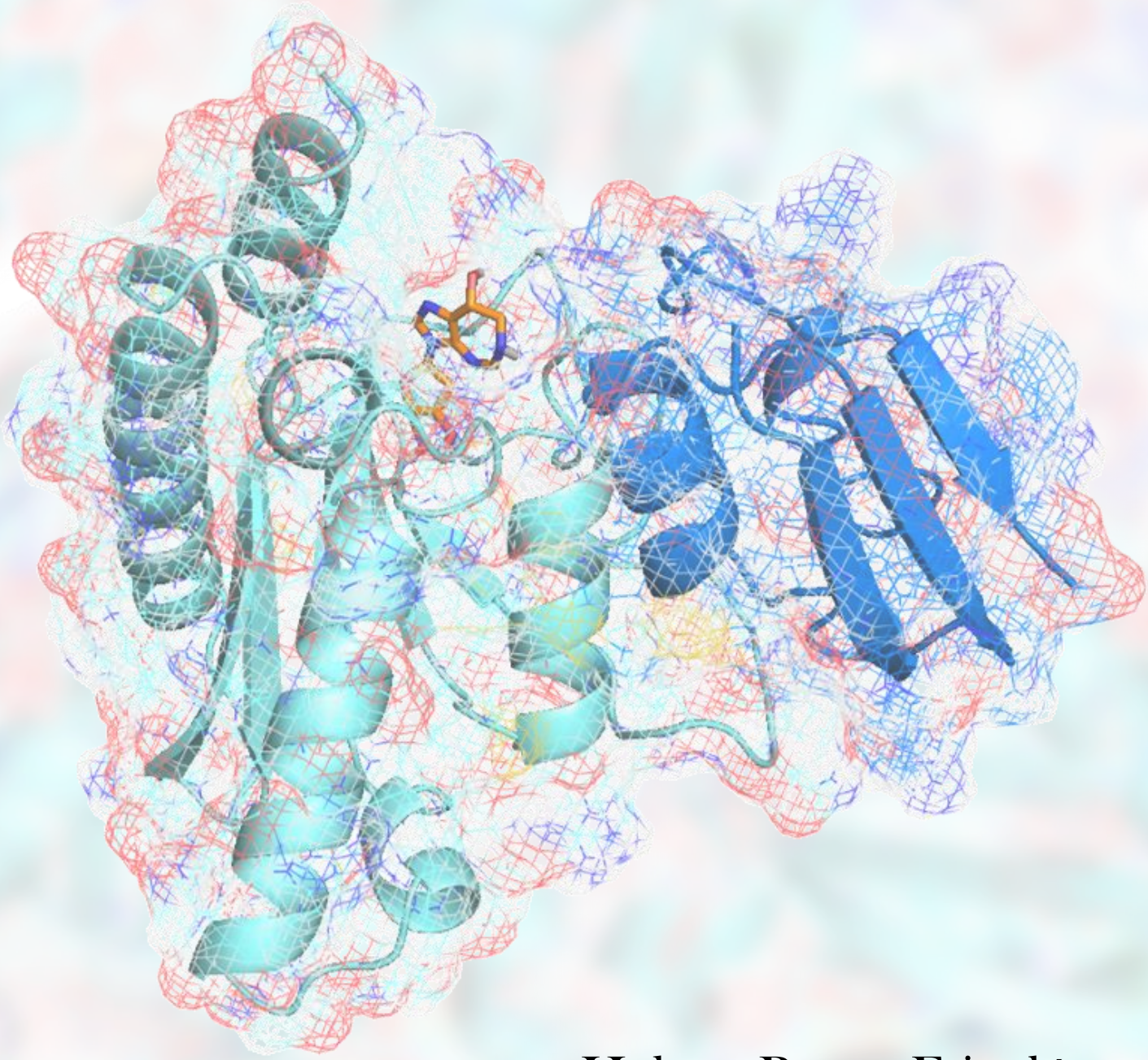
Biomedical studies of human adenosine deaminase acting on transfer RNA and related therapeutic strategies

Helena Roura Frigolé

Biomedical studies of human adenosine deaminase acting on transfer RNA and related therapeutic strategies



Helena Roura Frigolé

UNIVERSITAT DE BARCELONA

FACULTAT DE FARMÀCIA I CIÈNCIES DE L'ALIMENTACIÓ

Programa de doctorat en Biotecnologia

INSTITUTE FOR RESEARCH IN BIOMEDICINE

Biomedical studies of human adenosine deaminase acting on transfer RNA and related therapeutic strategies

Memòria presentada per **Helena Roura Frigolé** per optar al títol de doctor per la Universitat de Barcelona

Lluís Ribas de Pouplana	Helena Roura Frigolé	Elena Escubedo Rafa
Thesis director	PhD candidate	Tutor

Helena Roura Frigolé

2018

Our greatest weakness lies in giving up. The most certain way to succeed is always to try just one more time.

Thomas A. Edison

Summary

Adenosine deaminase acting on transfer RNA (ADAT) is a human heterodimeric enzyme that catalyzes the deamination of adenosine (A) to inosine (I) at the first position of the anticodon of transfer RNAs (tRNAs) (position 34, or wobble position); one of the few essential post-transcriptional modifications on tRNAs (1-5). Inosine 34 allows the recognition of three different nucleotides: cytidine, uridine and adenosine, at the third position of the codon, thus increasing the decoding capacity of tRNAs to more than one messenger RNA (mRNA) codon (adenosine 34 can in principle only pair with codons with uridine at the third position) (6, 7). This alters the tRNA pool available for each codon and it has been proved to align the correlation between codon usage and tRNA gene copy number (9). It has also been suggested to improve fidelity and efficiency of translation (9, 10), especially for mRNAs enriched in codons translated by modified tRNAs (11, 12).

Monitoring ADAT-mediated deamination is crucial for the characterization of the enzyme in terms of activity, substrates, regulation, as well as for drug discovery purposes. However, this analysis is often challenging, laborious and lacks quantitativeness. We developed an *in vitro* deamination assay based on restriction fragment length polymorphism (RFLP) analyses to monitor ADAT activity in an efficient, cost-effective, and semiquantitative manner (13). To overcome a limitation of the method being the need of reverse transcription and amplification of the tRNA, we designed a direct method to quantify I34 formation *in vitro* using the first fluorescent analogs of nucleic acids that have been reported to undergo enzymatic deaminations (14-16).

ADAT has been conserved over the evolution with the acquisition of multi-substrate specificity. Whereas its bacterial homolog TadA deaminates exclusively $tRNA^{Arg}$ (2), the human enzyme deaminates eight different tRNAs (3, 17). However, the mechanisms that drove this evolution remain unknown. While the substrate recognition in TadA has been well studied, in the eukaryotic ADAT is poorly understood. Through *in vitro* enzymatic activity assays with different variants of $tRNA^{Arg}$ and $tRNA^{Ala}$, we elucidated the most important features for efficient A34-to-I34 conversion and characterized the substrate recognition of the human enzyme. We also proposed a new potential mechanism of control of ADAT deamination activity by human tRNA-derived fragments, which provides new insights into the regulation of ADAT function and may open a door for the development of new strategies to modulate ADAT activity.

A missense mutation (V128M) in one of the two subunits of the human ADAT enzyme causes intellectual disability and strabismus, but the molecular bases of the pathology are

unknown (18, 19). We characterized human ADAT in terms of kinetics and structure, and investigated the effect of the V128M mutation. We found that this substitution decreases ADAT deamination activity, and severely affects the stability of the quaternary structure of the enzyme. In this regard, we discovered small molecules with the ability to activate the enzyme, which could potentially recover the defective tRNA editing caused by the mutation.

Resum

L'adenosina deaminasa específica per RNA de transferència (ADAT) és un enzim humà heterodimèric que catalitza la reacció de deaminació de l'adenosina (A) a inosina (I) a la primera posició de l'anticodó de RNAs de transferència (tRNAs) (també anomenada posició 34 o posició de balanceig); una de les poques modificacions post-transcripcionals essencials en tRNAs (1-5). La inosina 34 permet el reconeixement de tres nucleòtids diferents: citidina, uridina i adenosina, a la tercera posició del codó, augmentant per tant la capacitat de descodificació dels tRNAs a més d'un codó en l'RNA missatger (mRNA) (l'adenosina 34 en principi únicament pot aparellar-se amb uridina en la tercera posició) (6, 7). Això altera els nivells de tRNAs disponibles per cada codó i s'ha demostrat que alinea la correlació entre l'ús de codons i número de còpies gèniques de cada tRNA (9). També s'ha suggerit que millora la fidelitat i eficiència de la traducció (9, 10), especialment per mRNAs enriquits en codons traduïts per tRNA modificats (11, 12).

Monitoritzar la deaminació produïda per ADAT és clau per la caracterització de l'enzim en termes d'activitat, substrats, regulació, així com també pel disseny de fàrmacs. No obstant, aquest anàlisi és sovint complex, laboriós i poc quantitatiu. Per això, hem desenvolupat un assaig de deaminació *in vitro* basat en el polimorfisme de longitud dels fragments de restricció (RFLP) amb el propòsit de monitoritzar l'activitat d'ADAT de manera eficient, cost-efectiva i semiquantitativa (13). Per superar una limitació del mètode essent la necessitat de transcripció reversa i amplificació del tRNA, hem dissenyat un mètode directe per quantificar la formació d'I34 *in vitro* utilitzant els primers anàlegs fluorescents d'àcids nucleics que són substrats de deaminacions enzimàtiques descrits fins al moment (14-16).

ADAT ha estat conservat en l'evolució amb l'adquisició de multi-especificitat de substrat. Mentre que el seu homòleg bacterià TadA deamina exclusivament $tRNA^{Arg}$ (2), l'enzim humà deamina vuit tRNAs diferents (3, 17). Tot i així, els mecanismes que van conduir a aquesta evolució romanen desconeguts. Mentre que el reconeixement de substrat en TadA es coneix bé, en ADAT eucariòtic ha estat poc estudiat. A través d'assajos enzimàtics *in vitro* amb diferents variants de $tRNA^{Arg}$ i $tRNA^{Ala}$, hem elucidat els trets més importants per a una conversió eficient d'A34 a I34 i hem caracteritzat el reconeixement de substrat per part de l'enzim humà. També proposem un nou potencial mecanisme de control de l'activitat d'ADAT per part de fragments derivats de tRNAs humans, el qual ofereix noves perspectives en la regulació de la funció d'ADAT i que pot obrir la porta al desenvolupament de noves estratègies per modular-ne l'activitat.

Una mutació sense sentit (V128M) en una de les dues subunitats en l'enzim ADAT humà s'ha associat a retard mental i estrabisme, encara que les bases moleculars de la patologia es desconeixen (18, 19). Hem caracteritzat ADAT humà en termes de cinètica enzimàtica i estructura, i n'hem investigat l'efecte de la mutació V128M. Hem descobert que la substitució de valina 128 per metionina redueix l'activitat deaminatòria d'ADAT i que altera severament l'estabilitat de l'estructura quaternària de l'enzim. En aquest aspecte, hem descobert molècules amb l'habilitat d'activar l'enzim, la qual cosa podria revertir potencialment la reducció en l'activitat enzimàtica causada per la mutació.

Abbreviations

A	Adenine / Adenosine
aa	Amino acid
aaRS	Aminoacyl-tRNA synthetase
aa-tRNA	Aminoacyl-transfer RNA
ADA	Adenosine deaminase
ADAR	Adenosine deaminase acting on RNA
ADAT	Adenosine deaminase acting on transfer RNA
AGO	Argonaute
ALS	Amyotrophic lateral sclerosis
AMP	Adenosine monophosphate
ANG	Angiogenin
APAF1	Apoptotic protease-activating factor 1
APB	Acryloylaminophenylboronic
APOBEC	Apolipoprotein B mRNA editing enzyme catalytic peptide-like
ATP	Adenosine triphosphate
bp	Base pair
bt	Base triple
C	Cytosine / Cytidine
CDA	Cytidine deaminase
CDAR	Cytidine deaminase acting on RNA
cDNA	Complementary DNA
CDS	Coding sequence
CF	Cystic fibrosis
COPD	Chronic obstructive pulmonary disease
D	Dihydrouridine
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
dsRBD	Double-stranded RNA binding domain
dsRNA	Double-stranded RNA
EHNA	*Erythro*-9-(2-Hydroxy-3-nonyl)adenine
EMSA	Electrophoresis mobility shift assay
ER	Endoplasmic reticulum

FC	Fold change
FPLC	Fast protein liquid chromatography
G	Guanine / Guanosine
Hsp	Heat-shock protein
HTS	High-throughput screening
I	Inosine
IC50	Half maximal inhibitory concentration
ID	Intellectual disability
IPTG	Isopropyl β-D-1-thiogalactopyranoside
IRES	Internal ribosome entry site
i-tRF	Internal tRNA-derived fragment
LC/MS-MS	Liquid chromatography-tandem mass spectrometry
k_{cat}	Catalytic constant
K_m	Michaelis constant
k_{cat}/K_m	Catalytic efficiency
k_{off}	Dissociation constant
m^1I	1-methylinosine
$m^2{}_2G$	N2-N2-dimethyl-guanosine
MACCS	Molecular ACCess System
MALDI	Matrix-assisted laser desorption/ionization
MALS	Multiangle light scattering
mcm^5s^2U	5-methoxycarbonylmethyl-2-thiouridine
mcm^5U	5-methoxycarbonylmethyluridine
Mg	Magnesium
miRNA	MicroRNA
mRNA	Messenger RNA
MS	Mass Spectrometry
MSC	Multi-aminoacyl-tRNA synthetase complex
Mw	Molecular weight
ncRNA	Non-coding RNA
NSun2	NOP2/Sun RNA methyltransferase 2
nt	Nucleotide
NTP	Nucleotide triphosphate
PAGE	Polyacrylamide gel electrophoresis

PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PPi	Pyrophosphate
R	Purine
RBS	Ribosome binding site
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
rRNA	Ribosomal RNA
RSV	Respiratory syncytial virus
RT	Reverse transcription
scaRNA	Small cajal RNA
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
SG	Stress granule
siRNA	Small interfering RNA
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
SOD1	Superoxide dismutase 1
T	Thymine / Thymidine / Ribothymidine (in RNA)
T4PNK	T4 polynucleotide kinase
TadA	Transfer RNA adenosine deaminase A (bacterial)
Tc	Tanimoto coefficient
tiRNA	tRNA-derived stress-induced RNA
TLC	Thin-layer chromatography
tRF	tRNA-derived fragment
Trm	tRNA methyltransferase
TRMT1	tRNA methyltransferase 1
tRNA	Transfer RNA
TRNT1	CCA-adding enzyme
TSEN	tRNA splicing endonuclease
UV	Ultraviolet
Ve	Elution volume
Vo	Void volume

WB	Western blot
wt	Wild type
Y	Pyrimidine
Ψ	Pseudouridine

Introduction

Cells are basic units of biological structure and function, and constitute the building blocks of tissues, organs and organisms. A human body contains about 40 trillion cells (20) and around 200 different cell types (21). All cells in a particular individual have the same genetic information stored as deoxyribonucleic acid (DNA), which is divided in chromosomes that contain the genes. The complete set of genes in a cell is known as genome. What determines the cell identity and function is the gene expression profile. The products resulting from expression of genes can be either ribonucleic acid (RNA) or proteins. The synthesis of DNA, RNA and proteins are sequential and unidirectional processes, which is known as the Central Dogma of Molecular Biology (22) (**Figure 1**). The genomic information is expressed in two steps: transcription and translation. Transcription is the process by which a segment of DNA is copied into RNA. Translation converts the information contained in RNA into a functional protein.

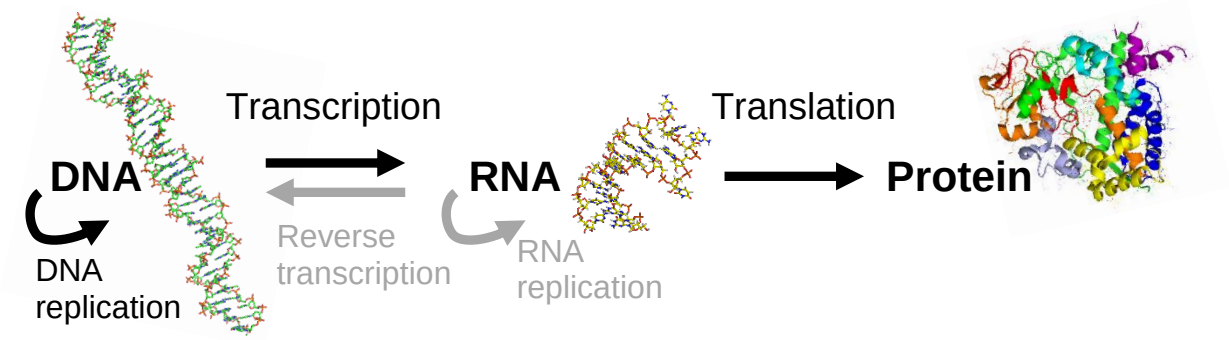


Figure 1. Central Dogma of Molecular Biology. Biological information (as bases in nucleic acids or amino acid residues in proteins) can only be transferred from nucleic acid to nucleic acid, or from nucleic acid to protein, but not from protein to protein, or from protein to nucleic acid. The general flow of biological information (black arrows) comprises: DNA can be copied to DNA (DNA replication), DNA can be copied into RNA (transcription), and proteins can be synthesized using the messenger RNA (mRNA) as a template (translation). Special transfers (grey arrows) include: RNA copied from RNA (RNA replication), and DNA synthesized using an RNA template (reverse transcription), which are only found in some viruses.

Nucleic acids

The nucleic acids RNA and DNA are linear polymers of repeating monomer units called nucleotides. Nucleotides are constituted by a sugar connected to a base through a β-glycosidic bond and to one to three phosphate groups through an ester bond (**Figure 2**). RNA differs from DNA in the sugar: in RNA it is ribose and in DNA deoxyribose. Both ribose and deoxyribose are five-ringed sugars (pentoses) with a 3'-OH group, but ribose has an additional 2'-OH group. The additional hydroxyl group in the ribose has two main consequences for the function of RNA compared to DNA: it increases the polarity of the

sugar and makes the RNA more chemically reactive than DNA, and influences on the secondary structures that RNA forms compared to DNA.

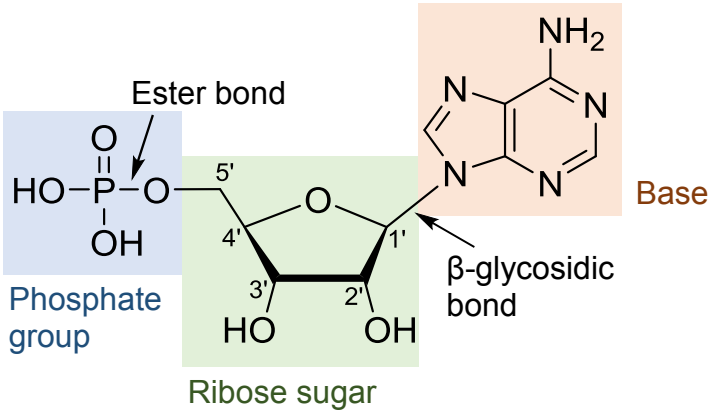


Figure 2. General structure of a nucleotide of RNA. The β-glycosidic bond is between the 1'-carbon atom of the sugar and a nitrogen atom in the base (N1 in pyrimidines, and N9 in purines). Phosphate groups are attached to the 5'-carbon of the sugar through an ester bond.

There are four different bases that can be attached to the sugar in both DNA and RNA: adenine (A), guanine (G), cytosine (C), and thymine (T) in DNA, and uracil (U) in RNA. All four bases are heterocycles containing carbon and nitrogen atoms with either one ring (U, C and T, called pyrimidines) or two rings (A and G, called purines) (**Figure 3**). As a result, both DNA and RNA are constituted by combinations of four different nucleotides that differ in the base.

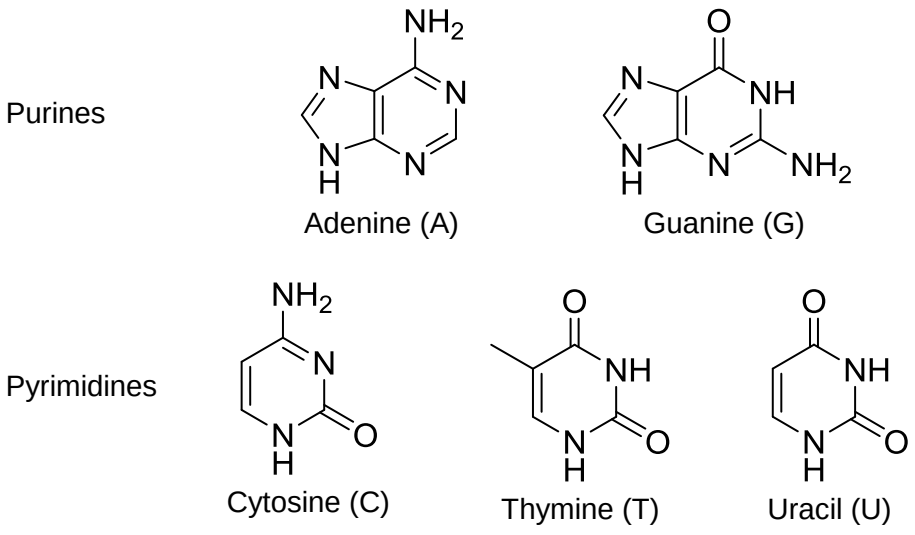


Figure 3. Structure of bases in nucleic acids. Thymine is exclusively found in DNA, and uracil in RNA.

Nucleotides are connected through phosphodiester bonds between the 3'-hydroxyl of a sugar group in a nucleotide and the phosphate group of the adjacent nucleotide. Nucleotides are thus joined together in a polynucleotide chain starting with a 5'-phosphate group and finishing with a 3'-OH group. This gives the chains a directional polarity 5'-3'. Although bases do not contribute to the phosphodiester backbone, they play a key role in interactions between nucleotide chains.

Complementary base pairs connect via hydrogen bonds according to Watson-Crick rules (23). G pairs with C (G:C), and A with T in DNA (A:T), and with U in RNA (A:U). A:T and A:U base pairs are connected by two hydrogen bonds, whereas G:C base pairs are connected by three, which makes the latter base pairing stronger (**Figure 4**).

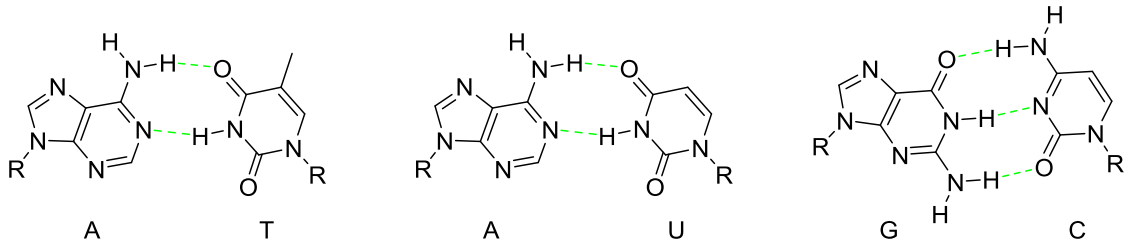


Figure 4. Watson-Crick base pairs. A pairs with T (U in RNA), and G with C. Dashed lines indicate hydrogen bonds. R represents the attachment site of the sugar (ribose or deoxyribose).

Even though both DNA and RNA have primary structures composed of nucleotide sequences, their structural organization is generally very different, which explains their distinct biological functions. DNA is generally found as a continuous double helix between two complementary DNA strands constituting the chromosomes. In contrast, RNA can adopt a range of different secondary and tertiary structures. The secondary structure is composed of helices formed through base pairing, either within a single RNA molecule (intramolecular) or between different RNA molecules (intermolecular). The tertiary structure is the highest level of organization, in which RNA molecules with secondary structure fold up into very compact and highly organized structures. This ability of RNA to fold into diverse structures plays a key role in its functions in the cell. Whereas the biological function of DNA is storing the genetic information, RNA's functions are substantially more diverse, including enzymatic catalysis, cell signaling, recognition, structural roles, scaffolding, and acting as a template for the protein synthesis.

Proteins

Proteins are marcromolecules composed of one or more polypeptides, which are linear polymers of amino acids (aa). There is a total of 20 natural amino acids that consist of a

central chiral carbon (except for glycine) linked to an amino group, a carboxylic acid group, a hydrogen atom, and a distinctive R group known as side chain (**Figure 5**). Side chains can vary in size, shape, hydrogen-bonding capacity, hydrophobic character, chemical reactivity, etc., thus influencing protein structure and function (**Figure 6**).

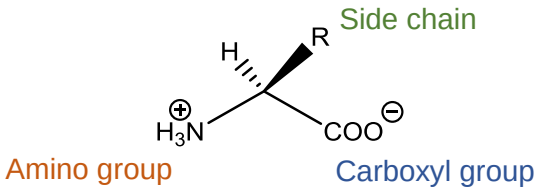


Figure 5. General chemical structure of an amino acid. This structure is common to all α-amino acids, except for proline which is a cyclic amino acid (the amino group is attached directly to the side chain). The R group or side chain attached to the α-carbon is different in each amino acid.

Amino acids are connected by peptide bonds between the α-carboxyl group of one amino acid to the α-amino group of another amino acid. This gives the polypeptides a directional polarity from N-terminus to C-terminus. The amino acid sequence of a polypeptide or protein is called primary structure. The spatial arrangement of amino acid residues that are nearby in the sequence interacting through local hydrogen bonds is called secondary structure. The two most common secondary structural elements are α-helix and β-sheet. These local features fold into a three-dimensional structure by non-covalent forces (mostly hydrogen bonds) constituting the tertiary structure. In the process of folding, amino acid residues with hydrophobic side chains tend to be located on the inside of the protein while those with hydrophilic side chains tend to be on the outside. Proteins containing more than one polypeptide chain (or subunit) exhibit a fourth level of structural organization. The spatial arrangement of subunits and the nature of their interactions constitutes the quaternary structure. Proteins are subjected to many reversible post-translational modifications, such as the covalent addition of a phosphate or an acetyl group to a specific amino acid side chain. The addition of these modifying groups is used to regulate the activity of a protein, changing its conformation, its binding to other proteins and its location inside the cell.

Proteins can form enormously sophisticated structures, whose functions largely depend on their three-dimensional structure and the detailed chemical properties of their surfaces. They have a large variety of functions, including structural, regulation, enzymatic, messenger, transport, storage, etc., constituting key elements in most of biological processes. For example, enzymes are catalytic proteins that greatly speed up reaction rates by binding the high-energy transition states for a specific reaction path. The rates of enzyme reactions are often so fast that they are limited only by diffusion; consequently, rates can be further increased if enzymes that act sequentially on a substrate are joined into a single multienzyme complex, or if the enzymes and their substrates are confined to the same compartment of the cell.

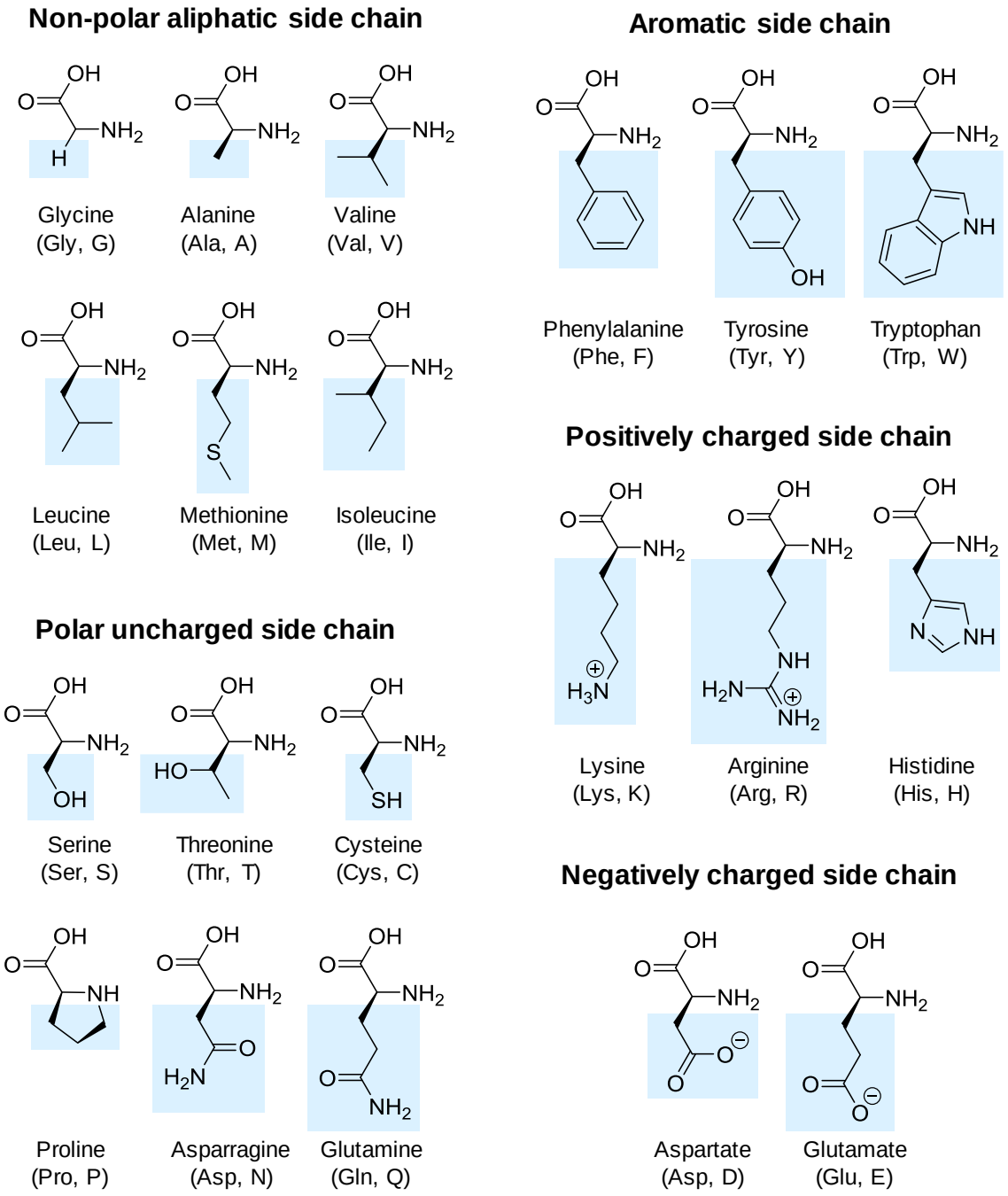


Figure 6. Chemical structures of the twenty natural amino acids. Amino acids are grouped according to the nature of the side chain. Three-letter and one-letter abbreviations for each amino acid are indicated.

The interactions between a protein and other molecules determines its biological properties. The binding always shows great specificity, and each protein molecule can usually bind just one or a few molecules out of the many thousands of different types it encounters. The ability of a protein to bind selectively and with high affinity to a ligand depends on the formation of noncovalent bonds: hydrogen bonds, electrostatic attractions, and Van der

Waals forces, plus favorable hydrophobic interactions. Since each individual bond is weak, effective binding occurs only when many of these bonds form simultaneously. Binding sites for ligands are formed as surface cavities in which precisely positioned amino acid side chains are brought together by protein folding. Proteins reversibly change their shape when ligands bind to their surface. The allosteric changes in protein conformation produced by one ligand can affect the binding of a second ligand, and this linkage between two ligand-binding sites provides a crucial mechanism for regulating cell processes. Whereas some small molecules inhibit enzymes, others activate them, and this is a mechanism to control enzyme function in response to changes in the concentrations of these ligands.

Transcription

Gene expression is the process whereby genes, encoded by DNA, express their inherited information usually in the form of proteins, albeit sometimes in the form of non-coding RNAs. Transcription is the first step of gene expression, in which a particular segment of DNA is copied into RNA. The enzyme RNA polymerase catalyzes the synthesis of the RNA molecule by adding free nucleotides that are complementary to the ones in the template strand of DNA. As a result of the transcription, a complementary antiparallel RNA strand called transcript is produced. In eukaryotic cells there are three RNA polymerases (I, II, and III), which are structurally similar but transcribe different types of genes. RNA polymerase I and III transcribe the genes encoding transfer RNA, ribosomal RNA, and various small RNAs, whereas RNA polymerase II transcribes most genes, including all those that encode proteins.

If the gene transcribed encodes a protein, the transcript is called messenger RNA (mRNA), which serves as a template for the protein synthesis. The RNA polymerase initiates transcription at the transcription start site of the chromatin template, with the intervention of chromatin-remodeling complexes, histone-modifying enzymes and transcription factors. The elongation phase of transcription in eukaryotes, in which nucleotides are incorporated to the growing RNA chain, is tightly coupled to RNA processing, and the nascent RNA undergoes three main processing events. First, a 7-methyl G is added to the 5'-end, a process called capping. As the polymerase moves along the DNA template, the spliceosome assembles on the RNA and intron sequences are removed from the pre-mRNA, which is known as splicing. Finally, when the polymerase reaches the end of the gene, the RNA is cleaved and the enzyme poly-A polymerase (PAP) adds approximately 200 A nucleotides to the 3'-end of the RNA (polyadenylation). Properly processed mRNAs are selectively exported from the nucleus to the cytosol through nuclear pore complexes, where they are translated into protein (**Figure 7**).

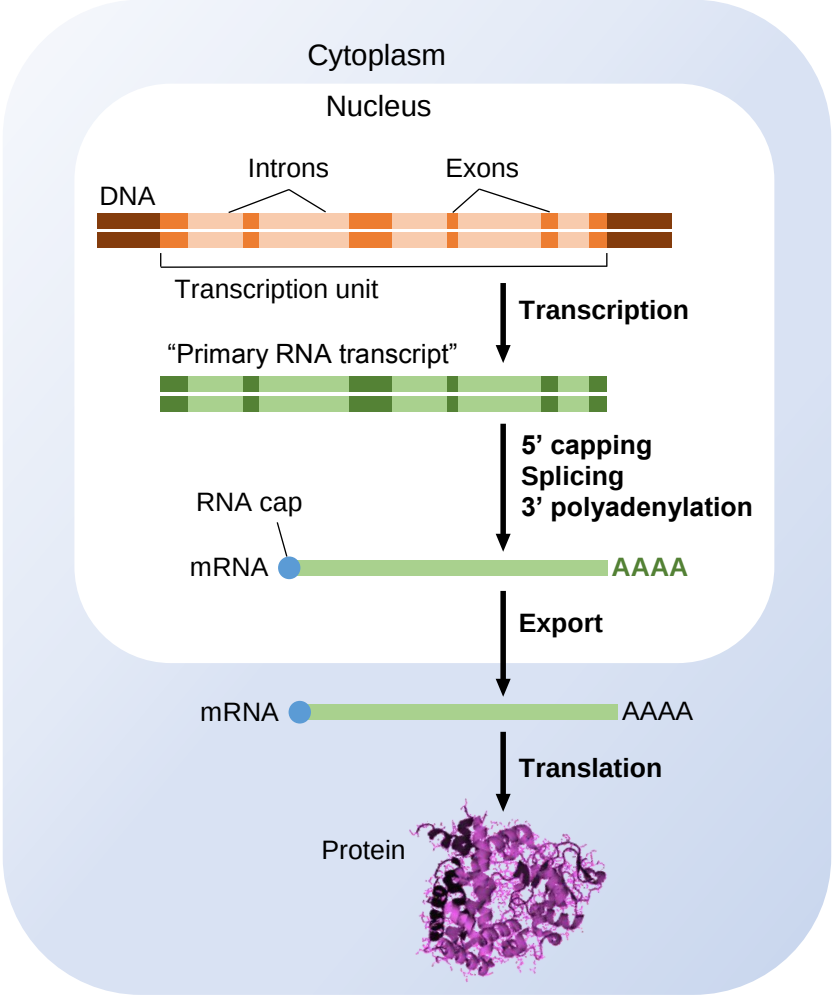


Figure 7. Schematic representation of the process of gene expression in eukaryotes, including transcription, RNA processing and translation. In eukaryotic cells, the RNA molecule resulting from transcription contains both coding (exon) and non-coding (intron) sequences. Before it can be translated into protein, the two ends of the RNA are modified, the introns are removed by an enzymatically catalyzed RNA splicing reaction, and the resulting mRNA is transported from the nucleus to the cytoplasm. Although the steps in this figure are depicted as occurring one after the other, they can occur concurrently. For example, the RNA cap is added and splicing typically begins before transcription has been completed. Because of the coupling between transcription and RNA processing, primary transcripts (unprocessed RNAs) are found only rarely.

For some genes, RNA is the final product. These non-coding RNAs serve as enzymatic and structural components for a wide variety of processes in the cell. The main types of non-coding RNAs produced in cells are:

- Ribosomal RNAs (rRNAs), which form the basic structure of the ribosome and catalyze protein synthesis.
- Transfer RNAs (tRNAs), acting as adaptors between mRNA and amino acids during protein synthesis.

