGG codon [15]. To explain the poor activity toward McTadA, they proposed that the enzyme would bind poorly to its tRNA substrates, based on the homology modeling of the SaTadA-Samini-RNA complex.

To validate this interesting hypothesis, we solved the structures of the WT and a mutant form of McTadA in this study. Our crystallographic studies and structural analyses offered support to Yokobori’s hypothesis in that multiple structural clashes against the RNA substrates would lead to a reduced association. In our \textit{in-vitro} activity assays, we did not observe any deamination activity from the McTadA in the recombinant form for unknown reasons. It is possible that the deamination process requires some pre-requisite tRNA modifications, which are not present on tRNA transcript. Additionally, we could not rule out the possibility that deamination catalyzed by McTadA may require an accessory protein or cofactor \textit{in vivo} yet to be identified.

Upon structural comparison to that of EcTadA or SaTadA, we introduced site-specific mutations to the enzyme and tried to restore its activity. However, the
TM mutant did not re-gain the deamination activity, and the expression failure of the quintuple mutant designed in the second-round mutagenesis prevented us from further engineering trials. Possible reasons for this result could come from the dimer interface differences. One should note that all our modeling efforts were merely based on the Samini-RNA substrate from the SaTadA complex (PDB 2B3J) [10], due to a lack of additional structural information on (t)RNA-bound form of TadAs. McTadA may adopt a distinct substrate binding mode, and in this case it calls upon the structure determination of McTadA in complex with its genuine RNA substrates.

The structural and sequence differences between McTadA and its orthologs are interesting. Among the three conflicts, the largest structural conflict comes from the Ser105-Ile110 fragment, in which Phe108 makes hydrophobic contacts with the terminal helix. These non-polar interactions fix this bulky and conserved phenylalanine residue in place, most likely to exclude the binding of RNAs as well as the base flipping of A34 and C35 (compare glycine or alanine in equivalent position in bacteria, Fig. 6D). These two nucleosides are the two most important recognition elements for the deaminase, especially for the yeast ScTad2/3 enzyme. Similarly, Asn22 is sterically incompatible with C35 and in many Mollicutes, and it is semi-conserved or replaced by iso-steric or larger residues in Mollicutes (glutamate or aspartate, Fig. 6D) but is substituted in bacteria. In contrast, glycine is found in this position in most other bacteria with a fully functional TadA enzyme. In addition to the visible hindrance from the complex model, the enzyme displays a wider dimer interface, which also weakens the association of RNA. All known bacterial TadA enzymes work as homodimers, and only one subunit carries out the catalytic function whereas the other subunit only provides a fraction of support to RNA substrates. On the other hand, euarkytic Tads work as heterodimers, but the Tad3 subunit is also a helper subunit due to the mutation of the key catalytic glutamate residue. As the two subunits of heterodimeric Tad each clearly have designated functions, the catalysis could not be achieved by a single subunit. In comparison, the second monomer of the McTadA contributes very little to the binding of RNA due to its loose association to the catalytic monomer, and it also creates structural conflicts around the Ser69 residue.
Consequently, by reducing the possible interactions and increasing destructive hindrance with RNA substrates, both the binding and catalytic efficiency of the McTadA deamination reaction would be quite low. In a sense, McTadA ingeniously employs the dimer assembly to impair its own deamination function. Consequently, *M. capricolum* retains a portion of non-deaminated tRNA<sup>Arg</sup><sub>ACG</sub>, which would make sure that all four arginine CGN codons be decoded.

