



Functional and structural investigation of N-terminal domain of the SpTad2/3 heterodimeric tRNA deaminase



Xiwen Liu ^{a,b}, Jie Zhou ^a, Ruiguang Ge ^{a,*}, Wei Xie ^{a,*}

^aMOE Key Laboratory of Gene Function and Regulation, State Key Laboratory for Biocontrol, School of Life Sciences, The Sun Yat-Sen University, Guangzhou, Guangdong 510006, People's Republic of China

^bDepartment of Colorectal Surgery, The Sixth Affiliated Hospital, Sun Yat-sen University, 26 Yuancun Erheng Rd, Guangzhou, Guangdong 510655, People's Republic of China

ARTICLE INFO

Article history:

Received 29 March 2021

Received in revised form 27 May 2021

Accepted 2 June 2021

Available online 05 June 2021

Keywords:

Adenosine deaminase acting on tRNA

Tad

Crystal structure

Kinase

ABSTRACT

Editing is a post-transcriptional process that changes the content of nucleic acids occurring on both DNA and RNA levels. Inosine at position 34 in tRNA is one such example, commonly produced via the deamination of A34, catalyzed by adenosine deaminase acting on tRNA (ADAT or Tad). The formation of inosine is essential for cell viability. The eukaryotic deaminases normally consist of the catalytic subunit Tad2 and the structural subunit Tad3, but the catalytic process is poorly understood. Despite the conservation of the (pseudo-) catalytic domains, the heterodimeric enzyme Tad2/3 also possesses additional domains that could exhibit novel functions. Here we present the structure of the N-terminal domain of the *Schizosaccharomyces pombe* Tad2/3 heterodimeric tRNA(A34) deaminase (N-SpTad2), which shares ~30% sequence identities with uridine-cytidine or pantothenate kinases, but lacks the predicted kinase functions. While biochemical assays indicated that the domain is not a nucleic-acid binder, it is able to significantly influence the A34-tRNA deamination activity of the holoenzyme. Through co-expression and purification analyses, we deduce that N-SpTad2 plays a role in mediating protein-protein contacts and enhancing the stability and solubility of SpTad2/3, without which the deaminase is not functional. Taken together, our structural and biochemical studies highlighted the importance of the additional domains to the intrinsic deaminase functions of heterodimeric Tad2/3 enzymes and promoted our understanding on this essential post-transcriptional tRNA modification.

© 2021 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

tRNA is the adaptor molecule in the translational process by linking the genetic information and the corresponding amino acid. As a part of its maturation process, it is heavily modified post-transcriptionally. Among over 100 modifications reported, nucleotides 34 and 37 in tRNA are frequently modified to ensure translational fidelity and efficacy in cells. The enzyme TadA or Tad2/3 catalyzes this modification by converting A34 to inosine, which is a critical and conserved event in both bacteria and eukarya. Inosine is capable of base pairing with U-, C- and A (i.e. “wobble”), and hence expands the usage of tRNA codons. Therefore, the I34 modification is essential for cell viability [1–6].

Inosine at position 34 (I34) is only found in tRNA^{Arg} in bacteria, but in 7–8 tRNA species in eukaryotes. The reason for more tRNA

substrates in the latter is that two different types of enzymes are employed by their respective kingdoms. The homodimeric enzyme TadA is the A34-tRNA deaminase in bacteria, which has strict substrate specificity due to its strict recognition of U33-G37 in the anticodon of tRNA^{Arg} [7]. In contrast, the eukaryotic counterpart of TadA is a heterodimeric deaminase (hetTad2/3), consisting of two subunits named Tad2 (or ADAT2) and Tad3 (or ADAT3) respectively, and has relaxed substrate specificities. Both types of enzymes belong to the cytidine deaminase superfamily. Sequence alignment revealed that hetTad2/3 possesses the conserved catalytic core of TadA and the zinc-binding signature sequences, but the key catalytic residue glutamate in the H(C)XE motif in Tad3 is missing [4]. As a result, Tad2 is the catalytic subunit, and Tad3 only plays a structural role. Additionally, Tad2 and Tad3 may contain additional domains that could exhibit novel functions.

hetTad2/3 enzymes have only been identified and investigated from several organisms: yeast [1], *Trypanosoma brucei* [3,8] and *Arabidopsis thaliana* [5], but the deamination mechanisms remain poorly understood. This key modification has been used to explain

* Corresponding author.

E-mail addresses: gerg@mail.sysu.edu.cn (R. Ge), xiewei6@mail.sysu.edu.cn (W. Xie).

the correlation between codon usage and tRNA pools. A recent study showed that the human I34 modification efficiency is regulated at a few developmental stages [6]. These findings indicated that both tRNA modifications and codon usage are implicated in gene expression regulation. More recently, the homozygous V144M mutation in human Tad3 was found to be causative for autosomal-recessive intellectual disability (ID) in 24 affected individuals [9].

We previously determined the crystal structure of the *Saccharomyces cerevisiae* Tad2/3 (ScTad2/3) heterodimer and revealed the assembly of eukaryotic Tad2/3 enzymes. The structure demonstrated that there exist extensive interactions between the two subunits, which explains the instability of the V144M mutant. Importantly, we discovered that the N-terminus of ScTad3 forms a stand-alone domain and contributes to the binding of tRNA. Moreover, ScTad3 employs its C-terminal cysteine to block the pseudo-catalytic pocket and eliminate its own deamination activity [10]. In this study, we first identified a potential nucleotide kinase located at the N-terminus of *Schizosaccharomyces pombe* Tad2 (N-SpTad2). However, due to the missing of some key residues, N-SpTad2 is not a functional kinase. To further study its function, we solved the crystal structure of this domain. Follow-up biochemical assays indicated that while this domain does not bind nucleic acids, it may mediate important contacts within the enzyme and influence the A34-tRNA deamination activity. This work highlights the importance of the accessory domain to the intrinsic deaminase functions of hetTad2/3s and promotes our understanding on this essential tRNA modification.

2. Methods

2.1. Gene cloning, protein expression and purification

The sequences of the *Schizosaccharomyces pombe* *tad2* (Gene ID: 2539634), *tad3* (Gene ID: 2541987), *Escherichia coli* *panK* (Gene ID: 948479) and *UCK* (Gene ID: 946597) genes were obtained from the NCBI database. *tad3* and *tad2* were amplified from *Schizosaccharomyces pombe* cDNA, and cloned into the multi-cloning sites I (MCSI) and MCS II of the pETDuet-1 vector respectively (Merck). The *EcUCK* and *EcPanK* genes were amplified from *Escherichia coli* strain DH5 α and subsequently cloned into the modified pET-28a vector (Merck). N-SpTad2 (Met1-Thr202) was cloned into the pET-28a vector as well. The primers used in this study were listed in Table S1. Both vectors employed the PreScission protease (PSP, Cytiva) site for tag removal. The mutants were obtained using the *Quikchange* method. After sequencing for the correct clones, the plasmids were transferred into the *E. coli* BL21 (DE3) strain for over-expression. Cells were cultured in 2L Luria-Bertani (LB) liquid medium containing the antibiotic until OD₆₀₀ reached 0.8. 0.2 mM IPTG was added to induce the expression of proteins of interest at 22 °C overnight. Then the cells were harvested by centrifugation at 3,000 g for 20 min and resuspended in pre-chilled nickel-nitrilotriacetic acid (Ni-NTA) buffer A containing 40 mM Tris-HCl (pH 8.0), 250 mM NaCl, 10 mM imidazole, 5 mM β -mercaptoethanol and 1 mM phenylmethylsulfonyl fluoride (PMSF). The resuspended cells were lysed by ultrasonication and the supernatant was obtained by centrifugation at 23,500 g for 1 h at 4 °C. The supernatant was then applied onto Ni-NTA affinity resin (Qiagen) pre-equilibrated with Ni-NTA buffer A. The target protein was eluted with Ni-NTA buffer B containing 40 mM Tris-HCl (pH 8.0), 250 mM NaCl, 250 mM imidazole, 5 mM β -mercaptoethanol and 1 mM PMSF. The 6 \times His-tag at the N-terminus was cleaved off by being treated with protease overnight at 4 °C in the presence of 5 mM β -mercaptoethanol, and the tag-free protein was loaded onto a Histrap HP column (Cytiva). The flow through was collected