- Small nuclear RNAs (snRNAs), functioning in a variety of nuclear processes, including splicing of pre-mRNA.
- Small nucleolar RNAs (snoRNAs), used to process and chemically modify rRNAs.
- Small cajal RNAs (scaRNAs), used to modify snRNAs and snoRNAs.
- MicroRNAs (miRNAs), which regulate gene expression typically by blocking translation of selective mRNAs.
- Small interfering RNAs (siRNAs), which turn off gene expression by directing degradation of selective mRNAs and the establishment of compact chromatin structures.
- Other non-coding RNAs (ncRNAs) that function in diverse cell processes, including telomere synthesis, X-chromosome inactivation, and the transport of proteins into the ER.

Overall, RNA molecules can have many different functions and help synthetize, regulate, and process proteins, playing a fundamental role in many biological processes within a cell.

Genetic code

Protein-coding genes are the only ones translated into proteins. For this, the four-letter alphabet of the nucleotide sequence of an mRNA is translated into the twenty-letter alphabet of proteins. The genetic code relates each amino acid to nucleotide triplets (codons) (**Table 1**). There are 64 possible combinations of three-nucleotide sequences that can be made from the four nucleotides. 61 of the 64 possible codons specify particular amino acids and 3 triplets designate the termination of translation (stop codons). The sequence of nucleotides is read sequentially without overlapping from a fixed starting point. Since the number of codons exceeds the number of amino acids, each amino acid is encoded by one or more specific codons. This is the reason why the genetic code is considered degenerate. In the opposite direction, the genetic code behaves unambiguously, and each codon encodes for on amino acid. Only tryptophan and methionine are encoded by just one codon each, the other 18 amino acids are each encoded by two or more. It has been shown that the number of codons for a particular amino acid correlates with its frequency of occurrence in proteins (24). Codons that specify the same amino acid are called synonyms. Synonym codons are grouped in the genetic code rather than randomly distributed, and in most cases share the first and second bases and only differ in the third one. Thus, changing one nucleotide to another at the third position of the codon frequently results in no change in the amino acid sequence and, in the cases that the amino acid changes, it usually has similar physical properties (25). In the same line, single-base substitutions in the first and second positions of the codon tend to result in changes in amino acids with related properties (e.g. codons with U as a second base encode hydrophobic amino acids, whereas codons with A as a second base encode for

hydrophilic amino acids) (26). As a result, point mutations or misreading are likely to preserve the amino acid specified or substitute it for another with related properties, suggesting that the genetic code would have evolved to minimize the effects caused by errors during translation.

First position (5'-end)	Second position: U	Second position: C	Second position: A	Second position: G	Third position (3'-end)
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Table 1. The genetic code identifies the amino acid encoded by each codon. The three termination codons are shaded in red, and the initiation codon AUG in green. All the amino acids except methionine and tryptophan are decoded by more than one codon. In most cases, codons that specify the same amino acid differ only at the third base.

Translation

Translation is the process by which the information contained in the nucleotide sequence of an mRNA molecule is used to synthetize a protein. For this, the transcribed mRNA is delivered to the ribosome where the translation takes place. First, aminoacyl-tRNA synthetases couple a particular amino acid to its corresponding tRNA. Second, aminoacylated tRNAs, whose anticodon base pairs with the complementary codon on the mRNA, deliver the amino acid to be incorporated to the growing polypeptide chain. The sequence of nucleotides in the mRNA is read from one end to the other in sets of three according to the genetic code. The polypeptide chain grows by the stepwise addition of

amino acids to its C-terminal end. Consequently, a protein is synthesized from its N-terminal end to its C-terminal end.

Aminoacyl-tRNA synthetases (aaRSs) are the enzymes that catalyze the aminoacylation reaction by covalently linking an amino acid to its cognate tRNA in the first step of protein translation. The amino acid is first activated through the covalent linkage of its carboxyl group to adenosine monophosphate (AMP), generating an adenylated amino acid. The reaction is driven by the hydrolysis of adenosine monophosphate (ATP), which donates the AMP, and pyrophosphate (PPi) is released. The adenylate-aaRS complex then binds the appropriate tRNA, and the carboxyl group of the amino acid is esterified to the 3- or 2-hydroxyl group of the ribose of the adenosine at the 3'-end of the tRNA, forming the aminoacyl-tRNA (aa-tRNA), also known as charged tRNA. Each aaRS attaches a particular amino acid to a specific tRNA. There are at least 20 different aaRSs, one for each amino acid. A total of 36 human aaRSs that charge all human tRNAs have been identified. 16 of them function entirely in the cytoplasm, 17 in the mitochondria, and 3 in both. Based on their structure, they have been classified in two classes: class I and class II, which can further be divided into subclasses based on sequence similarity. aaRSs discriminate cognate tRNAs mainly by their anticodon loop, the residue located before the 3'-CCA named discriminator base (N73), and other structural elements that are distinctive for each tRNA (27). In mammals, aaRSs form a large cytosolic complex composed of nine aaRSs and three auxiliary proteins known as the Multi-aminoacyl-tRNA Synthetase Complex (MSC) (28).

Protein synthesis is performed in the ribosome, a complex catalytic machine composed of more than 50 different proteins (the ribosomal proteins) and several RNA molecules, the ribosomal RNAs (rRNAs). Ribosomes are composed of two subunits. The small subunit (40 S) provides the framework on which tRNAs can be accurately matched to the codons of the mRNA, while the large subunit (60 S) catalyzes the formation of the peptide bonds that link the amino acid together into a polypeptide chain. Ribosomes have one binding site for mRNA and three binding sites for tRNA:

- A-site (aminoacyl-tRNA), which binds to the aminoacyl-tRNA that holds the new amino acid to be added to the polypeptide chain.
- P-site (peptidyl-tRNA), which bids to the tRNA holding the growing polypeptide chain of amino acids.
- E-site (exit), the final transitory strep before the tRNA without its amino acid is released by the ribosome.

A tRNA molecule is held tightly at the A- and P-sites only if its anticodon forms base pairs with a complementary codon on the mRNA molecule that is threaded through the ribosome. Ribosomes are considered ribozymes because the three ribosomal sites and the catalytic site are mainly formed by rRNA, which is responsible for the peptide bond formation

To initiate translation, the small ribosomal subunit binds to the mRNA molecule at a start codon (AUG) that is recognized by a unique initiator tRNA molecule (Met-tRNA$_i$). The large ribosomal subunit binds to complete the ribosome and protein synthesis begins. Aminoacyl-tRNAs, each bearing a cognate amino acid, bind sequentially to the complementary codons in mRNA. Each amino acid is added to the C-terminal end of the growing polypeptide in four sequential steps: aminoacyl-tRNA binding, peptide bond formation, and two ribosome translocation steps (of the large and small subunits, respectively). As a result of the two translocation steps, the entire ribosome moves three nucleotides along the mRNA and is positioned to start the next cycle. Elongation factors use GTP hydrolysis to drive these reactions forward and to improve the accuracy of amino acid selection (the GTP form of EF1α delivers the aminoacyl-tRNA to the A-site of the ribosome, and EF1βγ catalyzes the exchange of GTP for bound GDP). The mRNA molecule progresses codon by codon through the ribosome in the 5' to 3' direction until it reaches one of the three stop codons. A release factor then binds to the ribosome (eRF1), terminating translation by triggering the hydrolysis of the ester bond in the peptidyl-tRNA and releasing the completed polypeptide (**Figure 8**).

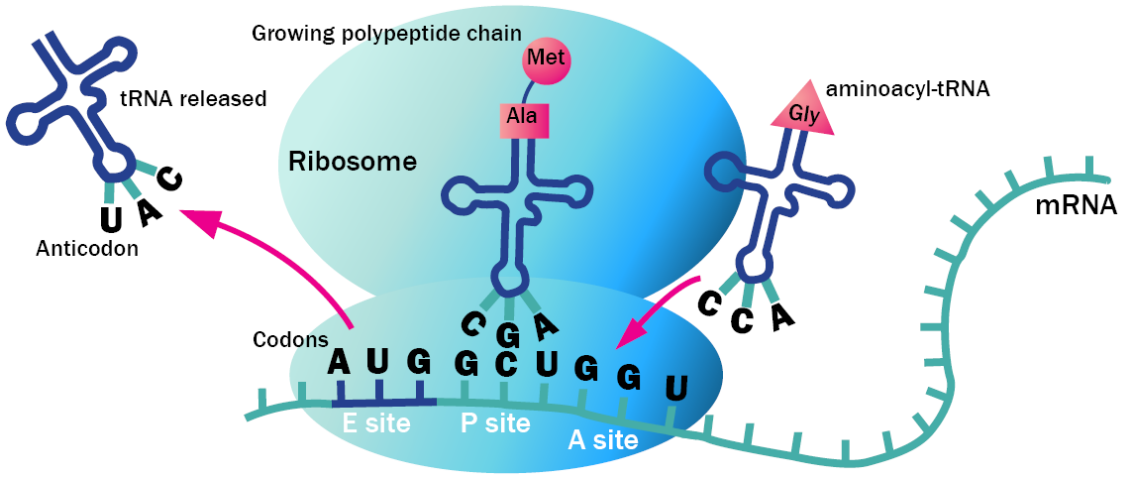


Figure 8. Translation of an mRNA molecule. Codons in the mRNA determine the specific amino acid to be added to the growing polypeptide chain. First, an aminoacyl-tRNA molecule binds to a vacant A-site on the ribosome so that the P-site and the A-site contain adjacent bound tRNAs. The carboxyl end of the polypeptide chain is then released from the tRNA at the P-site and joined to the free amino group of the amino acid linked to the tRNA at the A-site, forming a new peptide bond. The reaction is catalyzed by a peptidyl transferase contained in the large ribosomal subunit. Next, the large subunit translocates relative to the mRNA held by the small subunit. Finally, the small subunit translocates carrying its mRNA exactly three nucleotides through the ribosome. This results in the deacylated tRNA at the E-site, the peptidyl-tRNA at the P-site, and an empty A-site, ready for the next aminoacyl-tRNA molecule to bind.

The newly synthesized polypeptide chains must fold correctly into its three-dimensional conformation, bind any cofactors required for its function, be appropriately modified by protein-modifying enzymes (e.g. glycosylation, phosphorylation, acetylation, etc.), and assemble with other protein subunits. For some proteins, the folding begins as it emerges from a ribosome, starting from the N-terminal end (co-translational protein folding). Molecular chaperones such as Hsp60 and Hsp70 (Hsp stands for heat-shock protein, because their synthesis is highly increased upon exposure to high temperatures) can also guide the folding of polypeptide chains. These chaperones recognize exposed hydrophobic patches on proteins and serve to prevent the protein aggregation that would otherwise compete with the folding of newly synthetized proteins into their correct three-dimensional conformations. This protein folding process must also compete with an elaborate quality control mechanism that destroys proteins with abnormally exposed hydrophobic patches. In this case, ubiquitin is covalently added to a misfolded protein by a ubiquitin ligase, and the resulting polyubiquitin chain is recognized by the cap on a proteasome that moves the entire protein to the interior of the proteasome for proteolytic degradation (29).

Transfer RNAs

Transfer RNAs (tRNAs) act as adaptors to decode mRNAs into proteins. All tRNA molecules have common structural features that enable their interaction with ribosomes, mRNAs and protein factors that participate in translation. tRNAs are ubiquitously distributed and the most abundant of all small non-coding RNA molecules, constituting 4-10% of all cellular RNA (30). They are generally 70-90 nucleotides in length with about half of the nucleotides base paired. These auto-complementary regions are responsible for the cloverleaf secondary structure composed of three arms (T-arm, D-arm and anticodon arm) and the acceptor stem, which fold into the conserved L-shaped three-dimensional structure (31). Each arm has a stem with base paired nucleotides and a loop with non-base paired nucleotides at the end. Some tRNAs have an extra loop between the T-arm and the anticodon arm called variable loop. Each branch of the L is composed by the acceptor stem and the T-arm and the anticodon arm stacked over the D-arm, respectively. This architecture is maintained by a network of tertiary interactions mostly located in the 3D core of the tRNA, where the D- and T-loops are interdigitated, at the hinge of the L. In mature tRNAs, the conserved CCA terminal sequence at the 3'-end of the acceptor stem extends from one end of the L and contains the amino acid attachment site. The anticodon loop is at the other end of the L, making accessible the three bases that constitute the anticodon (positions 34, 35 and 36 of the tRNA). Thus, the anticodon can interact with the complementary codon on mRNA while the end that is linked to an amino acid is well positioned to participate in peptide bond formation. All tRNAs share common structural features to fit in the same ribosomal site during translation, still, they exhibit distinctive identity elements to selectively

interact with their biological partners, such as aaRSs and tRNA modification enzymes (32, 33) (**Figure 9**).

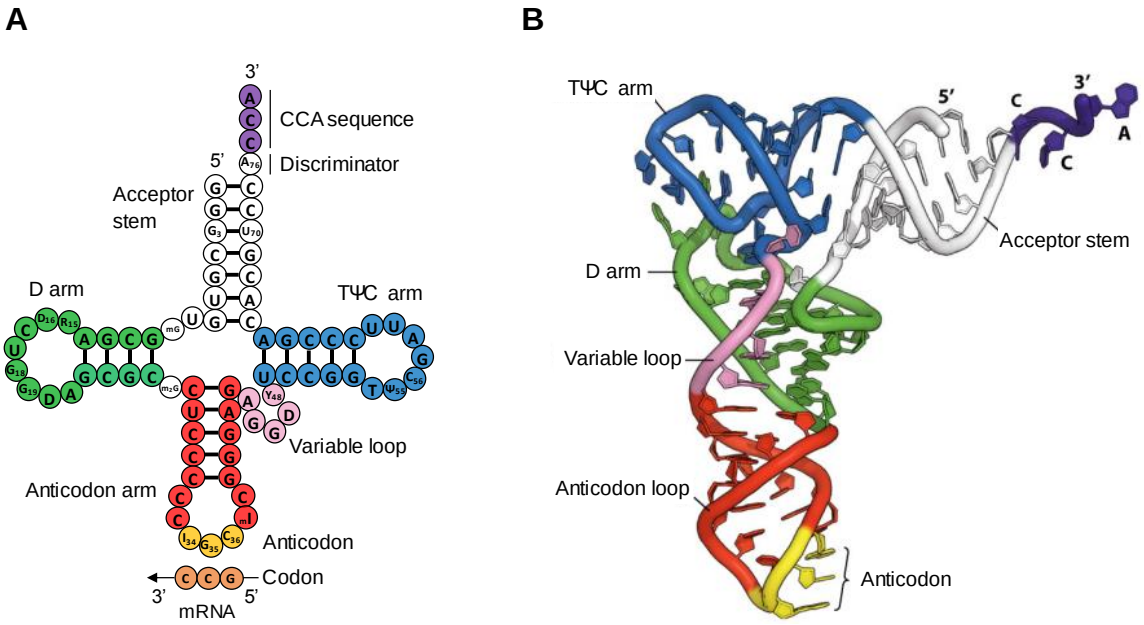


Figure 9. tRNA secondary (left) and tertiary (right) structure.

A The cloverleaf structure shows the complementary base-pairing that creates the double-helical regions of the molecule. The anticodon (yellow) is the sequence of three nucleotides that base-pairs with a codon in mRNA (orange). The amino acid matching the codon/anticodon pair is attached at the 3'-CCA of the tRNA (purple). Stems and loops constituting the arms of the tRNA are indicated: acceptor stem (white), D-arm (green), anticodon arm (red), variable loop (pink, not always present) and TΨC-arm (blue). tRNAs contain some unusual bases, which are produced by chemical modification after the tRNA has been synthesized. For example, some residues can be methylated (m); uridine can be modified to dihydrouridine (D), ribothymidine (T) or pseudouridine (Ψ); and adenosine to inosine (I). D- and T-arms are named for the presence of conserved D and T modifications in their loops, respectively (T-arm is also known as TΨC-arm due to the presence of T, Ψ and C residues).

B Representation of the L-shape of the tRNA molecule, based on x-ray diffraction analysis. The secondary structural elements of the tRNA are color coded as in **A**.

tRNAs are synthesized as precursors that are then processed by a sequence of maturation events including removal of the 5'-leader by RNase P, trimming of the 3'-trailer by RNase Z, splicing of introns by a nuclear tRNA splicing endonuclease (TSEN) complex, addition of the 3'-terminal CCA residues by the CCA-adding enzyme (also known as TRNT1), and covalent modification of multiple nucleotides (**Figure 10**). Only correctly processed tRNAs are exported from the nucleus through a nuclear receptor that serves as a checkpoint for sorting tRNAs with incorrectly processed termini (30).

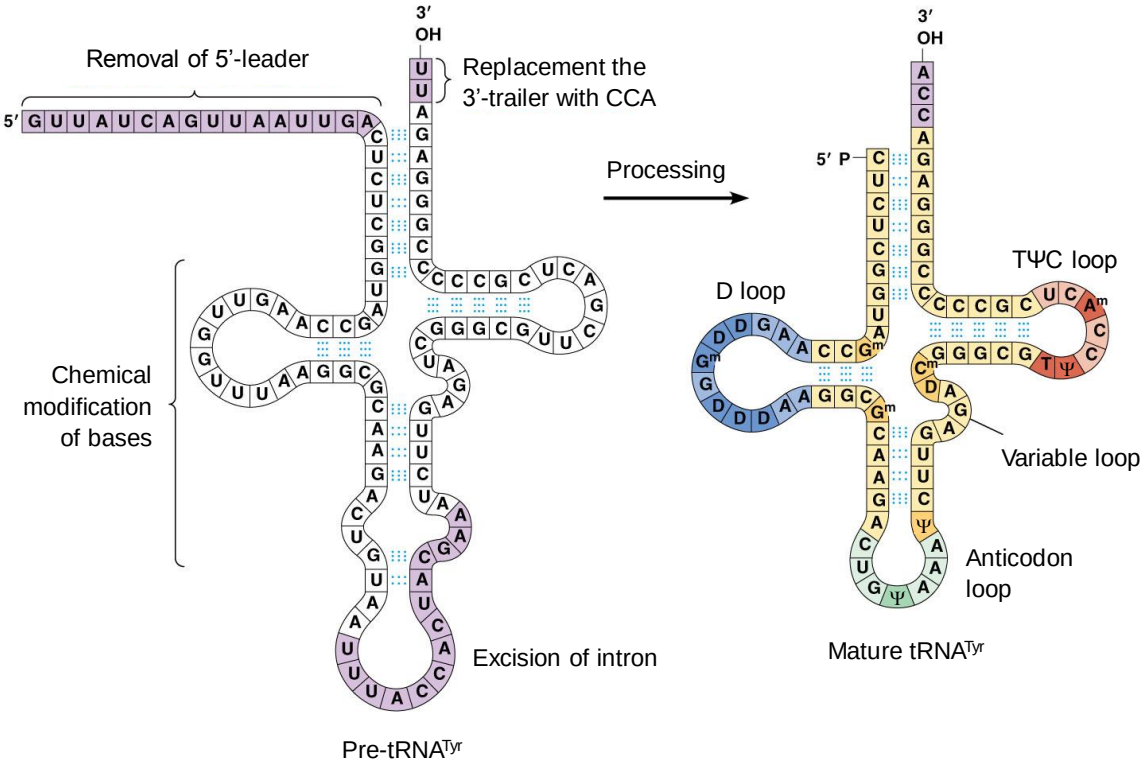


Figure 10. Processing of pre-tRNA^{Tyr}. The 14-nucleotide intron in the anticodon loop is removed by splicing (not all pre-tRNAs contain introns). The 5'-end leader sequence is cleaved by RNase P. U residues at the 3'-end are replaced with the CCA sequence. All cleaved sequences are shown in purple. Numerous bases chemically modified; D stands for dihydrouridine, Ψ for pseudouridine, and m for methyl group. Figure edited from (34).

tRNAs recognize specific codons in the mRNA through base pair complementarity. However, each codon in the mRNA does not have tRNA with a complementary anticodon, i.e. the number of tRNA molecules bearing different anticodons is always lower than the number of codons. To explain this, the tRNA wobble hypothesis proposes that the standard base pairing (Watson-Crick base pairing) may only occur in the first two positions of the codon, but there might be a higher pairing permissiveness between the third position of the codon and the first position of the tRNA anticodon (position 34 or wobble position) (6). Position 34 can wobble and pair with different nucleotides via non-Watson Crick interactions. U34 pairs with A and wobble pairs with G, and G34 pairs with C and wobble pairs with U. Whereas in standard base pairing the glycosidic bond always adopts the same position, in wobble pairing occupies an alternative position (6). Base modifications at position 34 can also be necessary for codon:anticodon wobbling. Common wobble modifications include methylation, hydroxylation or thiolation of U34, and deamination of A34 to form inosine (6, 8, 9). As a result, wobbling allows some tRNAs to decode different codons encoding the same amino acid, and some codons to be recognized by more than one

anticodon sequence. According to the wobble hypothesis, only 32 different tRNAs would be needed to translate all 64 codons (6).

The number of different tRNAs is highly variable between species, even among closely related species, and generally increases with the complexity of the organism. However, the number of tRNAs is always lower than 64. In humans, there are 49 different tRNAs (encoded by 513 nuclear tRNA genes, in addition to 22 mitochondrial tRNA genes) (35). Consequently, one tRNA can recognize more than one codon, and different tRNAs can code for the same amino acid. Different tRNA species carrying the same amino acid but with different anticodon sequences are called isoacceptors, whereas distinct tRNAs bearing the same amino acid and anticodon but with sequence variations in the tRNA body are called isodecoders. Isoacceptors and isodecoders are frequently encoded by a highly variable number of genes. tRNA gene copy number differs between species and even within the same species. In humans, tRNA gene copy number shows interindividual variability and can even differ among individuals who share the same ancestry (36, 37).

In less complex organisms (including unicellular eukaryotes), tRNA gene copy number correlates with tRNA abundance and it is often used as a proxy to predict cellular tRNA concentrations (38, 39). However, in higher eukaryotes the tissue-specific fluctuations in tRNA expression cannot be explained by tRNA gene copy number (40). Furthermore, in unicellular organisms tRNA abundance correlates with the codon usage of highly expressed genes (41). In higher eukaryotes this correlation is relatively poor, but it improves when tRNA modifications at the anticodon (deamination of A34 to inosine and methylation of U34) are considered (9). The expression of tRNAs changes depending on the cell type and state (42) and has a role in the regulation of the translation rate (43, 44). The tissue-specific expression of human tRNAs explains the codon bias of tissue-specific proteins (42), suggesting that tRNA expression is coordinated with the translational demand for each the cell type (40). Translation speed is precisely controlled through the selection of codons that are recognized by less abundant tRNAs at specific positions. The transitory pauses of ribosomes in the translation of these codons facilitate the proper co-translational folding and translocation. Consequently, synonymous substitutions can alter the kinetics of translation and affect protein structure and function (44).

In addition to their canonical role in translation, tRNAs also have key functions in a growing number of other non-translational activities, acting as signaling molecules in the regulation of numerous metabolic and cellular processes in both prokaryotes and eukaryotes. tRNAs have shown to regulate global gene expression in response to changes in amino acid pools in the cell. Under certain nutritional stresses, the aminoacylation levels of tRNAs change and the accumulated uncharged tRNAs participate in numerous biological pathways that regulated global gene expression levels, helping the organism to survive under adverse conditions (45). tRNAs contribute in energy metabolism and amino acid biosynthesis (46). $tRNA^{Glu}$ takes part in the first step of the heme and chlorophyll biosynthesis (47). aa-tRNAs

are used as amino acid donors for a variety of acceptor molecules, such as in antibiotic biogenesis (48). Specific cellular tRNAs or tRNA fragments are primers for reverse transcription of viruses (49, 50). aa-tRNAs can mediate specific degradation of proteins in the cytoplasm through the transfer of amino acids to the N-terminus of target peptides recognized by aa-tRNA transferases, which are then degraded by the cellular degradation machinery (N-recognins) (51, 52). $tRNA^{Arg}$ mediates amino-terminal arginylation, which is also a tag for degradation (53). tRNA genes act as insulators in the human genome, where they help to separate actively transcribed chromatin domains form silenced ones (54). tRNAs have also showed to inhibit apoptosis and promote cell survival by targeting cytochrome *c* (55, 56). Increased expression of tRNAs has been detected in a wide variety of cancerous cells, including ovarian and cervical cancer, carcinomas, and multiple myeloma cell lines, which might be related to regulation of cytochrome *c* (57-60). Finally, tRNAs can also serve as a source of tRNA derived fragments (tRFs), whose multiple regulatory roles in human cells are beginning to be discovered.

tRNA-derived fragments

tRNAs can be processed under certain conditions leading to the formation of tRNA-derived fragments (tRFs). tRFs are short non-coding RNAs generated by specific and regulated cleavage of tRNAs at specific sites. They are very heterogeneous in size (10-45 nucleotides), nucleotide composition and biogenesis. They can derive from precursor or mature tRNAs. The tRFs derived from pre-tRNAs contain sequences derived from the 5'-leader or the 3'-trailer. In contrast, if they derive from mature tRNAs, they will have the CCA at the 3'-end. Their abundance generally shows no correlation with the abundance of parent tRNAs, is highly variable and depends on the cell type and tissue, developmental stage, proliferative status, etc. (61). They also show interindividual variability, with fluctuations according to the gender, race, population origin, etc. (62) The functions of tRFs can be very diverse, playing important roles in cell signaling and regulating gene expression, including inhibition of protein synthesis (63). However, in many cases, their specific functions still remain unknown, and further research is needed to elucidate their molecular mechanisms and regulation.

tRFs can be classified in four structural types according to the cleavage site in the tRNA molecule: tRNA halves (including 5'-halves and 3'-halves), 5'-tRFs, 3'-tRFs, and internal-tRFs (i-tRFs). tRNA halves are produced after cleavage in the anticodon loop, 5'-tRFs and 3'-tRFs after cleavage in the D-loop and T-loop respectively, whereas i-tRNA, which comprise internal regions of tRNAs, are generated through a combination of cleavages in the anticodon loop and either D-loop or T-loop (62) (**Figure 11**).

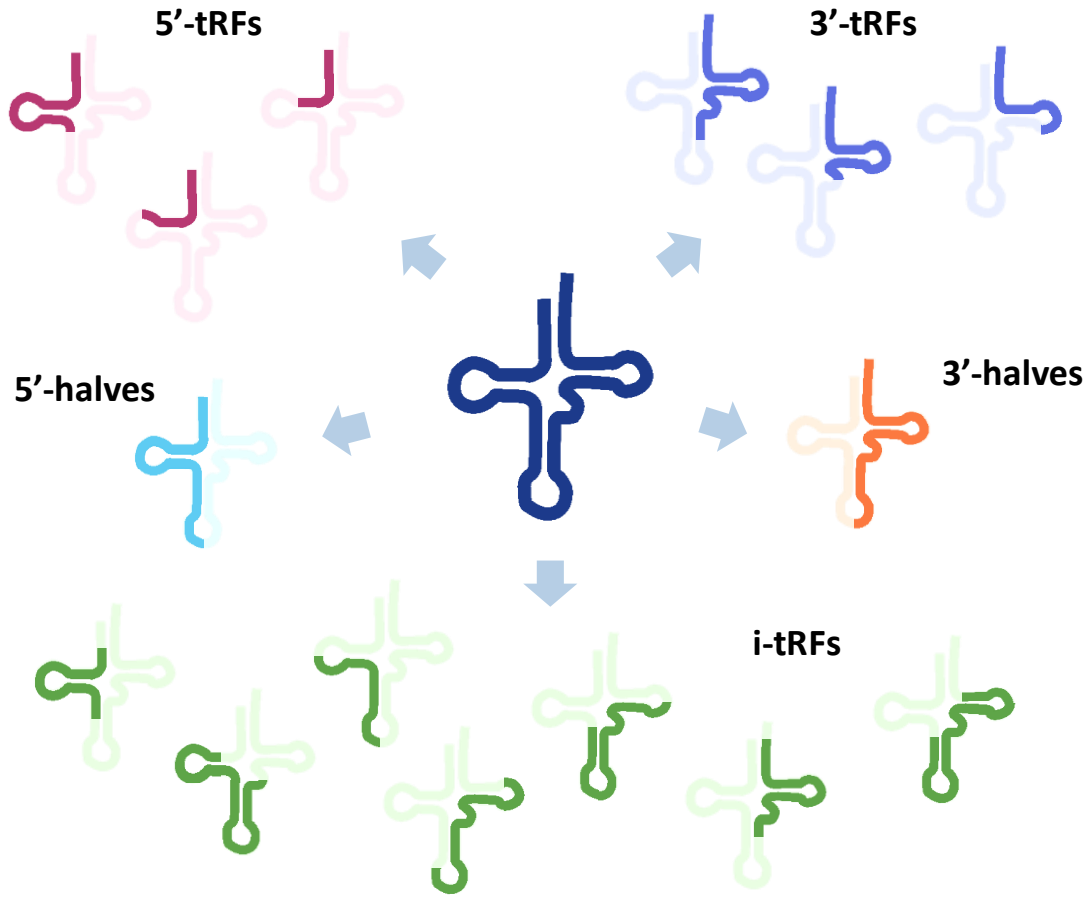


Figure 11. Structural types of tRNA fragments. Schematic representation of the five structural categories of tRNA fragments that are now known to arise from tRNAs, both mitochondrially and nuclearly encoded.

Whereas some tRFs are constitutive, others are produced under adverse conditions. In response to certain types of cellular stress (nutrient deprivation, oxidative and thermal stress, hypoxia, UV irradiation, etc.) angiogenin (ANG) cleaves mature cytoplasmic tRNAs in the anticodon loop, leading to the formation of tRNA halves (5'-halves of 30-35 nt, and 3'-halves of 40-50 nt), also known as tRNA-derived stress-induced RNAs (tiRNAs) (64, 65). In response to stress, ANG translocates from the nucleus to the cytoplasm and dissociates from its cytoplasmic inhibitor ribonuclease/angiogenin inhibitor 1 (RNH1, also known as RI or ANG inhibitor) and cleaves its tRNA targets. The primary targets for ANG are tRNAs with CA sequences in the anticodon loop (66). Post-transcriptional modifications play a role in the regulation of this tRNA cleavage by masking the nuclease recognition site (30). The generation of 5-mehtylcytidine (m^5C) modification at position 34 of tRNAs by NOP2/Sun RNA methyltransferase 2 (NSun2) has been shown to protect from ANG-mediated tRNA cleavage (67, 68). ANG cleavage has been associated to induction of stress granule (SG) formation (69) and inhibition of translation in mammalian cells (63). Interestingly, upon ANG cleavage, the levels of mature, full-length tRNAs do not significantly change (<5 % of

the tRNA pool is cleaved) (64, 70), meaning that it is unlikely to be a mechanism of translational repression through the depletion of the tRNA pool. This is consistent with the reduction in protein synthesis observed when isolated endogenous tRNA halves are transfected into cells (63). These effects have been associated to 5'-halves exclusively, and 5'-halves derived from $tRNA^{Ala}$ showed the strongest SG formation activity (69). The translational inhibition was associated to 5'-halves derived from both $tRNA^{Ala}$ and $tRNA^{Cys}$ (63), which seem to repress translation by interfering with the assembly of the cap binding complex eIF4F (a complex required for translation initiation). ANG-induced tiRNAs have also been shown to protect cells from apoptosis through binding Cyt *c* and competitively preventing its binding to the apoptotic protease-activating factor 1 (APAF1) protein (71), which is essential for the formation of the apoptosome and the downstream activation of cell death (72).

In addition to tRNA halves, tRFs have also been involved in global translational inhibition in human cells. 5'-tRFs are generated by specific cleavage by Dicer endoribonuclease after the conserved G18-G19 in the D-loop of tRNAs (73), whereas 3'-tRFs are formed by cleavage in the T-loop by either Dicer or other ribonucleases like ANG (74). A 5'-tRF derived from $tRNA^{Gln}$ decreased the expression of a reporter gene without sequence complementarity to that of the 5'-tRF, suggesting a non-specific translational repression (75). tRFs have been involved in various biological processes beyond translation and recent studies have shown an emerging role in gene expression regulation as miRNAs. Human 5'-tRFs and 3'-tRFs seem to repress translation via association with the components of the RNA interference machinery and have been reported to bind to Argonaute (AGO) proteins (74, 76-79).

Although many tRFs are formed in response to stress, other factors control their generation and some of them are constitutively produced. For example, tRNA halves are constitutive in hematopoietic tissues and bloodstream where they might have a role in immune signaling, and their levels change with age and calorie intake (80). Nevertheless, in many cases the mechanisms that control the tRNA cleavage and the molecular basis underlying the functions of tRFs in cell physiology remain unknown. Further research is needed to elucidate their specific rolls and dynamics in the cell.

tRFs and disease

Altered levels of tRFs have been associated to pathological conditions including cancer, infections, neurodegenerative diseases, or pathological stress injuries, representing potential biomarkers (61). However, it remains unclear whether these tRNA fragments contribute to the pathogenesis.

Consistent with the formation of tRFs upon stress, significantly high levels of tiRNAs were observed in tissue damage in animal models, and the amount of such tRFs correlated with the degree of damage. They reported that oxidative stress induces a change in tRNA conformation that promotes ANG-mediated cleavage and production of tiRNAs. Such conformational change caused by several injuries (renal ischemia/reperfusion injury, cisplatin-mediated nephrotoxicity, and brain and liver damage) was found to be preceding DNA damage and apoptosis *in vivo*. Quantification of the circulating levels of tiRNAs in patients was significantly higher upon tissue damage compared to healthy patients, thus, the levels of tiRNAs can be used as early biomarkers of oxidative stress and tissue damage (81).

Cancer cells demand elevated levels of protein synthesis for their increased proliferation and growth and, accordingly, they present higher levels of ribosomes, rRNAs and tRNAs (with selectivity for certain isoacceptors) compared to non-cancerous cells (82). Even though whether the rise in tRNA levels in cancer cells results in an increase in tRFs is unknown, high levels of tRFs have been associated to different types of cancer. For example, 3'U-tRFs derived from $tRNA^{Ser}$ are highly expressed in different cancer cell lines and are required for proliferation of prostate cancer cells (83), and accumulation of both 5'-tRFs and 3'CCA-tRFs has been observed in human B-cell lymphomas (74). In the same line, ANG is found to be overexpressed in almost all types of cancer (84, 85) and, consequently, it is possible that tiRNA levels also increase. Since ANG-induced tiRNAs cause translation re-programming and induce stress granule formation to survive under adverse conditions, they might be contributing to ANG-mediated angiogenesis and cell proliferation (61). On the other side, other tRFs have shown to decrease cancer growth and suppress metastasis through their binding to the RNA-binding protein Y box-binding protein 1 (YBX1 or YB-1). This binding displaces YB-1 from several of its substrate oncogenic transcripts, promoting its destabilization and down-regulation. It is the case of tRNA halves, which have been shown to play an important role in preventing breast cancer progression through the binding to YBX1 (86). Altogether, these results suggest that different types of tRFs can modulate diverse pathways and functions in the cell and play different roles in tumor growth and cancer progression.

Virus infections have also been linked to an increase of ANG-induced tRNA halves. Certain 5'-halves of mature tRNAs have been observed to be generated upon human respiratory syncytial virus (RSV) infection, and specifically 5'-halves derived from $tRNA^{Glu}_{CTC}$ showed to repress target mRNAs in the cytoplasm and promote RSV replication. Accordingly, the suppression of ANG by siRNA, blocked generation of 5'- $tRNA^{Glu}_{CTC}$ halves and significantly reduced the RSV infection (87). On this basis, inhibitors of ANG ribonuclease (RNase) activity might be potential treatments for RSV infections.

A number of neurological disorders have been linked to defects in tRNA metabolism and tRNA processing enzymes. ANG mutants with reduced RNase activity were associated to

amyotrophic lateral sclerosis (ALS) pathogenesis (88). Interestingly, subset of these ALS-associated mutants was also found in Parkinson's Disease, and all of them involved loss of RNase activity (89). Consistent with this data, recombinant ANG was observed to be neuroprotective for cultured motor neurons (90, 91), and administration of ANG to a mice model for ALS significantly increased life-span and motor function (92). The fact that most of ANG mutations associated to neurodegenerative diseases involve loss of RNase activity suggests an inability to form tiRNAs, which might contribute to neuron survival by modulating translation upon stress (63), inhibiting apoptosis (71), or inducing stress granule formation (69). In fact, it has been reported that transfection of 5'-tiDNA^{Ala} (the DNA equivalent of 5'-tiRNA^{Ala}) effectively rescues human motor neurons from stress-induced death (93). Not only this, but tiRNAs have also shown to modulate the activity of the enzyme superoxide dismutase 1 (SOD1), whose dysfunction has also been linked to ALS (94, 95).