The A34-tRNA deamination facilitates the codon recognition process and allows codon expansion according to the wobble hypothesis. This beneficial modification is widely present in both bacteria and eukarya. Consequently, Tad is an essential enzyme and the cells without this gene are not viable. However, *M. capricolum* and some Mollicutes choose to “sabotage” this modification event, since by retaining the A34-containing tRNA, they are equipped with the capability to decode yet an extra codon. The unmodified adenosine 34 in tRNA<sup>Arg</sup> was reported in as early as 1984 [29]. In this work, Sibter and co-workers identified a yeast mitochondria tRNA<sup>Arg</sup> belonging to the four CGN arginine codon family that contained unmodified A34. They tried to explain the decoding problem of the CGN codons with the classical ‘two out of three’ hypothesis, as the other two base pairs are of G-C combination. Moreover, the chloroplasts of *A. thaliana* encode only tRNA<sup>Arg</sup><sub>ACG</sub> and tRNA<sup>Arg</sup><sub>UCU</sub> to decode the arginine codon. However, the knockdown of the chloroplastic TadA expression was proved non-lethal to plants [30], similar to the *M. capricolum* scenario studied here. In Yokobori’s work, some species were found to go even further than *M. capricolum*, by losing the *tadA* gene (the Groups II and III Mollicutes) and relying completely on the ACG anticodon with an unmodified A34 to read the arginine four-codon box. This decoding strategy could be considered as another type of ‘superwobbling’ proposed by Rogalski *et al.*, in which a single base at the wobble position is employed to decode all degenerate codons and allows the viability of a species with a reduced set of tRNAs [31]. However, as Yokobori *et al.* have proposed, changing the decoding strategy through cellular evolution necessitates a series of pre-arranged events, such as complex point mutations in modification enzymes (probably also in tRNAs) etc. [15]. Whether the extra efforts to evolve the TadA
enzyme and corresponding tRNA genes are worthwhile in terms of cellular resources (energy, metabolite availability etc.) is an interesting question and warrants further investigation. For example, a possible scenario is that *M. capricolum* could retain a gene to encode tRNA^{Arg}_{CCG} to save going through the troubles of mutations. The current non-classical decoding strategy discovered by Yokobori *et al.* is a result of co-evolution of both the tRNA and *tadA* genes. In addition, the partial and complete “bypass” of the A34 deamination process provides an alternative solution to the predicament of organisms with reduced genomes (thus less tRNA species) and faster evolving speed like Mycoplasmas or Mollicutes. This strategy may not be unique among all the organisms, but similar approaches can be certainly adopted by others.
Database Depositions

The atomic co-ordinates and structure factors have been deposited in the RCSB Protein Data Bank with accession numbers 6L2L and 6L2M.

Abbreviations

Tad, tRNA-adenosine deaminase; ADAR, adenosine deaminase acting on RNA; CDA, cytidine deaminase; EcTadA, *Escherichia coli* TadA; AaTadA, *Aquifex aeolicus* TadA; SaTadA, *Staphylococcus aureus* TadA; McTadA, *M. capricolum* TadA; EcEndoV, *E. coli* Endonuclease V; *E. coli*, *Escherichia coli*; SeTadA, *Salmonella enterica* TadA; WT, wild type; McTadA-TM, the McTadA triple mutant; AfTadA, *Agrobacterium fabrum* TadA.

Author contribution

Wei Xie conceived and designed the research; Huijuan Liu, Saibin Wu and Dewei Ran performed the research; Wei Xie analyzed data and wrote the paper. All authors reviewed the manuscript.

Acknowledgements

We thank the BL19U1 beamline at Shanghai Synchrotron Radiation Facility for assistance during diffraction data collection. Funding for this work was provided by the National Natural Science Foundation of China 31870782, the Natural Science Foundation of Guangdong province 2018A030313313 and 2017A030310190.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.
References

1. Schmeing, T. M., Voorhees, R. M., Kelley, A. C., Gao, Y. G., Murphy, F. V., Weir, J. R. and Ramakrishnan, V. (2009) The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science*. **326**, 688-694.

2. Motorin, Y. and Helm, M. (2010) tRNA stabilization by modified nucleotides. *Biochemistry*. **49**, 4934-4944.

3. Hopper, A. K. (2013) Transfer RNA post-transcriptional processing, turnover, and subcellular dynamics in the yeast *Saccharomyces cerevisiae*. *Genetics*. **194**, 43-67.

4. Helm, M., Brule, H., Degoul, F., Cepanec, C., Leroux, J. P., Giege, R. and Florentz, C. (1998) The presence of modified nucleotides is required for cloverleaf folding of a human mitochondrial tRNA. *Nucleic Acids Res*. **26**, 1636-1643.

5. Zhang, Q., Zhang, L., Chen, D., He, X., Yao, S., Zhang, Z., Chen, Y. and Guan, M. X. (2018) Deletion of Mtu1 (Trmu) in zebrafish revealed the essential role of tRNA modification in mitochondrial biogenesis and hearing function. *Nucleic Acids Res*. **46**, 10930-10945.