and dialyzed against a final buffer containing 20 mM Tris-HCl pH 8.0, 250 mM NaCl, and 1 mM DTT. For size exclusion chromatography analysis, 0.5 mg of the sample was loaded to the Superdex 200 column (10/300, Cytiva), which was pre-balanced by a buffer containing 20 mM Tris-HCl pH 8.0, 250 mM NaCl, and 1 mM DTT with the flow rate of 0.5 ml/min. The proteins intended for enzyme activity assays were supplemented with 10% glycerol, flash cooled in liquid nitrogen, and stored at – 80 °C until further use.

2.2. Crystallization, data collection and structural analysis

The screens for crystals were set up at room temperature using the sitting-drop vapor diffusion method with the protein concentration of 4 mg/ml, and the best crystals were obtained from 24% PEG1500, 0.2 M (NH₄)₂SO₄, 0.1 M NaOAc, pH 5.0. Diffraction data were collected using Beamline 19U (BL19U1) at the Shanghai Synchrotron Radiation Facility (SSRF, Shanghai, P. R. China) and were processed with the program HKL-3000 [11]. The crystals diffracted X-rays to 2.60Å resolution and belong to the P1 space group. The asymmetric unit was predicted to contain four molecules. Molecular replacement (MR) was first performed with Phenix [12] using the human UCK2 (hUCK2) structure as the search probe (PDB 6PWZ). After a plausible solution was obtained, the model was manually built by COOT [13] according to the electron density map. The rebuilt model was fed to the phenix.refine [14] and multiple cycles of refinement were conducted, followed by model rebuilding. The R_{free}/R_{work} factors were 0.218 and 0.168 respectively and the final model was validated by Molprobit [15]. The structural representation of N-SpTad2 (PDB 7EEY) was prepared by PyMOL (<http://www.pymol.org/>).

2.3. Coupled deamination assay

The 10- μ l reaction systems contained 100 mM Tris-HCl pH 8.0, 25 mM KCl, 2.5 mM MgCl₂, 1 mM DTT, and 2 μ M tRNA. Full-length or truncated proteins were added to the reaction system to a concentration of 60 nM to start the reaction. The mixture was incubated at 30 °C for 30 min. then 0.5- μ l CHES (1 M, pH 9.5), and 2 μ M EcEndoV [16] were added and the reaction was incubated at 37 °C for 45 min. The reaction was stopped by 2 \times Urea loading buffer, and the product was analyzed by 15% urea-PAGE gel electrophoresis followed by staining with ethidium bromide.

2.4. EMSA

The 10- μ l reaction systems contained 20 mM Tris-HCl, pH 9.0, 1 mM DTT, and 2 μ M nucleic acid (tRNA or ssDNA). Full-length or truncated proteins (at different molar ratios to the nucleic acid) were added to the reaction system and mixed gently. The mixture was incubated on ice for 60 min. The reaction mixture was loaded to 6% native PAGE gel with a running buffer containing 40 mM Tris-HCl pH 9.0 and electrophoresed at 100 V for 60 min. The gel was stained with ethidium bromide.

2.5. Thin layer chromatography (TLC)

The 10- μ l reaction systems contained 100 mM KCl, 50 mM HEPES, pH 7.0, 5 mM MgCl₂, and 10 mM substrate (cytidine/uridine), and 30 μ M enzyme was added to start the reactions at 37 °C for 1 hour. 1 μ l of the reaction mix was spotted onto the TLC silica gel plates (Merck) and run for 20 minutes in a cylinder containing 5 mM HCl. Then it was transferred it to another cylinder containing 40 mM HCl for another 20-minute development. The plates were dried at room temperature for 5 min prior to analyses.

3. Results

3.1. Sequence homology with kinases

Through blast search, we found that the SpTad2 protein could be clearly divided into two distinct domains: while the sequence of C-terminal domain (Thr212–Gln360) suggests that it is a TadA-like domain, the N-terminal domain (Met1–Thr202) is either a putative uridine-cytidine or a pantothenate kinase (Fig. 1A–B). The sequence identities were 23.7–28.7% with these enzymes within their aligned regions respectively. This is quite different from its homolog ScTad2, which harbors the N-terminal catalytic domain (CD) and the C-terminal extension. This extension looks like a long string wrapping around the outside of CD (Fig. S1) and its deletion resulted in the expression of ScTad2 in inclusion body (data not shown). However, close inspection of the sequence alignments with the two kinases indicated that several important residues for the kinase functions are missing. Namely, crucial residues His117, Arg166, Arg169, Arg176, and Gln184 in the human uridine-cytidine kinase 2 [17] and Asp45, Tyr55, Thr103, and Lys303 in *Escherichia coli* pantothenate kinase do not exist [18]. Instead, they are replaced by residues that are unlikely to play catalytic roles (Fig. 1A–B).

3.2. Activity tests

To find out whether the N-terminal domain indeed has the putative kinase functions, we expressed and purified the recombinant protein in order to perform *in-vitro* activity assays. We first cloned the full-length *tad2* gene into both the pET-28a and the pET-21b vectors for expression tests. Both constructs were poorly expressed in *E. coli*, suggesting that the SpTad2 protein needs the

stabilization from the accessory subunit SpTad3, as was demonstrated in the ScTad2/3 case [10]. We also removed the first eleven residues from SpTad2, which was predicted to form a disordered region, but the truncated form was not expressed either (data not shown). In contrast, the separation of N-SpTad2 from the rest would generate an over-produced protein (Fig. S2), suggesting that this fragment forms a stable domain. In addition, N-SpTad2 could be well expressed and purified, as were both kinases of *E. coli*. As indicated by the thin layer chromatography (TLC) experiment shown in Figure 2A, while the *Escherichia coli* uridine-cytidine kinase (EcUCK) could transfer the terminal phosphate of ATP to either a cytidine or uridine to give the by-product ADP, N-SpTad2 failed to do so. The other homolog pantothenate kinase (EcPank) phosphorylates pantothenate using ATP as a donor, which plays an important role in the biosynthesis of CoA and the regulation of its intracellular concentration [18]. Similarly, while the *Escherichia coli* pantothenate kinase displayed robust kinase activity toward pantothenate, N-SpTad2 remained inactive (Fig. 2B).