Neurological disorders have also been associated to high levels of certain tRFs, and such increase has been associated to alterations in the tRNA modification pattern. Certain tRNA modifications have been shown to protect tRNAs from cleavage and the absence of such modifications results in increased tRNA processing and disease. Human cytosine-5 RNA methyltransferase NSUN2 methylates cytosine residues at positions 34 (in the anticodon loop), 48, 49 and 50 (at the intersection between the variable loop and the T-arm of tRNAs. The lack of NSUN2-mediated methylation increases ANG-mediated cleavage, resulting in the accumulation of 5'-halves, which repress translation, trigger stress response and cell death in neurons, leading to an increased sensitivity to various stress stimuli, decrease in neuron survival and impaired brain development. Consistent with this, some of the phenotypes in NSUN2 knockout were rescued by ANG inhibitors (68). Given that several mutations in NSUN2 have been associated to intellectual disability (96, 97) and to Dubowitz-like syndrome (98), ANG inhibitors might constitute potential drug candidates for patients with NSUN2 loss of function mutations. Even though the link between tRFs and neurodegenerative disorders needs to be further studied, the current data suggests a close relation between alterations in the levels of tRFs and the development of neurodegenerative diseases.

Overall, these discoveries shed light on the role of tRFs in the context of pathology, and may open a door for the development of new diagnostic and therapeutic tools based on tRNA fragments.

tRNA modifications

tRNAs need to be highly modified to be fully active. They are the RNA type that undergoes the most numerous and chemically diverse post-transcriptional modifications. There are almost 10 times more modified nucleotides in tRNA (with around 17 % of modified residues)

than in other abundant RNAs in the cell (the average of modified residues is 1-2 %). The type and number of modifications can vary between different tRNAs and even within similar isoacceptors. Over a hundred of chemically distinct modified nucleotides have been identified in RNAs (99). Such modifications increase the chemical diversity contained in the nucleotides of tRNAs and generally have two main purposes: they can be common in most tRNAs and help to maintain a constant general structure, or can provide specific features that discriminate any given tRNA from the rest and improve its interactions with one or more of its molecular partners (e.g. affecting the codon-anticodon interaction at the decoding center of the ribosome) (100).

The function of the modification generally depends on its position in the tRNA molecule (**Figure 12**). Modifications in the region of the anticodon loop can play a role in codon-anticodon recognition, preventing protein synthesis errors and affecting translational fidelity and efficiency. Modifications in the core of the tRNA are thought to predominantly affect tRNA structure and stability, rigidifying or increasing its flexibility. Modifications in the acceptor stem frequently serve as identity elements for the tRNA (e.g. important for the recognition by aaRS) (100). Nevertheless, modification of tRNAs is more complex, and the maintenance of optimal tRNA shape and functions often requires contributions from several modifications, suggesting the involvement of a network of modified nucleotides in coordinately promoting alternative tRNA conformations as needed.

tRNA modifications contribute to cellular homeostasis, and alterations in the modification levels are being increasingly linked to disease (101). Modification enzymes seem to play a role in the cellular response to stress, and the levels of certain modification enzymes have been shown to vary upon specific stresses (100). The methylation pattern of tRNAs can significantly change upon cell treatment with toxic chemical stimulants such as hydrogen peroxide (102). This is consistent with the observed phenotypes in different yeast Trm mutants under stress. Deletion of TRM9 (which catalyzes the last methylation step for mcm^5U or mcm^5s^2U at position 34 of different tRNAs) in *S. cerevisiae* results in increased sensitivity to stress (103), DNA damage (104) and to paromomycin, a translational inhibitor (105). Likewise, deletion of TRM4 (responsible for m^5C modification), confers hypersensitivity to oxidative stress (106). However, the specific roles that single modifications play in tRNA function and the dynamics of such modifications in tRNA populations still need to be further investigated. tRNAs lacking modifications are still generally functional albeit in some cases with reduced efficiency and/or fidelity. Only a few of the genes encoding tRNA modification enzymes are essential for viability. In most cases, deletion of single genes often causes little or no detectable growth defects, or only under particular stress conditions. Yet, the deletion of combinations of tRNA modification genes is generally not viable (99).

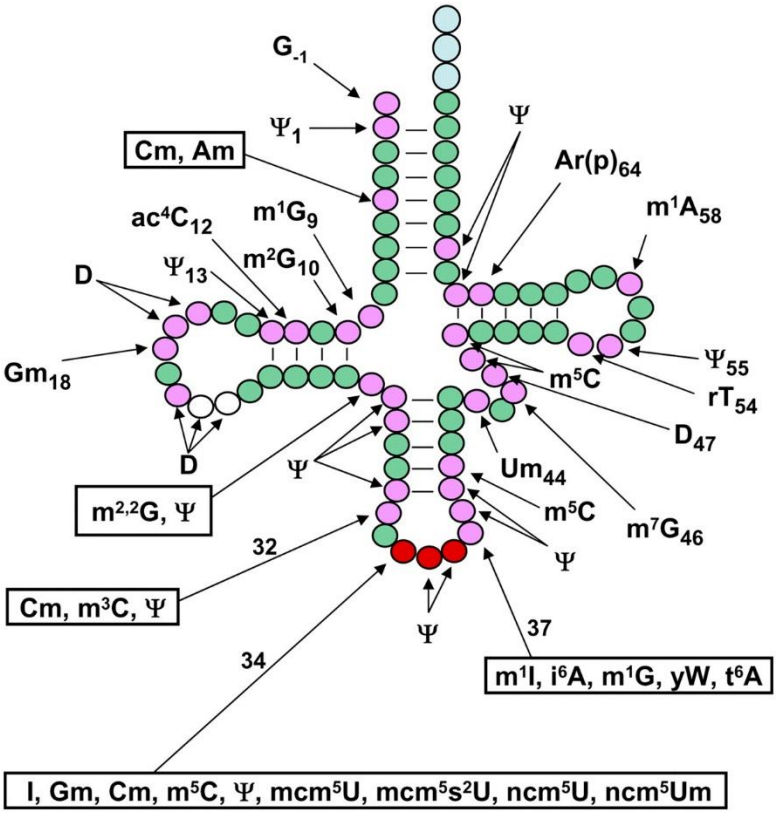


Figure 12. Post-transcriptional modifications in tRNA. Schematic representation of a tRNA secondary structure where circles correspond to nucleotides and lines to base pairs. Modifications found in cytoplasmic tRNA in *S. cerevisiae* are represented. Green circles indicate residues that are unmodified in all yeast tRNA species; pink circles, residues that are modified in some or all tRNA species; white circles, additional residues that are present in some, but not all, tRNAs and sometimes modified; red circles, anticodon residues, which are modified in some tRNAs; light blue circles, the CCA end. Conventional abbreviations are used; see the Modomics database (http://modomics.genesilico.pl). (Ψ) Pseudouridine; (Am) 2'-O-methyladenosine; (Cm) 2'-O-methylcytidine; (m^1G) 1-methylguanosine; (m^2G) 2-methylguanosine; (ac^4C) 4-acetylcytidine; (D) dihydrouridine; (Gm) 2'-O-methylguanosine; ($m^{2,2}G$) N2,N2-dimethylguanosine; (m^3C) 3-methylcytidine; (I) inosine; (m^5C) 5-methylcytidine; (mcm^5U) 5-methoxycarbonylmethyluridine; (mcm^5s^2U) 5-methoxycarbonylmethyl-2-thiouridine; (ncm^5U) 5-carbamoylmethyluridine; (ncm^5Um) 5-carbamoylmethyl-2'-O-methyluridine; (m^1I) 1-methylinosine; (i^6A) N6-isopentenyl adenosine; (yW) wybutosine; (t^6A) N6-threonylcarbamoyladenosine; (Um) 2'-O-methyluridine; (m^7G) 7-methylguanosine; (rT) ribothymidine; [Ar(p)] 2'-O-ribosyladenosine (phosphate). The pictured molecule starts at position −1 and is numbered consecutively from the next base (+1) to 76 (with the insertion of two residues (white circles)). Several tRNA species have a longer variable arm starting after residue 44, and some tRNAs have different numbers of residues in the D-loop and the variable arm, but the anticodon is always numbered residues 34, 35, and 36, and the CCA end is always numbered residues 74, 75, and 76. Figure from (8).

Two positions are found to be modified in almost all tRNAs and comprise the largest chemical diversity of modified nucleotides: position 37 and position 34. Modifications in position 37 help to maintain and open loop conformation, preventing base pairing with neighboring nucleotides on the other side of the loop (U33) and aiding the formation of the canonical U-turn structure important for the anticodon-codon pairing during decoding (107, 108). Depending on the base at position 37 (generally a purine), cells use different modification strategies to achieve the anticodon loop structure (109). In most cases, an encoded G37 is methylated at the base to form m^1G, which can be further modified to wybutosine and derivatives (110, 111). The lack of modifications at positon 37 leads to increased frameshifting and thus accuracy problems (112). Modifications in position 34 play a role in mRNA decoding, stabilize tRNA-mRNA interaction and fine-tune the wobble pairing with the third position of the mRNA codon and extend or restrict the number of codons that anticodons recognize. Thus, they ensure efficiency and fidelity during translation and, in some cases, prevent frameshifting (113).

Inosine 34 and ADAT

Inosine at the first position of the tRNA anticodon (position 34, wobble position) (I34) is one of the few essential post-transcriptional modifications on tRNAs. It is the result of the deamination of adenosine at this position (A34) (1-5) (**Figure 13**). I34 allows the recognition of three different nucleotides: C, U and A, at the third position of the codon, thus increasing the decoding capacity of tRNAs to more than one mRNA codon (6, 7) (**Figure 14**). This modifies the tRNA pool available for each codon and it has been proved to align the correlation between codon usage and tRNA gene copy number (9). It has also been suggested to improve fidelity and efficiency of translation (9, 10), especially for mRNAs enriched in codons translated by I34-tRNAs (11, 12).

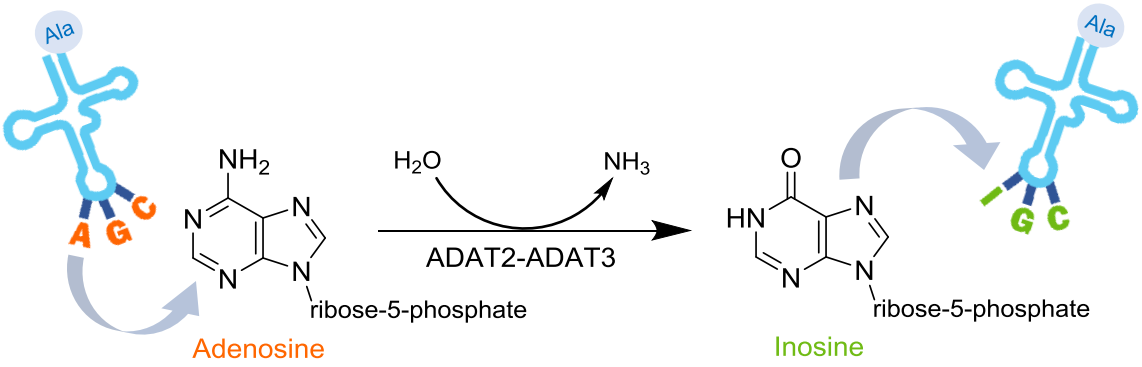


Figure 13. ADAT-mediated deamination of A34-tRNAs. The heterodimeric ADAT (ADAT2-ADAT3) catalyzes the hydrolytic deamination of adenosine to inosine at the first position of the anticodon (position 34, wobble position) of tRNAs. $tRNA^{Ala}_{AGC}$ is shown as an example. The chemical structures of adenosine and inosine are represented.

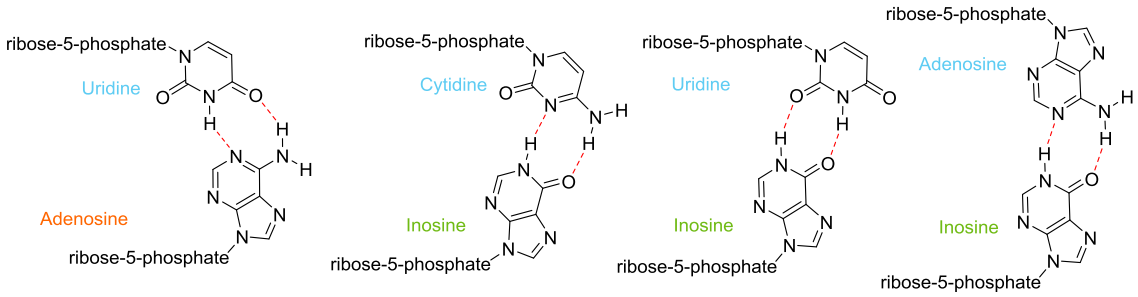


Figure 14. Inosine 34 increases the decoding capacity of tRNAs up to three different mRNA codons. Whereas A can in principle only pair with U, I can wobble pair with C, U, and A.

A-to-I conversion at position 34 of tRNAs is found in Bacteria and Eukarya, but not in Archaea (1, 4, 5, 7, 114). In eukaryotes, A34-to-I34 editing is catalyzed by the heterodimeric adenosine deaminase acting on transfer RNA (ADAT) (1, 3, 7), whereas in prokaryotes it is catalyzed by the homodimeric tRNA adenosine deaminase A (TadA) (2). Both ADAT subunits, namely ADAT2 and ADAT3 (1), as well as TadA, contain a deaminase domain that resembles to that of cytidine deaminase (CDA) superfamily and forms the catalytic center. The conserved deaminase motif signatures include a Zn-binding region having one histidine and two cysteine residues that coordinate a Zn ion (in addition to the water molecule responsible for the nucleophilic attack of the amino group); a glutamate residue that mediates the proton shuttling during catalysis; and a proline that binds the ammonia released with the deamination (1, 115) (**Figure 15**). Even though this core deaminase domain is present in both subunits of the heterodimeric enzyme, ADAT3 lacks the catalytic glutamate. On this basis, it has been suggested that ADAT2 is the catalytic subunit of the enzyme, while ADAT3 plays a role in tRNA substrate recognition. Nevertheless, both subunits have been shown to be essential for catalysis since the knock down of either subunit results in loss of activity (1). Welin and coworkers reported the crystal structure of a human ADAT2-ADAT2 homodimer (*Structural Genomics Consortium*; PDB ID: 3DH1), however, no structure for the heterodimeric form of ADAT has been resolved to date.

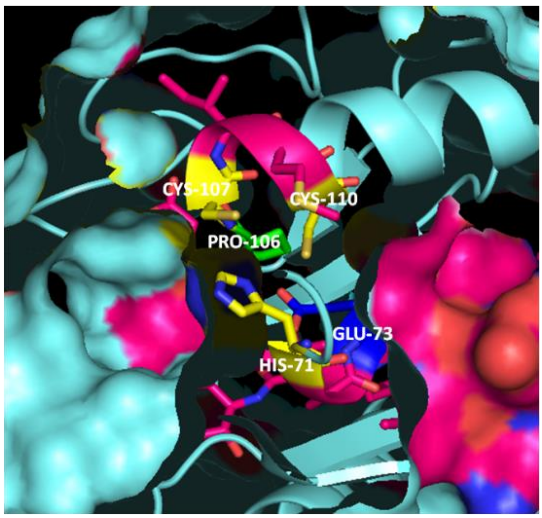


Figure 15. ADAT catalytic center (in ADAT2 subunit) contains conserved deaminase motif signatures. Conserved amino acids essential for the catalysis comprise His, Cys, Cys involved in Zn coordination (yellow), Glu involved in proton shuttling during catalysis (blue) and Pro that binds the NH_3 released (green).

The cellular localization of the heterodimeric ADAT has been reported to be mainly nuclear (3). Live cell imaging confocal microscopy experiments in human cells using overexpressed fluorescently tagged ADAT subunits showed that ADAT2 localizes mainly to nucleus and partially to the cytoplasm, while ADAT3 localizes mainly to cytoplasm and only partially to nucleus. However, upon co-expression of ADAT2 and ADAT3, both subunits co-localized in nucleus, suggesting that ADAT3 translocates to nucleus in an ADAT2-dependent manner (**Figure 16**). The different individual localization of the two subunits might be due to a mechanism of regulation of ADAT activity or independent functional roles of each subunit (3). Depletion of endogenous ADAT2 (the catalytic subunit of the heterodimeric ADAT) in cells showed to significantly reduce I34 levels in all possible substrate tRNAs. This suggests that, although a full knockout is not viable due to the essential nature of the enzyme, modulation of the levels of I34 can be tolerated (3).

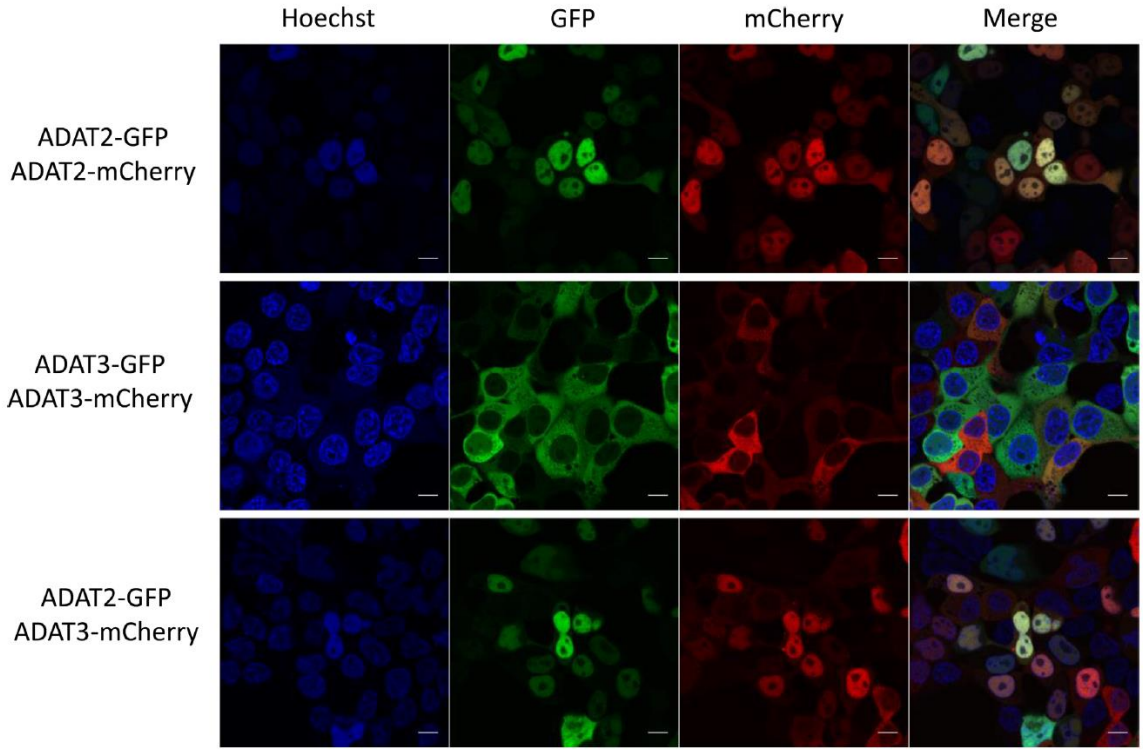


Figure 16. Live imaging confocal microscopy of HEK293T cells after co-expression of GFP- or mCherry-tagged HsADAT2 and HsADAT3 proteins. The cellular localization of ADAT2 is mainly nuclear, whereas for ADAT3 is mostly cytoplasmic. However, upon overexpression of both subunits, both ADAT2 and ADAT3 co-localize in the nucleus (3).

ADAT substrates

TadA, the bacterial homolog of ADAT, catalyzes the deamination of A34 almost exclusively in $tRNA^{Arg}_{ACG}$ (the only bacterial tRNA with A34) (2). In eukaryotes, TadA underwent a process of gene duplication and divergence to give rise to ADAT, whose substrate range is increased to include all A34-containing tRNAs acting in three-, four-, and six-box codon sets (with the exception of $tRNA^{Gly}$) (116). Eukaryotic heterodimeric ADAT enzymes are able to modify eight tRNAs (seven in yeast) (3, 17). In humans, eight tRNAs genetically encoded with A34 are modified to I34: $tRNA^{Thr}_{AGU}$, $tRNA^{Ala}_{AGC}$, $tRNA^{Pro}_{AGG}$, $tRNA^{Ser}_{AGA}$, $tRNA^{Leu}_{AAG}$, $tRNA^{Ile}_{AAU}$, $tRNA^{Val}_{AAC}$, and $tRNA^{Arg}_{ACG}$ (3, 117). This increase in the number of tRNA substrates in the eukaryotic enzyme advocates an evolution in substrate recognition. While TadA would predominantly recognize the local anticodon region, ADAT substrate recognition might require global features of the tRNA shared within all substrate tRNAs (118). It has been shown that the major determinants for TadA substrate recognition are the anticodon loop sequence and the anticodon stem loop structure. TadA is able to deaminate a $tRNA^{Arg}$ mini-substrate containing only the anticodon stem loop as efficiently as the full-length $tRNA^{Arg}$ (2). In contrast, the stem-loop RNA corresponding to the anticodon region of tRNA was not an active substrate for the *S. cerevisiae* ADAT heterodimer, suggesting that A34-to-I34 editing by the eukaryotic enzyme might require the full-length tRNA (118). In the same line, the activity of yeast ADAT depends on tRNA conformation and flexibility (especially within the anticodon stem and loop), and perturbations on its architecture resulted in loss of activity. For instance, disruption of critical base-pairs (i.e. residues G19:C56, A15:U48, A14:U8) or an extension in the anticodon loop in $tRNA^{Asp}_{AGC}$ resulted in complete loss of activity (119). Similarly, ADAT activity was abolished upon a single nucleotide deletion that severely altered the orientation of the variable arm on a synthetic pre-$tRNA^{Val}$ (120). Although the tRNA substrate recognition has been well studied for TadA, for the eukaryotic ADAT heterodimer it remains poorly understood. The lack of a crystal structure of the heterodimeric enzyme either alone or in complex with a substrate hampers the elucidation of the RNA-protein interaction.

ADAT has been shown to modify precursor tRNAs (pre-tRNAs), indicating that pre-tRNAs already have the necessary features to be substrates for ADAT and no other post-transcriptional modifications showed to be necessary for the deamination (119). This is consistent with the nuclear cellular localization of ADAT heterodimer (3), as pre-tRNAs are processed primarily in nucleus (121). Nevertheless, mature tRNAs also exhibit ADAT-mediated deamination, meaning that 5'-leader and 3'-trailer sequences (present in precursor tRNAs) are not an absolute requirement for ADAT (3). Consistently, cell studies of human ADAT reported that I34 is incorporated to human tRNAs at the precursor level and during maturation (3). However, the possibility that A34 deamination also occurs in mature tRNAs cannot be ruled out.

The divergence in substrate specificity between ADAT and TadA might be explained by their differences in structure (**Figure 17**). ADAT2 subunit displays a high identity with the bacterial counterpart TadA, mostly located in the N-terminal two-thirds of the protein, whereas ADAT3 shares sequence similarity at the C-terminal half (118). The additional amino acids in ADAT heterodimer that are missing in TadA might have a role in accommodating the different tRNAs, and could represent domains of interaction with parts of the tRNA molecule other than the anticodon arm. Computational docking models suggest that the additional residues in the C-terminus of ADAT2 and in the N-terminus of ADAT3 that are missing in TadA would be located on the same side to which the RNA substrate binds. Thus, they might be responsible for interactions with the full-length tRNA in regions other than the anticodon stem-loop (118). Consistent with this, the C-terminus of ADAT2 (which is composed of positively charged amino acids, mainly arginines and lysines) showed to be key for tRNA binding in *T. brucei* ADAT, and deletion of these residues leaded to defects in substrate binding and complete loss of enzymatic activity. Nevertheless, since ADAT2 alone exhibited a poor binding to tRNA, ADAT3 is likely to significantly contribute to substrate binding (122).

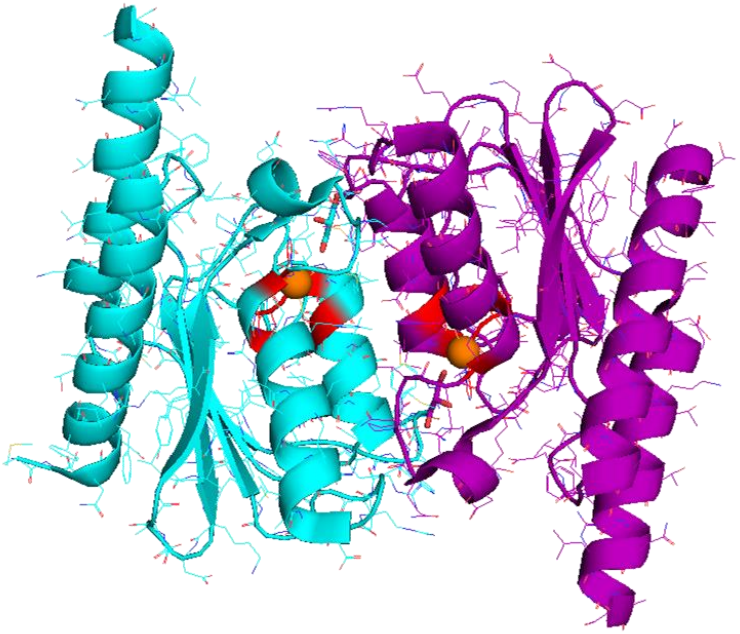


Figure 17. Crystal structure of TadA from *S. aureus* (PDB ID: 2B3J). The two subunits of the homodimeric enzyme are represented in blue and purple, respectively. Catalytic amino acids in each subunit are represented in red (His, Cys, Cys, Pro, and Glu), and Zn ions in orange.

The emergence of heterodimeric ADAT and the expansion of I34 in the eukaryotic tRNAome correlates with an increase in A34-tRNA genes coding for ADAT substrates (encoding for T, A, P, S, L, I, V and R amino acids). In bacterial genomes, tRNAs that code for these ADAT substrate tRNAs predominantly have G34 over A34, except for $tRNA^{Arg}_{ACG}$

(the only substrate of TadA). In contrast, in eukaryotic genomes, tRNA genes that encode for the same amino acids are biased towards A34-tRNAs rather than G34-tRNAs. The fact that genomes containing A34-tRNAs are devoid of tRNAs with G34, suggests that A-to-I conversion would allow the decoding of C-ended codons (which should be recognized by the absent G34-tRNAs). If I34 improves the codon-anticodon base pairing at the wobble position, A-to-I editing represents an advantage that might have been important for cell survival under selective conditions, and might have promoted the spreading of the A34 editing mechanism. On this basis, ADAT enzyme might have influenced in the evolution of eukaryotic genomes from a G34-tRNA enrichment to an A34 enrichment (9). Interestingly, $tRNA^{Gly}$ is an exception and did not follow this evolutionary trend. In contrast to ADAT substrates, $tRNA^{Gly}_{ACC}$ is absent in eukaryotic genomes, and $tRNA^{Gly}_{GCC}$ remains the most abundant $tRNA^{Gly}$ isoacceptor. Recent studies showed that the structural features of the $tRNA^{Gly}$ anticodon loop are incompatible with the presence of an adenosine at position 34, which would explain why A34-$tRNA^{Gly}$ could not evolve and become enriched in eukaryotes after the emergence of ADAT. On this basis, it has been proposed that these structural and functional constraints on tRNAs would have blocked evolution of new tRNA identities and, consequently, the incorporation of new amino acids into the genetic code (123).

ADAT and codon usage

I34 modification has been shown to balance codon usage and tRNA gene copy number in eukaryotes, and that highly translated genes tend to be enriched in codons recognized by tRNAs bearing I34 modification (9). Analysis of the codon usage in the human transcriptome revealed the presence of regions of coding sequences (CDSs) highly enriched in ADAT codons (> 83 %) (**Figure 18**). ADAT codons are those encoding for ADAT amino acids, which are amino acids recognized by ADAT substrate tRNAs (encoding T, A, P, S, L, I, V, R) (11). These regions, called ADAT stretches, tend to be enriched in codons recognized by ADAT-modified tRNAs, namely C-ended codons or ADAT-sensitive codons, over G-ended codons, which are not recognized by I34-tRNAs. This indicates that tRNAs with I34 modification are preferred for the translation of highly repetitive coding sequences, suggesting that I34 is a key modification for the synthesis of proteins with a highly skewed amino acid composition (124). Surprisingly, not all ADAT codons are equally enriched in ADAT stretches. Codons encoding for T, A, P, and S strongly predominate over codons encoding for L, I, V, and R (11). Within the list of genes with the highest enrichment in ADAT codons, mucins were identified as the most represented family (**Table 2**). This suggests a dependence on ADAT activity for mucins expression (data not published by Rafels-Ybern et al). However, the extent of this dependence and whether it might also affect other genes or specific cellular pathways remains unknown.

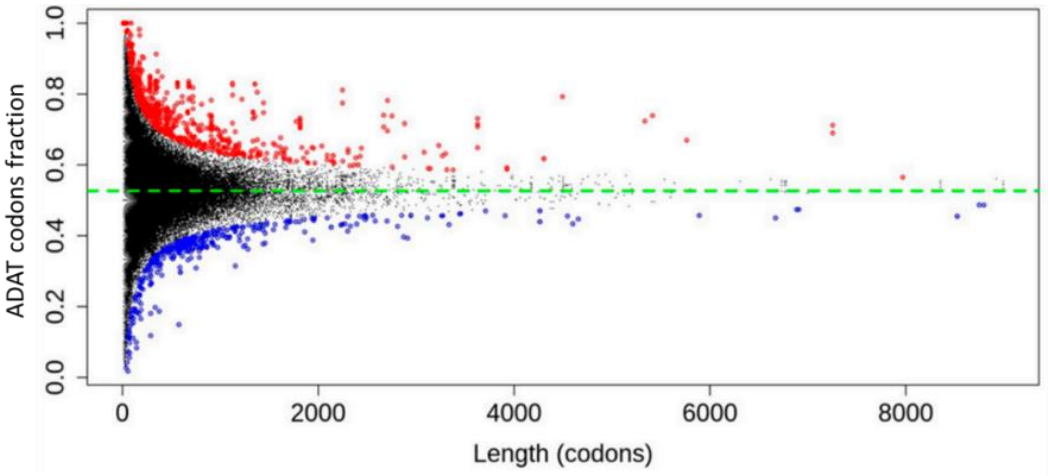


Figure 18. Distribution of ADAT codons in the human transcriptome. The dashed green line shows the mean of ADAT codons in the human transcriptome (0.527). Red and blue dots correspond to enriched and unenriched regions respectively with p-value <1e-11 based on a binomial distribution. Figure from (11).

Rank by size	Stretch size in codons	ADAT codon frequency	Gene Id
1	2247	0.81	**MUC17**
2	1376	0.81	SRRM2
3	1353	0.83	**MUC4**
4	720	0.82	**MUC5B**
5	453	0.80	**MUC16**
6	404	0.85	SRCAP
7	267	0.81	WNK1
8	215	0.81	FMN2
9	214	0.90	BCORL1
10	189	0.83	**MUC7**
11	176	0.97	PRRC2C
12	175	0.81	SH3BP1
13	174	0.89	SELV
14	169	0.81	AMOT
15	152	0.88	**MUC6**

Table 2. Human genes with the highest enrichment in ADAT codons. The top fifteen genes are listed with the respective stretch size and frequency in ADAT codons. Genes encoding for mucins (the most represented family) are in bold (data not published by Rafels-Ybern et al).

Comparison of ADAT stretches in Eukarya and Bacteria revealed that they are more abundant and have a longer size in the first group. In addition, Eukaryotic stretches have a higher relative concentration of ADAT-sensitive codons (namely higher ADAT enrichment) compared to the bacterial ones, and they show a positive correlation between ADAT enrichment and stretch size, whereas this correlation is not present in Bacteria (124) (**Figure 19**). This indicates a domain-specific bias in codon composition within ADAT stretches, suggesting a higher dependence on I34 modification in eukaryotes. This is in line with the different substrate-specificity between the bacterial and the eukaryotic enzyme. On this basis, I34 modification might have allowed ribosomal protein synthesis machineries to access new protein sequence spaces. Several examples of codon and amino acid compositions are known to impair ribosomal function and, in some cases, require additional cofactors to allow the ribosome to progress through these regions (125-128). Thus, ADAT enrichment might be an adaptation of the translational apparatus to allow the translation of proteins with high repetitions of specific amino acids (e.g. T, A, P, S).

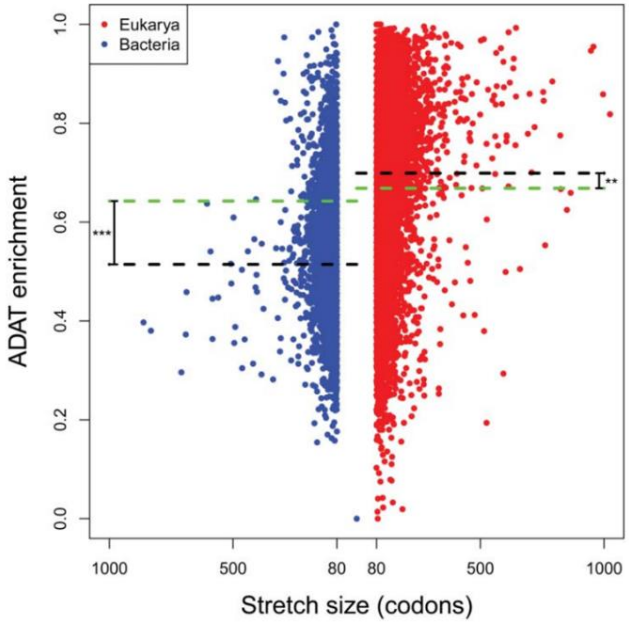


Figure 19. ADAT enrichment in proteins enriched in ADAT amino acids in Bacteria and Eukarya. Each dot (blue for Bacteria and red for Eukarya) represents a stretch of amino acids of length >80 with a composition of ADAT amino acids >83 %. Black dashed lines denote the mean enrichment in ADAT-sensitive codons in stretches of ADAT amino acids. Green dashed lines denote the mean enrichment in ADAT-sensitive codons for all the CDSs analyzed. There are significant differences between black and green dashed lines (p-value = 0.002 for Eukarya and p-value <2e-16 for Bacteria). Stretches longer than 1000 codons were not considered. Figure from (124).

V128M mutation in ADAT3 is linked to intellectual disability

Numerous mutations in tRNA modification enzymes have been linked to human diseases. In many cases, these mutations, instead of having generalized effects on protein synthesis, seem to affect organs unequally and, depending on the modification enzyme affecting the mutation, they can lead to a variety of pathologies (101). This is consistent with the tissue-specific tRNA expression and modification pattern, which favors the translation of mRNAs with a particular codon usage bias and allows the differential protein expression in each cell. In other words, the regulation of tRNA modification levels in different cell types and states could serve to modulate the synthesis of tissue-specific proteins and satisfy the translational demand in each case. Consequently, it is reasonable that alterations in the tRNA modification pattern affect each cell type differently.

Most of mutations in genes encoding for tRNA modification enzymes have been associated with neurological disorders, suggesting an important role for tRNA modifications in the nervous system. For instance, a homozygous frameshift mutation in the human tRNA methyltransferase 1 (TRMT1), in charge of the demethylation of G26 (m^2_2G) of tRNAs, results in cognitive disorders (129). Similarly, mutations in the gene encoding for a methyltransferase that catalyzes the formation of 5-methylcytosine (m^5G) at several positions of different tRNAs called NSUN2, have been related to autosomal-recessive intellectual disability (96, 97) and to Dubowitz-like syndrome (a rare disorder characterized by microcephaly, growth and mental retardation, eczema, and particular facial features) (98).

A single missense mutation in human ADAT3 has been linked to severe intellectual disability and strabismus (18). Further studies revealed that patients with the mutation also have microcephaly, brain abnormalities, growth failure, dysmorphic facial features, hypotonia, spasticity, epilepsy, behavioral problems, recurrent otitis media, and growth hormone deficiency (19) (**Figure 20**). The pathology has an autosomal recessive inheritance, meaning that only patients homozygous for the mutation exhibit ADAT3-related intellectual disability syndrome.

Figure 20. Facial features of individuals with ADAT3-related intellectual disability syndrome. Figure from (18).

The mutation in ADAT3 is a substitution of a conserved valine at position 128 for a methionine (V128M) (18). *In silico* analyses suggested that the mutation alters the 3D structure of ADAT3 (18) (**Figure 21**). Given the essential nature of ADAT (1-5, 114), V128M is likely to be a hypomorphic mutation that allows the enzyme to retain part of its activity, while null mutations would be incompatible with life. However, to what extent it affects ADAT function, leading to decreased I34 modification levels, remains unknown. As previously mentioned, I34 modification plays a key role in the efficiency and fidelity of mRNA decoding, nevertheless, the extent of such effect and whether it affects all genes equally or only specific cellular pathways is poorly understood. Thus, elucidating the molecular mechanisms of ADAT activity and how this affects protein translation are essential to reveal the molecular pathology underlying V128M ADAT3 mutation and develop new therapeutic strategies.