6. Wu, Y., Wei, F. Y., Kawarada, L., Suzuki, T., Araki, K., Komohara, Y., Fujimura, A., Kaitsuka, T., Takeya, M., Oike, Y., Suzuki, T. and Tomizawa, K. (2016) Mtu1-mediated thiouridine formation of mitochondrial tRNAs is required for mitochondrial translation and is involved in reversible infantile liver injury. *Plos Genet*. **12**, e1006355.

7. Björk, G. R. (1995) Biosynthesis and function of modified nucleosides. In tRNA: structure, biosynthesis, and function (Söll, D. and RajBhandary, U., ed.), pp. 165-205, ASM Press, Washington, DC.

8. Curran, J. F. (1998) Modified nucleosides in translation. In Modification and Editing of RNA (Grosjean, H. and Benne, R., ed.), pp. 493-516, ASM Press, Washington, DC.

9. 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**, 437-458.
10. Losey, H. C., Ruthenburg, A. J. and Verdine, G. L. (2006) Crystal structure of
Staphylococcus aureus tRNA adenosine deaminase TadA in complex with RNA.
Nat Struct Mol Biol. 13, 153-159.

11. Kuratani, M., Ishii, R., Bessho, Y., Fukunaga, R., Sengoku, T., Shirouzu, M.,
Sekine, S. and Yokoyama, S. (2005) Crystal structure of tRNA adenosine
deaminase (TadA) from Aquifex aeolicus. J Biol Chem. 280, 16002-16008.

12. Gerber, A. P. and Keller, W. (1999) An adenosine deaminase that generates inosine
at the wobble position of tRNAs. Science. 286, 1146-1149.

13. Tsutsumi, S., Sugiura, R., Ma, Y., Tokuoka, H., Ohta, K., Ohte, R., Noma, A.,
Suzuki, T. and Kuno, T. (2007) Wobble inosine tRNA modification is essential to
cell cycle progression in G(1)/S and G(2)/M transitions in fission yeast. J Biol
Chem. 282, 33459-33465.

14. Kim, J., Malashkevich, V., Roday, S., Lisbin, M., Schramm, V. L. and Almo, S. C.
(2006) Structural and kinetic characterization of Escherichia coli TadA, the
wobble-specific tRNA deaminase. Biochemistry. 45, 6407-6416.

15. Yokobori, S., Kitamura, A., Grosjean, H. and Bessho, Y. (2013) Life without
tRNA^Arg-adenosine deaminase TadA: evolutionary consequences of decoding the
four CGN codons as arginine in Mycoplasmas and other Mollicutes. Nucleic Acids
Res. 41, 6531-6543.

16. Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data
collected in oscillation mode. Methods Enzymol. 276, 307-326.

17. Elias, Y. and Huang, R. H. (2005) Biochemical and structural studies of A-to-I
editing by tRNA:A34 deaminases at the wobble position of transfer RNA.
Biochemistry. 44, 12057-12065.

18. Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. (2010) Features and
development of Coot. Acta Crystallogr D. 66, 486-501.

19. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular
graphics. Acta Crystallogr D. 60, 2126-2132.

20. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N.,
Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J.,
Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D. 66, 213-221.

21. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. and Adams, P. D. (2012) Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D. 68, 352-367.

22. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D. 66, 12-21.

23. Wang, C., Guo, Y., Tian, Q., Jia, Q., Gao, Y., Zhang, Q., Zhou, C. and Xie, W. (2015) SerRS-tRNA(Sec) complex structures reveal mechanism of the first step in selenocysteine biosynthesis. Nucleic Acids Res. 43, 10534-10545.

24. Zhang, Z., Jia, Q., Zhou, C. and Xie, W. (2015) Crystal structure of E. coli endonuclease V, an essential enzyme for deamination repair. Sci Rep. 5, 12754.

25. Guo, G., Ding, Y. and Weiss, B. (1997) nfi, the gene for endonuclease V in Escherichia coli K-12. J Bacteriol. 179, 310-316.

26. Gates, F. T. and Linn, S. (1977) Endonuclease V of Escherichia coli. J Biol Chem. 252, 1647-1653.

27. Weiss, B. (2001) Endonuclease V of Escherichia coli prevents mutations from nitrosative deamination during nitrate/nitrite respiration. Mutat Res. 461, 301-309.

28. Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol. 372, 774-797.

