A chiral selectivity relaxed paralog of DTD for proofreading tRNA mischarging in Animalia

Santosh Kumar Kuncha,[1,2]* Mohd Mazeed,[1]* Raghvendra Singh,[1] Bhavita Kattula,[1] Satya Brata Routh,[1] Rajan Sankaranarayanan[1]†

[1]CSIR–Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad – 500007, India.

[2]Academy of Scientific and Innovative Research (AcSIR), CSIR–CCMB campus, Uppal Road, Hyderabad – 500007, India.

*These authors contributed equally to this work

†Corresponding author. Email: sankar@ccmb.res.in (R.S.)

One Line Summary:

A switch from Gly-*cis*Pro in DTD to Gly-*trans*Pro in a unique editing factor that clears tRNA selection error by a synthetase

Abstract

D-aminoacyl-tRNA deacylase (DTD), a *trans*-editing factor found in all bacteria and eukaryotes, removes D-amino acids mischarged on tRNAs. A cross-subunit Gly-*cis*Pro motif forms the mechanistic basis of L-amino acid rejection from the catalytic site. Here, we present the identification of a DTD variant, ATD (Animalia-specific tRNA deacylase), that harbors a Gly-*trans*Pro motif. The *cis*-to-*trans* switch causes a "gain of function" through L-chiral selectivity in ATD resulting in the clearing of L-alanine mischarged on $tRNA^{Thr}$(G4•U69) by eukaryotic AlaRS. Animalia genomes enriched in $tRNA^{Thr}$(G4•U69) are in strict association with the presence of ATD, underlining the necessity for a dedicated factor to proofread tRNA misaminoacylation. The study highlights the role of ATD during tRNA gene expansion as a key event associated with the evolution of Animalia.

Translational quality control is a complex and tightly regulated process involving not only editing of errors, as observed in most scenarios, but also a targeted and selective compromise in fidelity under specific conditions such as oxidative stress (***1–8***). It ensures an optimum dynamic balance in the cellular proteome and hence overall cellular homeostasis. A key process in this orchestra includes decoupling of D-amino acids mischarged on tRNAs. This function, termed "chiral proofreading", is performed by a dedicated *trans*-editing factor called D-aminoacyl-tRNA deacylase (DTD) (***9–11***). DTD—present throughout Bacteria and Eukarya—employs an invariant cross-subunit active-site Gly-*cis*Pro motif for strict L-chiral rejection to ensure substrate stereospecificity (***11, 12***). DTD-like fold and DTD function are present in all three domains of life (***13–20***).

Using motif-based sequence analysis, we identified an uncharacterized protein, which we named ATD (as explained later), sharing <30% identity with DTD. This value is significantly lower than those between DTDs (>50%) or between ATDs (>45%) (**fig. S1**). Interestingly, ATD has –PQATL– and –TNGPYTH– as signature motifs, which are similar to though distinct from the corresponding motifs in DTD, *viz.*, –SQFTL– and –NXGPVT–, respectively (**Fig. 1A**). Strikingly, some key conserved residues near DTD's active-site, involved in a network of interactions and responsible for holding the Gly-*cis*Pro motif, are also different in ATD. The most notable among these is the highly conserved Arg7 in DTD from *Plasmodium falciparum* (PfDTD) which is replaced by the conserved Gln16 in ATD from *Mus musculus* (MmATD) (**Fig. 1A**). A thorough *in-silico* search for ATD sequences revealed that ATD is present in Eukarya, but absent in Bacteria and Archaea. Within Eukarya, ATD is restricted to kingdom Animalia, except for a few protozoa (four species of *Leishmania*, two of *Trypanosoma*, and one each of *Saprolegnia*, *Salpingoeca* and *Acanthamoeba*, whose genomes have been sequenced) (**data S1**).

More importantly, ATD is found all across phylum Chordata, whereas its distribution in non-chordate phyla is rather sparse (**Fig. 1B, fig. S2 and data S1**). Notably, most ATD-harboring protozoa are parasites of various vertebrates. Phylogenetic analysis of ATD and DTD showed that the two fall into two distinct groups (**Fig. 1B**).

To gain insights into ATD's function, we solved crystal structure of MmATD at 1.8 Å resolution by molecular replacement using PfDTD as the search model (**Fig. 1C, fig. S3 and table S1**). Structural superposition of MmATD on PfDTD and NTD from *Pyrococcus abyssi* (PabNTD) showed an rmsd of 1.68 Å over 141 Cα atoms and 3.34 Å over 77 Cα atoms, respectively (**Fig. 1C and fig. S4**). Interestingly, a Dali-based PDB search for ATD homologs identified a protein (ATD) from *Leishmania major* annotated as probable eukaryotic DTD, and deposited by Structural Genomics of Pathogenic Protozoa Consortium (***21***). MmATD and LmATD superimpose with an rmsd of 1.29 Å over 148 Cα atoms (**fig. S5**). The structure of ATD revealed that its characteristic motifs (–PQATL– and –TNGPYTH–) form the active-site at the dimeric interface just like the corresponding motifs of DTD (–SQFTL– and –NXGPVT–) (**Fig. 1D**). Besides the elements of DTD-like fold, specific interactions at the adenine-binding site to recognize 3′-terminal adenosine-76 of tRNA are also highly conserved in ATD (**fig. S6**).