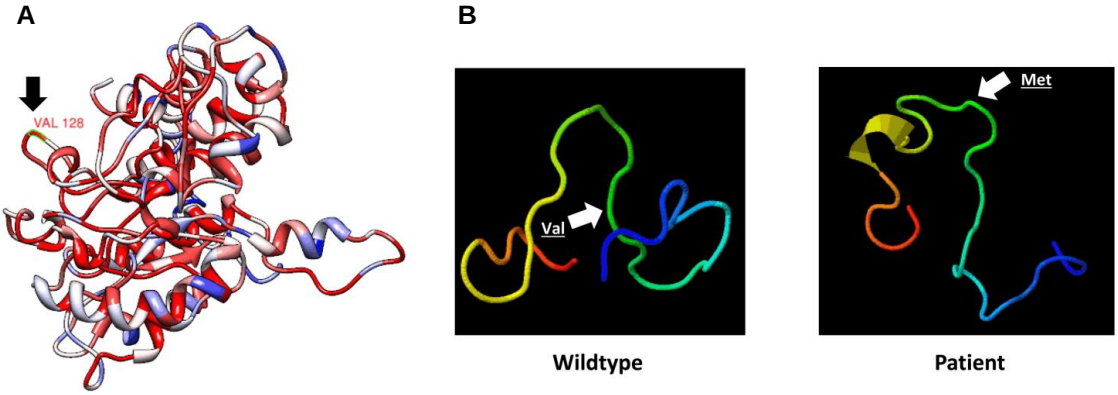


Figure 21. Computational analysis of the effect of V128M in ADAT3 structure. Figure from (18).

A Three-dimensional structure of the amino acid residues 2-350 of ADAT3. Conserved residues are represented in red, and semiconserved residues in blue. The conserved Val128 residue is indicated.

B Modelling of the local peptide region with Phyre2 (with 15 residues on either side of the affected amino acid) indicates that the valine forms one leg of a small hook jutting out from the protein (left), while substitution with methionine (right) generates a kink that prevents occurrence of this protrusion.

Apart from V128M mutation, another mutation has recently been described in ADAT3 subunit, consisting in a 8-bp duplication. Since ADAT3 gene has one coding exon, the frameshift mutation might produce a truncated protein. It has an autosomal-recessive mode of inheritance and shares many clinical features with V128M mutation (130). Another mutation has also been reported in the eukaryotic ADAT. In fission yeast *Schyzosaccharomyces pombe*, S257N mutation in TAD3 (the homolog of ADAT3) results in destabilization of the TAD2/TAD3 heterodimer, reduced deamination of $tRNA^{Ala}_{AGC}$, and cell cycle arrest in G_1 and G_2 (5). However, in the absence of a three-dimensional structure of ADAT, the molecular basis for the effects of these mutations remain unclear.

Additional inosines in RNA

Inosine is not only found at the first position of the anticodon, but also at positions 37 and 57 in the tRNA molecule. Inosine 57 is only present in Archaea as 1-mehtylinosine (m^1I57). However, its function and the enzyme catalyzing the modification are currently unknown (131). Inosine at position 37 is present only in eukaryotic $tRNA^{Ala}$, where it is also further modified to 1-methylinosine (m^1I37). I37 is catalyzed by the homodimeric enzyme ADAT1 (132), and the subsequent methylation is catalyzed by tRNA methyltransferase 5 (Trm5) (112). Unlike I34, I37 modification is not essential for cell viability (132, 133). Tad1 (ADAT1 homolog) mutant yeast strains did not show any particular slow growth phenotypes despite the absence of I37 and m^1I37 modifications (132). It has been proposed that m^1I37 might be playing a role in yeast during stress response, preventing frameshifts and improving translational fidelity (134). Consistent with this, in *A. thaliana*, Tad1 knockout lacking m^1I37 modification produced less biomass when grown under environmental stress, and respiration activity in the dark was reduced in Tad1 knockouts upon cold stress (133). Together with I34, I37 has been shown to be a major epitope for autoantibodies generated against the anticodon stem loop of $tRNA^{Ala}$ bearing the anticodon IGC in patients suffering from myositis, a chronic inflammatory muscle disorder (135).

In addition to tRNAs, I can also be present in eukaryotic mRNAs. Adenosine deaminases acting on RNA (ADARs) deaminate adenosine to inosine site-specifically in pre-mRNAs and non-specifically in long double-stranded RNA (dsRNA) (136, 137). Because I is read as G by the translational machinery (138), inosine in mRNAs can change codon specificity and thus the amino acid sequence of the encoded protein, in addition to alter splice sites during mRNA processing. Thus, RNA editing in mRNAs may result in multiple protein products with different biological functions from a single mRNA precursor, and can represent a mechanism to increase the genetic diversity in eukaryotes (136). A-to-I editing can also occur in non-coding RNAs and have regulatory functions, i.e. affecting the biogenesis, processing and target selection of siRNAs and miRNAs (139). Non-specific deamination has a role in defense against viruses that have dsRNA stages in their life cycle (140). ADARs are present only in metazoans, and three genes that encode for different ADAR proteins have been described in vertebrates. ADAR1 and ADAR2 are expressed in most tissues and catalyze non-specific deamination of dsRNA and site-specific deamination of mRNAs, respectively. ADAR3 (also called Red2), on the contrary, is only expressed in the central nervous system and its function is currently unknown, as it lacks deamination activity (137, 141).

ADARS have a conserved structure consisting in two or three dsRNA-binding domains (dsRBD) and a catalytic deaminase domain containing Zn-binding motifs (His, Cys, Cys residues coordinating Zn, and a Glu responsible for transition-state stabilization and proton shuttling). These catalytic residues are conserved among ADARs, ADATs, as well as cytidine deaminases acting on RNA (CDARs), including apolipoprotein B mRNA editing enzyme

catalytic peptide-like (APOBEC). Consistently, recent phylogenetic analysis concluded that all deaminases evolved from a common bacterial ancestor. Despite of the conservation of the catalytic amino acids and sharing mechanistic features, the spacing between Zn-coordinating residues is characteristically different among deaminases. In CDARs the second and third cysteines are separated by only two amino acids, whereas in ADARs at least 60 residues are inserted. As for ADATs, ADAT1 is organized like ADARs, whereas ADAT2 and ADAT3 share the configuration of CDARs (142). Based on this structural similarity and phylogenetic analysis, ADAT1 belongs to the clade containing ADARs, while ADAT2 and ADAT3 belong to the APOBEC clade (141). In this regard, it has been proposed that ADAT1 might be the ancestor of ADARs (ADARs have only been found in higher eukaryotes whereas ADAT1 is also present in yeast). On the other hand, ADAT2 and ADAT3 would have evolved from TadA, which in turn evolved directly from CDAs. Moreover, whereas ADAT2 and ADAT3 and CDARs are homologs, mononucleotide specific adenosine deaminases (ADA) (which deaminate free adenosine to inosine), despite being functionally related proteins, are not evolutionary related but rather the result of convergent evolution (142).

Analysis of post-transcriptional RNA modifications

tRNAs are among the most heavily modified RNA species. In recent years, post-transcriptional modifications on tRNAs have become a strong focus of research, as mounting evidence points to roles for tRNA modifications in human pathologies such as cancer (e.g., 5-methylcytosine), diabetes (e.g., 2-methylthio-N6-threonylcarbamoyladenosine), neurological disorders (e.g., 5-methoxycarbonylmethyl-2-thiouridine), and mitochondrial diseases (e.g., 5-taurinomethyluridine modifications on mitochondrial tRNAs), among others (101, 143). Unfortunately, the complex set of potential tRNA modifications makes their identification in any given tRNA extremely challenging.

Available methods for the analysis of posttranscriptional RNA modifications include (i) radiolabeling of RNA species followed by selective RNA digestion with specific RNases, and analysis of migration patterns by two-dimensional thin-layer chromatography (TLC) (144); (ii) the use of Liquid Chromatography coupled to Mass Spectrometry (LC−MS/MS) (102, 106); and (iii) methods that involve the reverse transcription (RT) of RNAs (145). Additional methods have also been developed to study specific modifications. These include the PHA6 assay in which isopentyl-N^6 -A37 can be detected on the basis of the differential hybridization of an oligonucleotide against the anticodon loop of tRNAs carrying the modification (146) and the acryloylaminophenylboronic (APB) acid-based boronic affinity electrophoresis for detecting queuosine modifications on tRNAs (147).

Methods based on radiolabeling or mass spectrometry are usually laborious and low-throughput and require specialized and expensive laboratory equipment and facilities. These methods may not be easily applied to studying modifications on specific tRNA residues or monitoring physiological levels of endogenous tRNA modifications *in vivo*. In addition, RNA samples need to be homogeneous; in some cases, large amounts of starting material are required, and a specific expertise about experimental optimization and data interpretation is often necessary. Therefore, there is still a need to develop novel techniques for tRNA modification detection that would overcome some of these limitations.

During RT, modified RNA bases either may be read by the reverse transcriptase as if unmodified, as a different base (resulting in a change in the resulting DNA sequence), or may block the advance of the reverse transcriptase, causing an arrest in the transcript. These possible outcomes of the RT reaction are dependent on the nature of the RNA modification (145). Restriction Fragment Length Polymorphism (RFLP) is a technique that studies the patterns of DNA cleavage after restriction enzyme treatments (148). It has been widely used in the past as a medium- to high-throughput technique to detect genetic polymorphisms, to define DNA marker loci, to perform genome mapping or gene tagging, and to construct genomic linkage relationships in human pedigrees (149). More recently, RFLP has been used to detect editing events on mRNAs such as adenosine-to-inosine (A-to-G mutation on the PCR product) and cytidine-to-uridine (C-to-T mutation on the PCR product) modifications (150-152), confirming that cDNA mutations caused by the reverse transcriptase can generate or abolish endonuclease restriction sites on PCR amplicons. We have applied this strategy for the detection of post-transcriptional tRNA modifications, and particularly to monitor A34-to-I34 conversion catalyzed by ADAT.

Objectives

General objective

The main goal of this project is to characterize the human ADAT enzyme in terms of kinetics, substrate recognition, regulation, structure, and to develop small molecules that modulate the enzyme activity as new potential therapeutic agents.

Specific objectives

- The development of an *in vitro* activity assay to monitor ADAT-mediated deamination.
- The characterization of the structural requirements of the human ADAT substrates.
- The analysis of the interaction between ADAT and its substrates to detect potential variations between the ancestral substrate $tRNA^{Arg}$ and the emergent tRNA substrates and elucidate how the substrate recognition would have evolved from TadA.
- The study of the mechanisms of regulation of ADAT activity.
- The evaluation of ADAT substrate binding and kinetics, and the effect of a mutation linked to intellectual disability and strabismus (V128M).
- The elucidation of ADAT quaternary structure, and the effect of a mutation linked to intellectual disability and strabismus (V128M).
- The discovery of small molecules targeting the human ADAT with the ability to selectively modulate the enzyme activity as new potential therapeutic agents.

Methods

Expression and purification of human ADAT wt and ADAT V128M

The cDNA sequences of ADAT2 and ADAT3 (NM_182503.2 and NM_138422.3, respectively) were codon optimized for expression in *E. coli* (Thermo Fisher Scientific) (**Supplementary Figure S1**). Note that two different transcript isoforms of ADAT3 are currently annotated in NCBI (NM_138422.3 and NM_001329533.1). We used the short ADAT3 isoform given that the resulting protein was previously validated for *in vitro* activity assays (123).

ADAT2 and ADAT3 were cloned into the bicistronic vector pOPINFS (Oxford Protein Production Facility) using the In-Fusion cloning system (Clontech) with specific primers (**Table 3**). The final construct encodes an N-terminal His_6-tagged ADAT2 and a C-terminal StrepII-tagged ADAT3 separated by an IRES. ADAT V128M was generated by QuikChange site-directed mutagenesis (Agilent Technologies) using primers specified in **Table 3**. Sequences for both constructs were confirmed by Sanger sequencing.

BL21(DE3)pLysS chemically competent cells (Thermo Fisher Scientific) transformed with pOPINFS-His_6-ADAT2-ADAT3-Strep (ADAT3 wt) or pOPINFS-His_6-ADAT2-ADAT3-V128M-Strep (ADAT3 mutant) were grown overnight in LB medium with 100 μg/mL ampicillin and 50 μg/mL chloramphenicol. The starting culture was diluted 1:20 in LB medium containing 100 μg/mL ampicillin, and grown at 37 ºC with shaking to an OD_{600} between 0.6 and 0.8. Protein expression was then induced with 0.5 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) at 37 ºC for 3 hours with shaking. Cells were harvested by centrifugation at 4,000 *g* at 4 ºC for 20 minutes; the cell pellet was washed once with phosphate-buffered saline (PBS) solution (40 mL of PBS per 1L of culture; centrifuged at 4,000 *g* at 4 ºC for 20 minutes) and stored at -80 ºC.

ADAT heterodimers were purified by Strep-tag affinity chromatography (ADAT3 is fused to Strep-tag). All procedures were done at 4 ºC in buffers containing a protease inhibitor cocktail (Roche). Cell pellets (stored at -80 ºC) were resuspended in StrepTrap binding buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 8), and 5 MU of bovine pancreas DNase I was added (Calbiochem). Cells were lysed with a cell disruptor (Constant Systems) at 25,000 psi twice. The cell lysate was clarified by centrifugation at 70,000 *g* for 1 hour, and the supernatant was filtered with a 0.45 μm syringe filter (Frisenette). The clarified lysate was then loaded onto a 5 mL StrepTrap HP column (GE Healthcare Life Sciences). The column was washed with 7 column volumes of binding buffer, and the bound protein was eluted with StrepTrap elution buffer (100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 5 mM *d*-Desthiobiotin, pH 8). Protein elution was monitored by measuring absorbance at 215, 260 and 280 nm. SDS-PAGE and Western blot (WB) analyses were used to confirm that ADAT2 and ADAT3 (both wt and mutant) co-purified at an equimolar ratio. The purified heterodimeric ADAT was stored at -80 ºC in elution buffer with 10 % glycerol. Protein

concentration was determined by measuring optical absorbance at 280 nm with a NanoDrop ND-1000.

Expression and purification of human ADAT2

For heterologous expression of human ADAT2, the sequence of ADAT2 subunit was re-cloned from the vector containing the WT enzyme (pOPINFS-His-ADAT2-ADAT3-Strep) into pETDuet vector using the In-Fusion cloning system (Clontech), and specific primers (**Table 3**), following the manufacturer's protocol. The sequence of the construct was confirmed by Sanger sequencing.

E. coli Rosetta2 (DE3) pLysS chemically competent cells (Novagen; Cat Number 71403-3) transformed with pETDuet-His-ADAT2 were grown overnight in LB medium with 100 μg/mL ampicillin and 50 μg/mL chloramphenicol. The starting culture was diluted 1:20 in OvernightExpress Instant LB Autoinduction Medium (Novagen; Cat Number 71757-3) containing 100 μg/mL ampicillin. Cells were incubated with shaking for 24 hours: first for 3 hours at 37 ºC, and then at 25 ºC (21 hours). Cells were harvested and stored as described for heterodimeric ADAT.

ADAT2 was purified by His-tag affinity chromatography (ADAT2 is fused to His-tag). All procedures were done at 4 ºC in buffers containing a protease inhibitor cocktail (Roche). Cell pellets (stored at -80 ºC) were resuspended in HisTrap binding buffer (10 mM sodium phosphate, 300 mM NaCl, 0.2% Tween 20, 2 mM DTT, 50 mM imidazole, pH 7.4), and 5 MU of bovine pancreas DNase I was added (Calbiochem). Cell lysis and clarification of the lysate were done as described for heterodimeric ADAT. The clarified lysate was then loaded onto a 5 mL HisTrap HP column (GE Healthcare Life Sciences). The column was washed with 80 column volumes of binding buffer, and the bound protein was eluted with HisTrap elution buffer (10 mM sodium phosphate, 300 mM NaCl, 0.2% Tween 20, 2 mM DTT, 500 mM imidazole, pH 7.4). Protein elution was controlled by measuring UV absorbance at 215, 260 and 280 nm. A 5 mL HiTrap Desalting column (GE Healthcare Life Sciences) was used to exchange HisTrap elution buffer for a buffer containing 50 mM Hepes-KOH, 100 mM KCl, 1 mM MgCl2, 0.1 mM EDTA, 2 mM DTT, pH 8. SDS-PAGE and WB analyses were used to confirm that ADAT2 was purified. The purified ADAT2 was stored at -80 ºC in elution buffer with 10% glycerol. Protein concentration was determined by measuring the absorbance at 280 nm with a NanoDrop ND-1000.

In vitro transcription of tRNAs and tRFs

The full-length variants of $tRNA^{Ala}_{AGC}$ and $tRNA^{Arg}_{ACG}$ were synthesized as previously described, using as a template for the *in vitro* transcription pUC19 plasmid cloned with the sequence to be transcribed (3) (**Tables 3 and 4**). For both tRNAs, six oligonucleotides

(**Table 3**) annealed and cloned into pUC19 vector and digested with BstNI (for $tRNA^{Ala}_{AGC}$) or NspI (for $tRNA^{Arg}_{ACG}$) restriction enzymes (NEB) were used as a template. The shorter tRNA variants and the chimeric tRNAs used in the study were *in vitro* transcribed using synthetic DNA templates with the T7 promoter. These DNA templates are partially single stranded and only the promoter region +1 is double stranded. Fifteen nucleotides were added upstream from the promoter to increase the yield of the transcription. Whereas the top strand contains the promoter region (plus the 15 nt upstream) but not the template (the same sequence was used for all transcripts), bottom strands contain both the promoter and template sequences (they are transcript-specific) (**Tables 3 and 4**) (153). All tRNAs were *in vitro* transcribed using T7 RNA polymerase and purified according to standard protocols.

The shortest tRNA fragments and also the human tRFs, whose *in vitro* transcription had very low yields (acceptor stem and the T-arm plus the anticodon arm derived from $tRNA^{Ala}_{AGC}$, the anticodon stem loop of both $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$), were chemically synthesized (Sigma-Aldrich). The RNA sequences of all tRNAs and fragments are listed in **Table 4**.

Synthesis of the emissive tRNAs containing thA, ^{tz}A, thG and ^{tz}G

Emissive nucleoside analogs (thA, ^{tz}A, thG and ^{tz}G) were synthesized by Tor laboratory at UCSD, as previously reported (16, 154). Two main strategies were used for the incorporation of the emissive nucleotides in the $tRNA^{Ala}_{AGC}$ molecule. The first strategy consisted in the division of the tRNA molecule in three fragments: two-non-emissive fragments (5'-end and 3'-end fragments), and one emissive fragment containing position 34 (where we want to incorporate thA). The three fragments were separately *in vitro* transcribed using three different double-stranded synthetic DNA templates with the T7 promoter as described before. Since the 3'-end fragment starts with A, we used a class II T7 promoter instead of the regular T7 promoter to increase the transcription yield (**Table 3**). The two non-emissive transcripts were transcribed using regular nucleotides triphosphate (NTPs) according to standard protocols. The emissive fragment containing position 34 was transcribed under the same conditions, but with 4 mM thATP instead of regular ATP. The three fragments were then mixed at equimolar concentrations together with of a complementary DNA oligonucleotide (20 μM of each oligonucleotide) (**Table 3**), annealed by heating 3 minutes at 90ºC and cooling down to room temperature, and ligated using 10 units of T4 RNA ligase 2 (New England Biolabs) in 50 mM Tris-HCl, 2 mM $MgCl_2$, 1 mM DTT, 400 μM ATP, pH 7.5, in a final volume of 20 μL for 2 hours at 37ºC.

The second strategy consisted in the division of the tRNA molecule in two fragments: an acceptor fragment and a donor fragment. The two fragments were separately *in vitro* transcribed using two different synthetic DNA templates with the T7 promoter as described before (**Table 3**). The non-emissive acceptor fragment was transcribed using regular NTPs according to standard protocols. The donor fragment was transcribed under the same

conditions but with a 10 mM excess of thA (or any other emissive nucleoside: ^{tz}A, thG or ^{tz}G). The donor fragment (6 μM) was then phosphorylated using 10 units of T4 polynucleotide kinase (T4PNK) (Takara) in 70 mM Tris-HCl, 10 mM $MgCl_2$, 5 mM DTT, 1 mM ATP, pH 7.6, in a final volume of 30 μL for 1 hour at 37ºC. The phosphorylated RNA was then purified using the miTotal RNA Extraction Miniprep System (Viogene) kit according to the manufacturer's protocol. 30 μM of phosphorylated donor fragment were added 300 μM of acceptor fragment and 100 μM of DNA splint (**Table 3**), annealed by heating 3 minutes at 90ºC and cooling down to room temperature, and ligated using 10 units of T4 DNA ligase 2 (New England Biolabs) in 50 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM DTT, 500 μM ATP, 3% PEG 4000, pH 7.8, in a final volume of 20 μL for 2 hours at 37ºC.

Ligated tRNAs were separated from the DNA splint and potential unligated RNA fragments by 8 M urea 10% PAGE and visualized at 254 nm, where nucleic acids are detected, and 312 nm, where fluorescent molecules absorb. The band corresponding to ligated emissive tRNA was purified from the gel according to standard protocols.

In vitro deamination assays and evaluation of inosine formation

Prior to all deamination reactions, all tRNAs (including tRNA fragments) were folded by heating to 95ºC for 3 minutes and cooling down to room temperature over 30 minutes (fast folding). The slow folding method used in **Figure 33** consisted in heating the tRNA in a boiling water bath and cooling to room temperature for 2 hours.

Liquid Chromatography-tandem Mass Spectrometry (LC-MS/MS) method

All deamination reactions were done using 50 μM of *in vitro* transcribed tRNA substrate in deamination buffer (50 mM Hepes-KOH, 100 mM KCl, 1 mM $MgCl_2$, 0.1 mM EDTA, pH 8) and 800 nM of purified human ADAT enzyme in a final volume of 30 μL, and incubated at 37ºC during the indicated time (1 hour or overnight). The RNA products were purified using the miTotal RNA Extraction Miniprep System (Viogene) kit according to the manufacturer's protocol. The tRNA was eluted from the purification column with 50 μL of nuclease-free water, and quantified by measuring absorbance at 260 nm with a NanoDrop ND-1000. tRNAs were then digested into 5'-NMPs with 100 units of Nuclease P1 from *Penicillium citrinum* (Sigma-Aldrich) in 30 mM ammonium acetate and 0.1 mM zinc chloride buffer, in a final volume of 60 μL at 50 ºC for 2 hours. The 5'-NMPs were extracted using centrifugal filters with 10 kDa MWCO (Merck Millipore) following the manufacturer's protocol. IMP levels were then evaluated by LC-MS/MS. Sample was loaded to a 2.1 x 150 mm, 5μm, XTerra MS C18 2.1 x 150 mm; 3.5 μm (Waters Corporation) column at a flow rate of 100 μL/min using a Quaternary pump, Finnigan, Mod. Surveyor MS (Thermo Electron Corporation) and Finnigan, Mod. Micro AS (Thermo Electron Corporation) as autosampler. Compounds were separated with a 10-min run by isocratic elution at 0 % B,

followed by linear gradients from 0 % to 75 % B in 1 min and isocratic elution at 75 % B in 3 min, other linear gradients from 75 % to 0 % B in 1 min and stabilization to initial conditions (A= H2O + 0.1 % FA, B= ACN + 0.1 % FA). The column outlet was directly connected to ESI source fitted on an LTQ-FT Ultra mass spectrometer (Thermo). The mass spectrometer was operated in in FT wide SIM acquisition mode (m/z 310-370). MS scans were acquired in the FT with the resolution (defined at 400 m/z) set to 100,000. Spray voltage in the ESI source voltage was set to 3.7 kV. Capillary voltage and tube lens on the LTQ-FT were tuned to -48 V and -94 V, respectively. The spectrometer was used in negative polarity mode. At least one blank run before of analysis was performed in order to ensure the absence of cross contamination from previous samples.

Restriction Fragment Length Polymorphism (RFLP) method

All deamination reactions were done using 3 μM of *in vitro* transcribed tRNA substrate in deamination buffer and the indicated concentration of affinity-purified human ADAT enzyme (15 nM ADAT was used in all assays with tRNA fragments) in a final volume of 30 μL, and incubated at 37°C for the indicated time. The RNA products were purified using the miTotal RNA Extraction Miniprep System (Viogene) kit according to the manufacturer's protocol. The tRNA was eluted from the purification column with 30 μL of nuclease-free water, and quantified by measuring absorbance at 260 nm with a NanoDrop ND-1000. RNA was reverse transcribed with the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, catalog no. 4368814). RT of tRNA species was performed using 100 ng of *in vitro* transcribed tRNA as a RT template in a final reaction volume of 20 μL. For experiments depicted in **Figure 25**, fully modified and fully unmodified *in vitro* transcribed $tRNA^{Ala}_{AGC}$ were mixed in the indicated proportions for a total amount of 100 ng of RNA (e.g., 80 (I34):20 (A34) proportions correspond to 80 ng of in *in vitro* transcribed $tRNA^{Ala}_{IGC}$ mixed with 20 ng of in *in vitro* transcribed $tRNA^{Ala}_{AGC}$, prior to RT reaction). 10 pmol of specific RT primers (**Table 3**) were used for each tRNA species. The RT cycle was as follows: 10 min at 25°C, 120 min at 37°C, and 5 min at 85°C. 2 μL of *in vitro* transcribed tRNA-derived cDNA were used for a standard 20 μL PCR. Forward and reverse primers used for PCR amplification of each target are detailed in **Table 3**. The PCR cycle was as follows: 3 min at 94°C (1 cycle); 45 s at 94°C, 30 s at the target-specific annealing temperature (60°C, for both $tRNA^{Ala}_{AGC}$ and $tRNA^{Arg}_{ACG}$), and 30 s at 72°C (35 cycles); and 10 min at 72°C (one cycle). When required, PCR amplicons were purified using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel, catalog no. 740609.250) and sequenced using their respective forward and reverse PCR primers as sequencing primers (GATC Biotech). PCR amplicons were digested with 10 units of the appropriate restriction enzyme in 1× recommended reaction buffer, following the manufacturer's protocol. $tRNA^{Ala}_{AGC}$-derived PCR amplicons were digested with Bpu1102I (Thermo Scientific, catalog no. ER0091) in Tango buffer (Thermo Scientific) for 12 hours at 37°C, and $tRNA^{Ala}_{AGC}$-derived PCR amplicons with AciI (NEB) in CutSmart buffer (NEB) for 3 hours at 37°C. Digested samples were then resolved in 2%

agarose gels with DNA Stain G (Serva) (run for 1.5 h at 120 V) and visualized under UV shadowing. Gel bands were quantified using ImageJ.

In all RFLP assays, potential DNA contamination was ruled out by using as PCR template: i) material from an RT reaction without reverse transcriptase (-RT); ii) material from an RT reaction without tRNA template (-tRNA template); and iii) nuclease-free H_2O water (-cDNA) (**Figure 22**). Primers used for reverse transcription and PCR reactions are listed in Table S4.

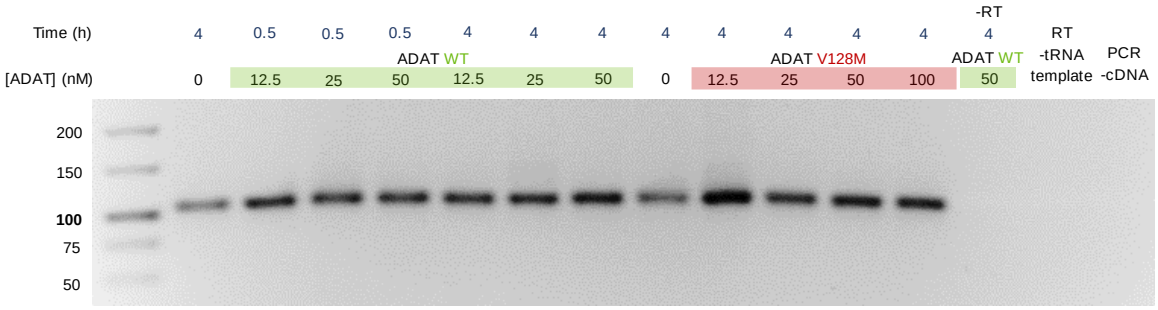


Figure 22. PCR amplicons and RT-PCR controls for RFLP-based activity assays. Agarose gel electrophoresis of representative samples of PCR amplicons derived from $tRNA^{Ala}_{AGC}$ used in the RFLP-based activity assays. A single band of the expected 100 bp was detected for all samples (lanes 1-12). Lanes 13-15 correspond to RT and PCR controls: RT reaction without reverse transcriptase (-RT); RT reaction without tRNA template (-tRNA template); PCR reaction without cDNA template (-cDNA). In the three cases, nuclease free-H_2O was added instead of the missing reagents.

All tRNA fragments added to the deamination reactions were dissolved in nuclease-free water. All small molecules added used for the *in vitro* assays were dissolved in DMSO. The final concentration of DMSO in the deamination reaction was 5% in all cases (negative and positive controls also contained 5% DMSO).

Size-exclusion chromatography (SEC)

For the evaluation of the stability of affinity-purified heterodimeric ADAT enzymes, ADAT2-ADAT3 wt and ADAT2-ADAT3 V128M, 100 μg of protein dissolved in 500 μL of SEC buffer (20 mM Tris-HCl, 1 M NaCl, 1 mM EDTA, pH 7.5) were injected onto a Superdex 200 10/300 GL column (GE Healthcare Life Sciences). SEC was performed as per manufacturer instructions. Protein was eluted with 1.5 column volume of buffer. The collected fractions were analyzed by SDS-PAGE and WB for protein detection, and 8 M urea PAGE for RNA detection. Protein and RNA content of the size exclusion-eluted fractions was quantified in each gel using ImageJ software. Protein standards were used to estimate the molecular weight (Mw) of the eluted fractions (**Figure 43 and Table 5**). SEC of protein standards was done under the same conditions, injecting 500 μL of 8.9 mM

lysozyme from chicken egg white (Fluka), 8.4 mM ovalbumin from chicken egg white (Sigma-Aldrich), 12.5 mM lectin from *Phaseolus vulgaris* (Sigma-Aldrich), 1.5 mM catalase from bovine liver (Sigma-Aldrich), and 0.8 mM ferritin from equine spleen (Calbiochem) in StrepTrap elution buffer.

SEC-MALS was performed using the same column and experimental conditions as in SEC, in a Prominence liquid chromatography system (Shimadzu) connected to a DAWN HELEOS II multi-angle light scattering (MALS) detector and a Optilab T-REX refractive index (RI) detector (Wyatt Technology). ASTRA 7 software (Wyatt Technology) was used for data processing and result analysis. A dn/dc value of 0.185 ml/g (typical for proteins) was assumed for calculations.

The quaternary structure of affinity-purified human ADAT2 was also assessed by SEC. For this, 100 μg of protein dissolved in 500 μL of 100 mM phosphate buffer, 50 mM NaCl, pH 7.5 were injected onto a Superdex 75 10/300 GL column (GE Healthcare Life Sciences). SEC was performed as per manufacturer instructions. Protein was eluted with 1.5 column volume of the same buffer. The collected fractions were analyzed by 12 % SDS-PAGE and BlueSafe staining (NZY Tech) and WB. Protein standards were used to estimate the molecular weight (Mw) of the eluted fractions (**Figure 48 and Table 8**). SEC of protein standards (conalbumine, ovalbumin, carbonic anhydrase, and ribonuclease R) was done under the same conditions.

Western blot

Protein extracts were resolved on 12 % SDS-PAGE, transferred to polyvinylidene difluoride (PVDF) membranes (Immobilion-P, Millipore) at 250 mA for 90 minutes at 4°C and immunoblotted with antibodies specific to Strep-tag (dil. 1:1000; ab76949, Abcam), and His-tag (dil. 1:1000; ab18184, Abcam). HRP-conjugated anti-mouse (GE Healthcare) was used as secondary antibody. Visualization was performed using ECL (GE Healthcare) in an Odyssey Fc Imager (Li-Cor).

Statistical analyses

Curve fittings and statistical analyses were performed with GraphPad Prism v6.

Homology Modeling

The homology model of ADAT2 homodimer used as molecular docking receptor was created using MOE (155). The sequence of human ADAT2 (PDB ID: 3DH1) was used as query sequence. The structure of *Staphylococcus Aureus* TadA (2B3J) was used as a template due to its sequence and structural similarity, but with an optimized structure for binding

ligands. 10 different homology models were created with the default parameters of MOE. A visual inspection followed by fpocket (156) calculations allowed us to select the model with the most druggable cavity.

Molecular Docking

All molecular docking calculations were run with rDock software (157). The cavity was generated using the results of fpocket as input, with default parameters. For the virtual screening, the commercial library of Specs was prepared using LigPrep (158) and the standard protocol of rDock was used, as previously published (157). Two specific interactions with a water molecule in the catalytic center (around residues His53 and Glu55) and with Ala60 were defined as pharmacophoric restraints in the rDock protocol. The results were filtered by docking score and structural diversity and a visual inspection was used for final selection.

Table 3. DNA oligonucleotides used in the study. All synthetic DNA oligonucleotides were from Sigma-Aldrich.

DNA oligonucleotide	Sequence
Heterologous protein expression	
Forward primer for PCR amplification of human ADAT gene for heterologous expression	AAGTTCTGTTTCAGGGCCCGATGGAAGCAAAAGCAGCACCGAAACC
Reverse primer for PCR amplification of human ADAT gene for heterologous expression	GGTGGCTCCAGCTAGCGGTATCCGGATCCAGCCAACGACAC
Forward primer for human ADAT3 QuikChange mutation V128M (GTG → ATG) for heterologous expression	CCGTTTCTGGTTCCGATGCCTGCCCGTCCGC
Forward primer for human ADAT3 QuikChange mutation V128M (GTG → ATG) for heterologous expression	GCGGACGGGCAGGCATCGGAACCAGAAACGG
Forward primer for In-Fusion cloning of ADAT2 for heterologous expression	GCCAGGATCCGAATTCATGGAGGCGAAGGCGGCACCCAAGC
Reverse primer for In-Fusion cloning of ADAT2 for heterologous expression	CGCCGAGCTCGAATTCTCAAGATTTCTGACATTCCTTTTTCCG
Templates for *in vitro* transcription and RT-PCR amplification in RFLP assay	
$tRNA^{Ala}_{AGC}$ oligo 1 – specific forward primer for PCR	AGCTTAATACGACTCACTATAGGGGGTGTGGCT
$tRNA^{Ala}_{AGC}$ oligo 2	CAGTGGTAGAGCGCGTGCTTAGCATGCACGAG
$tRNA^{Ala}_{AGC}$ oligo 3	GcCCCGGGTTCAATCCCCGGCACCTCCACCAGGG
$tRNA^{Ala}_{AGC}$ oligo 4	CGCTCTACCACTGAGCCACACCCCCTATAGTGAGTCGTATTA
$tRNA^{Ala}_{AGC}$ oligo 5	TTGAACCCGGGGCCTCGTGCATGCTAAGCACG
$tRNA^{Ala}_{AGC}$ oligo 6 – specific reverse primer for RT and PCR	GATCCCCTGGTGGAGGTGCCGGGGA
$tRNA^{Arg}_{ACG}$ oligo 1 – specific forward primer for PCR	AGCTTAATACGACTCACTATAGGGCCAGTGGCGCAA
$tRNA^{Arg}_{ACG}$ oligo 2	TGGATAACGCGTCTGACTACGGATCAGAAGATTCC
$tRNA^{Arg}_{ACG}$ oligo 3	AGGTTCGACTCCTGGCTGGCTCGCCAACATGTG
$tRNA^{Arg}_{ACG}$ oligo 4	GTTATCCATTGCGCCACTGGCCCTATAGTGAGTCGTATTA

$tRNA^{Arg}_{ACG}$ oligo 5	TCGAACCTGGAATCTTCTGATCCGTAGTCAGACGC
$tRNA^{Arg}_{ACG}$ oligo 6 – specific reverse primer for RT and PCR	GATCCACATGTTGGCGAGCCAGCCAGGAG
15 nt + T7 promoter sequence for *in vitro* transcription (double stranded part of the template)	GCGCAGTCAAGTCAGTAATACGACTCACTATAG
15 nt + T7 promoter + $tRNA^{Arg}_{ACG}$ template for *in vitro* transcription	TGGCGAGCCAGCCAGGAGTCGAACCTGGAATCTTCTGATCCGT AGTCAGACGCGTTATCCATTGCGCCACTGGCCCTATAGTGAGT CGTATTACTGACTTGACTGCGC
15 nt + T7 promoter + $tRNA^{Ala}_{AGC}$ template for *in vitro* transcription	TGGTGGAGGTGCCGGGGATTGAACCCGGGGCCTCGTGCATGC TAAGCACGCGCTCTACCACTGAGCCACACCCCCTATAGTGAGT CGTATTACTGACTTGACTGCGC
15 nt + T7 promoter + $tRNA^{Arg}_{ACG}$ lacking the CCA template for *in vitro* transcription	CGAGCCAGCCAGGAGTCGAACCTGGAATCTTCTGATCCGTAGT CAGACGCGTTATCCATTGCGCCACTGGCCCTATAGTGAGTCGTA TTACTGACTTGACTGCGC
15 nt + T7 promoter + $tRNA^{Ala}_{AGC}$ lacking the CCA template sequence for *in vitro* transcription	TGGAGGTGCCGGGGATTGAACCCGGGGCCTCGTGCATGCTAAG CACGCGCTCTACCACTGAGCCACACCCCCTATAGTGAGTCGTAT TACTGACTTGACTGCGC
15 nt + T7 promoter + $tRNA^{Arg}_{ACG}$ lacking the acceptor stem template for *in vitro* transcription	CAGCCAGGAGTCGAACCTGGAATCTTCTGATCCGTAGTCAGACG CGTTATCCATTGCGCCACTGCTATAGTGAGTCGTATTACTGACTT GACTGCGC
15 nt + T7 promoter + $tRNA^{Ala}_{AGC}$ lacking the acceptor stem template for *in vitro* transcription	TGCCGGGGATTGAACCCGGGGCCTCGTGCATGCTAAGCACGCG CTCTACCACTGAGCCACACTATAGTGAGTCGTATTACTGACTTGA CTGCGC
15 nt + T7 promoter + $tRNA^{Arg}_{ACG}$ T-arm and anticodon arm template for *in vitro* transcription	CCAGGAGTCGAACCTGGAATCTTCTGATCCGTAGTCAGACTATA GTGAGTCGTATTACTGACTTGACTGCGC
Synthesis of tRNAs with emissive nucleotide analogs	
T7 promoter + 5'-end fragment template for *in vitro* transcription (forward)	AGCTTAATACGACTCACTATAGGGGGTGTGGCTCAGTGGTAGAC CAGGG
T7 promoter + 5'-end fragment template for *in vitro* transcription (reverse)	GATCCCCTGGTCTACCACTGAGCCACACCCCCTATAGTGAGTCG TATTA
T7 promoter + anticodon fragment template for *in vitro* transcription (forward)	AGCTTAATACGACTCACTATAGCGCGTGCTTAGCGCTG
T7 promoter + anticodon fragment template for *in vitro* transcription (reverse)	GATCCAGCGCTAAGCACGCGCTATAGTGAGTCGTATTA
T7 promoter (class II) + 3'-end fragment template for *in vitro* transcription (forward)	AGCTTTAATACGACTCACTATTAATGCACGAGGCCCCGGGTTCA ATCCCCGGCACCTCCAGGG
T7 promoter (class II) + 3'-end fragment template for *in vitro* transcription (reverse)	GATCCCCTGGAGGTGCCGGGGATTGAACCCGGGGCCTCGTGC ATTAATAGTGAGTCGTATTAA
T7 promoter sequence for *in vitro* transcription (double stranded part of the template)	TAATACGACTCACTATAG
T7 promoter (class II) sequence for *in vitro* transcription (double stranded part of the template)	TAATACGACTCACTATTA
T7 promoter + acceptor fragment (thA/^{tz}A) template for *in vitro* transcription	AAGCACGCGCTCTACCACTGAGCCACACCCCCTATAGTGAGTC GTATTA
T7 promoter (class II) + donor fragment (thA/^{tz}A) template for *in vitro* transcription	TGGTGGAGGTGCCGGGGATTGAACCCGGGGCCTCGTGCATGCT AATAGTGAGTCGTATTA
T7 promoter + acceptor fragment (thG/^{tz}G) template for *in vitro* transcription	TAAGCACGCGCTCTACCACTGAGCCACACCCCCTATAGTGAGTC GTATTA
T7 promoter + donor fragment (thG/^{tz}G) template for *in vitro* transcription	TGGTGGAGGTGCCGGGGATTGAACCCGGGGCCT CGTGCATGCTATAGTGAGTCGTATTA
DNA splint 1	GAACCCGGGGCCTCGTGCATGCTAAGCACGCGCTCTACCACTG AGCCACACCC
DNA splint 2	GATTGAACCCGGGGCCTCGTGCATGCTAAGCACGCGCTCTACC ACTGAGCCACACCCCC

Table 4. Sequences of the transcripts used in the study.