29. Sibler, A. P., Dirheimer, G. and Martin, R. P. (1986) Codon reading patterns in Saccharomyces cerevisiae mitochondria based on sequences of mitochondrial tRNAs. FEBS Lett. 194, 131-138.

30. Delannoy, E., Le, R. M., Faivre-Nitschke, E., Estavillo, G. M., Bergdoll, M., Taylor, N. L., Pogson, B. J., Small, I., Imbault, P. and Gualberto, J. M. (2009)
Arabidopsis tRNA adenosine deaminase arginine edits the wobble nucleotide of chloroplast tRNA<sub>Arg</sub>(ACG) and is essential for efficient chloroplast translation. *Plant Cell.* 21, 2058-2071.

31. Rogalski, M., Karcher, D. and Bock, R. (2008) Superwobbling facilitates translation with reduced tRNA sets. *Nat Struct Mol Biol.* 15, 192-198.
### Table 1. Sequences of the primers and tRNA used in this study.

The restriction sites of the cloning primers were underlined in the amplifying primers while the mutated bases were in italic in the Quikchange primers.

| Primers name | Sequence (5’-3’) |
|--------------|-----------------|
| TadA-pET28a-F | CGGCAGCCATATGGACGACTTAAACAAATAC | |
| TadA-pET28a-R | GGGCCGACTCGAGATGGACGACTTAAACAAATAC | |
| TadA-pCOLD-MBP-F | GGGCCGACCTCGAGATGGACGACTTAAACAAATAC | |
| TadA-pCOLD-MBP-R | CCCGCAGCCATATGGACGACTTAAACAAATAC | |
| TadA-pET21b-F | CGGCAGCCATATGGACGACTTAAACAAATAC | |
| TadA-pET21b-R | GGGCCGACCTCGAGATGGACGACTTAAACAAATAC | |
| S105-I110(DPK)-pET21b-F | GGCACGATGGACGACTTAAACAAATAC | |
| S105-I110(DPK)-pET21b-R | GGGCCGACCTCGAGATGGACGACTTAAACAAATAC | |
| I19-N22 (AQLG)-pET21b-F | CTTGAGACTCGAGATGGACGACTTAAACAAATAC | |
| I19-N22 (AQLG)-pET21b-R | CTTGAGACTCGAGATGGACGACTTAAACAAATAC | |
| K146A-pET21b-F | ATCTTCCCTCATCAACGCGACGCTCCGAGC | |
| K146A-pET21b-R | ATCTTCCCTCATCAACGCGACGCTCCGAGC | |
| R45P-Y46I-pET21b-F | ATGACGCTCCGAGATGGACGACTTAAACAAATAC | |
| R45P-Y46I-pET21b-R | ATGACGCTCCGAGATGGACGACTTAAACAAATAC | |
| EcTadA-E59Q-pET28a-F | CCCACGCACATGCACAATTACGCGGCTTCGGGC | |
| EcTadA-E59Q-pET28a-R | CCCACGCACATGCACAATTACGCGGCTTCGGGC | |
| McTadA-E56Q-pET21b-F | ATCGACGATGGACGACTTAAACAAATAC | |
| McTadA-E56Q-pET21b-R | ATCGACGATGGACGACTTAAACAAATAC | |
| SaTadA-E55Q-pET28a-F | GGGCTTTTTCGATTGCAATATG | |
| SaTadA-E55Q-pET28a-R | GGGCTTTTTCGATTGCAATATG | |
| BsTadA-E55Q-pET28a-F | GGGCTTTTTCGATTGCAATATG | |
| BsTadA-E55Q-pET28a-R | GGGCTTTTTCGATTGCAATATG | |
| Tad3-F1 | GCTTAGAATTCGATGGTTAGGAAAGT | |
| Tad3-R1 | ATAGGAAGAGGCTTTCGACGCAAGATCATTTGAT | |
| Tad2-F1 | CATATGGCAGATCTCCAGACGATATTAAACAAATAC | |
| Tad2-R1 | ATAGGAAGAGGCTTTCGACGCAAGATCATTTGAT | |
| MetRNA<sup>Arg</sup> | CCCCGGTAGATCAATTTGGATAGATCGCTTGGACTACG | |
| Samini-RNA | GGUGACUACGGGAUCACC | |
Table 2. Data collection and refinement statistics.