3.3. Structure of N-SpTad2

To further study the functions of N-SpTad2, we crystallized this protein for hints from the structure. The structure of the protein was solved by molecular replacement using the human UCK (hUCK) structure as the search probe (PDB 6PWZ, to be published), which shares 26% sequence identity with N-SpTad2. There are four protein molecules in the P1 asymmetric unit and six sulfate molecules as well (Fig. 3A, Table 1), introduced by the co-crystallization experiments. In addition, there is a disulfide bond between Cys83 and Cys89. Of note, the four copies of the protein made plenty of intermolecular contacts. Particularly, the two α 6 helices form the interface between chains B and C, while two α 1 helices form the

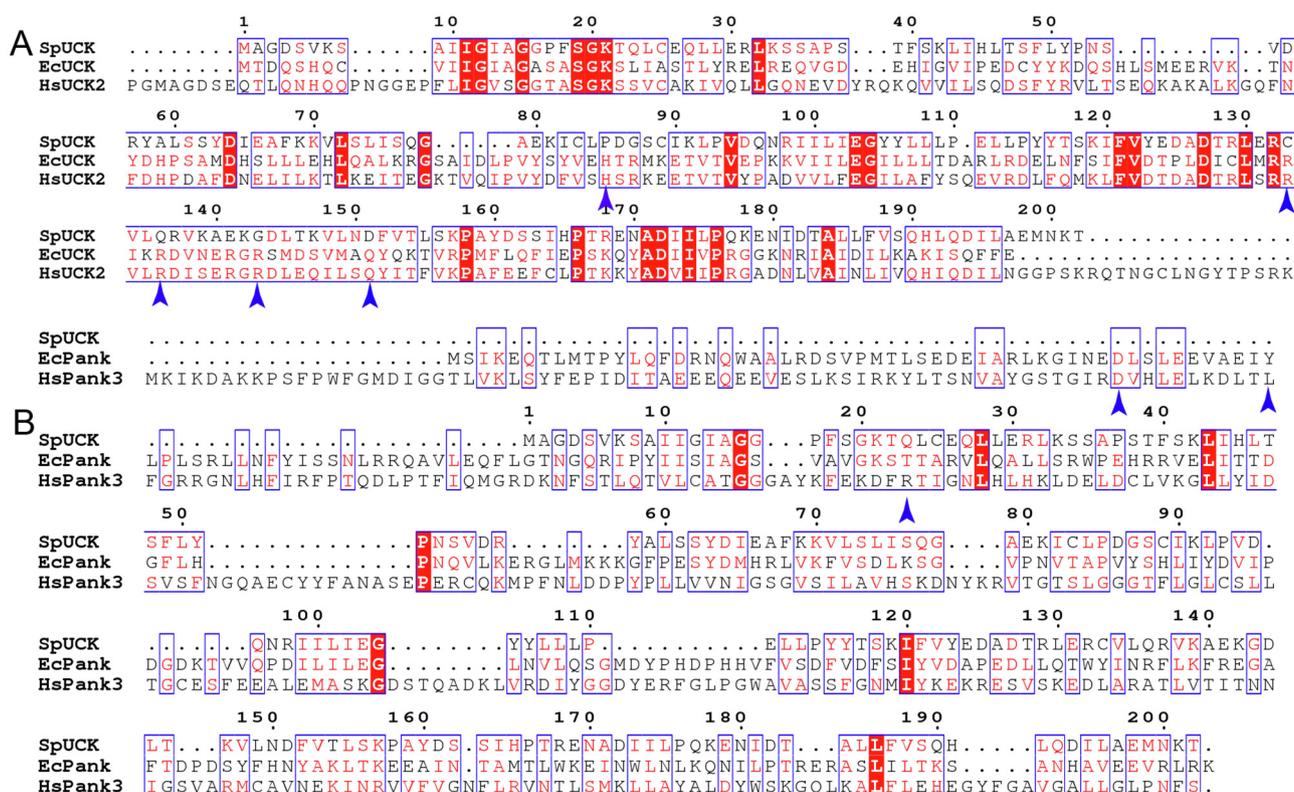


Fig. 1. Sequence alignments of the N-SpTad2 and homologs. (A–B) Sequence alignment with the *Escherichia coli* and human uridine-cytidine kinases (A); and with the *Escherichia coli* and human pantothenate kinases (B). The blue up-arrows indicated key residues responsible for substrate binding or catalysis, and the numbering was according to the sequence of N-SpTad2.

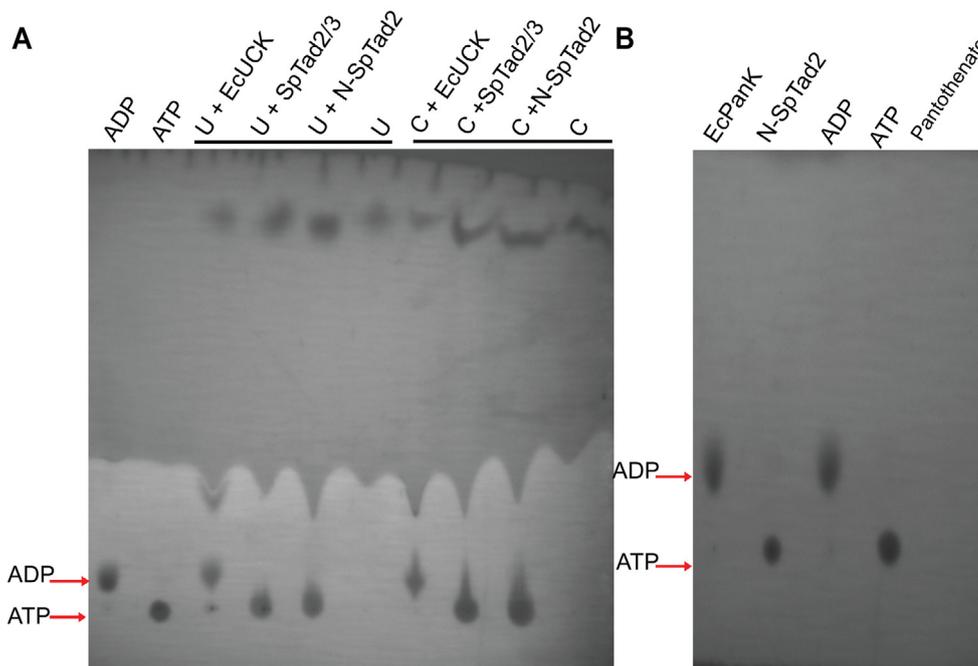


Fig. 2. The kinase activity assays for N-SpTad2 and related enzymes. (A) The uridine-cytidine kinase activity analysis. (B) The pantothenate kinase activity analysis. The by-product ADP and the phosphate donor ATP were indicated by the arrows on the side, and both uridine and cytidine substrates were tested. SpTad2/3, EcUCK or the EcPank were used as positive controls respectively. The kinase reaction products were assayed by TLC and detected by UV exposure.

interface between chains C and D (Fig. 3B). Their calculated buried surface areas are 1616 Å² and 2090 Å² respectively. The PDBePISA server predicted that the protein exists as a possible tetramer in solution [19], but the size exclusion chromatography analysis of the protein indicated that it is a monomer. The retention volume of N-SpTad2 (with a molecular weight of 23.2 kDa) on a Superdex 200 column (10/300, Cytiva) was ~15.8 ml, very close to the retention volume of the reference protein YerA from *Bacillus subtilis* with a known molecular weight of 23.0 kDa (15.9 ml, Fig. 3C).