Surprisingly, MmATD's Gly-Pro motif occurs in *trans* conformation, unlike DTD's Gly-Pro motif which always exists in *cis* conformation as observed in 107 protomers of 19 crystal structures from 5 different organisms (**Fig. 2, A and B, and fig. S7**). Notably, LmATD also possesses a cross-subunit Gly-*trans*Pro motif like MmATD (**fig. S8**). B-factor analysis reveals that ATD's Gly- *trans*Pro motif is rigidly held like DTD's Gly-*cis*Pro motif (**fig. S3**). Additionally, ATD's Gly-Pro residues exhibit a dramatic ~180° change in ψ torsion angle when compared to DTD's Gly-Pro due to remodelling of the local network of interactions (**Fig. 2, C**

and D). DTD's Gly-*cis*Pro carbonyl oxygens are parallel and protrude into the active-site pocket, *i.e.*, "outward parallel" orientation. ATD's Gly-*trans*Pro carbonyl oxygens are also parallel, but they face away from the active site toward the protein core, *i.e.*, "inward parallel" orientation (**Fig. 2, A and B**).

Interestingly, Arg7 of DTD seems to have migrated to a new position in ATD, *viz.*, Arg151 (**Fig. 1A and Fig. 2D**). The side chain of highly conserved Arg7 in DTD interacts with the main chain of Gly-*cis*Pro motif from the same monomer, thereby locking the motif rigidly in *cis* conformation (**Fig. 2D and fig. S10**). In contrast, the interaction of ATD's invariant Arg151 side chain with the main chain of Gly-*trans*Pro motif from the dimeric counterpart pulls the motif's backbone outwards, thus placing the motif rigidly in *trans* conformation (**Fig. 2D and fig. S10**). In DTDs, the "outward parallel" Gly-*cis*Pro carbonyl oxygens act as a "chiral selectivity filter" to strictly reject all L-amino acids from the pocket through steric exclusion (***11, 12***). Comparative analysis of active-sites of DTD and ATD further revealed that the inward flip of ATD's carbonyl oxygens due to *trans* conformation of its Gly-Pro motif has created "additional" space in its active-site when compared to DTD. Consequently, ATD can easily accommodate a larger group in that space as opposed to just hydrogen in DTD (**fig. S9 and table S2**). This clearly suggests that ATD can cradle smaller L-amino acids. The "additional" space created as a consequence of Gly-*trans*Pro switch in ATD prompted us to check its biochemical activity profile on L-aminoacyl-tRNAs, in addition to D-aminoacyl- and glycyl-tRNAs.

MmATD shows significant activity at 50 nM concentration on D-Tyr-tRNA^{Tyr}, but fails to act on the L-counterpart even at 100-fold higher concentration (**Fig. 2, E and F**). It also deacylates Gly-tRNA^{Gly} at 500 nM concentration (**Fig. 2G**). Thus, like DTD, ATD acts on both D-chiral and achiral substrates, albeit with significantly less efficiency. Strikingly, when tested with L-Ala-

tRNA^{Ala}, 500 nM MmATD displayed noticeable activity (**Fig. 2H**). By contrast, *E. coli* DTD (EcDTD) or PfDTD does not act on L-chiral substrates even at 100-fold higher concentration than required for D-chiral substrate (***11, 12***). Therefore, biochemical probing suggested that ATD is an aminoacyl-tRNA deacylase with a relaxed specificity for substrate chirality, primarily due to the *trans* conformation of its active-site Gly-Pro motif. It is for this reason that we named this protein ATD, which stands for Animalia-specific tRNA deacylase. ATD thus has a "gain of function" in L-chiral activity when compared to DTD due to the switch from Gly-*cis*Pro to Gly-*trans*Pro. Furthermore, biochemical data in conjunction with structural data indicate that ATD's active-site pocket, because of the space created by the inward movement of the Gly-Pro carbonyl oxygens, can accommodate only smaller L-amino acids like L-alanine but not the bigger ones like L-tyrosine.

While we were in the process of identifying the physiological substrate for ATD, a recent study reported that eukaryotic AlaRS has acquired the function of L-alanine mischarging on tRNAs harboring G4•U69 wobble base pair in the acceptor stem, such as tRNA^{Thr}(G4•U69) (***22***). This is in addition to its canonical recognition of AlaRS-specific universally occurring G3•U70 in tRNA^{Ala} (***23, 24***). The above finding prompted us to test ATD's role in proofreading tRNA^{Thr}(G4•U69) selection error made by eukaryotic AlaRS, which is non-discriminating (AlaRSND) when compared to the bacterial discriminating AlaRS (AlaRS^{D}). Strikingly, biochemical assays with MmATD showed significantly higher selectivity for L-Ala-tRNA^{Thr}(G4•U69) as the enzyme acts on the substrate at just 1 nM compared to its similar activity on L-Thr-tRNA^{Thr}(G4•U69) at 50 nM (**Fig. 3, B and C**). Thus, MmATD displays a 50-fold difference in biochemical activity on these two substrates.

To rule out any species-specific phenomenon regarding ATD's biochemical activity, we tested ATDs from zebrafish (*Danio rerio*, DrATD) and human (*Homo sapiens*, HsATD). It was observed that these ATDs can act on L-Ala-tRNA^{Thr}(G4•U69) more efficiently than on L-Thr-tRNA^{Thr}(G4•U69) (**Fig. 3, D and E**). Therefore, not only ATD's activity on the non-cognate substrate but also its ability to discriminate between L-Ala-tRNA^{Thr}(G4•U69) and L-Thr-tRNA^{Thr}(G4•U69) is conserved across diverse eukaryal taxa. We then analyzed tRNA^{Thr}(G4•U69) gene sequences from diverse species which revealed that the first five base pairs in the acceptor stem are invariant (**fig. S11**). As ATD belongs to the DTD-like fold, its interaction with the tRNA may not go beyond the first three or four base pairs in the acceptor stem. Hence, lack of variation in the acceptor stem of tRNA^{Thr}(G4•U69) further confirms that ATD's biochemical profile is conserved across diverse taxonomic groups.