RNA transcript	Sequence
$tRNA^{Arg}_{ACG}$	GGGCCAGTGGCGCAATGGATAACGCGTCTGACTACGGATCAG AAGATTCCAGGTTCGACTCCTGGCTGGCTCGCCA
$tRNA^{Ala}_{AGC}$	GGGGGUGUGGCUCAGUGGUAGAGCGCGUGCUUAGCAUGCA CGAGGCCCCGGGUUCAAUCCCCGGCACCUCCACCA
$tRNA^{Arg}_{ACG}$ lacking the CCA	GGGCCAGTGGCGCAATGGATAACGCGTCTGACTACGGATCAG AAGATTCCAGGTTCGACTCCTGGCTGGCTCG
$tRNA^{Ala}_{AGC}$ lacking the CCA	GGGGGUGUGGCUCAGUGGUAGAGCGCGUGCUUAGCAUGCA CGAGGCCCCGGGUUCAAUCCCCGGCACCUCCA
$tRNA^{Arg}_{ACG}$ lacking the acceptor stem	CAGUGGCGCAAUGGAUAACGCGUCUGACUACGGAUCAGAAG AUUCCAGGUUCGACUCCUGGCUG
$tRNA^{Ala}_{AGC}$ lacking the acceptor stem	GUGUGGCUCAGUGGUAGAGCGCGUGCUUAGCAUGCACGAG GCCCCGGGUUCAAUCCCCGGCAC
$tRNA^{Arg}_{ACG}$ T-arm and anticodon arm	UCUGACUACGGAUCAGAAGAUUCCAGGUUCGACUCCUGG
$tRNA^{Ala}_{AGC}$ T-arm and anticodon arm	CGUGCUUAGCAUGCACGAGGCCCCGGGUUCAAUCCCCGG
$tRNA^{Arg}_{ACG}$ anticodon stem loop	UCUGACUACGGAUCAGA
$tRNA^{Ala}_{AGC}$ anticodon stem loop	CGUGCUUAGCAUGCACG
Acceptor stem analog derived from $tRNA^{Pro}_{AGG}$ (with loop)	GGCUCGUUGGGACGAGCCCCCA
Acceptor stem analog derived from $tRNA^{Pro}_{AGG}$ without loop (5'-end strand)	GGCUCGU
Acceptor stem analog derived from $tRNA^{Pro}_{AGG}$ without loop (3'-end strand)	ACGAGCCCCCA
Chimeric $tRNA^{Arg}_{ACG}$ with $tRNA^{Ala}_{AGC}$ anticodon stem loop	GGGCCAGUGGCGCAAUGGAUAACGCGCGUGCUUAGCAUGCA CGAGAUUCCAGGUUCGACUCCUGGCUGGCUCGCCA
Chimeric $tRNA^{Ala}_{AGC}$ with $tRNA^{Arg}_{ACG}$ anticodon stem loop	GGGGGUGUGGCUCAGUGGUAGAGCGUCUGACUACGGAUCA GAAGGCCCCGGGUUCAAUCCCCGGCACCUCCACCA
5'-half from $tRNA^{Ala}_{AGC}$	GGGGGUGUGGCUCAGUGGUAGAGCGCGUGCUU
5'-tRF from $tRNA^{Ala}_{AGC}$	GGGGAUGUAGCUCAGUGG
3'-tRF from $tRNA^{Ala}_{AGC}$	UCCCCGGCACCUCCACCA
5'-tRF from $tRNA^{Arg}_{ACG}$	GGGCCAGUGGCGCAAUGGA
3'-tRF from $tRNA^{Arg}_{ACG}$	CUCCUGGCUGGCUCGCCA
5'-half from $tRNA^{Cys}_{GCA}$	GGGGGUAUAGCUCAGUGGUAGAGCAUUUGACUG

Results

Development of an *in vitro* activity assay to monitor ADAT-mediated deamination

Abstract

tRNAs are among the most heavily modified RNA species. I34 is an essential post-transcriptional modification in tRNAs catalyzed by ADAT enzyme. Monitoring ADAT-mediated deamination is crucial for the characterization of ADAT enzyme in terms of activity, substrates, regulation, as well as for drug discovery purposes. However, this analysis is often challenging, laborious and lacks quantitativeness. We developed an *in vitro* deamination assay based on restriction fragment length polymorphism (RFLP) analyses to monitor ADAT activity in an efficient, cost-effective, and semiquantitative manner. To overcome a limitation of the method being the need of reverse transcription and amplification of the tRNA, we designed a direct method to quantify I34 formation *in vitro* using the first fluorescent analogs of nucleic acids that have been reported to undergo enzymatic deaminations.

Restriction fragment length polymorphism (RFLP) method

Wulff TF, Arguello RJ, Molina Jordan M, ***Roura Frigole H****, Hauquier G, Filonava L, Camacho N, Gatti E, Pierre P, Ribas de Pouplana L, Torres AG. Detection of a Subset of Posttranscriptional Transfer RNA Modifications in Vivo with a Restriction Fragment Length Polymorphism-Based Method. Biochemistry. 2017;56(31):4029-38. Epub 2017/07/14.*

RFLP analyses study the patterns of DNA cleavage after restriction enzyme treatment and have been used for the qualitative detection of modified bases in RNA. Some post-transcriptional modifications in RNA induce specific and reproducible base "mutations" when RNAs are reverse transcribed. It is the case of inosine, which is detected as guanosine when an inosine-containing tRNA is reverse transcribed, PCR-amplified and sequenced (3). These point mutations on RT-PCR amplicons generated in the RT reaction can create or abolish endonuclease restriction sites. RFLP analyses study the patterns of DNA cleavage after restriction enzyme treatments (148). In this context, RFLP method confers some advantages by avoiding the need of RNA sequencing, allowing the detection of the modification in standard laboratory facilities in an efficient and cost-effective manner, and providing semiquantitative results. Here we have applied RFLP to study I34 modification in human tRNAs (13).

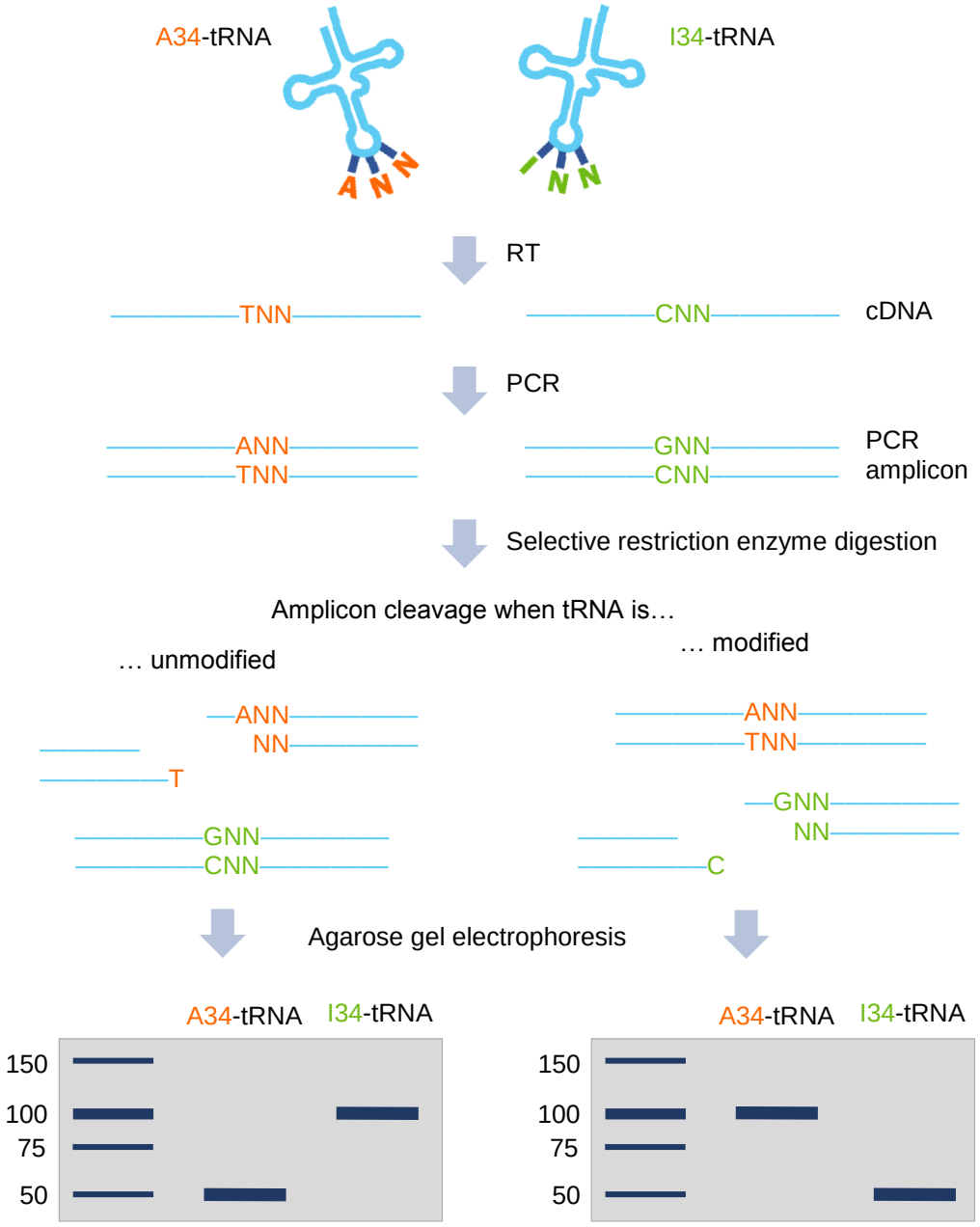


Figure 23. Schematic representation of the RFLP method for I34 detection. During RT, the reverse transcriptase reads I34 as a G and incorporates a C on the cDNA strand. Therefore, after PCR amplification, the original I34 position is read as G34. The obtained amplicon is then digested with a specific restriction enzyme that would recognize and cleave the target in a modified/unmodified residue-dependent manner. The products of digestion are then resolved by gel electrophoresis, and the proportion of cleaved to uncleaved amplicon can be quantified to estimate the levels of modified tRNA.

The RFLP method for I34 detection in tRNAs has four steps: RT, PCR amplification with specific primers, incubation of the PCR product with a specific restriction enzyme that targets position 34, and detection of the cleavage patterns by agarose gel electrophoresis. During RT, modified tRNA containing I34 is read as guanosine due to the high structural similarity between the two bases. Hence, due to the A-to-I conversion, an A-to-G mutation is produced on the PCR product and, consequently, an endonuclease restriction site is created or abolished. If the modification creates a restriction site, the PCR amplicon will be cleaved only when it is derived from modified tRNA containing I34, but not A34-tRNA. Conversely, if the modification abolishes a restriction site, the PCR amplicon will be cleaved only if it is derived from unmodified A34-tRNA, but not I34-tRNA. The analysis of the modified tRNA population versus the unmodified, is based on the quantification of the bands of the restriction fragments in the agarose gel (**Figure 23**) (13).

We first studied I34 modification in human $tRNA^{Ala}_{AGC}$, a natural substrate of the human ADAT enzyme. For this, *in vitro* transcribed $tRNA^{Ala}_{AGC}$ was incubated with affinity-purified human ADAT. The reaction product was purified, reverse transcribed and amplified by PCR. Deamination was confirmed by sequencing the PCR amplicon, a commonly used strategy (1-5, 159) (**Figure 24A**). As expected, A34 was detected on the *in vitro* transcribed tRNA (**Figure 24A, upper panel**) but was fully converted to I34 (detected as G34) upon incubation with ADAT (**Figure 24A, lower panel**). An *in silico* search for restriction enzymes revealed that BlpI (Bpu1102I) would recognize and cleave a PCR amplicon derived from unmodified $tRNA^{Ala}_{AGC}$ (A34) but would not cleave the $tRNA^{Ala}_{IGC}$-derived counterpart (G34) (**Figure 24B**). Indeed, upon incubation with BlpI, only the PCR amplicon derived from the unmodified tRNA was cleaved (**Figure 24C**). Importantly, no cleavage was observed for the modified tRNA, consistent with the 100% modification status of the tRNA upon incubation with ADAT observed by sequencing (**Figure 24A**). Thus, the results obtained with the RFLP method are in agreement with those obtained by sequencing (13).

To further characterize this method, we tested its sensitivity toward I34 detection and its potential to be used in a semiquantitative manner. Unmodified and fully modified *in vitro* transcribed $tRNA^{Ala}_{AGC}$ were mixed in known proportions. The levels of I34 modification were next evaluated by RFLP, and the observed gel bands were quantified. **Figure 25A** shows a representative gel obtained for this type of experiment. Gel quantification of three independent replicates revealed a linear correlation ($R^2 = 0.99$) between the expected and observed levels of I34 (**Figure 25B**), suggesting that the method can be semiquantitative. Importantly, the method proved to be sensitive enough to reliably detect low levels of modification (expected, 1% I34; observed, 2.80 ± 1.05% I34) (**Figure 25B**).

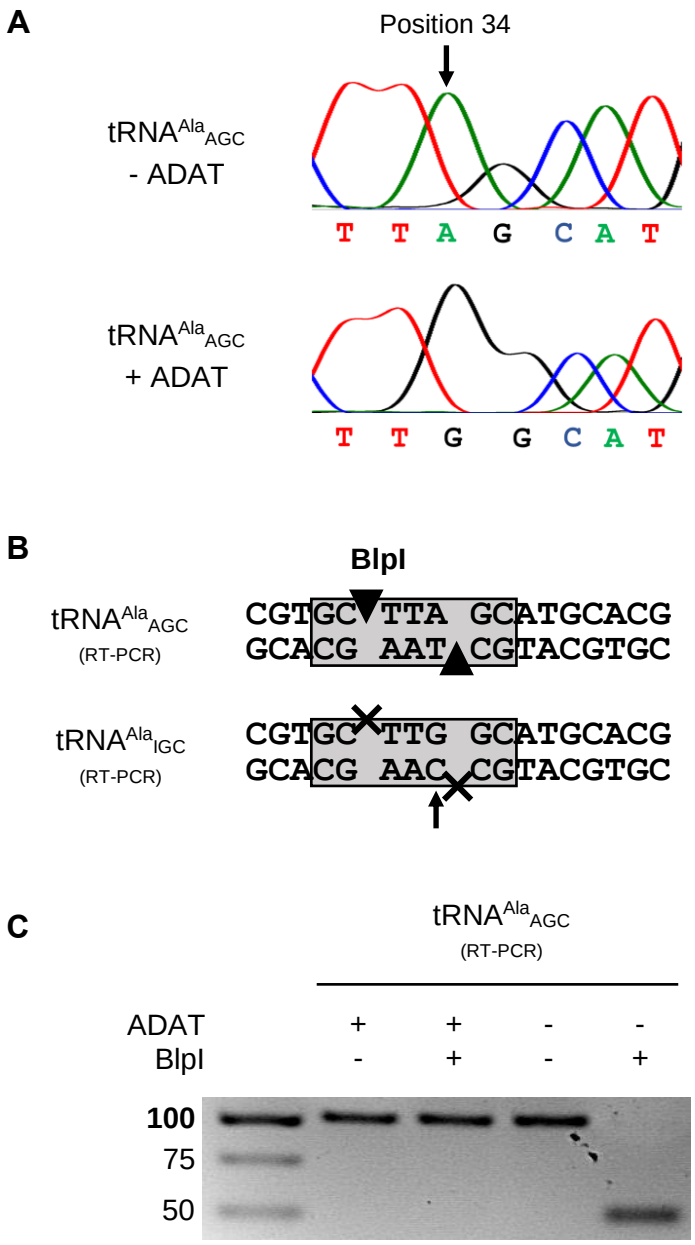


Figure 24. Evaluation of *in vitro* deamination by human ADAT on *in vitro* transcribed tRNA$^{Ala}_{AGC}$. Figure from (13).

A Sequencing spectrum for *in vitro* transcribed tRNA$^{Ala}_{AGC}$ incubated with or without human ADAT. I34 is detected as G34.

B Schematic representation of the recognition site for BlpI for amplicons derived from unmodified or modified tRNA. The arrow indicates the modification site (position 34 on the tRNA). BlpI will cleave the amplicon only when it is derived from unmodified tRNA$^{Ala}_{AGC}$.

C Evaluation of *in vitro* deamination by human ADAT on *in vitro* transcribed tRNA$^{Ala}_{AGC}$ using the RFLP method. After cDNA synthesis, PCR amplification, and digestion with BlpI, samples were resolved in a 2% agarose gel. As expected, the full-length amplicon is fully cleaved only when A34 is not modified.

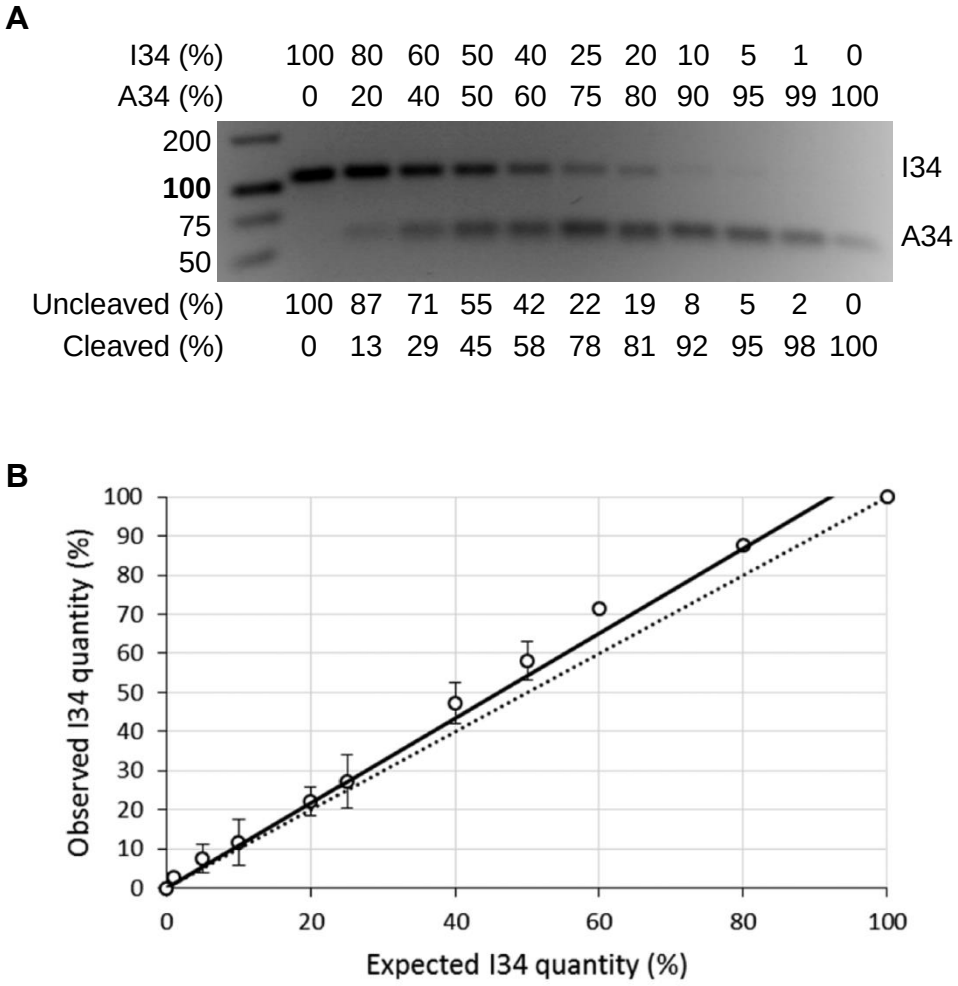


Figure 25. Evaluation of the sensitivity and accuracy of the RFLP method for I34 detection. Fully modified (I34) and fully unmodified (A34) *in vitro* transcribed $tRNA^{Ala}_{AGC}$ were mixed in the indicated proportions. The RNA mix was then reverse transcribed, PCR amplified, digested with BlpI, and run in a 2% agarose gel. The cleavage pattern was quantified using ImageJ. Figure from (13).

A Representative gel showing the obtained quantification results. Band quantification was normalized to total lane intensity (uncleaved + cleaved band).

B Plot of the estimated levels of I34 on *in vitro* transcribed $tRNA^{Ala}_{AGC}$ determined by RFLP (solid regression line, $y = 1.084x$; $R^2 = 0.99$) against the theoretically expected levels of modification (dotted regression line, $y = x$; $R^2 = 1$). The mean and standard deviations are shown for three independent replicates. Curve fitting was performed following a linear regression model.

We next applied the RFLP method to monitor the effect of different reaction conditions on ADAT activity by varying the pH and magnesium (Mg) concentration of the reaction buffer. Optimal deamination of $tRNA^{Ala}_{AGC}$ was achieved at pH 8.0 and 1 mM Mg (**Figure 26**). These results suggest that ADAT activity *in vitro* is sensitive to pH and Mg concentration and illustrate the potential use of the method in reaction optimization.

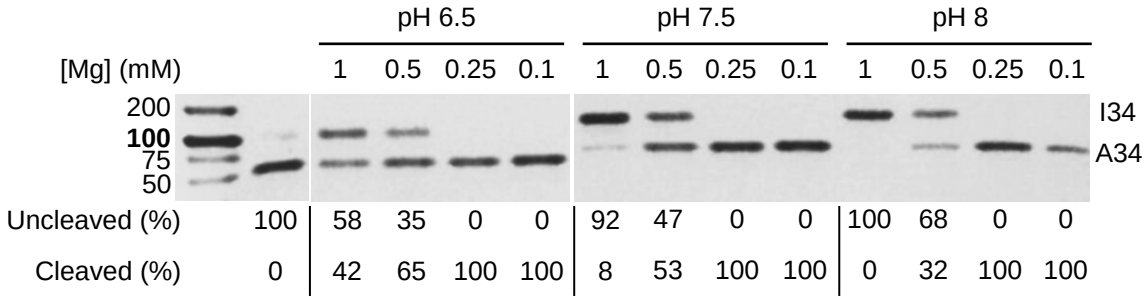


Figure 26. Evaluation of *in vitro* deamination efficiency by human ADAT on *in vitro* transcribed tRNA$^{Ala}_{AGC}$. Different reaction conditions (variable pH and Mg) were evaluated using the RFLP method. Samples were resolved in a 2% agarose gel. Band quantification was normalized to total lane intensity (uncleaved + cleaved band). Optimal reaction conditions were observed at pH 8.0 and 1 mM Mg.

Fluorescence-based assay

Project performed during a three-month stay at Dr. Tor laboratory at the University of California San Diego (UCSD).

One of the limitations of the RFLP method to monitor ADAT-mediated A34-to-I34 conversion is the need of RT and PCR amplification of the tRNA after the deamination, and the consequent lack of high-throughputness. To overcome this, we designed a direct and fully quantitative *in vitro* assay to monitor ADAT activity without the need of tRNA processing, which would allow high-throughput analysis of the deamination. The method consists in monitoring ADAT editing by measuring changes in fluorescence due to the A34-to-I34 conversion.

Dr. Tor laboratory at the University of California San Diego (UCSD) developed the first emissive adenosine analogs undergoing enzymatic deaminations. thA and ^{tz}A are fluorescent adenosine analogs, derived from thieno[3,4-d]-pyrimidine and isothiazolo[4,3-d]pyrimidine respectively, that are converted to the corresponding thI and ^{tz}I upon deamination (14, 16, 154). It has been reported that they can be deaminated both as free nucleosides by adenosine deaminase (ADA) (14, 16) and in the context of an mRNA molecule by adenosine deaminases acting on RNA (ADARs) (160). These conversions result in changes in emission that can be quantified to estimate the levels of deamination (**Figure 27**) (14-16, 160).

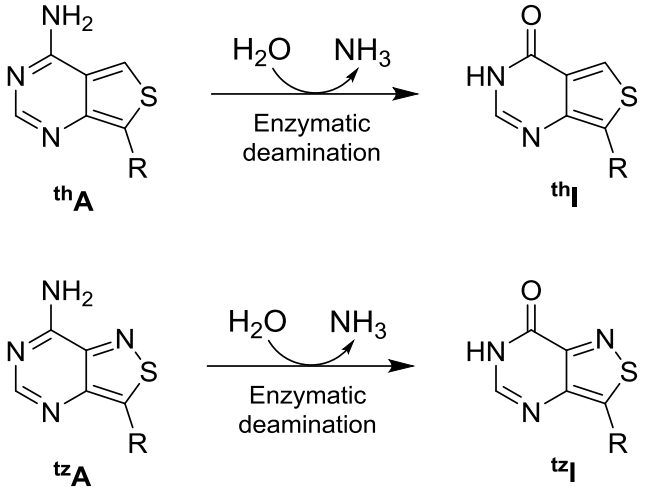


Figure 27. Chemical structures of emissive RNA nucleosides thA, ^{tz}A, thI and ^{tz}I. thA and ^{tz}A are fluorescent isomorphic analogs of A, and thI and ^{tz}I correspond to their respective deamination products. R groups indicate the position where the ribose sugar is attached.

We selectively incorporated the fluorescent adenosine analog thA at position 34 of $tRNA^{Ala}_{AGC}$. For this, the tRNA molecule was divided in three fragments so that the fragment that contained the anticodon had one single A, which corresponded to position 34 (where we want to incorporate the fluorescent nucleotide). The three fragments were *in vitro* transcribed separately using the corresponding synthetic DNA templates. The anticodon fragment was *in vitro* transcribed using the triphosphate from of the fluorescent A analog, thATP, instead of regular ATP, so that an emissive A was incorporated at position 34. The other two fragments (the 5'-end and the 3'-end fragments, respectively) were *in vitro* transcribed using regular (non-emissive) nucleotides triphosphate (NTPs). The three fragments were subsequently hybridized (using a complementary DNA splint) and ligated, so that a full-length tRNA molecule containing thA at position 34 was obtained (**Figure 28**).

To simplify and increase the yield of the synthesis of the emissive tRNA, we developed a second strategy that does not require the triphosphate form of the emissive nucleoside analog but the nucleoside instead. We selectively incorporated the fluorescent adenosine analogs thA and ^{tz}A at position 34 of $tRNA^{Ala}_{AGC}$. In this case the tRNA sequence was divided in two fragments: an acceptor fragment and a donor fragment. The acceptor fragment corresponded to the 5'-end half of the tRNA, whereas the donor fragment corresponded to the 3'-end half and its first base was position 34 (where we want to incorporate the emissive A). While the acceptor fragment was *in vitro* transcribed using standard conditions, the donor fragment was *in vitro* transcribed using regular NTPs and an excess of thA or ^{tz}A nucleosides (without phosphate groups). Since thA and ^{tz}A have no phosphate groups, they can only be incorporated at the first position of the transcript. The donor fragment with the emissive base at the first position was then phosphorylated at the 5'-end and ligated to the acceptor fragment (using a complementary DNA splint) (**Figure 29**).

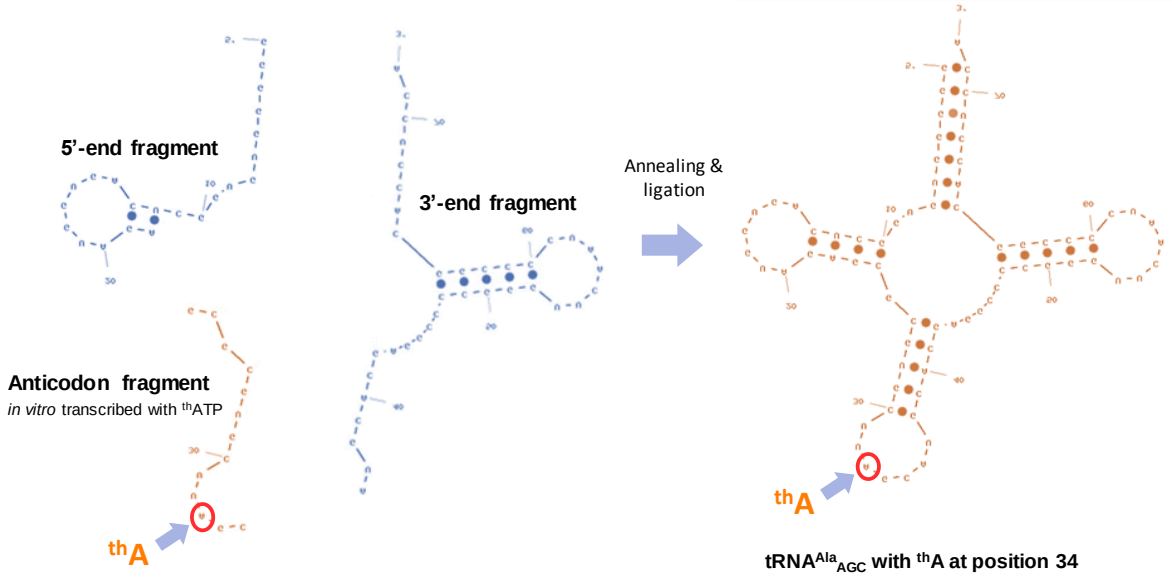


Figure 28. Schematic representation of the selective incorporation of the emissive nucleotide thA at position 34 of tRNA$^{Ala}_{AGC}$. After the *in vitro* transcription of the three fragments separately, they were annealed to a complementary DNA splint and ligated, resulting in an emissive tRNA$^{Ala}_{AGC}$ with thA34.

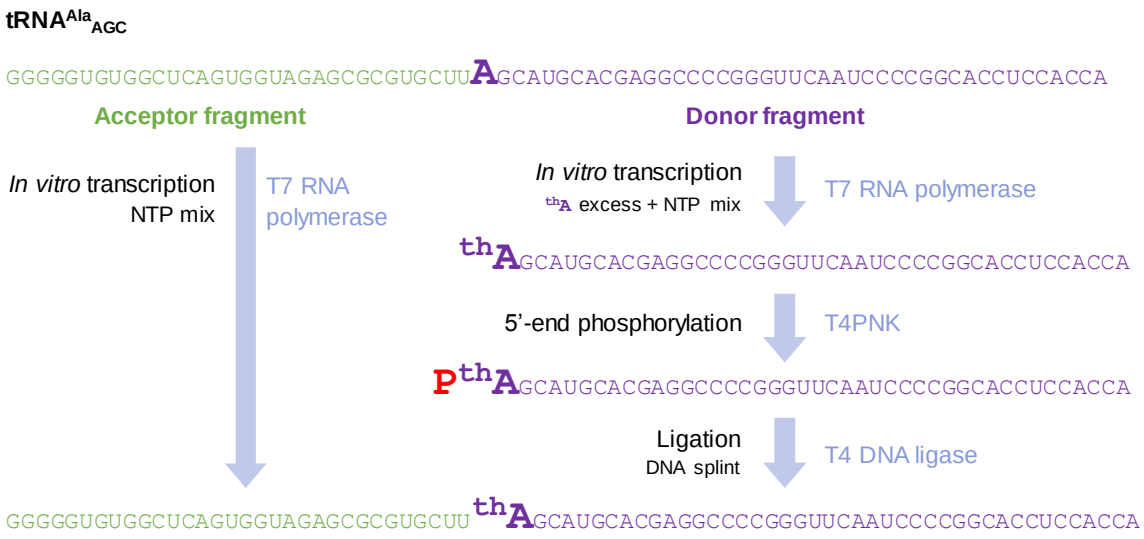


Figure 29. Schematic representation of the alternative strategy to selectively incorporate of the emissive nucleoside thA at position 34 of tRNA$^{Ala}_{AGC}$. After the *in vitro* transcription of the acceptor and donor fragments (the donor fragment with an excess of thA), the donor fragment was 5'-phosphoryated and ligated to the acceptor fragment using a complementary DNA splint. The enzymes used in each step are indicated.

We designed a second approach of the fluorescence-based assay consisting in monitoring A34-to-I34 conversion by measuring changes in the fluorescence of an emissive nucleotide adjacent to position 34, being G35, the second position of the anticodon. An emissive G analog being thG or ^{tz}G, derived from thieno[3,4-d]-pyrimidine and isothiazolo[4,3-d]pyrimidine, respectively, was incorporated at position 35 of $tRNA^{Ala}_{AGC}$ (**Figure 30**) (16, 154). The strategy of the donor and acceptor fragments was used to selectively incorporate thG or ^{tz}G at position 35.

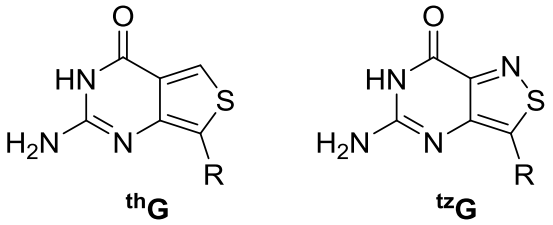


Figure 30. Chemical structures of emissive RNA nucleosides thG and ^{tz}G. thG and ^{tz}G are fluorescent isomorphic analogs of G. R groups indicate the position where the ribose sugar is attached (16, 154).

After the ligation of the RNA fragments, the emissive full-length tRNAs were separated from the DNA splint and potential unligated fragments by SDS-PAGE and purified. For all the emissive tRNAs, a band of the expected Mw was detected at both 254 nm and 312 nm (at which nucleic acids and fluorescent molecules are detected, respectively). The final products were confirmed by matrix-assisted laser desorption/ionization (MALDI) analysis (**Figure 31 and Supplementary Figures S2 and S3**).