Among the four copies of the protein, chain D displays the lowest temperature factor, which we will focus our following discussion on. Chain D resolved Ser5-Asn200 with no internal disorders. The protein displays a general globular shape, with a central five-strand β-sheet sandwiched by helices on both sides. Between β6 and β7, there is a long helical region. Surface representation indicated several places of positive patches, one of which forms a central cavity (Fig. 3D). Given the function of Tad as a tRNA deaminase, we wondered whether the domain is involved in the binding of tRNA. However, EMSA immediately excluded the possibility for N-SpTad2 as a tRNA-binder. When present at two-fold excess, N-SpTad2 barely generated any supershift, while the full-length SpTad2/3 showed reasonable binding (Fig. 3E). Consequently, we speculate that N-SpTad2 barely contributes to tRNA deamination, the intrinsic function of the deaminase. We subsequently created the SpTad2-CD/Tad3 truncation mutant (the deletion of N-SpTad2 from the holoenzyme), and conducted the deamination assays. The deamination product was confirmed by the cleavage action of the enzyme EcEndoV on the resultant I34-containing tRNA^{Ala}. As shown by Figure 3F, the SpTad2/3 heterodimer was very efficient in catalyzing the reaction, and finished the deamination in 30 minutes, as indicated by the complete disappearance of the tRNA band. In contrast, removal of the N-terminal domain resulted great loss in activity, and the cleaved tRNA product was hardly observed. Additionally, the complementation of the truncated enzyme with purified N-SpTad2 were unsuccessful in salvaging the activity (Fig. 3F). Therefore, this domain indeed

contributes to tRNA deamination in a manner yet to be investigated.

3.4. Comparison to homologs

The Dali search indicated that the overall structure of the protein was closest to that of hUCK crystallized in various forms [20]. Structural superposition of N-SpTad2 with the hUCK-CTP complex (PDB 1UDW) [17] revealed that the RMSD was 1.7 Å over 205 aligned Cαs. Of interest, the terminal phosphate of CTP coincides in space with the sulfate ion bound near the α1 helix (S1, Fig. 4A). Therefore, the sulfate actually mimics the phosphate group. Pantothenate kinase was cocrystallized with the ATP analog AMPNP (PDB 1ESN) [21]. The enzyme employs a magnesium ion to activate the catalytic water, but it was not present in our structure. The superimposition indicated that its phosphate also occupies the same position as the sulfate molecule in our structure (S1 in Figs. 4B–D), but its base ring points the opposite direction to that of CTP. In addition, both the ribose rings of the ligands pose steric clashes with N-SpTad2, suggesting that a nucleotide phosphate would not be accommodated in a similar manner.

The arrangement of the sulfate ions is very interesting. One sulfate ion was located at the N-terminal end of the α1 helix of each subunit (S1' and S1 in chains C and D respectively), whereas the rest two sulfate ions (S2 and S2') were bound next to each other at the interface formed by these helices running to the opposite directions. In the equivalent positions (the N-terminal end of α1), the two sulfates S1'' and S1''' were also bound. Therefore these sulfate molecules bound strongly with the N-SpTad2 as if they were a part of the proteins. By contrast, the binding of S2 and S2' at the C/D interface was much weaker and they did not have counterparts in the A/B subunits. Taken together, the six sulfates represent two binding patterns: the tightly bound sulfates S1 (S1', S1'', S1''') and loosely bound sulfates S2 (S2') (Fig. 3A). Particularly, S1 not only makes hydrogen bonds with the backbone nitrogen atoms of Ser19-Thr22, it also forms a hydrogen bond and a salt bridge with