We then checked whether EF-Tu can protect L-Thr-tRNA^{Thr}(G4•U69) from ATD, as EF-Tu occurs abundantly in the cell and most aminoacyl-tRNAs are expected to exist in complex with EF-Tu, ready for delivery to the ribosome. On the basis of thermodynamic compensation, EF-Tu is expected to bind the non-cognate L-Ala-tRNA^{Thr}(G4•U69) with significantly lower affinity compared to the cognate L-Thr-tRNA^{Thr}(G4•U69) (***25***). Competition assays demonstrated that L-Ala-tRNA^{Thr}(G4•U69) undergoes significant deacylation with 50 nM MmATD in the presence of EF-Tu (**Fig. 3F**). By contrast, L-Thr-tRNA^{Thr}(G4•U69) is completely protected by EF-Tu even at 500 nM enzyme, whereas its protection is marginally relieved at 5 µM MmATD (**Fig. 3G**). Hence, the discrimination potential of MmATD for these two substrates enhances from ~50-fold in the absence of EF-Tu to >100-fold in the presence of EF-Tu (**Fig. 3, B, C, F and G**). The above biochemical data clearly suggest that L-Ala-tRNA^{Thr}(G4•U69) is ATD's physiological substrate.

To identify any correlation between the presence of ATD and $tRNA^{Thr}$(G4•U69), we performed a thorough bioinformatic analysis which showed that many Animalia genomes are enriched in $tRNA^{Thr}$(G4•U69) genes, with the latter comprising about one-third of total $tRNA^{Thr}$ genes. Such enrichment occurs throughout phylum Chordata as well as in one organism from phylum Echinodermata whose tRNA sequences are available (**Fig. 3A, Fig. 4A, table S3 and data S1**). Additionally, amongst all the (G4•U69)-containing tRNA genes found in Chordata, $tRNA^{Thr}$(G4•U69) genes comprise the major fraction, occurring in significantly more number of organisms than any other (G4•U69)-containing tRNA gene (**Fig. 4B**). Strikingly, a survey for the presence of ATD revealed its strict association with the enrichment of $tRNA^{Thr}$(G4•U69) genes (**Fig. 4A**). Remarkably, organisms (*e.g.*, *Drosophila melanogaster*) that lack $tRNA^{Thr}$(G4•U69) genes simultaneously lack ATD, including archaea which also seem to possess eukaryotic-type $AlaRS^{ND}$. Although many bacteria possess $tRNA^{Thr}$(G4•U69) genes, they lack $AlaRS^{ND}$ altogether and hence, the problem of L-alanine misacylation of $tRNA^{Thr}$(G4•U69) does not arise at all. Thus, the problem of mischarging of $tRNA^{Thr}$(G4•U69) with L-alanine arises only when $tRNA^{Thr}$(G4•U69) is present alongside $AlaRS^{ND}$. Such a strong as well as strict correlation between ATD and the problem of $tRNA^{Thr}$(G4•U69) mis-selection by $AlaRS^{ND}$, in terms of either concomitant occurrence or concomitant absence, clearly points toward a functional link between ATD and proofreading of $tRNA^{Thr}$(G4•U69) mis-selection.

The possibility of ATD's involvement in proofreading L-alanine mischarged on $tRNA^{Cys}$(G4•U69) or other G4•U69-carrying tRNAs can be ruled out on the following bases. Firstly, the distribution of other such tRNA genes is restricted to only a small subset of organisms. For instance, $tRNA^{Cys}$(G4•U69) genes occur in only 19 of 62 chordate species whose tRNA sequences are available, whereas $tRNA^{Thr}$(G4•U69) genes exist in all 62 species (**Fig. 4B**).

Secondly, the relative abundance of G4•U69 is significantly less in other tRNA genes compared to $tRNA^{Thr}$ genes in Chordata. For example, $tRNA^{Cys}$(G4•U69) genes constitute only 0.69–11.9% (average ~5%) of total $tRNA^{Cys}$ genes, whereas $tRNA^{Thr}$(G4•U69) genes form 20–40% (average ~30%) of total $tRNA^{Thr}$ genes. Thirdly, in the proteomics study of HeLa cells, misincorporation of L-alanine was observed at cysteine positions but not at threonine positions (***22***). The latter observation is striking, since the relative abundance of G4•U69 is significantly more in $tRNA^{Thr}$ genes than in $tRNA^{Cys}$ genes. Overall, our in-depth *in-silico* analysis in combination with extensive biochemical probing confirms that ATD serves as a unique and dedicated factor for correcting $tRNA^{Thr}$(G4•U69) selection error committed by eukaryotic $AlaRS^{ND}$ (**Fig. 4C**).