|                           | Apo-WT (6L2L)       | TM (6L2M)          |
|---------------------------|---------------------|--------------------|
| **Data collection**       | Home X-ray          | SSRF-BL19U         |
| Resolution (Å)            | 21.10-2.40 (2.53-2.40) | 50.00-2.30 (2.38-2.30) |
| Space group               | P4321               | P321               |
| Cell dimensions (Å)       |                     |                    |
| a, b, c (Å)               | 49.1, 49.1, 147.9   | 55.6, 55.6, 109.2  |
| α, β, γ (°)               | 90, 90, 90          | 90, 90, 90         |
| R<sub>merge</sub><sup>b</sup>| 0.10 (0.46)         | 0.105 (0.51)       |
| CC<sub>1/2</sub>          | 0.998 (0.884)       | 0.990 (0.969)      |
| Redundancy                | 6.6 (7.1)           | 18.7 (18.1)        |
| Completeness (%)          | 99.8 (99.9)         | 100 (100)          |
| I/σ(I)                    | 15.6 (4.5)          | 38.5 (4.8)         |
| **Refinement**            |                     |                    |
| Resolution (Å)            | 21.10-2.40 (2.75-2.40) | 48.15-2.30 (2.64-2.30) |
| No. reflections           | 7649                | 9119               |
| R<sub>work</sub>/R<sub>free</sub><sup>d</sup>(%) | 19.8/21.9 | 19.8/25.1 |
| No. atoms                 |                     |                    |
| Protein                   | 1160                | 1147               |
| Ligand                    | 1 (Zn<sup>2+</sup>) | 1 (Zn<sup>2+</sup>) | 2(Cl<sup>-</sup>) |
| Water                     | 59                  | 38                 |
| B-factor (Å<sup>2</sup>)  |                     |                    |
| Protein                   | 36.7                | 48.1               |
| Ligand                    | 27.9                | 31.3 (Zn<sup>2+</sup>) | 54.8 (Cl<sup>-</sup>) |
| Water                     | 32.5                | 46.0               |
| RMS (bonds) (Å)           | 0.005               | 0.007              |
| RMS (angles) (°)          | 0.81                | 0.81               |
| Ramachandran favored (%)  | 96.40               | 95.71              |
| Outliers (%)              | 0                   | 0                  |

<sup>a</sup>Values in parentheses are for the highest-resolution shell.  
<sup>b</sup>R<sub>merge</sub> = Σ|I - <I>|/σ(I), where I is the observed intensity.  
<sup>c</sup>R<sub>work</sub> = Σ||Fo| - |Fc||/Σ|Fo|, calculated from working data set.  
<sup>d</sup>R<sub>free</sub> is calculated from 5.0 % of data randomly chosen and not included in refinement.
Figure 1. Biochemical characterization of WT McTadA. (A) The RNA substrates used in the assays. A34 was colored in red. (B) The coupled deaminase activity assay on McTadA. The red arrow indicates the cleaved product at correct size from MctRNAArg. (C) The activity assays with supplemented metals. Two concentrations were employed for each metal.