Although we were able to synthesize and purify all the described emissive tRNAs, due to timing constraints, we were not able to finish developing the assay. The subsequent step to develop the fluorescence-based assay is to determine the absorption and emission spectra unmodified (with thA34/ ^{tz}A34) and modified tRNAs (with thI34/ ^{tz}I34), respectively, and to establish the correlation between A-to-I and absorption/emission. For the second approach, in which changes in the emission of thG35/ ^{tz}G35 are recorded, the absorption and emission spectra of the unmodified (A34) and modified (I34) thG35/^{tz}G35-tRNAs are compared instead. Finally, ADAT activity *in vitro* can be estimated by recording changes in the emission over the reaction course. This assay will allow high-throughput screening for the discovery of active molecules against ADAT, as well for enzyme kinetic studies.

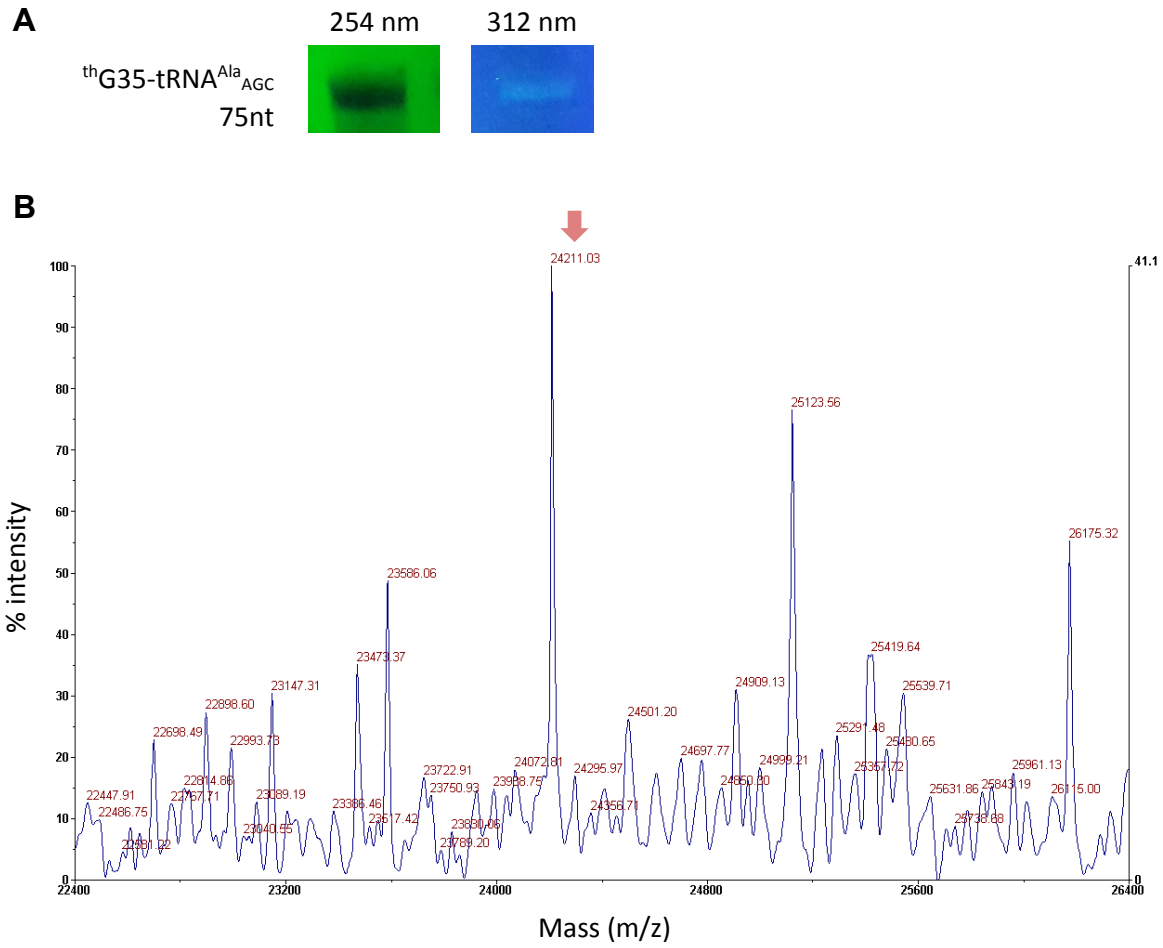


Figure 31. Representative SDS-PAGE (A) and MALDI (B) analyses of thG35-tRNA$^{Ala}_{AGC}$.

A SDS-PAGE was visualized at 254 nm, where nucleic acids are detected, and at 312 nm, where only fluorescent molecules are detected. The band observed at 254 nm and 312 nm coincided, confirming thus that the tRNA was emissive.

B The main peak of the MALDI spectra (red arrow) coincided with the expected molecular weight for thG35-tRNA$^{Ala}_{AGC}$ (23984.66 g/mol).

New insights into substrate recognition and regulation of human ADAT

***Roura Frigolé H**, Castellví Coma M, Fernandez Lozano C, Rafels-Ybern À, García-Lema J, Camacho N, Canals A, Coll M, Ribas de Pouplana L. New insights into substrate recognition and regulation of human ADAT. RNA. 2018. Manuscript in preparation.*

Abstract

ADAT is an essential enzyme conserved over the evolution with the acquisition of multi-substrate specificity. Whereas its bacterial homolog TadA deaminates exclusively $tRNA^{Arg}$, the eukaryotic enzyme deaminates eight different tRNAs. The mechanisms that drove this evolution remain unknown and substrate recognition of the eukaryotic ADAT is poorly understood. Here we show that tRNA architecture is the most important feature for efficient A34-to-I34 conversion by human ADAT enzyme. We report that both unfolded tRNAs and tRNA variants reduced in length result in loss of deamination. We find that the enzyme-substrate interaction mode varies between different tRNAs, and that parts of the tRNA far from the anticodon arm such as the acceptor stem play a role in ADAT binding. Interestingly, the recognition of $tRNA^{Arg}$ seems to have evolved as the other emergent substrate tRNAs, and is not restricted to the anticodon (the substrate of the bacterial ancestor of ADAT, TadA). Finally, we present a new potential mechanism of control of ADAT deamination activity by human tRNA-derived fragments, which provides new insights into the regulation of ADAT function and may open a door for the development of new strategies to modulate ADAT activity.

ADAT substrates do not share a common signature sequence

We explored the possibility that a distinctive sequence signature was present in ADAT substrate tRNAs, which might allow ADAT enzyme to discriminate the two tRNA groups. Multiple sequence alignments of the eight ADAT substrates were compared to those of non-ADAT substrates (**Figure 32, and Supplementary Figures S4, S5 and S6**). No differences in the conserved and semiconserved residues could be observed in T, A, P, S, L, I, V, and R tRNA isoacceptors between ADAT substrates (with A34) and non-ADAT substrates (with C34 or U34) (**Figure 32A and 32B**). Moreover, the great majority of conserved residues corresponded to nucleotides that have been reported to be essential for the tRNA secondary and tertiary structure. This suggests that A34 is probably the main determinant for the ADAT recognition among T, A, P, S, L, I, V, and R tRNA isoacceptors.

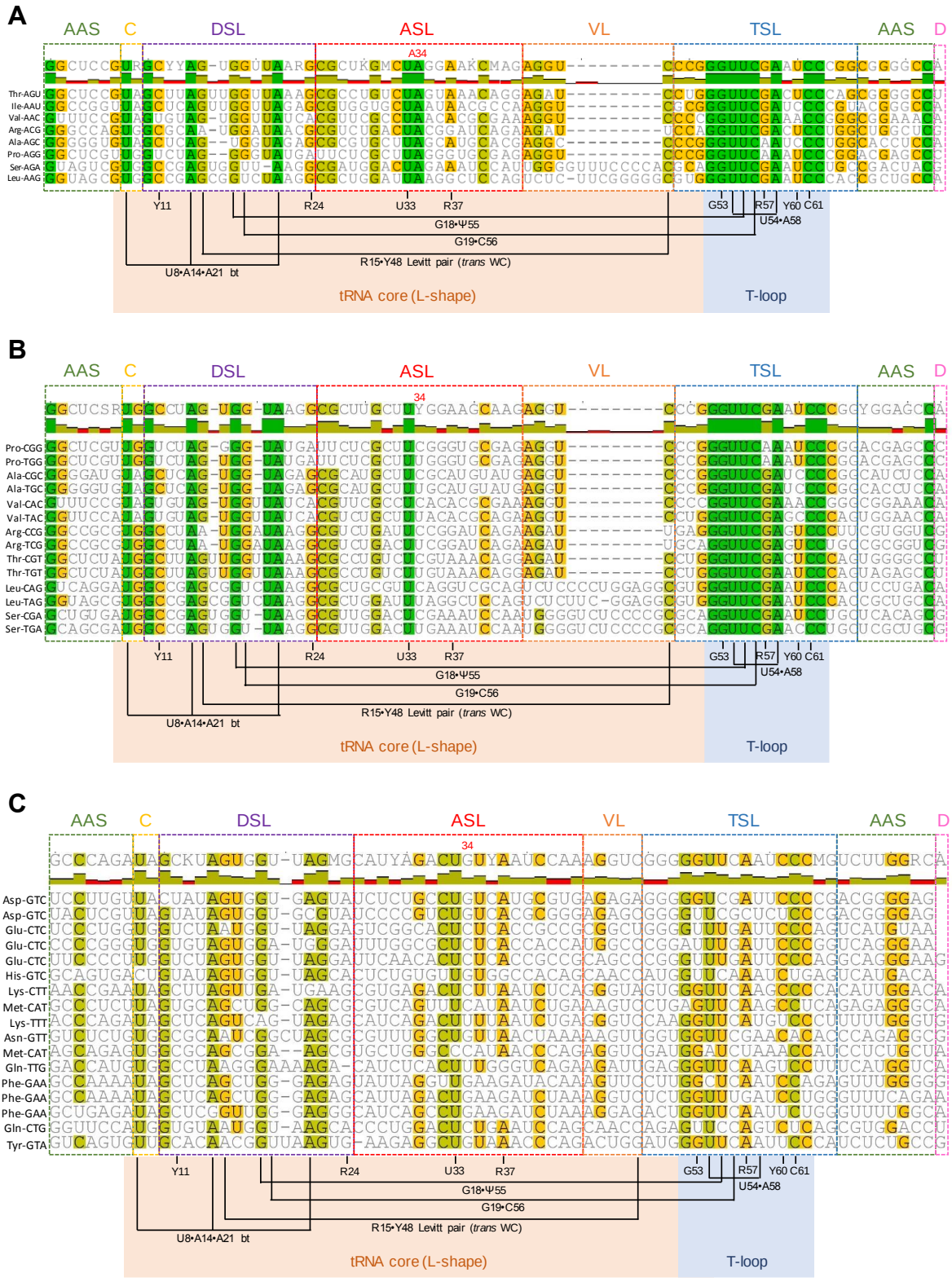


Figure 32. Sequence alignment of a representative set of human tRNAs including ADAT substrates (A) and non-ADAT substrates (B, C). The cognate amino acid and anticodon for each tRNA are specified. Secondary structural motifs are indicated. AAS stands for amino acid acceptor stem (green); C, connector between AAS and DSL (yellow); DSL, D-stem loop (or D-arm) (purple); ASL, anticodon stem loop (or anticodon arm) (red); VL, variable loop (orange); TSL, T-stem loop (or T-arm) (blue); and D, discriminator nucleotide (pink). Conserved and semiconserved residues are

color coded: green for 100% similar; light green for 80 to 100% similar, yellow for 60 to 80% similar; and white for less than 60% similar. Structure-based relationships between conserved and semiconserved nucleotides that control the 3D-structure formation are indicated. R stands for purine; Y for pyrimidine; and bt for base triple. Interactions in the 3D core of the tRNA for L-shape formation have orange background, and interactions for T-loop formation have blue background.

A The most abundant isodecoder sequence of each of the eight human tRNAs substrates of ADAT is represented.

B, C Representative sets of non-ADAT substrate tRNAs encoding for T, A, P, S, L, I, V, R (**B**) and the rest of aa (**C**).

Sequence alignment of a representative set of non-T, A, P, S, L, I, V, R tRNAs showed higher sequence variability (**Figure 32C**). However, the residues conserved in ADAT substrate tRNAs, which have been reported to be essential for the tRNA structure, were also conserved. Overall, the absence of a nucleotide signature shared among ADAT substrate tRNAs suggests that the enzyme substrate recognition might depend on the tRNA 3D architecture rather than on specific nucleotide sequences.

Unfolded tRNAs are not ADAT substrates

To study the importance of the tRNA three-dimensional structure for ADAT recognition, we analyzed ADAT-mediated deamination of human $tRNA^{Ala}_{AGC}$ prepared under different folding conditions by the Restriction Fragment Length Polymorphism (RFLP) method (13) (**Figure 33**). *In vitro* transcribed human $tRNA^{Ala}_{AGC}$ was subjected to either fast folding, slow folding or absence of folding, in the presence or absence of 1 mM magnesium (Mg) buffer, and incubated with affinity-purified human ADAT. tRNA deamination by ADAT is seen as a decrease in the intensity of the band corresponding to the digested product (50 bp), and an increase of undigested product (100 bp) (13).

In the absence of ADAT, fast-folded $tRNA^{Ala}_{AGC}$ with and without Mg buffer, showed a single 50-band, consistent with absence of deamination (lanes 1 and 2). Upon addition of 20 nM of ADAT, the same tRNA was almost fully deaminated, both in presence and absence of Mg buffer (lanes 3 and 4; 97.4 ± 1.4 % and 97.6 ± 0.6 %, respectively). Slow folding the tRNA slightly reduced the deamination rate to around 75 %, with slightly higher levels upon Mg addition (lanes 5 and 6; 75.4 ± 1.0 % and 78.7 ± 1.4 %, respectively). In the absence of tRNA folding and Mg buffer, ADAT activity was nearly abolished (lane 7; 8.9 ± 2.2 %). In contrast, adding Mg buffer to the unfolded tRNA recovered the deaminase activity to similar levels to the slow-folded tRNA (lane 8; 65.1 ± 2.6 %), consistent with the contribution of Mg ions to RNA assembly and tertiary structure stabilization. Overall, these results suggest that a proper tRNA architecture is required for ADAT recognition and deamination activity (**Figure 33**).

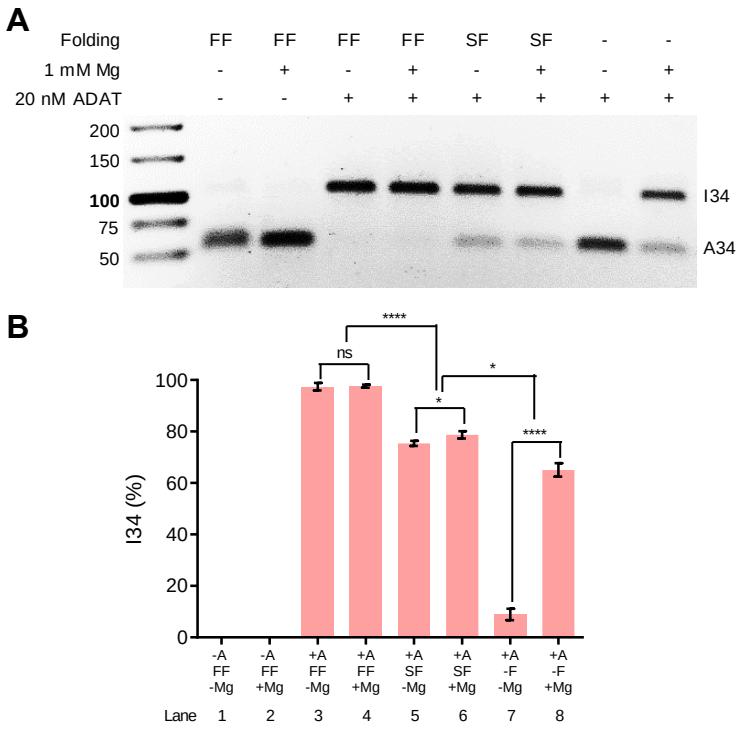


Figure 33. Effect of tRNA folding conditions upon the deamination of $tRNA^{Ala}_{AGC}$ by ADAT.

A Representative RFLP experiment to assess ADAT-mediated deamination of $tRNA^{Ala}_{AGC}$ under different tRNA folding conditions: fast folding (FF), slow folding (SF) or absence of folding (-F), with or without 1 mM Mg buffer (±Mg). -A indicates absence of ADAT, and +A addition of 20 nM ADAT.

B Levels of I34-modification as in A. All deamination reactions were done in duplicate and averaged. Error bars reflect the standard deviation. Asterisks indicate statistical significance between samples.

Short tRNA variants are poor ADAT substrates

We assessed the possibility that the recognition mode of $tRNA^{Arg}$ by ADAT enzyme might be similar to that of TadA (which mainly relies on the anticodon arm (2)), and investigated whether new ways to recognize the rest of tRNA substrates may have emerged with the evolution of the enzyme and the acquisition of multi-substrate specificity. To resolve these questions, we examined ADAT-mediated deamination of several tRNA mini-substrates derived from $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$ respectively. All tested tRNA fragments preserved the anticodon stem loop of the original tRNA as shown in **Figure 34A**. Since it was not possible to determine the deamination of the tRNA fragments by RFLP method (the shortest fragments could not be PCR-amplified), I34 levels were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) (161). For this, after incubation of the tRNA fragments with human ADAT enzyme, the reaction product was purified and treated with Nuclease P1 to completely digest the RNA sequence into 5'-nucleotides monophosphate (NMPs), which were analyzed by LC-MS/MS (**Figure 34B**).

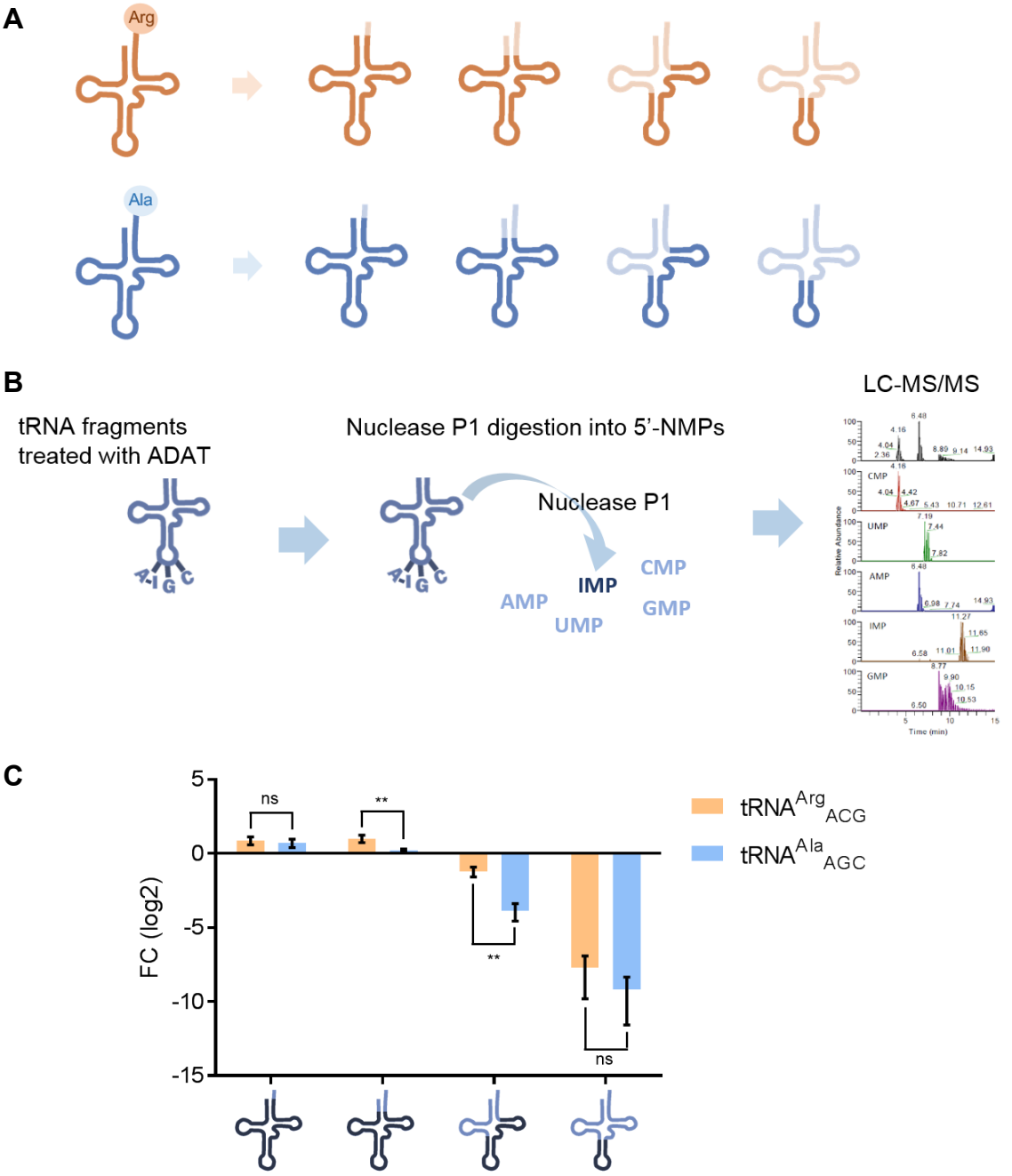


Figure 34. Analysis of ADAT-mediated deamination of tRNA mini-substrates derived from $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$.

A Schematic representation of the $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$ fragments analyzed: a tRNA lacking the CCA (which is added during the maturation of tRNAs at the 3'-end), a tRNA lacking the acceptor stem, a fragment with the T-arm and the anticodon arm, and the anticodon stem loop.

B Workflow of the assay used to detect I34 modification in the tRNA fragments by LC-MS/MS.

C Fold change in I34 levels represented in log2 scale between the full-length tRNA and the different mini-substrates derived from $tRNA^{Arg}_{ACG}$ (orange) and $tRNA^{Ala}_{AGC}$ (blue) after overnight incubation with human ADAT. All deamination reactions were done in duplicate and averaged. Error bars reflect the standard deviation. Asterisks indicate statistical significance between samples.

Incubation of $tRNA^{Ala}_{AGC}$-derived fragments with ADAT for one hour resulted in the deamination of the full-length tRNA and the tRNA lacking the CCA (**Supplementary Figure S7**). The deamination of the tRNA without the 3'-terminal CCA to the same extent as the full-length tRNA is consistent with the observation that ADAT acts at the precursor tRNA level (3). The absence of I34 modification in the shorter fragments suggests that tRNA recognition by ADAT is not limited to the anticodon region.

To detect if any of the tRNA mini-substrates could still be deaminated albeit with reduced efficiency, we treated the tRNA mini-substrates derived from both $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$ with ADAT and extended the incubation time to 12 hours. The tRNAs lacking the CCA and the acceptor stem showed deamination levels comparable to the full-length tRNAs. In contrast, the fragments containing the T-arm plus anticodon arm and the anticodon stem loop alone exhibited a dramatic decrease in I34 levels, especially the acceptor stem loop alone (**Figure 34C**). Hence, the shortest tRNA fragments showed to be poor substrates and only the tRNA fragments containing at least the three RNA arms were noticeably deaminated. This suggests that ADAT substrate recognition might not be limited to the anticodon arm, and that additional parts of the tRNA molecule are required. Significant differences between $tRNA^{Arg}$ and $tRNA^{Ala}$ fragments were observed in the tRNA fragments lacking the acceptor stem and D-arm. In both cases the fragments derived from $tRNA^{Arg}$ showed higher levels of deamination, which might indicate that the deamination of $tRNA^{Ala}$ is more sensitive to alterations in the tRNA structure.

The substrate recognition mode of ADAT varies between different tRNAs

To evaluate the recognition of $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$ by ADAT, we analyzed the effect of the aforementioned tRNA variants reduced in length to the ADAT-mediated deamination of $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$. Using the RFLP method (13), we examined I34 formation in $tRNA^{Arg}_{ACG}$ upon addition of 50 μM $tRNA^{Ala}_{AGC}$ fragments (**Figure 35A**), and in $tRNA^{Ala}_{AGC}$ upon addition of 50 μM $tRNA^{Arg}_{ACG}$ fragments (**Figure 35B**). For the digestion of the $tRNA^{Arg}_{ACG}$-derived PCR amplicons, we used AciI restriction enzyme, which cleaves the amplicons derived from I34-$tRNA^{Arg}_{ACG}$ (in contrast to BlpI, which cleaves the PCR amplicons derived from A34-$tRNA^{Ala}_{AGC}$).

All $tRNA^{Ala}_{AGC}$ fragments inhibited $tRNA^{Arg}_{ACG}$ deamination, in contrast, only the longest $tRNA^{Arg}_{ACG}$ fragments showed to inhibit $tRNA^{Ala}_{AGC}$ deamination. Whereas an anticodon stem loop analog is sufficient to inhibit the deamination of $tRNA^{Arg}_{ACG}$, longer fragments containing at least three arms of the tRNA are required to inhibit $tRNA^{Ala}_{AGC}$ deamination. This suggests that the major interaction between ADAT and $tRNA^{Arg}_{ACG}$ would be through the anticodon arm, while in $tRNA^{Ala}_{AGC}$ the highest contribution may rely on other parts in the top half of the tRNA molecule. Hence, the parts that contribute the most to the interaction with ADAT vary among different tRNAs.

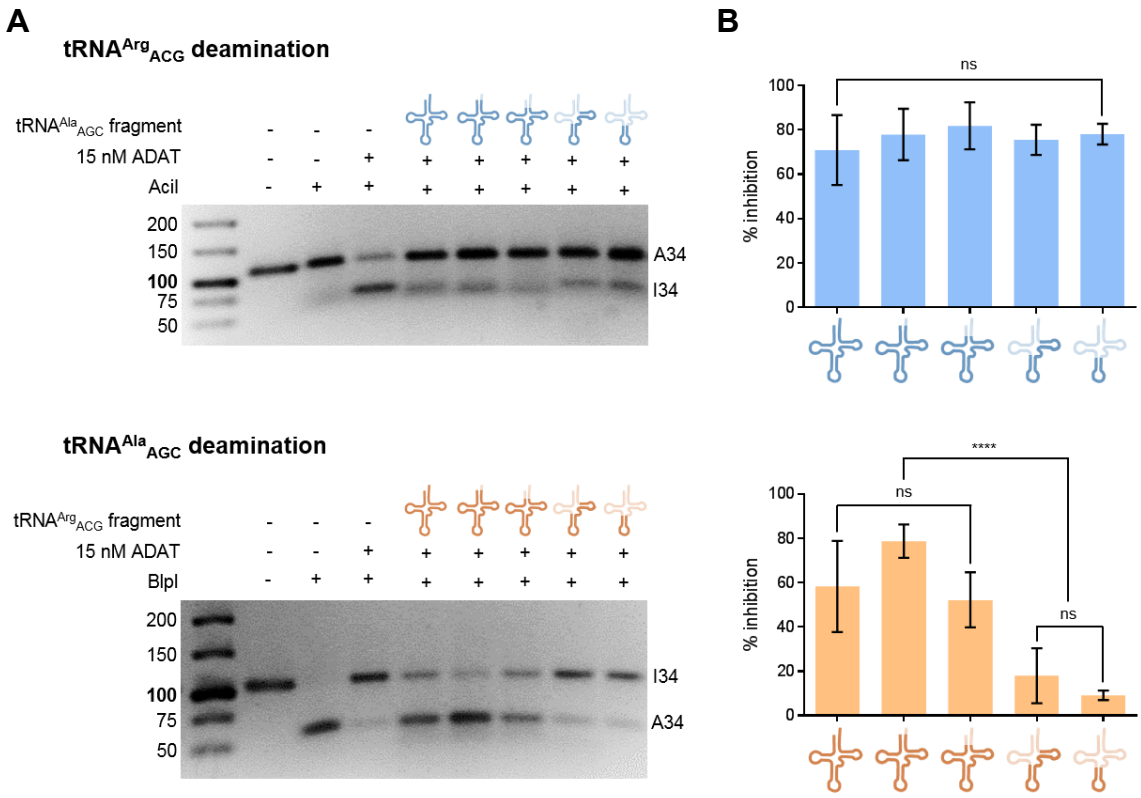


Figure 35. Analysis of the effect of tRNA fragments on ADAT-mediated deamination of tRNA$^{Arg}_{ACG}$ and tRNA$^{Ala}_{AGC}$.

A Representative RFLP experiments to assess ADAT-mediated deamination of tRNA$^{Arg}_{ACG}$ in the presence of 50 μM tRNA$^{Ala}_{AGC}$ fragments (upper panel) and tRNA$^{Ala}_{AGC}$ in the presence of 50 μM tRNA$^{Arg}_{ACG}$ fragments (lower panel).

B Percentage of inhibition caused by the tRNA fragments calculated from three independent experiments. All deamination reactions were done in triplicate and averaged. Error bars reflect the standard deviation. No statistical differences were observed between different tRNA fragments (P value = 0.6814).

ADAT substrate recognition might involve the acceptor stem of some tRNAs

We further explored the possibility that ADAT substrate recognition might depend on the top half of the tRNA molecule. To prove this hypothesis, we analyzed the effect of an acceptor stem analog derived from tRNA$^{Pro}_{AGG}$ (a known ADAT substrate) on ADAT-mediated deamination of tRNA$^{Arg}_{ACG}$ and tRNA$^{Ala}_{AGC}$. We used two variants of the acceptor stem analog: one consisting in a mini-helix (with the stem and a mini-loop), and the other one with the stem only (two strands annealed without the loop). Using the RFLP method

(13), we examined I34 formation in $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$ upon addition of 50 μM of both $tRNA^{Pro}_{AGG}$ acceptor stem analogs (**Figure 36**). The deamination of both tRNAs was significantly inhibited (around 80% of inhibition) by both acceptor stem analogs. No significant differences were observed between the two acceptor stem analogs in the deamination of $tRNA^{Arg}_{ACG}$. As for $tRNA^{Ala}_{AGC}$, slightly lower inhibition levels were observed in the analog with the mini-loop compared to the analog with the stem only. The fact that the analog lacking the loop shows similar levels of inhibition (or even higher) than the analog with the mini-loop reinforces the idea that the inhibition would be because of the interaction between the acceptor stem and ADAT, instead of due to similarity to the anticodon arm.

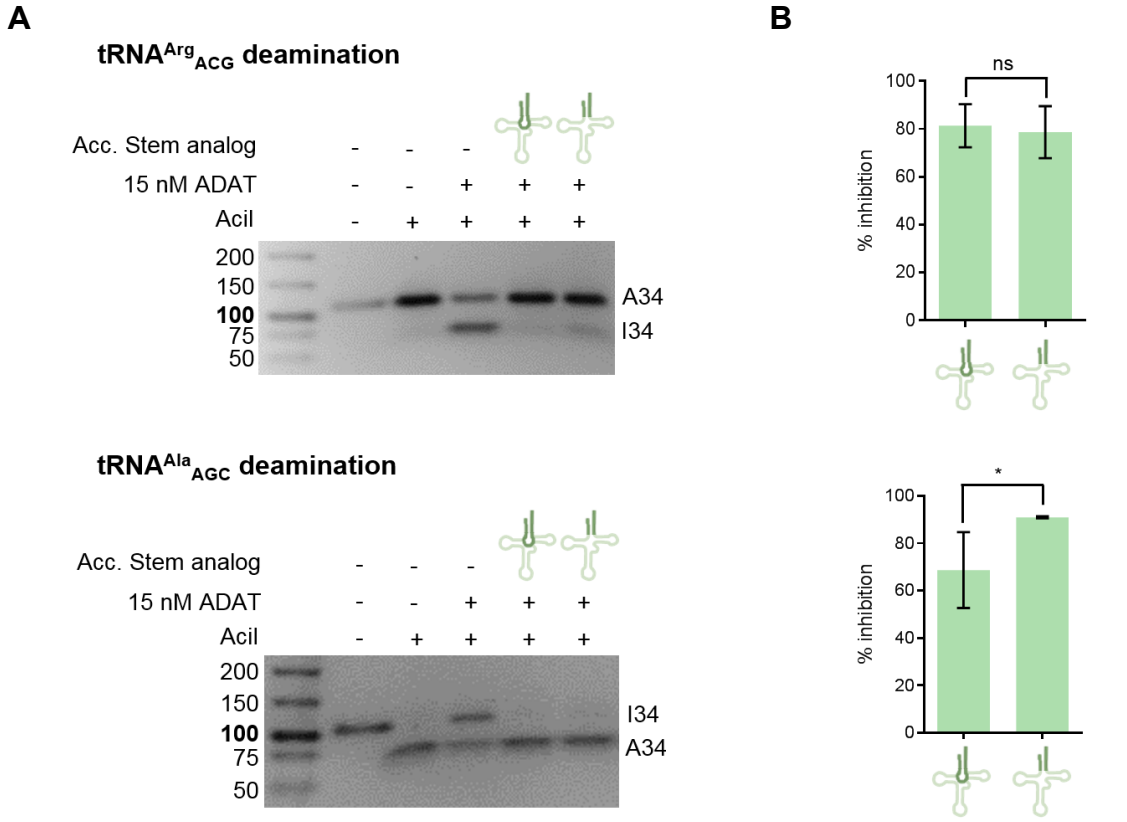


Figure 36. Analysis of the effect of an analog of the acceptor stem of $tRNA^{Pro}_{AGG}$ on ADAT-mediated deamination of $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$.

A Representative RFLP experiments to assess ADAT-mediated deamination of $tRNA^{Arg}_{ACG}$ (upper panel) and $tRNA^{Ala}_{AGC}$ (lower panel) upon addition of 50 μM of $tRNA^{Pro}_{AGG}$ acceptor stem analgos (with and without loop).

B Percentage of inhibition caused by the acceptor stem anlogs on $tRNA^{Arg}_{ACG}$ (upper panel) and $tRNA^{Ala}_{AGC}$ (lower panel) deamination calculated from three independent experiments. All deamination reactions were done in triplicate and averaged; the error bars reflect the standard deviation. Asterisks indicate statistical significance between samples.

The whole tRNA architecture is the main determinant for ADAT recognition

To further examine the interaction between ADAT and its substrates $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$, we designed chimeric tRNAs that combined the parts with the putative highest and lowest binding affinity. We synthesized a $tRNA^{Ala}_{AGC}$ with $tRNA^{Arg}_{ACG}$ anticodon arm and a $tRNA^{Arg}_{ACG}$ bearing $tRNA^{Ala}_{AGC}$ anticodon arm. Using RFLP method, we analyzed the deamination of the two chimeric tRNAs upon 15 nM of ADAT, and compared it to the deamination of the native ADAT substrates $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$ (**Figure 37**). The deamination results revealed that only $tRNA^{Arg}_{ACG}$ with $tRNA^{Ala}_{AGC}$ anticodon arm had deamination levels comparable to those for the native substrates $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$ (52.5 ± 5.1 % versus 65.5 ± 14.1 % and 77.5 ± 18.9 % in the native substrates). On the contrary, $tRNA^{Ala}_{AGC}$ with $tRNA^{Arg}_{ACG}$ anticodon arm showed extremely low levels of deamination (4.4 ± 1.8 %).

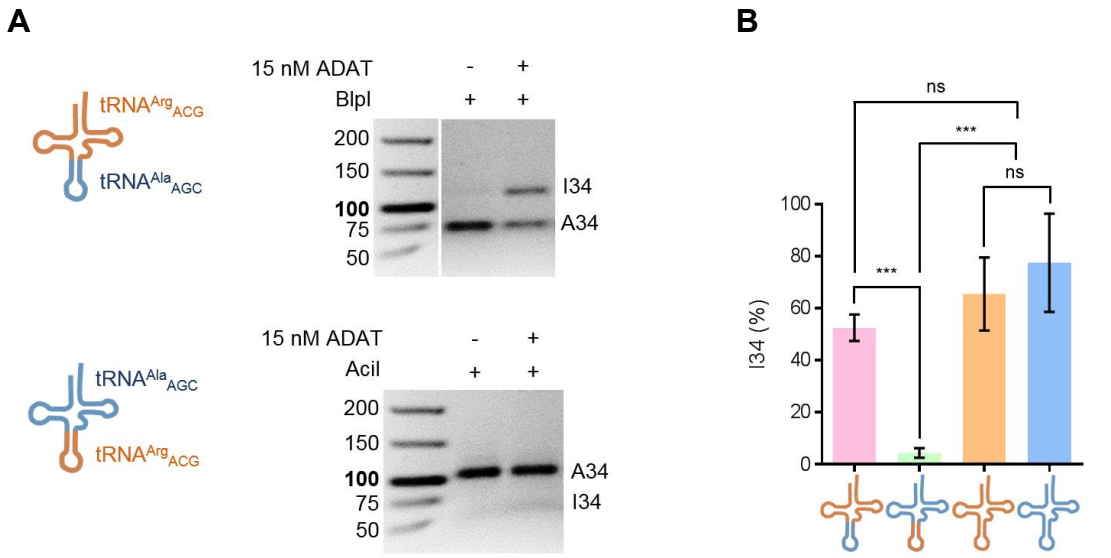


Figure 37. Analysis of ADAT-mediated deamination of chimeric tRNAs.