The current study has identified ATD as a novel Animalia-specific proofreading module belonging to the DTD-like fold, wherein the *trans* conformation of its active-site Gly-Pro motif has led to a "gain of function" by relaxing its substrate chiral specificity. ATD rectifies a unique tRNA mis-selection rather than a mistake in amino acid selection by a synthetase that has been studied so far in the context of proofreading (***26–28***). Such an error correction capability has not been attributed to any known editing domain, although ambiguous tRNA selection happens in several instances, wherein the ambiguity imparts selective advantage to the system (**Fig. 4C**) (***7***). Besides, it also suggests that evolutionary gain of function by $AlaRS^{ND}$ in charging (G4•U69)-bearing tRNAs with L-alanine may be advantageous, but may also require factors like ATD to keep such "errors" below precarious levels, thus avoiding global misfolding. The regulatory function, if any, of ATD in generating proteome diversity, thereby providing selective advantage to a cell or tissue type, needs to be explored through up/down-regulation as well as by using knockout approaches in multiple systems. It has been recently suggested that the size of tRNA

limits the identity determinants required for faithful translation without cross-reacting with non-cognate synthetases (***29***). As has been noted in the current work, the expansion of genome size (from around 100 million base pairs in non-Chordata to >1 billion base pairs in Chordata) has led to such a cross-reactivity and enhancement in tRNA mis-selection, thereby necessitating recruitment of dedicated factors for error correction (**Fig. 4A**). Identification of ATD in the present study provides the first instance of such a scenario. The advent of ATD thus marks a key event associated with the appearance of Animalia, and more specifically of Chordata about 500 million years ago, to ensure translational quality control.

References

1. M. Ibba, D Söll, Aminoacyl-tRNA synthesis. *Annu. Rev. Biochem.* **69**, 617–650 (2000).
2. J. M. Ogle, V. Ramakrishnan, Structural insights into translational fidelity. *Annu. Rev. Biochem.* **74**, 129–177 (2005).
3. M. Guo, P. Schimmel, Structural analyses clarify the complex control of mistranslation by tRNA synthetases. *Curr. Opin. Struct. Biol.* **22**,119–126 (2012).
4. A. Moghal, K. Mohler, M. Ibba, Mistranslation of the genetic code. *FEBS Lett.* **588**, 4305–4310 (2014).
5. L. Ribas de Pouplana, M. A. Santos, J. H. Zhu, P. J. Farabaugh, B. Javid, Protein mistranslation: friend or foe? *Trends Biochem. Sci.* **39**, 355–362 (2014).
6. M. V. Rodnina, The ribosome in action: Tuning of translational efficiency and protein folding. *Protein Sci.* **25**, 1390–1406 (2016).

7. M. H. Schwartz, T. Pan, Function and origin of mistranslation in distinct cellular contexts. *Crit. Rev. Biochem. Mol. Biol.* **52**, 205–219 (2017).
8. C. L. Simms, E. N. Thomas, H. S. Zaher, Ribosome-based quality control of mRNA and nascent peptides. *Wiley Interdiscip. Rev. RNA* **8**, e1366 (2017).
9. R. Calendar, P. Berg, D-Tyrosyl RNA: formation, hydrolysis and utilization for protein synthesis. *J. Mol. Biol.* **26**, 39–54 (1967).
10. J. Soutourina, P. Plateau, F. Delort, A. Peirotes, S. Blanquet, Functional characterization of the D-Tyr-tRNA^{Tyr} deacylase from *Escherichia coli*. *J. Biol. Chem.* **274**, 19109–19114 (1999).
11. S. Ahmad, S. B. Routh, V. Kamarthapu, J. Chalissery, S. Muthukumar, T. Hussain, S. P. Kruparani, M. V. Deshmukh, R. Sankaranarayanan, Mechanism of chiral proofreading during translation of the genetic code. *Elife* **2**, e01519 (2013).
12. S. B. Routh, K. I. Pawar, S. Ahmad, S. Singh, K. Suma, M. Kumar, S. K. Kuncha, K. Yadav, S. P. Kruparani, R. Sankaranarayanan, Elongation factor Tu prevents misediting of Gly-tRNA(Gly) caused by the design behind the chiral proofreading site of D-aminoacyl-tRNA deacylase. *PLoS Biol.* **14**, e1002465 (2016).
13. M. L. Ferri-Fioni, E. Schmitt, J. Soutourina, P. Plateau, Y. Mechulam, S. Blanquet, Structure of crystalline D-Tyr-tRNA^{Tyr} deacylase. A representative of a new class of tRNA-dependent hydrolases. *J. Biol. Chem.* **276**, 47285-47290 (2001).
14. M. L Ferri-Fioni, M. Fromant, A. P. Bouin, C. Aubard, C. Lazennec, P. Plateau, S. Blanquet, Identification in archaea of a novel D-Tyr-tRNA^{Tyr} deacylase. *J. Biol. Chem.* **281**, 27575–27585 (2006).

15. S. Wydau, M. L. Ferri-Fioni, S. Blanquet, P. Plateau, GEK1, a gene product of *Arabidopsis thaliana* involved in ethanol tolerance, is a D-aminoacyl-tRNA deacylase. *Nucleic Acids Res.* **35**, 930–938 (2007).

16. S. Wydau, G. van der Rest, C. Aubard, P. Plateau, S. Blanquet, Widespread distribution of cell defense against D-aminoacyl-tRNAs. *J. Biol. Chem.* **284**, 14096–14104 (2009).

17. S. Dwivedi, S. P. Kruparani, R. Sankaranarayanan, A D-amino acid editing module coupled to the translational apparatus in archaea. *Nat. Struct. Mol. Biol.* **12**, 556–557 (2005).

18. T. Hussain, S. P. Kruparani, B. Pal, A. C. Dock-Bregeon, S. Dwivedi, M. R. Shekar, K. Sureshbabu, R. Sankaranarayanan, Post-transfer editing mechanism of a D-aminoacyl-tRNA deacylase-like domain in threonyl-tRNA synthetase from archaea. *EMBO J.* **25**, 4152–4162 (2006).

19. T. Hussain, V. Kamarthapu, S. P. Kruparani, M. V. Deshmukh, R. Sankaranarayanan, Mechanistic insights into cognate substrate discrimination during proofreading in translation. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22117–22121 (2010).