Biochemical Journal. This is an Accepted Manuscript. You are encouraged to use the Version of Record that, when published, will replace this version. The most up-to-date version is available at https://doi.org/10.1042/BCJ20190858
Figure 2. The overall structure of apo-McTadA. (A) The structure was shown in the ribbon rendition in two orthogonal views. The N- and C-termini were indicated and the secondary structure elements of McTadA were labeled. The blue dot indicated the zinc ion. Only one monomer was drawn before the rotation, which was colored according to its secondary structure. The right panel: the protein was drawn in the dimeric form, where the second monomer was shown in gray. (B) The close-up view of the zinc-coordination state. The side chains of the key residues were shown as sticks and labeled. The water molecule for catalysis (Wat) was represented by the pink sphere and its distance from the general base Glu56 was indicated. The distances of the ligands from the metal were shown (units in Å). (C) The hydrophobic interactions of Phe107 with residues Ile136, Leu139 and Phe143 from the terminal helix.
Figure 3. Structure and sequence comparison to homologs. (A) The structural superposition of the apo-McTadA (PDB 6L2L, color scheme as in figure 1A) with that of tRNA-bound SaTadA (PDB 2B3J, cyan, RNA was omitted for clarity), AaTadA (PDB 1WWR, magenta), and SeTadA (PDB 3OCQ, yellow), respectively. The proteins were represented in backbone traces and in cross-eyed stereoview. (B) Multiple sequence alignments of the structural homologs. The secondary structure elements were labeled above the sequences. The red dots indicated key residues responsible for zinc binding, and the up-arrow indicated the catalytic Glu56 residue. The regions for the first- and second-round engineering were represented by the black and green broken boxes respectively.
Figure 4. Modeling studies showing potential steric hindrance between McTadA and the Samini-RNA substrate. (A) Cartoon showing the overall hypothetical McTadA-Samini-RNA complex. The coordinates of Samini-RNA (in orange) were from PDB 2B3J and only nucleotides U33-C35 were drawn. The two most severe structural conflicts were indicated by the broken red circles. P5P34: the hypoxanthine base. (B) The close-up of the clashes posed by the Ser105-Ile110 fragment; and (C) by the Ile19-Asn22 fragment as well as Lys146. The key residues were shown as sticks.
Figure 5. Modeling studies showing the relief of the potential steric hindrance between the McTadA-TM mutant and the Samini-RNA substrate (from PDB 2B3J). The orientation of each panel was the same as that in Figure 4. (A) Cartoon showing the close-up view of the structural rearrangements by the AQLG fragment (the replacement of the original IKHN fragment at positions of 19-22); and (B) by the 105DPK107 fragment as well as Ala146. The key residues were shown as sticks.
Figure 6. Further explanation of the inability of the McTadA-TM mutant to catalyze. (A) The superposition of chain A of the McTadA dimer (with chain A and B colored in cyan and gray respectively) with that of the EcTadA (PDB 1Z3A, blue), AaTadA (PDB 1WWR, magenta) and AfTadA (PDB 2A8N, orange) dimers in the ribbon representation. Note the translation of α4 as indicated by the arrow. (B) The modeled structure of RNA-bound McTadA dimer, showing more structural clashes between the RNA substrate and the McTadA-TM mutant. Unresolved structural conflicts were shown by the red box. (C) The close-up view showing the proximity of C35 and G37, to Tyr46 of the catalytic subunit and Ser46 from the non-catalytic subunit respectively. (D) The multiple sequence alignment with a focus on the region of Asn22 and Phe108, indicated by the red arrow head and red arrow respectively. The bacteria and Mollicute species were divided by the red horizontal line. *Nitrosomonas, Nitrosomonas europaea; Aquifex, A. aeolicus; Streptomyces, Streptomyces avermitilis; Synchococcus, Synchococcus elongatus; Staphylococcus, S. aureus; Listeria, Listeria monocytogenes; Oenococcus, Oenococcus oeni; Acholeplasma, A. laidlawii; Aster-yellow, Aster yellows witches’-broom phytoplasma; Phyto-aust, Candidatus Phytoplasma australiense; Phyto-mali, Candidatus Phytoplasma mali; Onion-yellow, Onion yellows phytoplasma; Mesoplasma, Mesoplasma florum L1.* The secondary structures for EcTadA and MacTad were labeled at the top and bottom of alignment.
Figure 7. The ITC binding studies. The binding affinities of BsTadA/E55Q (A) and WT McTadA (B) to Samini-RNA were tested.
Supplementary Material

Structure of a tRNA-specific deaminase with compromised deamination activity

Huijuan Liu¹, Saibin Wu¹, Dewei Ran¹ and Wei Xie¹,*

¹MOE Key Laboratory of Gene Function and Regulation, State Key Laboratory for Biocontrol, School of Life Sciences, The Sun Yat-Sen University, 135 W. Xiangang Rd., Guangzhou, Guangdong, 510006, People's Republic of China

*Correspondence: Wei Xie (xiewei6@mail.sysu.edu.cn)

Contents:
Supplementary figures S1-2
Figure S1. The far-UV CD analysis on the secondary structure of the McTadA WT and mutant.
Figure S2. The EMSA studies performed using McTadA and ScTad2/3. For each binding pair, the protein or the tRNA was loaded as a blank control. Various molar ratios (protein:tRNA) ranging from 1:1-8:1 were employed.
200 μM Samini-RNA to 30 μM BsTadA-E55Q mutant

\[ K_d (V) = 2.82e-6 = 414e-9 \]

300 μM Samini-RNA to 35 μM WT McTadA