A Left panel: schematic representation of the two chimeric tRNAs: $tRNA^{Arg}_{ACG}$ with $tRNA^{Ala}_{AGC}$ anticodon arm (upper panel) and $tRNA^{Ala}_{AGC}$ with $tRNA^{Arg}_{ACG}$ anticodon arm (lower panel). Right panel: representative RFLP experiment to assess ADAT-mediated deamination of the two chimeric tRNAs.

B Levels of I34-modification from three independent experiments as in **A** compared to I34-levels in two native ADAT substrates: $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$, in the same reaction conditions. All deamination reactions were done in triplicate and averaged; the error bars reflect the standard deviation. Asterisks indicate statistical significance between samples.

The almost absence of deamination of $tRNA^{Ala}_{AGC}$ with $tRNA^{Arg}_{ACG}$ anticodon arm (the tRNA with putatively the highest binding affinity to ADAT) might be due to a very low dissociation constant (K_{off}) (i.e. the dissociation rate of the tRNA-ADAT complex is very low). On this basis, although the tRNA bound to the enzyme would be deaminated, since it represents a very small percentage of total substrate, the majority of the tRNA in the reaction mix would be unable to bind to the catalytic center (it is occupied) and would remain unmodified. Alternatively, might be that, the combination of two parts from different tRNAs, despite preserving the typical cloverleaf secondary structure (**Supplementary Figure S8**), would result in the loss of important interactions for the 3D structure that would alter the tRNA architecture and lead to an inactive substrate. We are currently performing binding assays to evaluate the affinity between the tRNA and ADAT.

Overall, the fact that combining the parts with the putative highest and lowest binding affinity does not lead to the expected increase and reduction in the respective deamination rates suggests that the recognition of the tRNAs is not based on specific regions of the tRNA but it probably involves the whole tRNA structure. Thus, the entire architecture of the tRNA is possibly a key feature for ADAT recognition and activity.

Some of the most abundant tRNA-derived fragments (tRFs) in human cells can modulate ADAT activity

The inhibition of ADAT deamination by tRNA fragments opened the possibility that ADAT activity might be regulated by small RNA fragments. We searched for the most abundant tRFs in human cells reported at MINTbase (62, 162) and we selected a total of six tRFs, two derived from $tRNA^{Arg}_{ACG}$, three from $tRNA^{Ala}_{AGC}$, and one from $tRNA^{Cys}_{GCA}$ (**Figure 38A**). Using RFLP method (13), we analyzed the effect of 50 μM of each tRF on ADAT-mediated deamination of $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$. We did not analyze the effect of 5'-half derived from $tRNA^{Ala}_{AGC}$ on $tRNA^{Ala}_{AGC}$ deamination since we could not selectively amplify $tRNA^{Ala}_{AGC}$ without amplifying the $tRNA^{Ala}_{AGC}$ 5'-half. For the same reason, we did not analyze the effect of $tRNA^{Arg}_{ACG}$ tRFs on $tRNA^{Arg}_{ACG}$ deamination. 5'-half and 3'-tRF derived from $tRNA^{Ala}_{AGC}$ and 5'-half derived from $tRNA^{Cys}_{GCA}$ showed to significantly inhibit the deamination of $tRNA^{Arg}_{ACG}$. In the case of $tRNA^{Ala}_{AGC}$, it was the 5'-tRF and 3'-tRF derived from $tRNA^{Ala}_{AGC}$ and the 5'-half derived from $tRNA^{Cys}_{GCA}$ the ones that showed the highest inhibitory effect (**Figure 38B and 38C**).

We considered the possibility that the inhibition was due to annealing between the tRF and the tRNA substrate, which might disrupt the tRNA architecture and inhibit deamination. This could be possible in the cases in which the tRFs derived from the same tRNA used as a substrate, e.g. when adding $tRNA^{Ala}_{AGC}$ tRFs in $tRNA^{Ala}_{AGC}$ deamination reaction. However, in case of 5'-half and 3'-tRF derived from $tRNA^{Ala}_{AGC}$, in which we obtained significantly high levels of inhibition of $tRNA^{Arg}_{ACG}$ deamination, the identity between RNA sequences is not sufficient for the annealing to occur. The same applies for the inhibition of $tRNA^{Ala}_{AGC}$ deamination by 5'-half derived from $tRNA^{Cys}_{GCA}$.

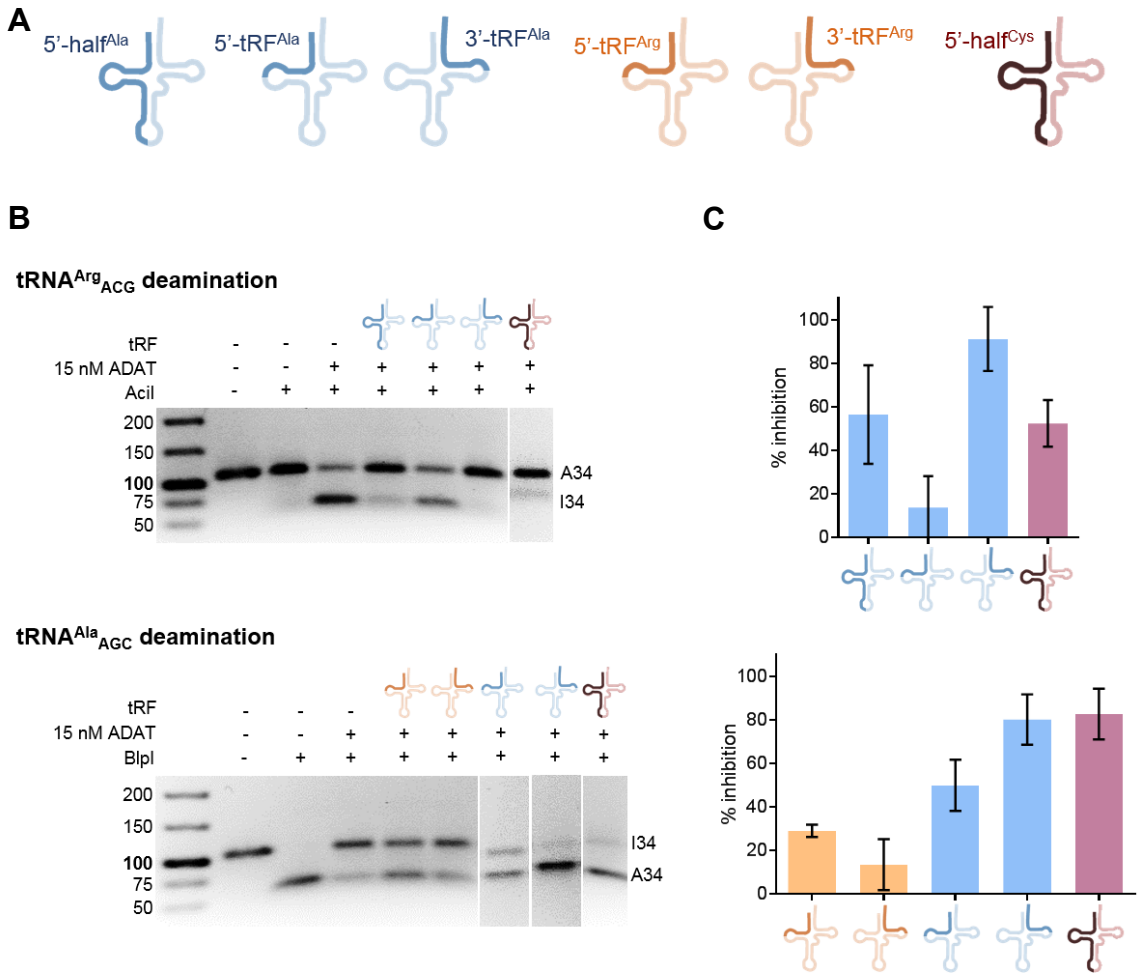


Figure 38. Analysis of the effect of human tRFs on ADAT-mediated deamination of $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$.

A Schematic representation of the human tRFs analyzed: a 5'-tRF and a 3'-tRF derived from $tRNA^{Arg}_{ACG}$; a 5'-half, a 5'-tRF and a 3'-tRF derived from $tRNA^{Ala}_{AGC}$; and a 5'-half derived from $tRNA^{Cys}_{GCA}$.

B Representative RFLP experiment to assess ADAT-mediated deamination of $tRNA^{Arg}_{ACG}$ (upper panel) and $tRNA^{Ala}_{AGC}$ (lower panel) upon addition of 50 μM human tRFs.

C Percentage of inhibition of ADAT-mediated deamination of $tRNA^{Arg}_{ACG}$ (upper panel) and $tRNA^{Ala}_{AGC}$ (lower panel) caused by tRFs calculated from three independent experiments as in **B** and averaged. Error bars reflect the standard deviation.

To further characterize ADAT inhibition by the tRFs, we built dose-response curves of the tRFs that exhibited the highest inhibitory activity and estimated the corresponding half maximal inhibitory concentrations (IC50 values) (**Figure 39 and Supplementary Figure S9**). Dose-response curves showed a dose-dependent inhibitory effect, with the highest inhibition level observed at a 500 μM concentration. The highest inhibition of $tRNA^{Arg}_{ACG}$ (and lowest IC50 value) was observed with the 3'-tRF derived from $tRNA^{Ala}_{AGC}$, whereas in the case of $tRNA^{Ala}_{AGC}$ it was the 5'-half derived $tRNA^{Cys}_{GCA}$ the fragment with the highest inhibitory activity.

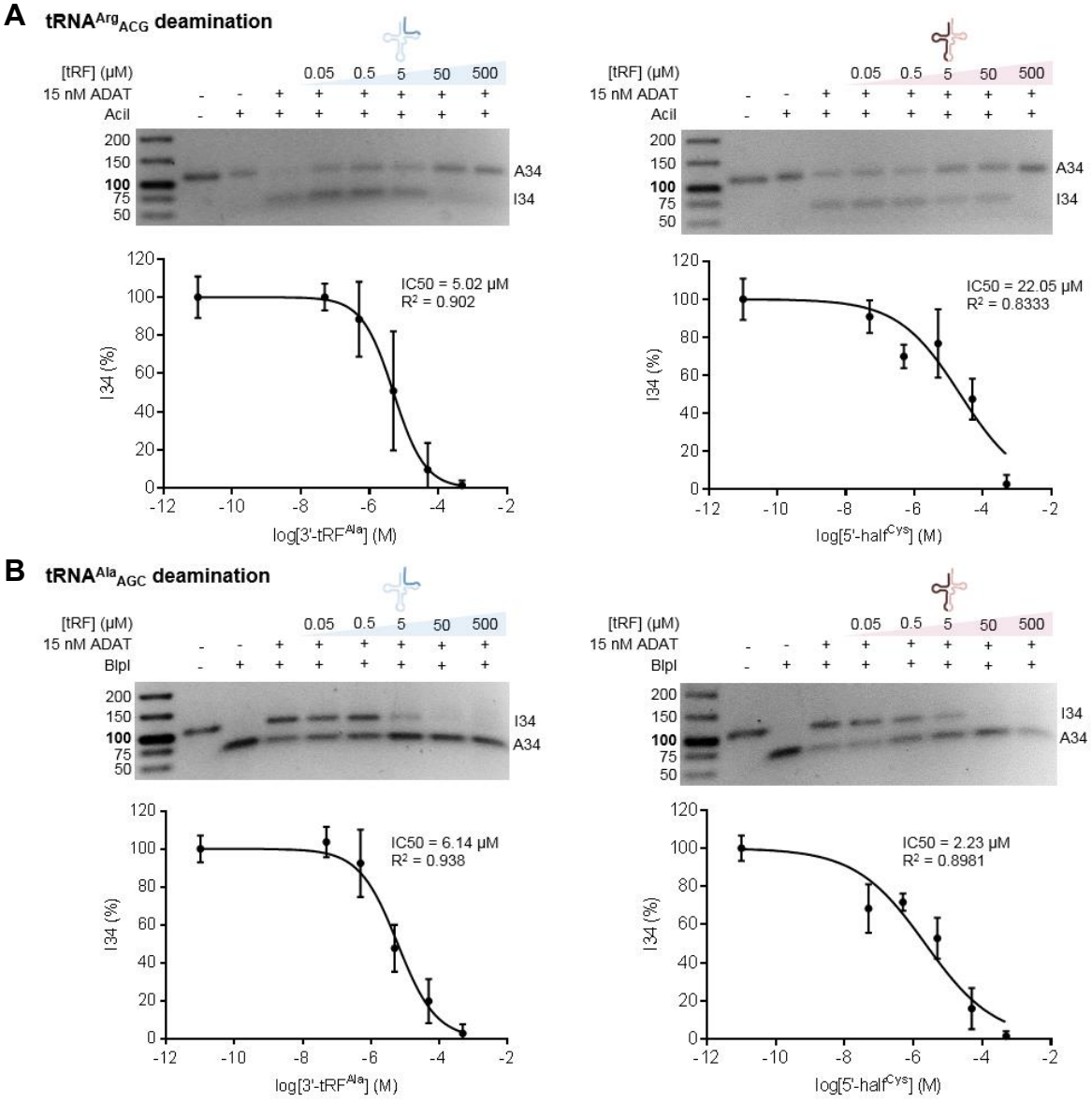


Figure 39. Dose-response curves for the tRFs with the highest inhibitory effect on ADAT-mediated deamination of tRNA$^{Arg}_{ACG}$ (A) and tRNA$^{Ala}_{AGC}$ (B). Upper panels: representative RFLP experiments. Lower panels: percentage of I34 modification upon varying concentrations of tRFs from three independent experiments and averaged. Error bars reflect the standard deviation. IC50 values are indicated.

Study of ADAT structure and kinetics and effect of a mutation linked to intellectual disability

Rodríguez-Escribà M, **Roura Frigolé H***, Camacho N, García-Lema J, Filonava L, Castellví Coma M, Skaros A, Canals A, Coll M, Stracker TH, Gallego O, Torres AG, Ribas de Pouplana L. Activity and stability of ADAT are compromised by a mutation linked to human intellectual disability. J Biol Chem. 2018. Manuscript under review.*

**** First co-authors (written in alphabetical order).***

Abstract

The eukaryotic heterodimeric ADATs are enzymes that catalyze the deamination of adenosine to inosine at the first position of the anticodon; one of the few essential tRNA modifications described to date. They are composed of two subunits: ADAT2 and ADAT3. A missense mutation (V128M) in human ADAT3 has been linked to intellectual disability and strabismus and other clinical manifestations named ADAT3-related intellectual disability syndrome. However, the molecular bases of the pathology are unknown. Here we characterize human ADAT in terms of quaternary structure and kinetics, and investigate the effect of the V128M mutation. We report that this substitution reduces the stability of ADAT3 and decreases the tRNA editing activity of ADAT. Altogether these results shed light on the biochemical repercussions of the mutation and might open a door for the development of new therapeutic strategies.

Human ADAT structure and effect of V128M mutation

We studied the quaternary structure of human ADAT by size-exclusion chromatography (SEC) analysis of affinity-purified enzyme. ADAT wild type (wt) eluted as a double peak in fractions 28-34 (**Figure 40A**). The elution volume (Ve) of the first peak was consistent with the expected molecular weight (Mw) of the ADAT heterodimer (Ve = 14.8 mL) (**Figures 40A and 43 and Table 5**), and the corresponding fractions displayed an optical absorption maximum at 280 nm, indicating a predominant protein content (260/280 ratio: 0.67). On the other hand, fractions corresponding to the second peak absorbed mainly at 260 nm, suggesting a predominant presence of nucleic acids (260/280 ratio: 1.86). SDS-PAGE and Western blot (WB) analyses confirmed the presence of both ADAT subunits in the fractions of the two peaks, albeit a higher concentration in the first peak (**Figure 40B, 40C and 40E**).

Urea-PAGE analysis confirmed the presence of nucleic acids of tRNA-like size in fractions 30-34 (**Figure 40B and 40D**). SEC coupled to multiangle light scattering (MALS) analysis assigned a Mw of around 60 kDa to the first peak, consistent with the Mw of ADAT heterodimer, and a Mw of 22 kDa to the predominant molecular species in the second one, consistent with that of a tRNA (**Supplementary Figure S10**). These results indicate that the affinity purification of ADAT wt results in the enrichment of a tRNA-ADAT complex that dissociates in two elution peaks after SEC. Unexpectedly, the peak of ADAT coeluting with RNA had a higher Ve than the peak of ADAT alone. This suggests that ADAT might adopt a more compact conformation upon binding to tRNAs, which results into a smaller size and a higher Ve for ADAT-RNA complexes (bigger particles elute faster than small ones).

To eliminate the RNA coeluting with ADAT in the SEC, we treated the affinity-purified enzyme with RNase prior to the injection to the size-exclusion column. Heterodimeric ADAT treated with RNase eluted as a single peak (fractions 28-32, Ve = 14.5 mL), consistent with the expected Mw of ADAT heterodimer. The second and main peak in the chromatogram corresponded to the RNase (fractions 34-36, Ve = 17.5 mL) (**Figures 41A and 43 and Table 5**). SDS-PAGE and WB analyses confirmed the presence of both ADAT subunits in the fractions corresponding to the first peak, and of RNase in the second peak (**Figure 41B and 41D**). Urea-PAGE of the fractions corresponding to the peaks showed no bands, indicating that the RNA coeluting with ADAT was digested (**Figure 41C**).

We then examined whether ADAT quaternary structure might be affected by V128M. SEC analysis of affinity-purified ADAT V128M generated a different elution profile to that of ADAT wt (**Figure 42A**). First, a protein peak was observed in the void volume (Vo = 7.7 mL, fractions 16-17). SDS-PAGE analysis of these fractions revealed the presence of a protein of molecular weight similar to ADAT3, suggestive of ADAT3 aggregates (**Figure 42C**). The high molecular nature of these aggregates was confirmed by SEC-MALS (**Supplementary Figure S10**). Unfortunately, WB and mass spectrometry analyses of these fractions were unable to ascertain the identity of this aggregated protein.

The mutant heterodimeric enzyme was detected in fractions 28-30 (Ve = 14.9 mL) by SDS-PAGE and WB analyses (**Figure 42A, 42C and 42E**). Finally, a third peak in fractions 31-34 (Ve = 16.5 mL) contained ADAT2 and a truncated form of ADAT3. The short ADAT3 had a Mw similar to ADAT2 since the two bands overlapped in the SDS-PAGE (**Figure 42C**), however, WB analysis confirmed the presence of a truncated form of ADAT3 (**Figure 42E**). This cleaved ADAT heterodimer coeluted with nucleic acids of tRNA size, as determined by urea-PAGE analysis (**Figure 42D**). This might indicate that the conformational change due to tRNA binding, also observed in the wt enzyme, exposes a proteolytic cleavage site exclusively found in ADAT3-V128M. Altogether, these results suggest that V128M mutation promotes the cleavage of ADAT3 subunit, generating a truncated ADAT heterodimer.

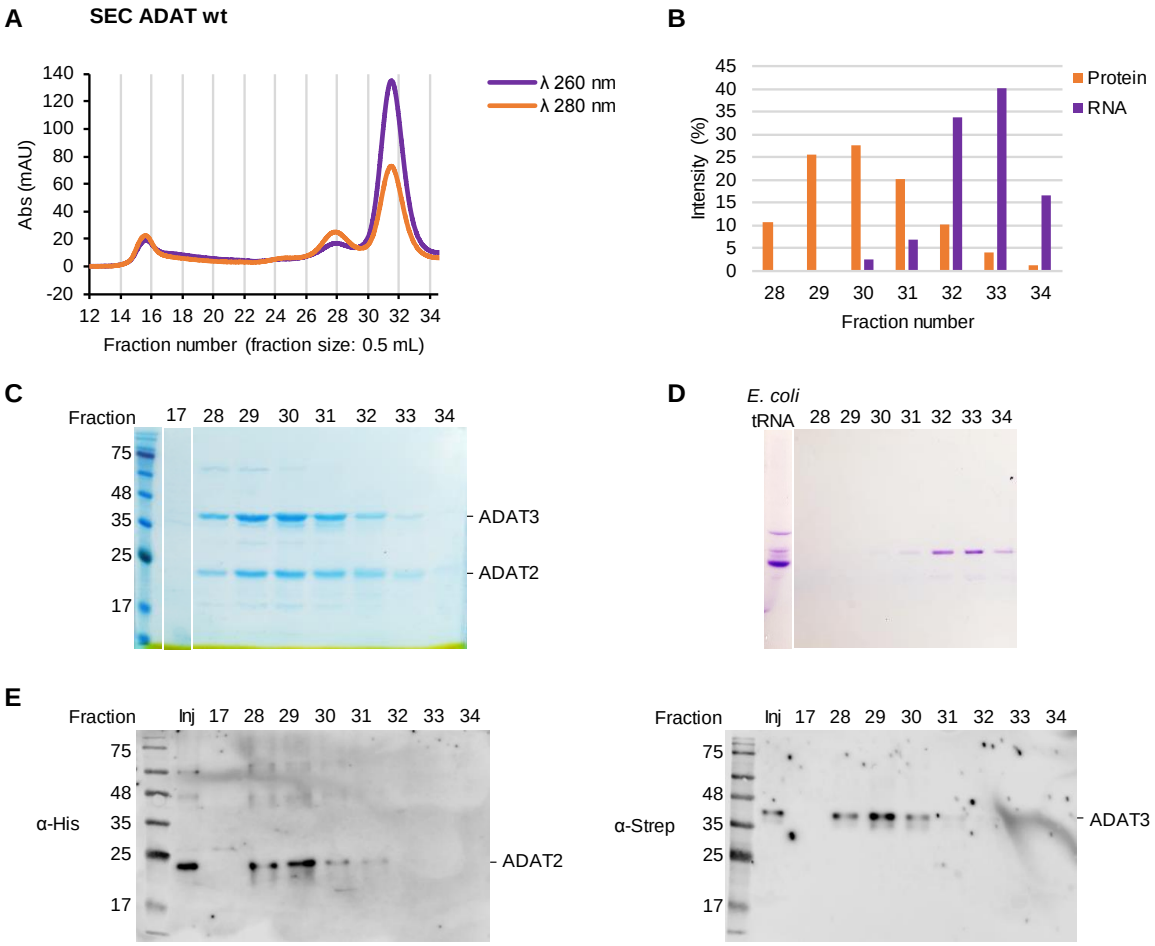


Figure 40. Analysis of ADAT quaternary structure by SEC.

A Representative size-exclusion chromatogram of ADAT wt**.** Elution fraction numbers (fraction size: 0.5 mL) are represented on the x-axis, and the UV absorbance at 260 nm (purple) and 280 nm (orange) is represented on the y-axis.

B Variation in protein and RNA content of the size exclusion-eluted fractions for ADAT wt. Values are normalized to the total intensity of the gel.

C Representative SDS-PAGE and BlueSafe staining of the size-exclusion eluted fractions corresponding to absorption peaks for ADAT wt. Lanes are labelled with their elution fraction number. ADAT2 (22 kDa) and ADAT3 (39 kDa) bands are indicated.

D Representative urea-PAGE and toluidine blue staining of the protein-containing fractions for ADAT wt. Lanes are labelled with their elution fraction number. *E. coli* total tRNA was added as a control (first lane).

E Representative Western blot of the injected sample and SEC-eluted fractions corresponding to absorption peaks for ADAT wt. ADAT2 subunit (fused to His-tag) is detected with an anti-His primary antibody. ADAT3 subunit (fused to Strep-tag) is detected with an anti-Strep primary antibody. The observed bands are consistent with those detected by BlueSafe staining in **C**.

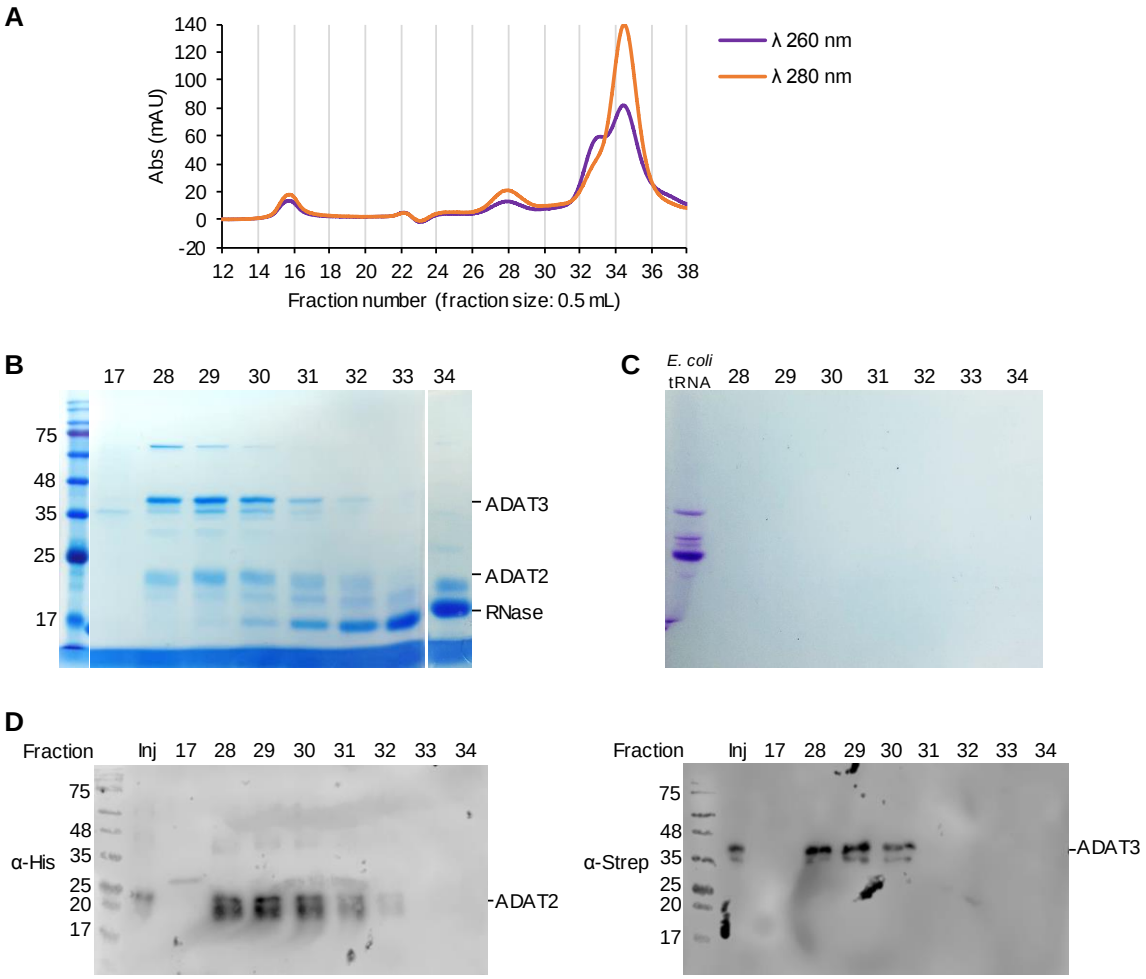


Figure 41. Analysis of ADAT quaternary structure by SEC upon RNase treatment.

A Representative size-exclusion chromatogram of ADAT wt after RNase treatment**.** Elution fraction numbers (fraction size: 0.5 mL) are represented on the x-axis, and the UV absorbance at 260 nm (purple) and 280 nm (orange) is represented on the y-axis.

B Representative SDS-PAGE and BlueSafe staining of the size-exclusion eluted fractions corresponding to absorption peaks for ADAT wt. Lanes are labelled with their elution fraction number. ADAT2 (22 kDa) and ADAT3 (39 kDa) bands are indicated.

C Representative urea-PAGE and toluidine blue staining of the protein-containing fractions for ADAT wt. Lanes are labelled with their elution fraction number. *E. coli* total tRNA was added as a control (first lane).

D Representative Western blot of the injected sample and SEC-eluted fractions corresponding to absorption peaks for ADAT wt after RNase treatment. ADAT2 subunit (fused to His-tag) is detected with an anti-His primary antibody. ADAT3 subunit (fused to Strep-tag) is detected with an anti-Strep primary antibody. The observed bands are consistent with those detected by BlueSafe staining in **B**.

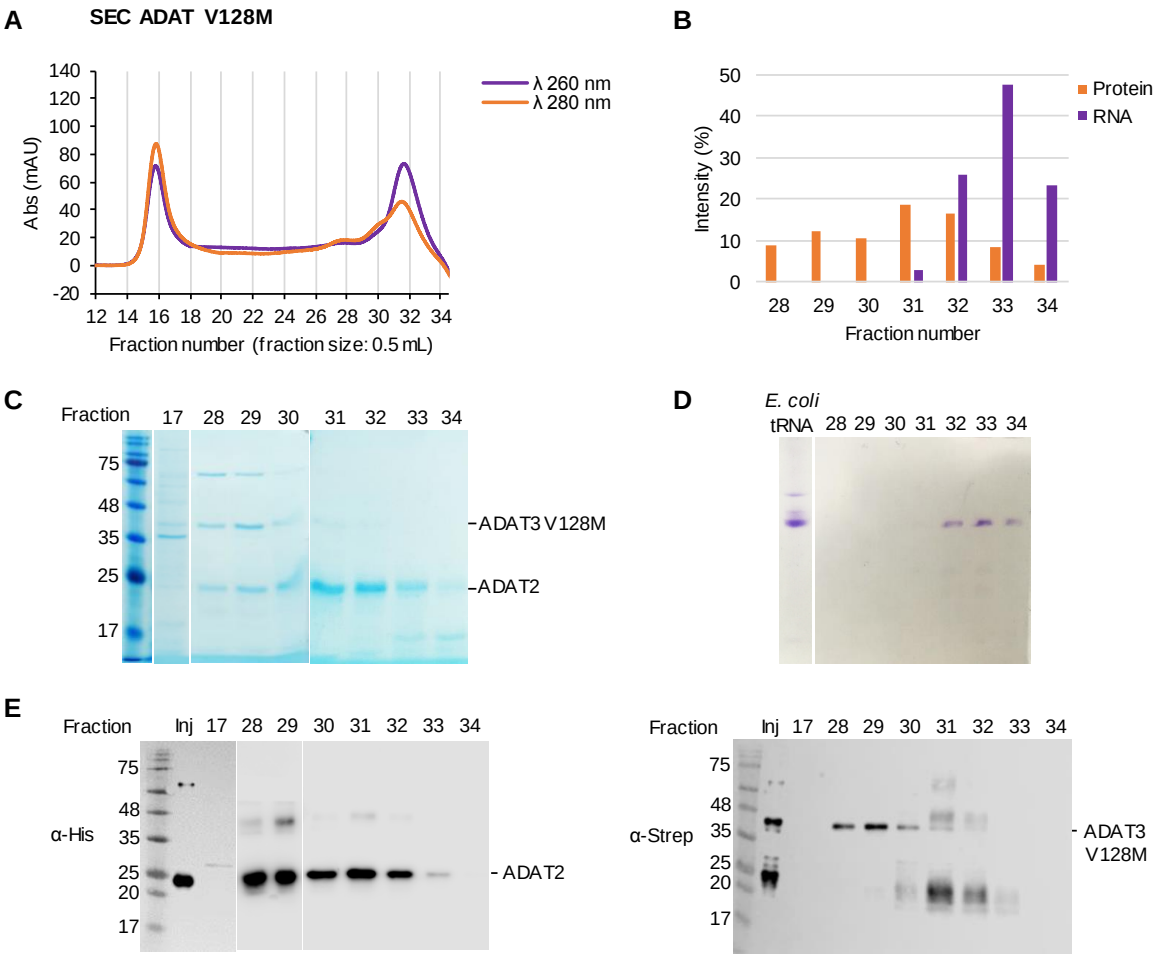


Figure 42. Analysis of ADAT V128M quaternary structure by SEC.

A Representative size-exclusion chromatogram of ADAT V128M**.** Elution fraction numbers (fraction size: 0.5 mL) are represented on the x-axis, and the UV absorbance at 260 nm (purple) and 280 nm (orange) is represented on the y-axis.

B Variation in protein and RNA content of the size exclusion-eluted fractions for ADAT V128M. Values are normalized to the total intensity of the gel.

C Representative SDS-PAGE and BlueSafe staining of the size-exclusion eluted fractions corresponding to absorption peaks for ADAT V128M. Lanes are labelled with their elution fraction number. ADAT2 (22 kDa) and ADAT3 (39 kDa) bands are indicated.

D Representative urea-PAGE and toluidine blue staining of the protein-containing fractions for ADAT V128M. Lanes are labelled with their elution fraction number. *E. coli* total tRNA was added as a control (first lane).

E Representative Western blot of the injected sample and SEC-eluted fractions corresponding to absorption peaks for ADAT V128M. ADAT2 subunit (fused to His-tag) is detected with an anti-His primary antibody. ADAT3 subunit (fused to Strep-tag) is detected with an anti-Strep primary antibody. The observed bands are consistent with those detected by BlueSafe staining in **C**.

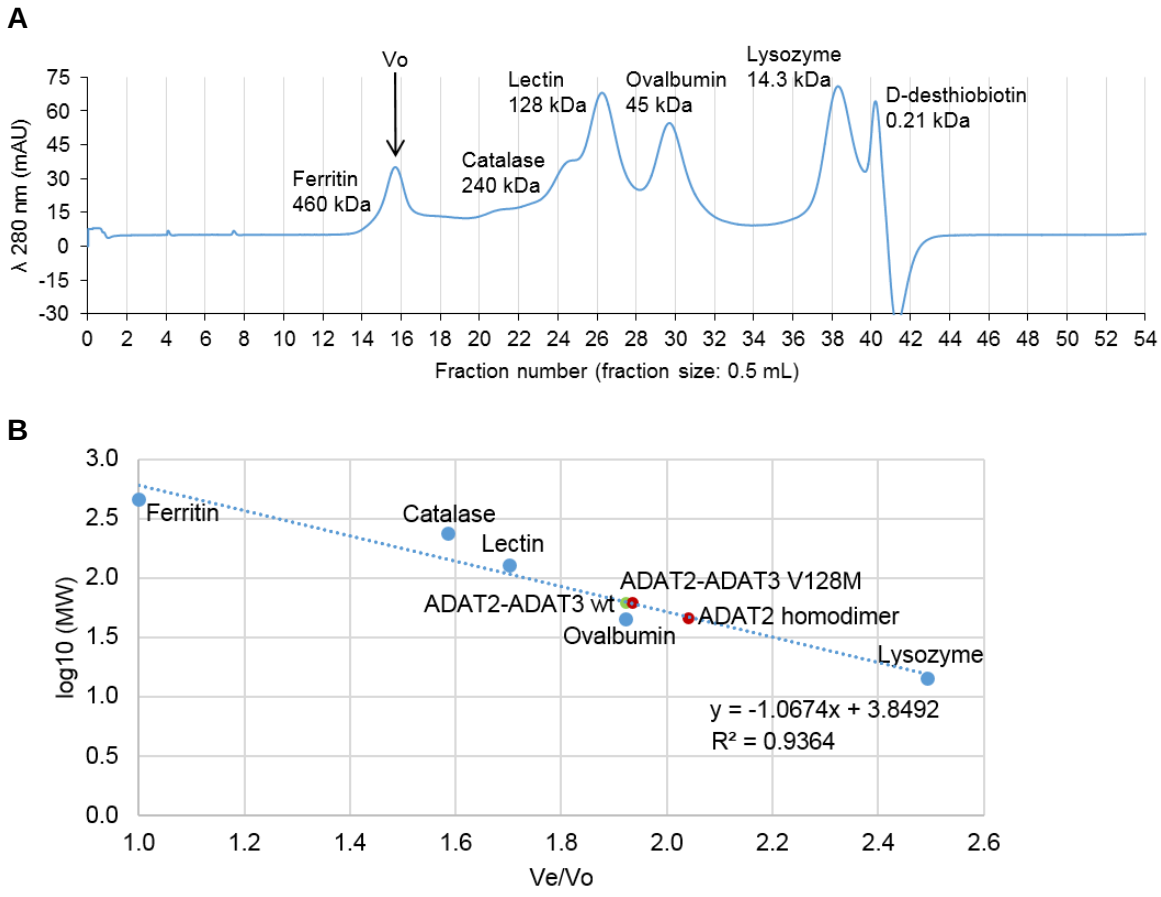


Figure 43. SEC analysis of protein standards. Protein standards of known molecular weights (ferritin, catalase, lectin, ovalbumin and lysozyme) were injected onto the same size-exclusion column, and their elution peaks were used to determine the relationship between the elution volume (Ve) and molecular weight (Mw). The Ve of each protein standard was determined by UV absorbance at 280 nm. The void volume (Vo) of the column was taken as the Ve of ferritin (Vo=7.7 mL).

A Size-exclusion chromatogram of protein standards. Elution fraction numbers (fraction size: 0.5 mL) are shown on the x-axis and the UV absorbance at 280 nm is shown on the y-axis.