20. S. Ahmad, S. Muthukumar, S. K. Kuncha, S. B. Routh, A. S. Yerabham, T. Hussain, V. Kamarthapu, S. P. Kruparani, R. Sankaranarayanan, Specificity and catalysis hardwired at the RNA–protein interface in a translational proofreading enzyme. *Nat. Commun.* **6**, 7552 (2015).

21. E. Fan, D. Baker, S. Fields, M. H. Gelb, F. S. Buckner, W. C. Van Voorhis, E. Phizicky, M. Dumont, C. Mehlin, E. Grayhack, M. Sullivan, C. Verlinde, G. Detitta, D. R. Meldrum, E. A. Merritt, T. Earnest, M. Soltis, F. Zucker, P. J. Myler, L. Schoenfeld, D. Kim, L. Worthey, D. Lacount, M. Vignali, J. Li, S. Mondal, A. Massey, B. Carroll, S.

Gulde, J. Luft, L. Desoto, M. Holl, J. Caruthers, J. Bosch, M. Robien, T. Arakaki, M. Holmes, I. Le Trong, W. G. Hol, Structural genomics of pathogenic protozoa: an overview. *Methods Mol. Biol.* **426**, 497–513 (2008).

22. L. Sun, A. C. Gomes, W. He, H. Zhou, X. Wang, D. W. Pan, P. Schimmel, T. Pan, X. L. Yang, Evolutionary gain of alanine mischarging to noncognate tRNAs with a G4:U69 base pair. *J. Am. Chem. Soc.* **138**, 12948–12955 (2016).
23. Y. M. Hou, P. Schimmel, A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature*. **333**, 140–145 (1988).
24. W. H. McClain, K. Foss, Changing the identity of a tRNA by introducing a G-U wobble pair near the 3′ acceptor end. *Science*. **240**, 793–796 (1988).
25. F. J. LaRiviere, A. D. Wolfson, O. C. Uhlenbeck, Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science* **294**, 165–168 (2001).
26. L. Pauling, "[The probability of errors in the process of synthesis of protein molecules]" in *Arbeiten aus dem Gebiet der Naturstoffe* (Birkhäuser Verlag, Basel, 1958), pp. 597–602.
27. A. R. Fersht, Sieves in sequence. *Science* **280**, 541 (1998).
28. O. Nureki, D. G. Vassylyev, M. Tateno, A. Shimada, T. Nakama, S. Fukai, M. Konno, T. L. Hendrickson, P. Schimmel, S. Yokoyama, Enzyme structure with two catalytic sites for double-sieve selection of substrate. *Science* **280**, 578–582 (1998).
29. A. Saint-Léger, C. Bello, P. D. Dans, A. G. Torres, E. M. Novoa, N. Camacho, M. Orozco, F. A. Kondrashov, L. Ribas de Pouplana, Saturation of recognition elements blocks evolution of new tRNA identities. *Sci. Adv.* **2**, e1501860 (2016).

30. F. van den Ent, J. Löwe J, RF cloning: a restriction-free method for inserting target genes into plasmids. *J. Biochem. Biophys. Methods* **67**, 67–74 (2006).

31. Z. Otwinowski, W. Minor, "Processing of X-ray diffraction data collected in oscillation mode" in *Methods in Enzymology, Macromolecular Crystallography, Part A*, C. Carter Jr., R. Sweet, Eds. (Elsevier, 1997), 1st ed., vol. 276, pp. 307–326.

32. CCP4, The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760–763 (1994).

33. G. N. Murshudov, A. A. Vagin, E. J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).

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

35. R. A. Laskowski, M. W. MacArthur, D. S. Moss, J. M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).

36. S. Ledoux, O. C. Uhlenbeck, [3′-^{32}P]-labeling tRNA with nucleotidyltransferase for assaying aminoacylation and peptide bond formation. *Methods* **44**, 74–80 (2008).

37. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–12 (2004).

Acknowledgements

The authors acknowledge Dr. P. Chandra Shekar for kindly providing mouse cDNA. S.K.K. thanks DST-INSPIRE, India, for research fellowship. M.M. thanks Department of Biotechnology, India for research fellowship. S.B.R. and R.S. acknowledge funding from 12th Five Year Plan Project BSC0113 of CSIR, India. R.S. also acknowledges funding from J.C. Bose Fellowship of SERB, India, and Centre of Excellence Project of Department of Biotechnology, India.