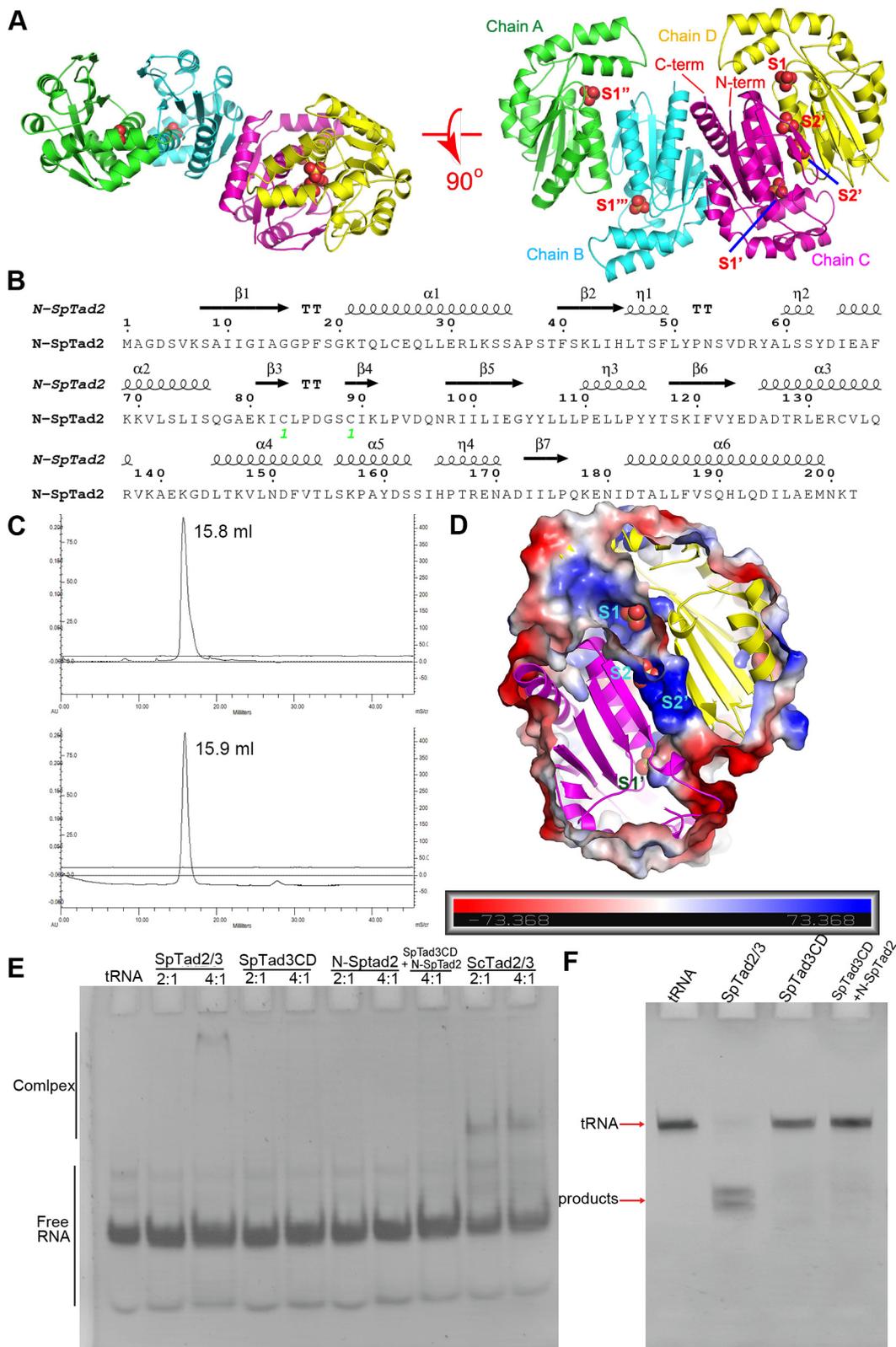


Fig. 3. The structural and biochemical characterization for N-SpTad2. (A) The overall structure of N-SpTad2. The structure was shown in the ribbon rendition in two orthogonal views. The N- and C-termini of chain C were indicated. The four monomers were colored differently and the red spheres represented the sulfate ions. The tightly and loosely bound sulfates were named S1 and S2 respectively, and the sulfates belonged to different monomers (S1', S1'', S1''' or S2, S2'). (B) Secondary structure of N-SpTad2. (C) The gel-filtration analysis of N-SpTad2. top: The elution profile of N-SpTad2; bottom: The elution profile of a reference protein of 23.0 kDa. (D) The surface representation of chains C and D of N-SpTad2. The positively charged patch was colored blue while the negatively charged patch was colored red. (E) The EMSA analysis of possible complex formation with tRNA^{Ala}. 2 μ M tRNA was used to bind different amounts of protein (molar ratios of 2:1 or 4:1). The complex was indicated at the left side by the arrow. (F) The coupled deaminase activity assay of N-SpTad2. The red arrow indicates the cleaved products (35 and 41 nts respectively) from tRNA^{Ala}. 2 μ M tRNA was loaded in each lane and 60 nM enzyme was added.

Table 1
Data collection and refinement statistics.

PDB IDs	7EEY
Data collection	SSRF BL19U1
Wavelength	0.979
Space group	P1
Cell dimensions	
a, b, c (Å)	62.53, 65.60, 65.64
α, β, γ (°)	63.97, 69.96, 73.56
Resolution (Å)	50.0–2.59 (2.68–2.59) ^a
R _{mea} (%)	7.3 (15.8)
R _{merge} ^b (%)	6.8 (14.6)
CC _{1/2}	0.994 (0.986)
I/ σ (I)	32.8 (15.7)
Completeness (%)	98.7 (97.4)
Redundancy	7.1 (7.1)
Refinement	
Resolution (Å)	37.70–2.60 (2.70–2.60)
No. reflections	26299
R _{work} ^c /R _{free} ^d	0.168/ 0.218
No. atoms	
Protein	6032
Ligand (sulfate)	30
Water	167
B-factors (Å ²)	
Protein	38.5
Ligand (sulfate)	48.6
Water	37.4
R.m.s deviations	
Bond lengths (Å)	0.004
Bond angles (°)	0.78
Ramachandran favored (%)	99.6
Allowed (%)	0.4
Outliers (%)	0

^a : Values in parentheses are for the highest-resolution shell.

^b : $R_{\text{merge}} = \sum (|I - \langle I \rangle|) / \sum I$, where I is the observed intensity.

^c : $R_{\text{work}} = \sum_{\text{hkl}} (|F_o| - |F_c|) / \sum_{\text{hkl}} |F_o|$, calculated from working data set.

^d : R_{free} is calculated from 5.0% of data randomly chosen and not included in refinement.

the side chains of Thr22 and Lys21 respectively. On the other hand, the loosely bound S2 makes two salt bridges and a hydrogen bond with the surrounding residues: Lys33', Lys42' (chain C) and His45 (Fig. 4E). The four sulfates in close distances along the dimer interface were interesting as well: they were separated from each other by 9.5, 5.5, and 9.6 Å respectively (closest distances, Fig. 4B). We wondered if the sulfate groups mimicked the phosphates along a DNA chain. Therefore, we tried the EMSA assay with single- or double-stranded DNAs of different sequences. The results indicated that N-SpTad2 could not bind ss- or dsDNAs (Fig. S3).

3.5. Possible role in catalysis

Next, we tested whether N-SpTad2 serves the purpose of promoting the solubility/stability of the individual domains of the SpTad2/3 heterodimeric enzyme. There are four annotated domains in this enzyme: SpTad3-PCD (pseudo-catalytic domain, Thr153-Val315), SpTad2-CD (Ser203-Lys389), N-SpTad2 (Met1-Thr202), and N-SpTad3 (Met1-Phe152). The first two make the catalytic dimer core, which resembled the TadA dimer. In comparison, ScTad2/3 lacks the N-terminal domain at the ScTad2 subunit, but displays a C-terminal extension instead (Fig. S4). We subsequently tried expression of the individual or different combinations of the four domains. Specifically, expression trials of single domains indicated that both N-terminal domains were soluble but neither C-terminal domain was. The two-domain co-expression tests using the pET-Duet-1 vector generated either the full-length SpTad3, or both the N-SpTad2 and N-SpTad3 in soluble forms out of the four combination possibilities, but in the latter scenario, the two domains did not form a complex (Fig. S5). The three-domain co-

expression tests only produced the SpTad3/Tad2-CD complex (i.e. the SpTad2-CD/Tad3 truncation mutant described in section 3.3 above), but not the SpTad2/Tad3-PCD complex. Lastly, the co-expression of both proteins in fulllengths would generate a stable holoenzyme (Table 2). To our surprise, the PCD/CD combination (SpTad2-CD and SpTad3-PCD) did not yield a complex or a “pseudoenzyme” like the TadA dimer. It suggested that both PCD and CD have poor solubilities. They are incapable of burying the hydrophobic surface areas of each other and thus could not be co-expressed as a soluble complex. Consequently, the N-SpTad3 domain indeed plays a structural role, without which the heterodimer would not be a stable assembly. On the other hand, N-SpTad2, and N-SpTad3 are highly soluble, with the latter being more hydrophilic, considering the fact that the full-length SpTad3 could be expressed but SpTad2 could not. Additionally, the co-expression of N-SpTad3 prevents SpTad2-CD from forming inclusion bodies, but not vice versa.

Besides the role as the solubility enhancer, N-SpTad2 has addition unusual features. We found quite a few positively charged residues in the N-SpTad2 domain, namely Arg57, Arg137, and Lys209, etc. These residues are located in the vicinities of the bound sulfates and may play a functional role. Furthermore, there is an intramolecular disulfide bond between Cys83 and Cys89, which fixes a pair of antiparallel strands in N-SpTad2. To check the influence of these residues, we made the R57A, R137A, K21A/T22A and C83S mutations. However, the co-expression of these individual SpTad2 mutants with SpTad3 all led to the sole expression of SpTad3, except for the Tad2-C83S/SpTad3 co-expression (Fig. 5A-E). Nevertheless, the Ni-NTA affinity purification of C83S (Tad2)/Tad3 produced unbalanced proportions of each subunit as indicated by the slightly weaker band of C83S(Tad2) (upper band in Fig. 5D), and this band would gradually disappear in further purification steps (Fig. 5E). Taken together, the full-length SpTad2 and the mutants were mostly expressed as inclusion bodies, underscoring the stabilizer role played by N-SpTad2. However, the stabilizing effects were not strong enough to prevent SpTad3-PCD from forming the inclusion body as the full-length SpTad2 is insoluble, suggesting that the contacts between the two domains were not abundant enough. Moreover, in the co-expression test, both the N-SpTad3 and N-SpTad2 domains were expressed, but only N-SpTad3 was isolated because the 6 × His tag was fused at its N-terminus. This observation also suggested that there were very limited interactions between the two N-terminal domains.