List of Supplementary Materials

Figures

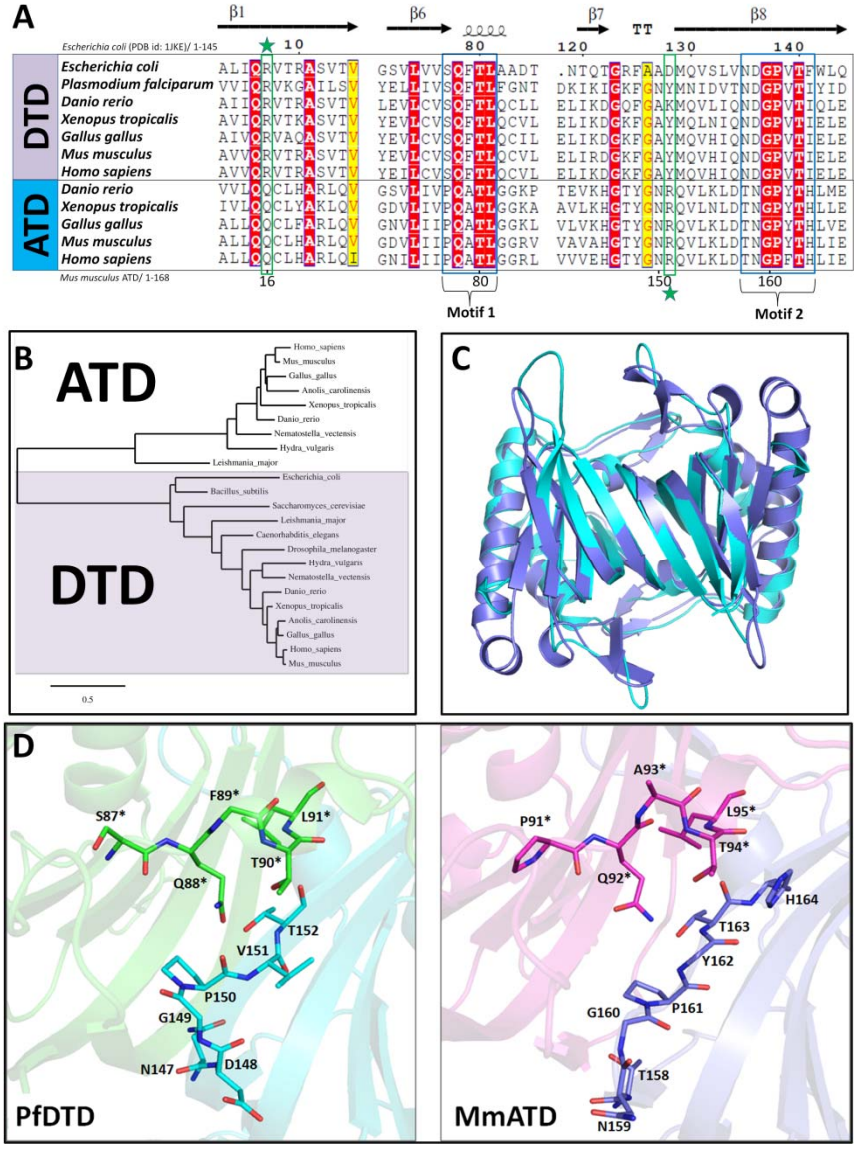

Fig. 1. ATD is an Animalia-specific protein belonging to DTD-like fold. (**A**) Multiple sequence alignment showing similar but distinct and characteristic sequence motifs in DTD and ATD (motifs 1 and 2). The highly conserved arginine in DTD (Arg7, EcDTD) is indicated by a star above, whereas the invariant arginine in ATD (Arg151, MmATD) is highlighted by a star below. (**B**) Phylogenetic classification of DTD and ATD showing their grouping into two

separate categories. (**C**) Crystal structure of MmATD homodimer (blue) superimposed on that of PfDTD homodimer (cyan; PDB id: 4NBI). (**D**) Crystal structures of PfDTD (PDB id: 4NBI) and MmATD showing that motifs 1 and 2 form the active site at the dimeric interface in both. Residues from the dimeric counterpart are indicated by *.