4. Discussion

The deamination at A34 in tRNA is essential to all organisms, which is a great example that supports the Wobble theory. This phenomenon is conserved in both bacteria and eukarya, but the mechanism is distinct in the two kingdoms. Although bacterial Tadas are homodimeric enzymes, their eukaryotic counterparts are more complex and consist of the Tad2/3 heterodimers. Except for limited species including human, *Saccharomyces cerevisiae*, *Trypanosoma brucei* etc., the tRNA-A34 deaminases and corresponding catalytic mechanisms in eukaryotes are poorly understood.

In this study, we investigated the structure and function of the unusual N-terminal domain of SpTad2. The domain was initially predicted to be a putative kinase by bioinformatic studies but the key residues are missing, which raises doubts about its genuine function. Only three other *Schizosaccharomyces* organisms possess this domain, and it is always located at the N-termini of related proteins (data not shown). The domain could be well expressed, but activity assays indicated that it was incapable of transfer the phosphate group to either uridine/cytidine or pantothenate, which motivated us to carry on the study. We first attempted to obtain

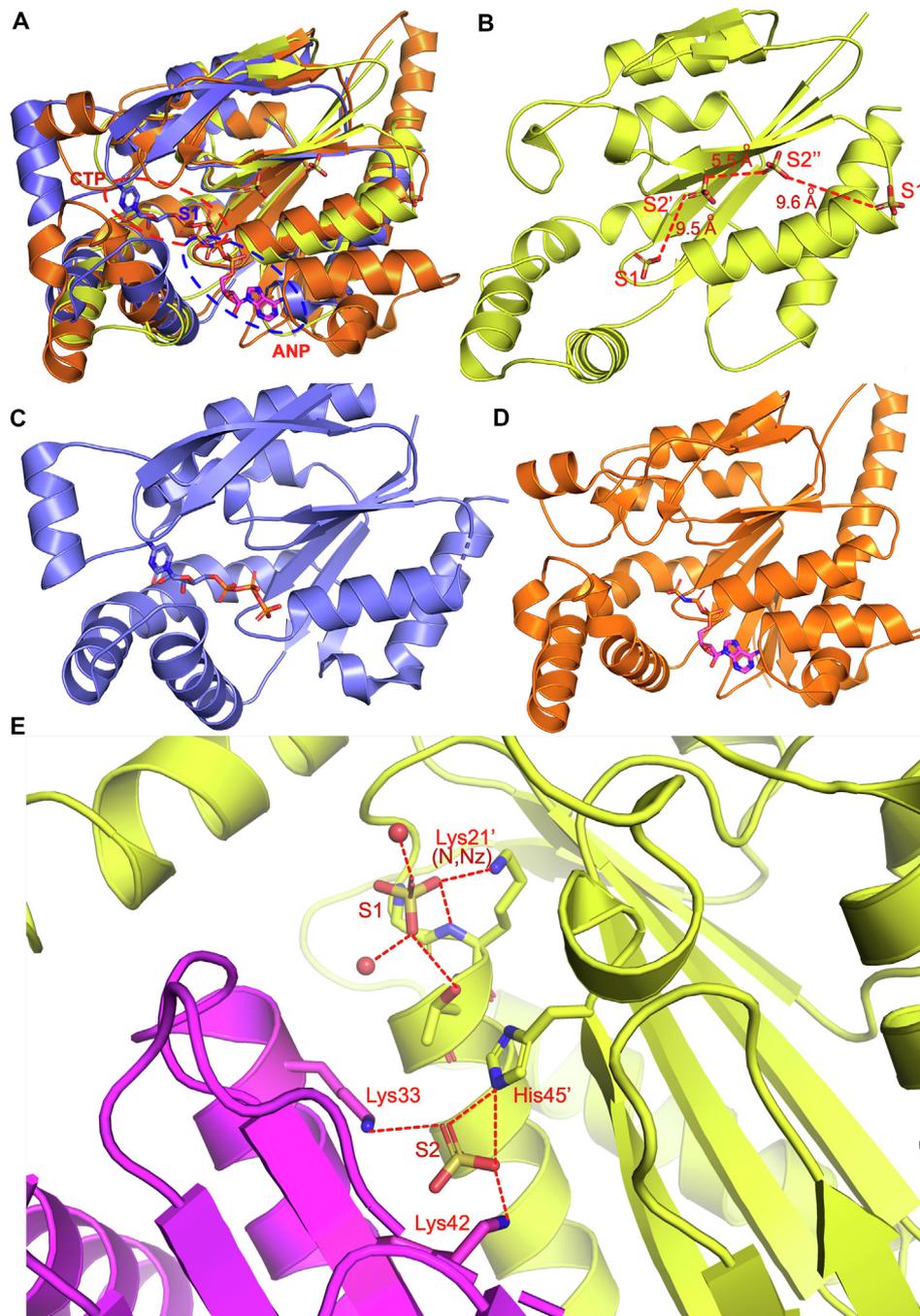


Fig. 4. Structure comparison to homologs. (A) The structural superposition of the N-SpTad2-sulfate complex (PDB 7EEY, color yellow) with that of CTP-bound hUCK (PDB 1UDW, violet), and the AMPPNP-bound EcPanK (PDB 1ESN, orange), respectively. The sulfate, CTP, and AMPPNP ligands were shown in sticks and the latter two were indicated by the broken ovals. (B–D) The separate view of the N-SpTad2-sulfate complex bound by chain D (B); and the CTP-hUCK complex (C); and the AMPPNP-PanK complex (D). Note that the base rings were pointing to the opposite directions. The distances between the sulfates were shown. (E) The close-up view of the interaction patterns of the S1 and S2 sulfates, bound between the C–D chains. Chains C and chain D were colored magenta and yellow, respectively. The side chains of the key residues were shown as sticks and labeled. The water molecule for sulfate interactions were represented by the red spheres.

hints through structural studies and managed to solve the crystal structure of the N-SpTad2 domain. Interestingly, the crystals were grown in the presence of sulfate, and six molecules of sulfate were found located at various sites of the four monomers of N-SpTad2 in the asymmetric unit, with four of them mimicking the phosphate positions of CTP or pantothenate. This finding promoted us to test whether N-SpTad2 could function as a nucleic acid binder. Although our EMSA assays indicated that this domain was unlikely to bind tRNA, the loss of N-SpTad2 from the heterodimer eliminated its tRNA deamination capability, suggesting that this domain

still mediated important protein-protein interactions in the deamination process. Follow-up co-expression and mutational studies indicated that the truncation/mutations of N-SpTad2 would remarkably interfere with the expression of the catalytic subunit SpTad2. Additionally, many mutations within N-SpTad2 also led to similar results, evidenced by the sole expression of the latter, thus suggesting its cross-talks with SpTad3.

SpTad2/3 is a modular protein. It is closely related to ScTad2/3 but still exhibited quite a few differences. Despite the conserved SpTad3-PCD and SpTad2-CD domains, the extra N-SpTad2 domain

Table 2
The (co)-expression profiles of various combinations of the individual domains in the heterodimeric SpTad2/3 deaminase.