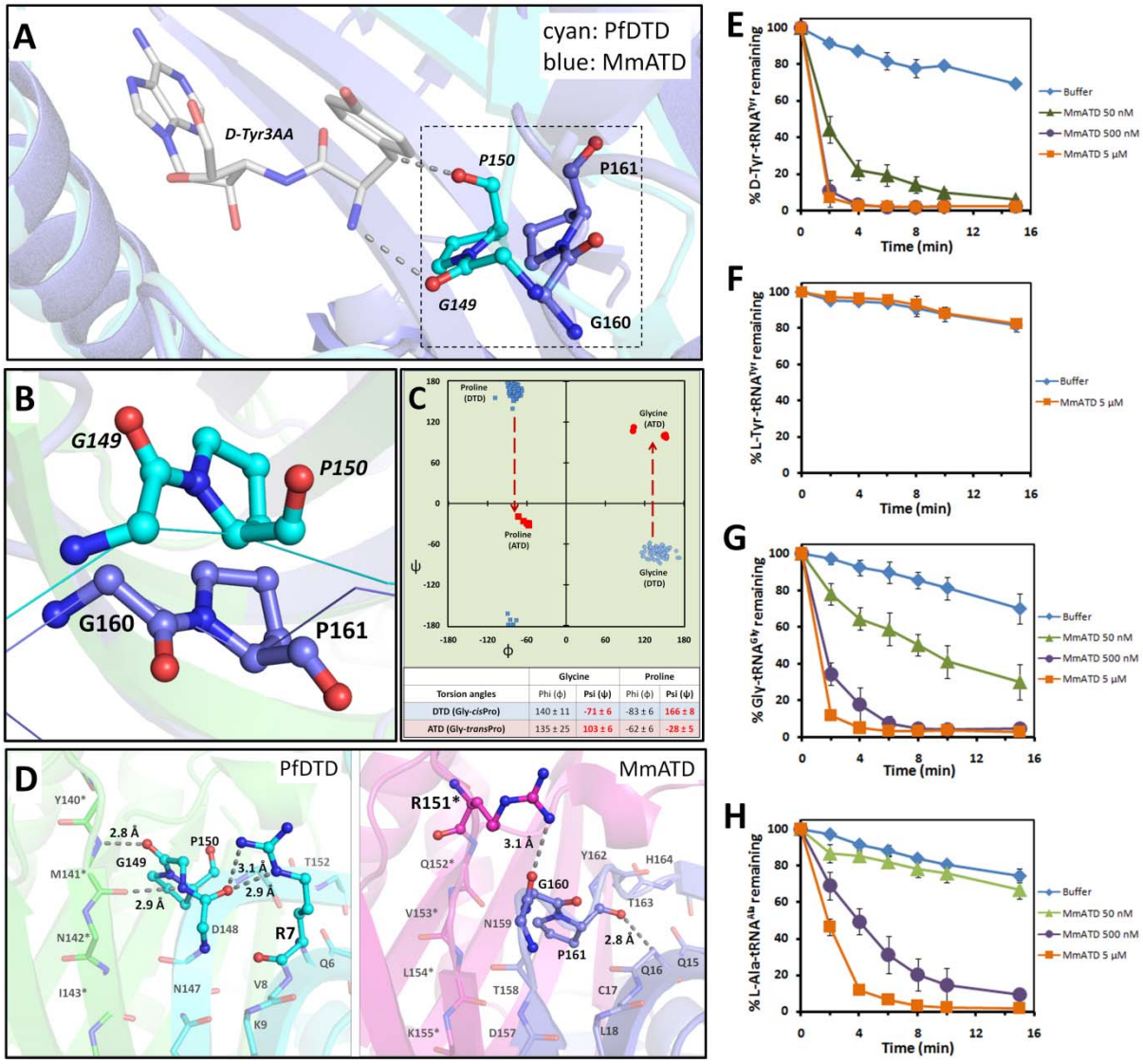

Fig. 2. Relaxation of substrate chiral specificity in ATD. (**A**) A comparison of Gly-*trans*Pro motif in MmATD with Gly-*cis*Pro motif in PfDTD (PDB id: 4NBI) after structural superposition of the two proteins. (**B**) The comparison shown in (**A**) depicted from a different angle, highlighting the opposite orientation of carbonyls of the two enzymes. (**C**) Ramachandran plot of glycine and proline residues of the Gly-Pro motif of all the available crystal structures of DTD (blue) and ATD (red), highlighting the change of ~180° in the ψ torsion angle. (**D**) Interaction of the side chain of Arg7 with the Gly-*cis*Pro motif in PfDTD (PDB id: 4NBI), and the side chain

of Arg151 with the Gly-*trans*Pro motif in MmATD. Residues from the dimeric counterpart are indicated by *. (**E–H**) Deacylation of D-Tyr-tRNA^{Tyr}, L-Tyr-tRNA^{Tyr}, Gly-tRNA^{Gly} and L-Ala-tRNA^{Ala} by different concentrations of MmATD. Error bars denote one standard deviation from the mean of three independent readings.

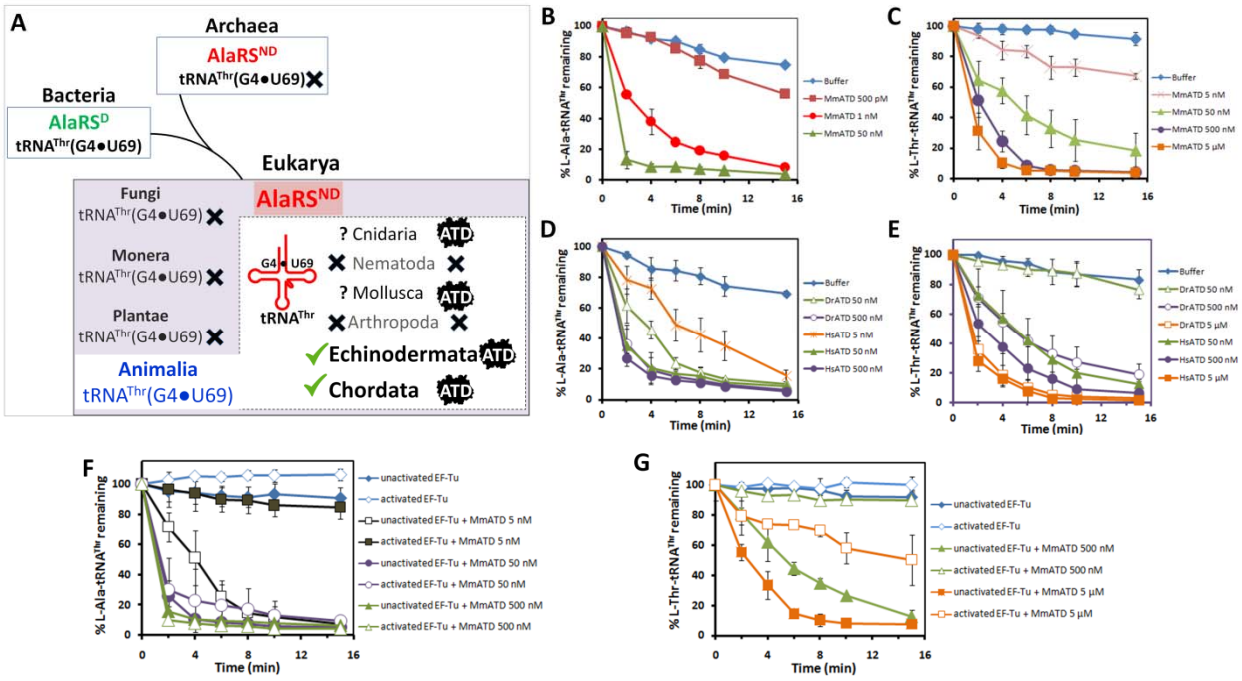

Fig. 3. Proofreading of tRNA^{Thr}(G4•U69) selection error, made by eukaryotic AlaRSND, by ATD. (**A**) Distribution of AlaRSND, tRNA^{Thr}(G4•U69), and ATD in different domains of life. tRNA sequences of Cnidaria and Mollusca are not available in the database. Data for occurrence of AlaRSND and tRNA^{Thr}(G4•U69) have been taken from (***22***). (**B** and **C**) Deacylation of L-Ala-tRNA^{Thr}(G4•U69) and L-Thr-tRNA^{Thr}(G4•U69) by different concentrations of MmATD. (**D** and **E**) Deacylation of L-Ala-tRNA^{Thr}(G4•U69) and L-Thr-tRNA^{Thr}(G4•U69) by different concentrations of DrATD and HsATD. (**F** and **G**) Deacylation of L-Ala-tRNA^{Thr}(G4•U69) and L-Thr-tRNA^{Thr}(G4•U69) by MmATD in the presence of EF-Tu, showing protection of the cognate substrate by EF-Tu against MmATD. Error bars denote one standard deviation from the mean of three independent readings.