N-Tad2	Tad2-CD	N-Tad3	Tad3-PCD	Expression profiles
+				√
	+			×
		+		√
			+	×
+	+			×
+		+		√ ^a
		+	+	√
	+		+	×
+	+		+	×
	+	+	+	√
+	+	+	+	√

a: Both domains were expressed but only SpTad3-CD was isolated.

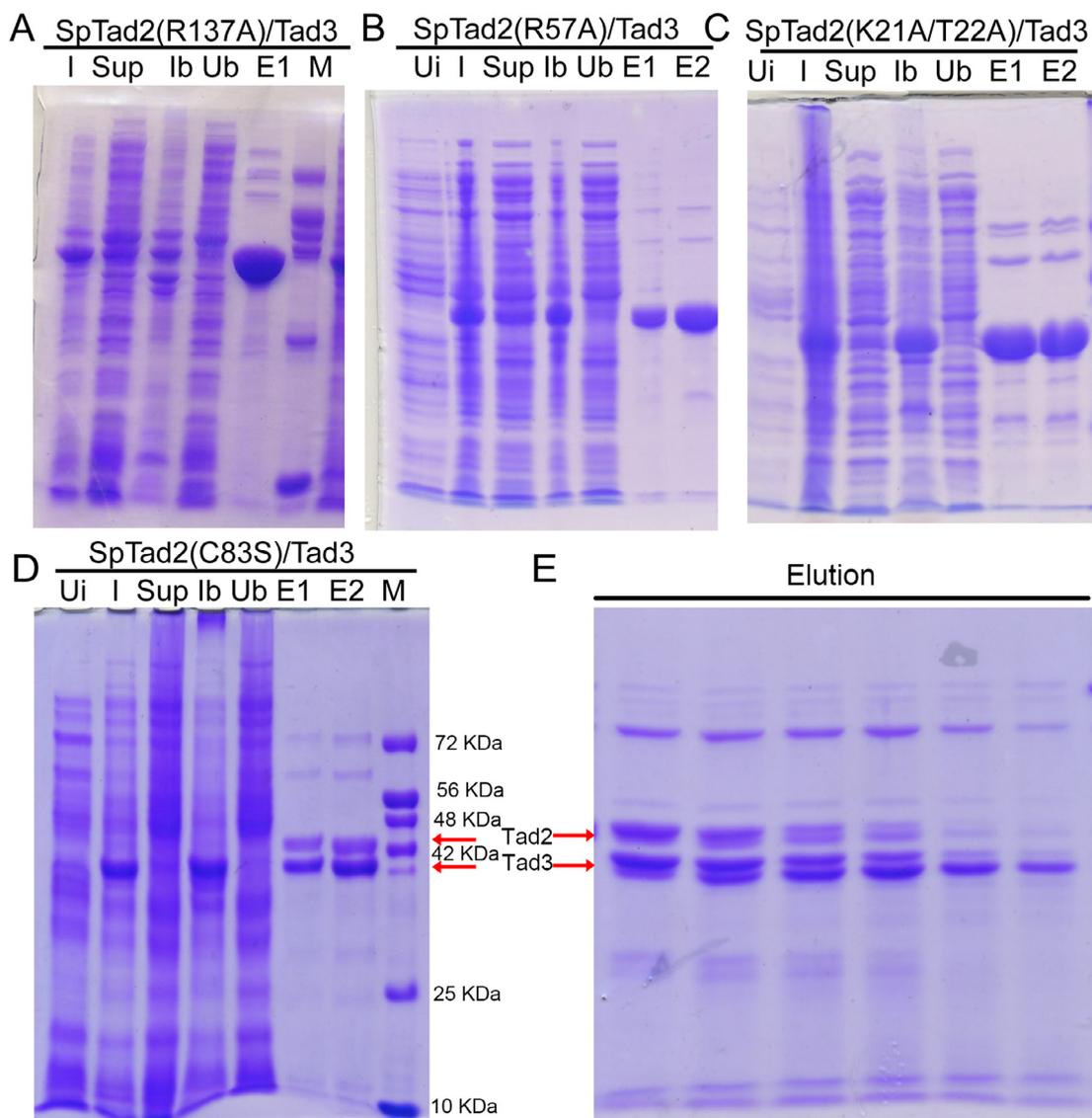


Fig. 5. The expression and purification profiles of the SpTad2/3 mutants. (A–D) The SDS-gel analyses of the co-expression and Ni-NTA affinity purification results of the SpTad2(R137A)/Tad3 (A); and SpTad2(R57A)/Tad3 (B); and SpTad2(K21A/T22A)/Tad3 (C); and SpTad2(C83S)/Tad3 (D). (E) Re-purification of the SpTad2(C83S)/Tad3 complex by the IMAC affinity purification chromatography after overnight at 4 °C. Different lanes represent the elution fractions of the affinity chromatography. Ui: uninduced; I: induced; Sup: supernatant; Ib: Inclusion body; Ub: unbound; E: elution; M: marker. The respective positions of the SpTad2 and SpTad3 subunits were indicated by the red arrows on the side.

helps the enzyme’s stability and solubility. In ScTad2/3, there is no such an extra domain at the N-terminus of Tad2, and the N-lobe of ScTad3 is responsible for tRNA binding. Moreover, ScTad3 utilizes

the C-terminal cysteine to coordinate to the zinc ion in ScTad3-PCD, which is lacking in SpTad3. We conducted structural studies on SpTad2/3 as a whole or in part, but the expressed proteins or

fragments failed to crystallize due to their instabilities or poor solubilities. Taken together, the SpTad2/3 deaminase is built around the core Tad2-CD/Tad3-PCD heterodimer, with both N-terminal domains promoting the stability and solubility through protein-protein interactions.

Other than the binding of multiple sulfates, the oligomerization state of N-SpTad2 is intriguing. Pantothenate kinase forms a tetramer for function. The predicated homotetrameric state of N-SpTad2 does not appear to be biologically relevant because the size exclusion chromatography analysis indicated that it is a monomer. Moreover, the large size of the tetramer would disrupt the assembly of the holoenzyme and block the entry of tRNA. In addition, the dimerization between subunits C and D was apparently induced by the S1 (S1') and S2 (S2') sulfates. The line-up of the sulfates along the dimer interface points to the possibility of single- or double-stranded DNA assisted dimerization. Some adenosine deaminases like APOBEC family members have been reported to bind ssDNA. However, our follow-up EMSA experiment ruled out this possibility.

Lastly, it is not known the original source of the N-terminal domain of SpTad2. The Dali search suggested that the overall structure of the protein was similar to that of UCKs. Blast search suggested the genome *Schizosaccharomyces pombe* owns two putative uridine kinases and one pantothenate kinase. The functions of these proteins were inferred from sequence homologies and await further functional validation. While N-SpTad2 does not show strong sequence homology with the *Schizosaccharomyces pombe* pantothenate kinase, it does show 26–28% sequence identities with the catalytic domains of the putative uridine kinases (Fig. S6A–B). The possible scenario is that the core regions of SpTad2/3 consisting of Tad2-CD/Tad3-PCD were quite unstable, and therefore it acquired additional domains through horizontal gene transfer to fuse at its N-termini in both subunits. N-SpTad2 could share ancestor genes with both genes, but underwent divergent evolution later. By accumulating some mutations over the course of its evolution, N-SpTad2 gradually lost its original function. Lastly, Rubio *et al.* reported that under certain circumstances, the *Trypanosoma brucei* Tad2/3 enzyme could act as a C-to-U deaminase [22] or show deamination activity toward ssDNA [3]. In our case, N-SpTad2 was not found to show binding abilities to DNA or tRNA. In our previous studies, we consider species with Tad3s that end with two cysteine residues were more ancestral species [10]. Therefore, the *Schizosaccharomyces pombe tad3* gene may occur later than its *Saccharomyces cerevisiae* counterpart. Correspondingly, N-SpTad3 replaces the C-terminal extension present in ScTad2 and becomes more complicated in sequence and structure. Whether the presence of this unusual domain confers novel functions to the enzyme remains to be investigated.

5. Conclusions

Here we investigated the structure and function of the unusual N-terminal domain of SpTad2, which was predicted to be a putative kinase by bioinformatic studies. Due to the missing of some key residues, N-SpTad2 did not display the kinase activities. While EMSA assays indicated that this domain was unlikely to bind tRNA or ssDNA, the loss of N-SpTad2 from the heterodimer eliminated the tRNA deamination capability of the holoenzyme, suggesting that N-SpTad2 may still mediate important protein-protein interactions in the deamination process.

Funding

This work was supported by the National Natural Science Foundation of China [31870782], Natural Science Foundation of the Guangdong Province [2020A1515010965], the Fundamental

Research Funds for the Central Universities [19ykpy08], and China Postdoctoral Science Foundation [2020M672950].

Availability of Data and Materials

The atomic coordinates and structure factors have been deposited in the Protein Data Bank with the accession code 7EEY. All relevant data are available from the authors.

Author contributions

WX conceived and designed research; XL, JZ performed research; WX, GR and XL analyzed data and wrote the paper. All authors reviewed the results and approved the final version of the manuscript.

Ethics approval and consent to participate

Not applicable.

Author Statement

Wei Xie conceived and designed research; Xiwen Liu, Jie Zhou performed research; Wei Xie, Ruiguang Ge and Xiwen Liu analyzed data and wrote the paper. All authors reviewed the results and approved the final version of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We thank the staff members from BL19U1 beamlines (National Center for Protein Sciences Shanghai (NCPSS)) at Shanghai Synchrotron Radiation Facility, for assistance during data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2021.06.008>.

References

- [1] Tsutsumi S, Sugiura R, Ma Y, Tokuoaka H, Ohta K, Ohte R, et al. Wobble inosine tRNA modification is essential to cell cycle progression in G(1)/S and G(2)/M transitions in fission yeast. *J Mol Biol.* 2007;282:33459–65.
- [2] Wolf J, Gerber AP, Keller W. tadA, an essential tRNA-specific adenosine deaminase from *Escherichia coli*. *EMBO J.* 2002;21:3841–51.
- [3] Rubio MAT, Pastar I, Gaston KW, Ragone FL, Janzen CJ, Cross GAM, et al. An adenosine-to-inosine tRNA-editing enzyme that can perform C-to-U deamination of DNA. *Proc. Natl Acad. Sci. USA* 2007;104(19):7821–6.
- [4] Gerber AP, Keller W. An adenosine deaminase that generates inosine at the wobble position of tRNAs. *Science* 1999;286:1146–9.
- [5] Zhou W, Karcher D, Bock R. Identification of enzymes for adenosine-to-inosine editing and discovery of cytidine-to-uridine editing in nucleus-encoded transfer RNAs of *Arabidopsis*. *Plant Physiol.* 2014;166(4):1985–97.
- [6] Torres, A.G., Pineyro, D., Rodriguez-Escriba, M., Camacho, N., Reina, O., Saint-Leger, A., Filonava, L., Batlle, E. and Ribas de Pouplana, L. (2015) Inosine modifications in human tRNAs are incorporated at the precursor tRNA level. *Nucleic Acids Res.* 43, 5145–5157.
- [7] Losey HC, Ruthenburg AJ, Verdine GL. Crystal structure of *Staphylococcus aureus* tRNA adenosine deaminase TadA in complex with RNA. *Nat Struct Mol Biol* 2006;13(2):153–9.
- [8] Rubio MA, Ragone FL, Gaston KW, Ibba M, Alfonso JD. C to U editing stimulates A to I editing in the anticodon loop of a cytoplasmic threonyl tRNA in *Trypanosoma brucei*. *J Biol Chem* 2006;281:115–20.
- [9] Alazami AM, Hijazi H, Al-Dosari MS, Shaheen R, Hashem A, Aldhmesh MA, et al. Mutation in ADAT3, encoding adenosine deaminase acting on transfer

- RNA, causes intellectual disability and strabismus. *J Med Genet* 2013;50:425–30.
- [10] Liu X, Chen R, Sun Y, Chen R, Zhou J, Tian Q, et al. Crystal structure of the yeast heterodimeric ADAT2/3 deaminase. *BMC Biol* 2020;18(1). <https://doi.org/10.1186/s12915-020-00920-2>.
- [11] Minor W, Cymborowski M, Otwinowski Z, Chruszcz M. HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr. Sect D: Biol Crystallogr* 2006;62(8):859–66.
- [12] Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect D: Biol Crystallogr* 2010;66(2):213–21.
- [13] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr. Sect D: Biol Crystallogr* 2010;66(4):486–501.
- [14] Afonine PV, Mustyakimov M, Grosse-Kunstleve RW, Moriarty NW, Langan P, Adams PD. Joint X-ray and neutron refinement with phenix.refine. *Acta Crystallogr. Sect D: Biol Crystallogr* 2010;66(11):1153–63.
- [15] Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. Sect D: Biol Crystallogr* 2010;66(1):12–21.
- [16] Zhang Z, Jia Q, Zhou C, Xie W. Crystal structure of *E. coli* endonuclease V, an essential enzyme for deamination repair. *Sci Rep* 2015;5:12754.
- [17] Suzuki NN, Koizumi K, Fukushima M, Matsuda A, Inagaki F. Structural basis for the specificity, catalysis, and regulation of human uridine-cytidine kinase. *Structure* 2004;12(5):751–64.
- [18] Ogata Y, Chohnan S. Prokaryotic type III pantothenate kinase enhances coenzyme A biosynthesis in *Escherichia coli*. *J Gen Appl Microbiol*. 2015;61:266–9.
- [19] Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 2007;372(3):774–97.
- [20] Holm L, Kääriäinen S, Wilton C, Plewczynski D. Using Dali for structural comparison of proteins. *Curr Protoc Bioinformatics*. **Chapter** 2006;5:Unit 5.5.
- [21] Yun M, Park CG, Kim JY, Rock CO, Jackowski S, Park HW. Structural basis for the feedback regulation of *Escherichia coli* pantothenate kinase by coenzyme A. *J Biol Chem* 2000;275:28093–9.
- [22] Rubio MAT, Gaston KW, McKenney KM, Fleming IMC, Paris Z, Limbach PA, et al. Editing and methylation at a single site by functionally interdependent activities. *Nature* 2017;542(7642):494–7.