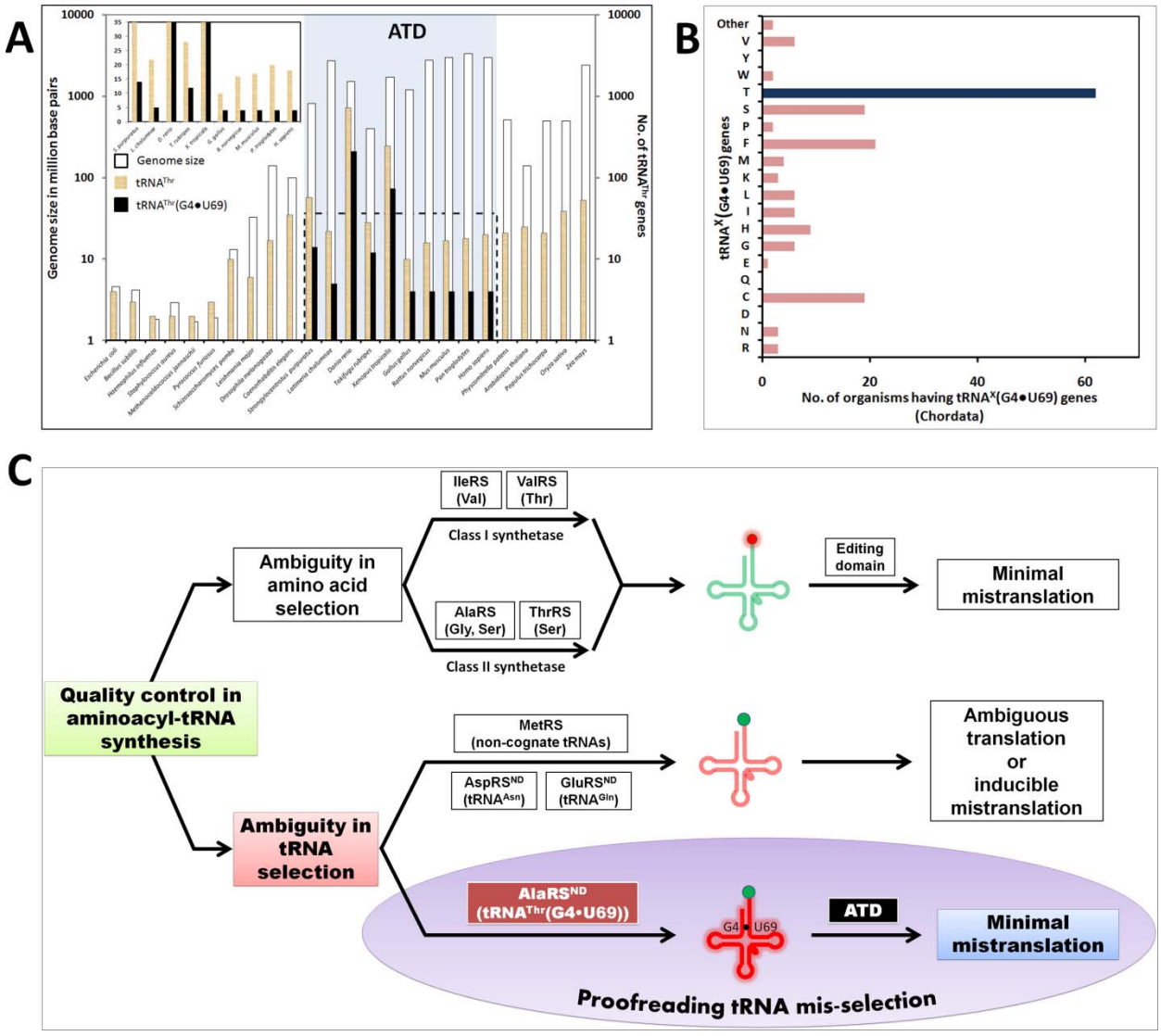

Fig. 4. Strict association between enrichment of tRNA^{Thr}(G4•U69) and presence of ATD. (**A**) Bar graph (logarithmic scale) depicting genome size, total number of tRNA^{Thr} genes, and number of tRNA^{Thr}(G4•U69) genes occurring in representative organisms belonging to all the three domains of life. Inset showing the number of total tRNA^{Thr} genes and tRNA^{Thr}(G4•U69) genes in normal scale; genome size has not been shown for the sake of clarity. Presence of ATD is highlighted in light blue box. Data for occurrence of AlaRSND and tRNA^{Thr}(G4•U69) has been taken from (***22***). (**B**) Bar graph showing the number of genomes (organisms) containing

(G4•U69)-containing tRNA genes which code for tRNAs specific for various proteogenic amino acids. (**C**) Model for mis-selection and consequent misacylation of $tRNA^{Thr}$(G4•U69) with L-alanine by $AlaRS^{ND}$, and its proofreading by ATD. Cognate and non-cognate tRNAs (clover leaf model) are coloured in green and red, respectively. Likewise, cognate and non-cognate amino acids (circle) are rendered in green and red, respectively.