B Plot of log10(Mw) against Ve/Vo ratios and standard curve, log10(Mw)=-1.0674Ve/Vo+3.8492, R^2=0.9364. Protein standards and the standard curve are represented in blue, ADAT wt-eluted fractions in green, and ADAT V128M-eluted fractions in dark red.

	Mw (kDa)	log10 (Mw)	Ve/Vo
SEC standards			
Lysozyme	14.3	1.2	2.5
Ovalbumin	45.0	1.7	1.9
Lectin	128.0	2.1	1.7
Catalase	240.0	2.4	1.6
Ferritin	460.0	2.7	1.0
ADAT wt			
ADAT2-ADAT3 wt	62.5	1.8	1.9
ADAT V128M			
ADAT2-ADAT3 V128M	62.5	1.8	1.9
ADAT2 homodimer	46.4	1.7	2.0
ADAT2 monomer	23.2	1.4	

Table 5. Mw and Ve/Vo values of the protein standards, ADAT wt, and ADAT V128M samples.

Characterization of the enzymatic activity of human ADAT and effect of V128M mutation

We assessed the deamination activity of human ADAT using $tRNA^{Ala}_{AGC}$ as a substrate by the RFLP method (13). All deamination reactions were done using affinity-purified ADAT enzyme and *in vitro* transcribed tRNA. After PCR amplification of the tRNAs, a single ~100 bp band was observed in all samples (first lane in **Figure 44A** is shown as a representative example). Deamination of $tRNA^{Ala}_{AGC}$ results in a decrease in BlpI digestion products (~50 bp) and an increase of undigested amplicon (~100 bp), which can be visualized by agarose gel electrophoresis. **Figure 44** shows the deamination activity of ADAT wt at different enzyme concentrations over time. As expected, deamination increased as higher amounts of enzyme and longer incubation times were used. At a 50 nM concentration of ADAT wt, nearly complete deamination (92.3 ± 10.8 %) was obtained after 4 hours of incubation (**Figure 44 and Table 6**).

We next estimated the *in vitro* catalytic efficiency of ADAT V128M mutant by the same method, using affinity-purified mutant enzyme and under the same reaction conditions. We observed a significant decrease in the deamination activity for the mutant as compared to the wt enzyme. At a concentration of 50 nM, ADAT V128M was only able to modify 42.0 ± 29.2 % of the tRNA substrate after 4 hours of incubation. Still, increasing the concentration of mutant enzyme resulted in higher levels of I34. Full deamination of $tRNA^{Ala}_{AGC}$ with ADAT mutant was obtained at an enzyme concentration of 800 nM (**Figure 45 and Table**

6). Overall, we estimated a 6-fold decrease in activity for the mutant enzyme (the reduction in fold change (FC) between wt and mutant at higher enzyme concentrations can be explained because the wt is close to the saturation plateau whereas the mutant is at the initial linear phase) (**Figure 46**).

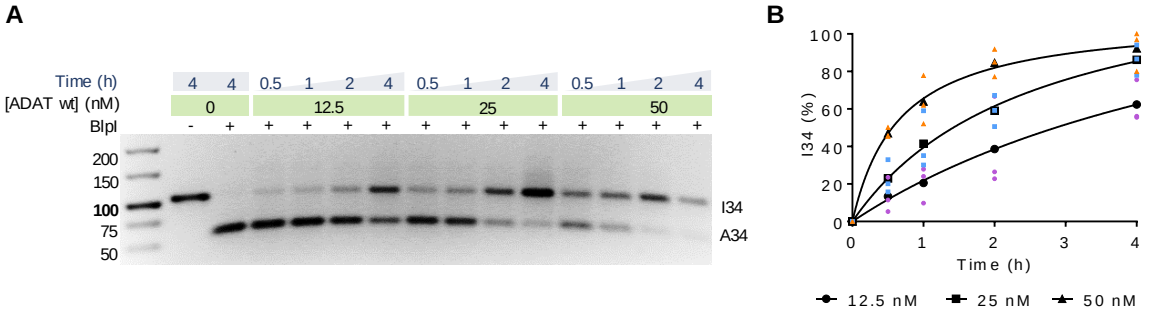


Figure 44. ***In vitro*** **enzymatic activity assays of human ADAT wt for tRNA$^{Ala}_{AGC}$ by RFLP method.**

A Representative RFLP experiment to assess the deamination activity of ADAT wt. Inosine formation was monitored at the indicated enzyme concentrations and time points.

B Levels of I34-modification from three independent experiments as in **A**. Mean values are shown in black and the values for each replicate are color-coded (orange for 50 nM ADAT wt, blue for 25 nM ADAT wt and purple for 12.5 nM ADAT wt). Curve fitting was performed following a nonlinear hyperbolic regression model (R^2=0.7950 for 12.5 nM ADAT wt; R^2=0.9331 for 25 nM ADAT wt; R^2=0.9570 for 50 nM ADAT wt).

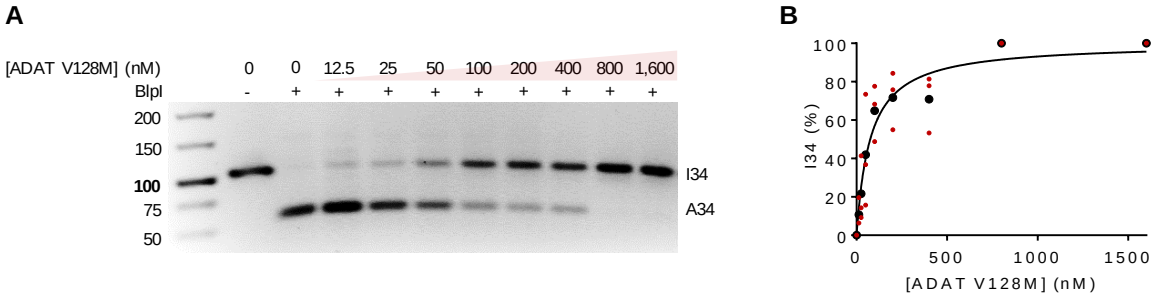


Figure 45. ***In vitro*** **enzymatic activity assays of human ADAT V128M mutant for tRNA$^{Ala}_{AGC}$ by RFLP method.**

A Representative RFLP experiment to assess the deamination activity of ADAT V128M. Inosine formation was monitored at the indicated enzyme concentrations upon 4 hours of incubation.

B Levels of I34-modification from three independent experiments as in **A**. Mean values are shown in black and the values for each replicate in dark red. Curve fitting was performed following a nonlinear hyperbolic regression model (R^2=0.8502).

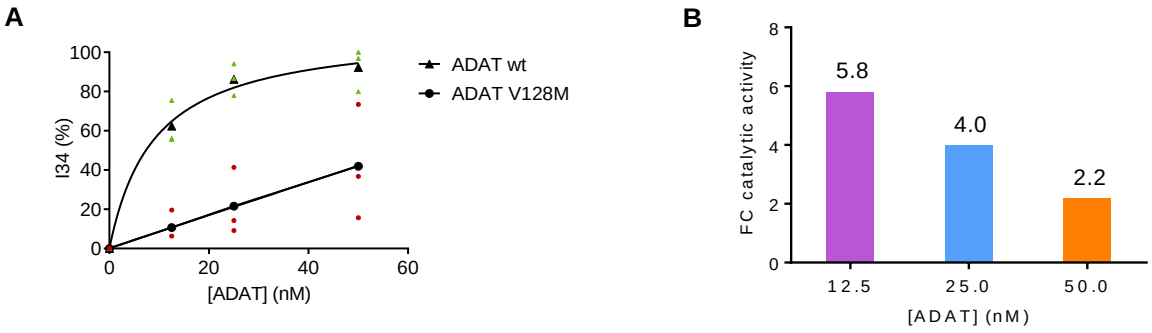


Figure 46. Differential catalytic activity between ADAT wt and ADAT V128M.

A Levels of I34-modification against enzyme concentration, for ADAT wt and ADAT V128M, after 4 hours of incubation. Mean values are shown in black and the values for each replicate are color-coded (green for ADAT wt and dark red for ADAT V128M). Curve fittings were performed following a nonlinear hyperbolic regression model for ADAT wt (R^2=0.9574) and a linear regression model for ADAT V128M (R^2=0.9997).

B Fold change in catalytic activity between ADAT wt and ADAT V128M at the studied concentrations.

[ADAT] (nM)	Time (h)	ADAT wt		ADAT V128M	
		% I34	SD	% I34	SD
12.5	0.5	13.4	9.3	-	-
	1	20.6	9.5	-	-
	2	38.7	24.3	-	-
	4	62.4	11.4	10.7	7.7
25	0.5	23.0	8.9	-	-
	1	41.4	15.5	-	-
	2	59.0	8.3	-	-
	4	86.2	8.1	21.6	17.3
50	0.5	47.1	2.7	-	-
	1	63.9	13.0	-	-
	2	84.6	7.3	-	-
	4	92.3	10.8	42.0	29.2
100	4	-	-	64.9	14.7
200	4	-	-	71.7	15.1
400	4	-	-	70.8	15.3
800	4	-	-	100.0	0.0
1,600	4	-	-	100.0	0.0

Table 6. Levels of I34-modification for ADAT wt and mutant obtained by RFLP at the indicated enzyme concentrations and time points. The represented values correspond to the average of triplicates. The associated standard deviations (SD) are indicated.

To further study the kinetics of the wt and mutant ADATs we analyzed the effect of the substrate concentration on the enzyme activity. All *in vitro* deamination reactions were done using a fixed enzyme concentration (20 and 100 nM of affinity-purified wt and V128M ADAT, respectively) and varying substrate concentrations (from 3 to 80 µM of *in vitro* transcribed $tRNA^{Ala}_{AGC}$). We observed an increase in the reaction rate with increasing substrate concentration until reaching the saturation plateau in both enzymes (**Figure 47**). Determination of the kinetic constants of ADAT wt (K_m: 2.98 µM, k_{cat}: 0.17 min^{-1}) produced values comparable to those reported for *T. brucei* ADAT and *E. coli* TadA (115, 163). Comparison between wt and mutant enzymes revealed a 10-fold decrease in the catalytic efficiency (k_{cat}/K_m) for ADAT mutant with respect to ADAT wt, a reduction mainly due to a 5-fold reduction in the k_{cat} value for ADAT V128M (K_m: 5.81 µM, k_{cat}: 0.03 min^{-1}). V_{max} values remained similar between the two enzymes (**Table 7**).

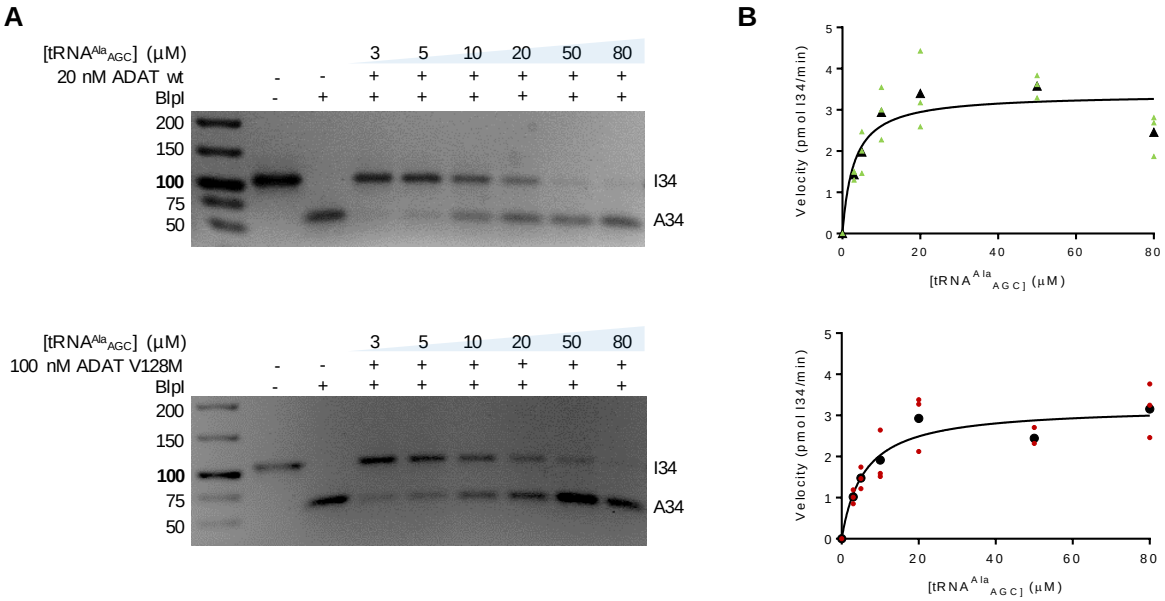


Figure 47. Kinetics of ADAT wt and mutant for $tRNA^{Ala}_{AGC}$ by RFLP method.

A Representative RFLP experiments to monitor the deamination activity of ADAT wt (upper panel) and ADAT V128M (lower panel) at varying $tRNA^{Ala}_{AGC}$ substrate concentrations. Inosine formation was monitored after 1 hour of incubation with 20 nM of ADAT wt and 100 nM of ADAT V128M respectively.

B Deamination rate (velocity) against substrate concentration for ADAT wt (upper panel) and ADAT V128M (lower panel) from three independent experiments as in **A**. Mean values are shown in black and the values for each replicate are color-coded (green for ADAT wt and dark red for ADAT V128M). Curve fitting was performed following a nonlinear hyperbolic regression model (R^2=0.7687 for ADAT wt; R^2=0.8347 for ADAT V128M).

	ADAT wt	ADAT V128M	FC
k_{cat} (min^{-1})	0.17	0.03	5.27
K_m (μM)	2.98	5.81	0.51
V_{max} (pmol/min)	3.39	3.21	1.05
k_{cat}/K_m	0.06	0.01	10.28

Table 7. Kinetic constants of ADAT wt and ADAT V128M for the deamination of $tRNA^{Ala}_{AGC}$. FC: fold change ADAT wt/ADAT V128M.

Analysis of the catalytic activity of ADAT2 homodimers

We explored the possibility that ADAT2 subunit alone (which contains ADAT catalytic center) might be catalytically active. For this, human ADAT2 was heterologously expressed in *E. coli* and purified by affinity chromatography. SEC of the purified ADAT2 resulted in a single peak with an elution volume consistent with ADAT2 homodimers (**Figure 48 and Table 8**), which is in agreement with previously reported ability of ADAT2 to homodimerize (1, 115).

Using RFLP method (13), we *in vitro* analyzed the deamination of $tRNA^{Ala}_{AGC}$ and $tRNA^{Arg}_{ACG}$ by ADAT2 homodimers. All *in vitro* deamination reactions were done using the affinity-purified ADAT2 at the indicated concentrations and *in vitro* transcribed tRNAs. The reaction conditions were the same as for the activity assays with the heterodimeric ADAT. No deamination was observed in any of the two tRNAs by ADAT2 at the nanomolar concentration range (**Figure 49 left panel and Supplementary Figure S11**). By contrast, when increasing ADAT2 concentration to the micromolar range, both $tRNA^{Ala}_{AGC}$ and $tRNA^{Arg}_{ACG}$ showed deamination. Full deamination was observed at 800 μM of ADAT2 (**Figure 49 right panel and Supplementary Figure S11**).

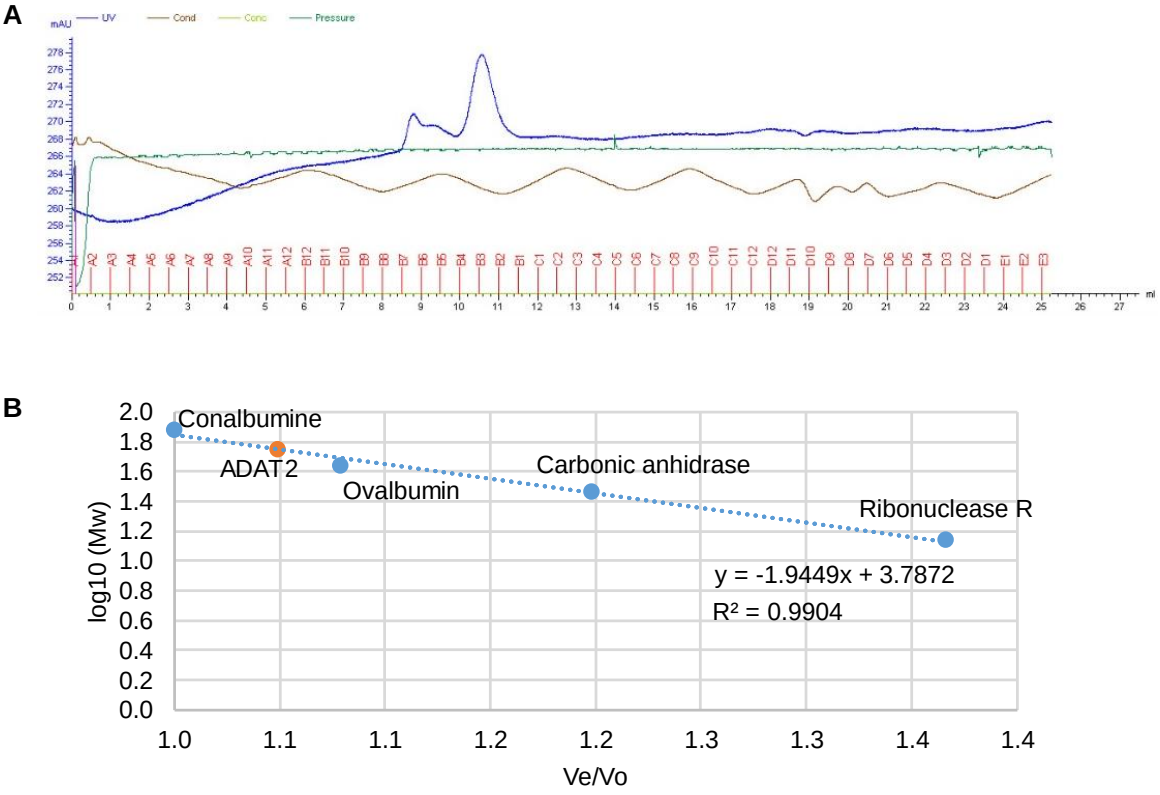


Figure 48. SEC analysis of ADAT2 homodimer and standardization of the size-exclusion column.

A Size-exclusion chromatogram of affinity-purified ADAT2. Elution volume (mL) is represented on the x-axis and the UV absorbance at 280 nm is represented on the y-axis.

B Plot of log10(Mw) against Ve/Vo ratios and standard curve, log10(Mw)=-1.9449Ve/Vo+3.7872, R^2=0.9892. Protein standards (conalbumine, ovalbumin, carbonic anhydrase, and ribonuclease R) and standard curve are represented in blue, and ADAT2 is represented in orange.

	Mw (kDa)	log10 (Mw)	Ve/Vo
SEC standards			
Ribonuclease R	13.7	1.1	1.4
Carbonic anhidrase	29	1.5	1.2
Ovalbumin	44	1.6	1.1
Conalbumine	75	1.9	1.0
ADAT2 homodimer	46.4	1.7	1.0
ADAT2 monomer	23.2	1.4	

Table 8. Mw and Ve/Vo values of the protein standards and ADAT2. ADAT2 elution volume is consistent with homodimers but not monomers.

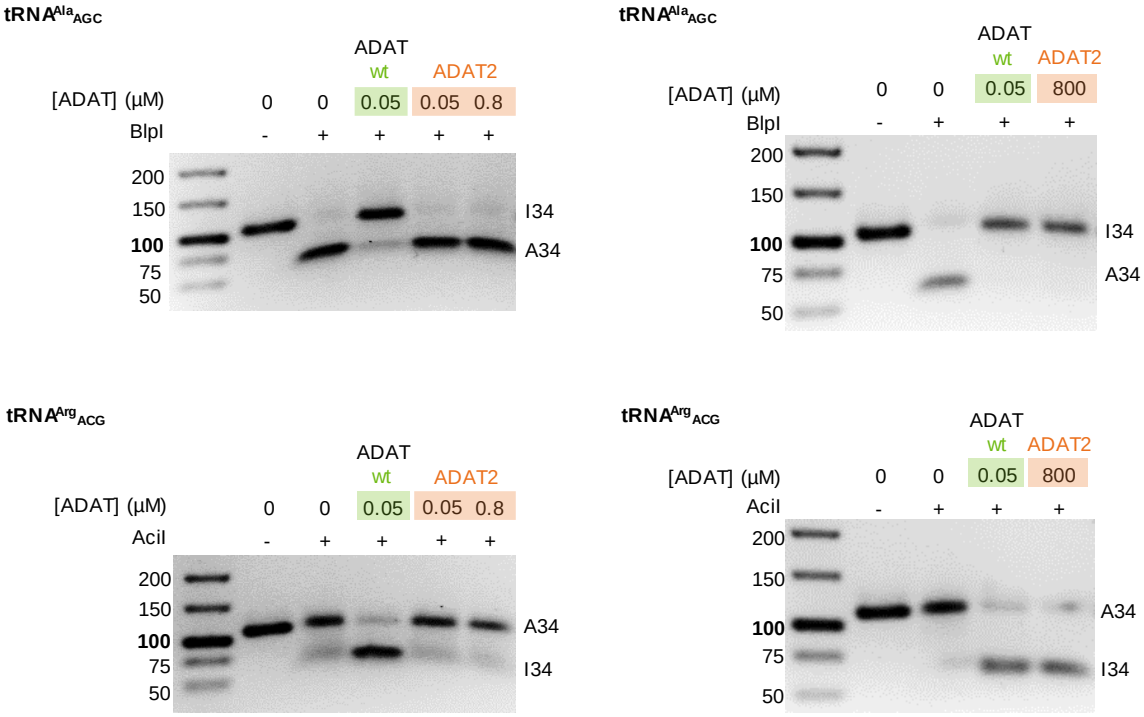


Figure 49. Analysis of the deamination of tRNA$^{Ala}_{AGC}$ and tRNA$^{Arg}_{ACG}$ by human ADAT2 homodimer at the nanomolar range (left panel) and micromolar range (right panel) (RFLP method). RFLP experiment to assess the deamination activity of ADAT2 homodimer on tRNA$^{Ala}_{AGC}$ (upper panel) and tRNA$^{Arg}_{ACG}$ (lower panel). Inosine formation was monitored at the indicated enzyme concentrations after 4 hours of incubation. 50 nM ADAT wt were used as a positive control. BlpI cuts A34-tRNA$^{Ala}_{AGC}$-derived PCR amplicons. AciI cuts I34-tRNA$^{Arg}_{ACG}$-derived PCR amplicons.

Small molecules targeting ADAT

In collaboration with the Computational Biology and Drug Design group (Dr. Xavier Barril) from the University of Barcelona. Computational analysis were done by Dr. Sergi Ruiz.

Abstract

ADAT deamination activity of adenosines at the first position of the anticodon of tRNAs plays a role in the regulation of translation. Thus, modulation of ADAT activity by small molecules targeting the enzyme is a promising strategy to control translation and, particularly, of transcripts with a high enrichment in ADAT-dependent codons (codons translated by modified tRNAs). Molecules acting as ADAT inhibitors might be used to control the expression of specific proteins encoded by genes enriched in such codons at the translational level. Given that some of these genes are found to be overexpressed in certain pathologies, ADAT inhibitors might be a promising therapeutic strategy. On the other hand, ADAT activators might potentially be used to recover the defective deaminase activity caused by V128M mutation. Combining *in silico* and *in vitro* approaches we have discovered the first active compounds against human ADAT.

ADAT-mediated deamination of adenosine 34 in eight human tRNAs changes the tRNA pool available for each mRNA codon, which has a role in the regulation of translation (9, 10), especially for mRNAs enriched in codons translated by I34-tRNAs (named ADAT-dependent codons) (11, 12). On this basis, the translation of these transcripts would be dependent on ADAT activity, and thus small molecules targeting ADAT constitute potential modulators of their expression. Interestingly, the analysis of the distribution of ADAT-dependent codons in the human transcriptome revealed that genes encoding for mucins tend to be highly enriched in such codons. This suggests a dependence on ADAT activity for mucins expression (data not published by Rafels-Ybern et al).

Mucins are large proteins heavily glycosylated in Ser and Thr repeats present in epithelial tissues of the respiratory and gastrointestinal tracts and major components of gel-like secretions. Alterations in mucins production and glycosylation patterns have been related to disease. Mucus hypersecretion, airway obstruction and mucociliary clearance impairment found in certain respiratory conditions like asthma, chronic obstructive pulmonary disease (COPD) or cystic fibrosis (CF), have been associated to mucins overexpression (164, 165).

Moreover, invasive proliferation of some cancers is linked to mucins overexpression or aberrant glycosylation (166). Due to the putative need for I34-modification in mucins expression, the inhibition of ADAT activity would be a promising strategy to block its overexpression in the context of these pathologies. Thus, the development of selective ADAT inhibitors might constitute a promising therapeutic strategy with a novel mechanism of action for the treatment of these diseases. Nonetheless, the dose of such compounds should be finely controlled to block mucins overexpression but also to allow regular protein synthesis.

On the other hand, small molecules with the ability to activate ADAT enzyme might also have a potential therapeutic application. We reported that a mutation in ADAT3 linked to ID and strabismus causes a reduction of ADAT deamination activity. Thus, ADAT activators could potentially recover the defective tRNA editing caused by the mutation.

Overall, small molecules acting as ADAT modulators constitute potential regulators of protein synthesis and have promising applications in therapeutics. Additionally, molecules targeting ADAT could also be used in research, i.e. to further study the role of ADAT, crystallographic purposes, etc. No ADAT modulators had been reported to date, and here we show the first active compounds against the human enzyme we have developed using a combination of computational and experimental approaches.

Deaminase inhibitors as potential ADAT inhibitors

As a first approach for the discovery of ADAT modulators, we tested whether small molecules that have been reported to inhibit adenosine deaminases were also able to inhibit ADAT enzyme. The structural similarity between ADAT and other deaminases, whose catalytic centers share conserved deaminase domains (1, 115), let us to envisage that such compounds might also be potential ADAT inhibitors. Combining computational and experimental analyses, we studied the activity of four deaminase inhibitors against ADAT enzyme: *Erythro*-9-(2-Hydroxy-3-nonyl)adenine hydrochloride (EHNA hydrochloride), an inhibitor of ADA and cyclic nucleotide phosphodiesterase- 2 with anticancer activity (167, 168); Pentostatin, an ADA inhibitor used in cancer treatment (169, 170); 3-deazaadenosine, an inhibitor of ADA and of leukocyte adhesion (171, 172); and Zebularine, an inhibitor of cytidine deaminase and DNA methyltransferase (DNMT) with antitumor activity (173-175) (**Figure 50**).

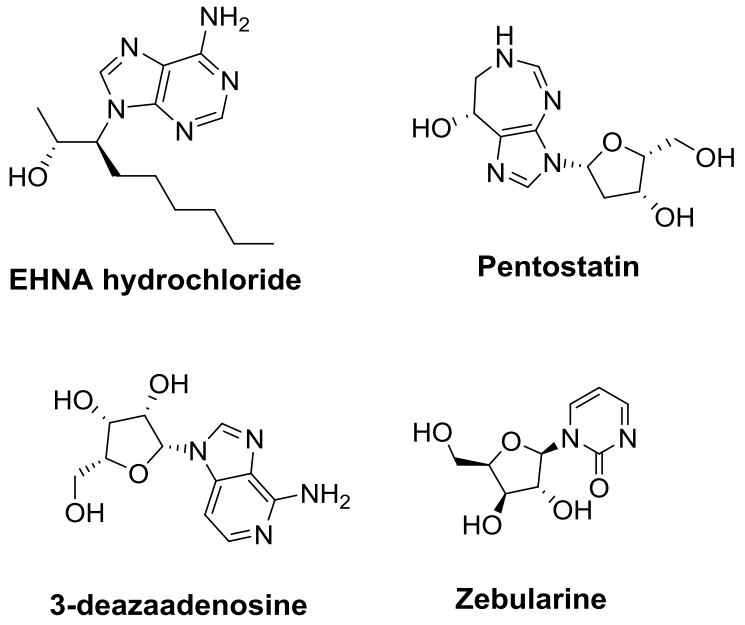


Figure 50. Chemical structures of the four small molecules with reported activity against adenosine deaminases.

The crystal structure of the human heterodimeric ADAT2-ADAT3 has not been resolved to date. However, ADAT2 subunit of the human enzyme has been crystalized as a homodimer (**Figure 51**). For the *in silico* analysis, we built a computational model of the human ADAT2-ADAT3 heterodimer using the protein homology-modeling server SWISS-MODEL (**Figure 52**). The docking analysis was performed using the crystal structure of nebularine bound to TadA from *Staphylococcus aureus* (176) as a model for the binding of the small molecules to ADAT heterodimer. Even though the catalytic center of the eukaryotic enzyme has conserved deaminase motifs as the bacterial homolog, the computational analysis revealed noteworthy structural differences (**Figure 53**). Consequently, constraining the small molecules to adopt the same pose as nebularine in TadA active site resulted in low score values for all the candidates (**Table 9**).

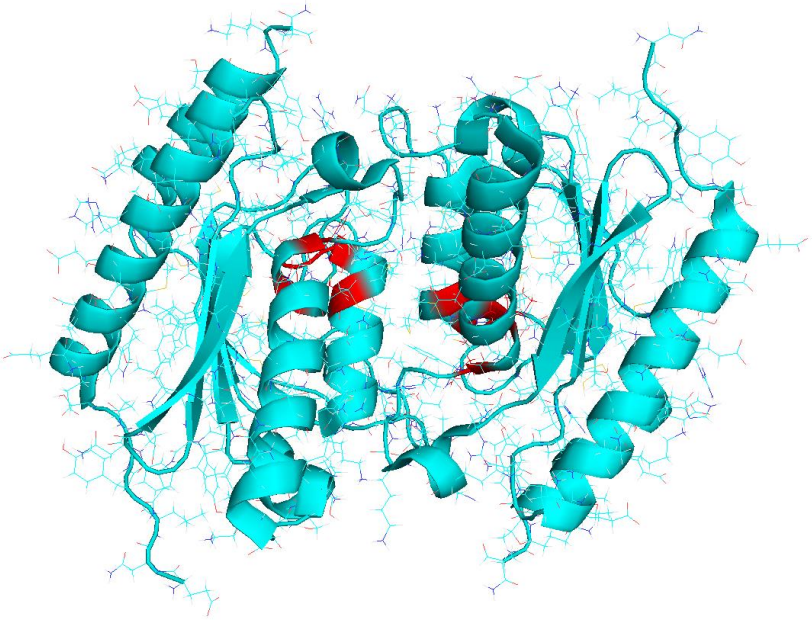


Figure 51. Crystal structure of the human ADAT2 homodimer (PDB ID: 3DH1). The two ADAT2 subunits are represented in blue and catalytic amino acids in red.

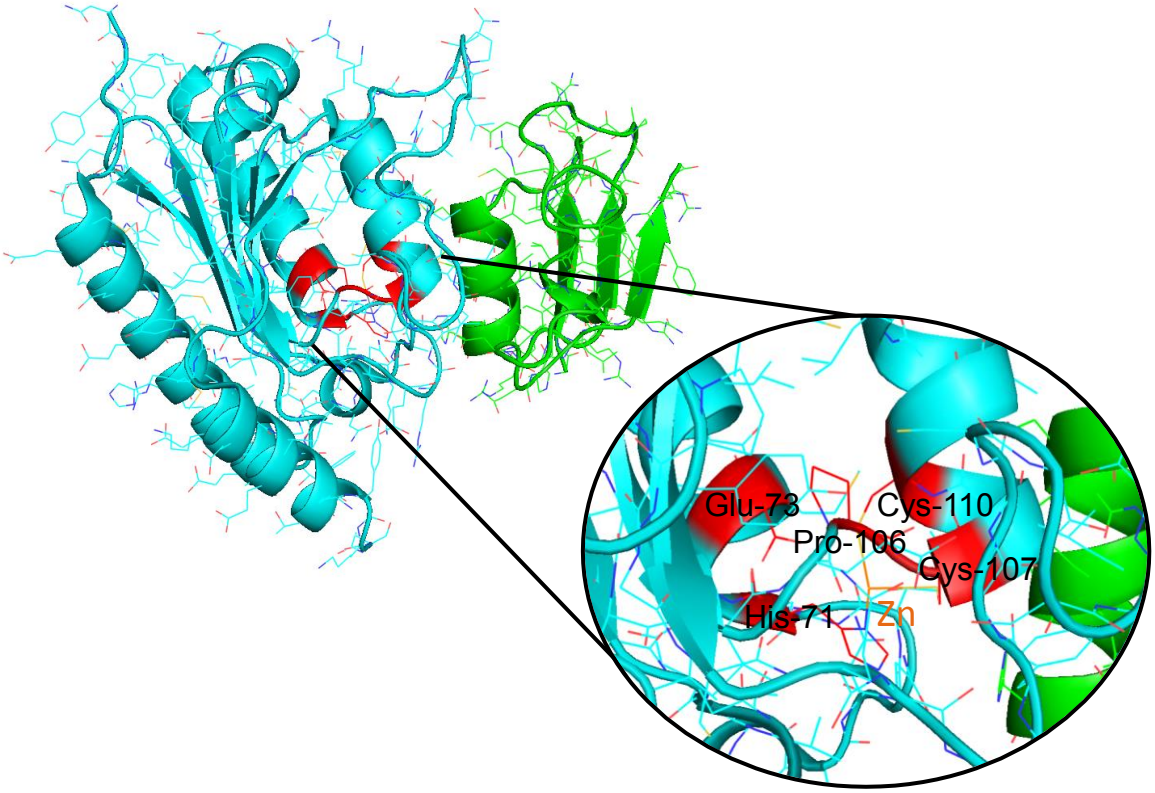


Figure 52. Computational model of the human ADAT2-ADAT3 heterodimer. ADAT2 subunit is represented in blue, and ADAT3 in green. ADAT3 homology model (47% of sequence identity) lacks the N-terminal end (which has low identity to ADAT2 and TadA) since the structure could not be accurately predicted. Active site is zoomed in, catalytic amino acids are represented in red and catalytic zinc ion in orange.

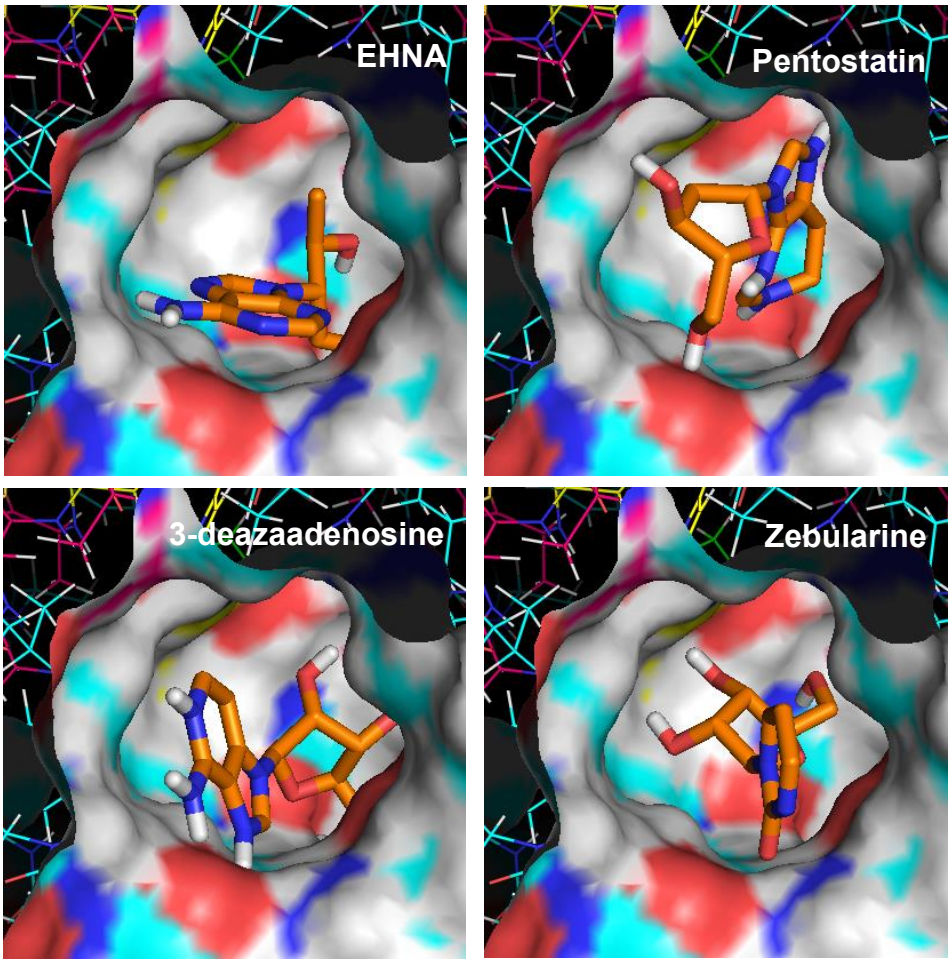


Figure 53. Molecular docking analysis of the binding of the four deaminase inhibitors to the catalytic center of ADAT2-ADAT3 computational model.

	Score
3-deazaadenosine	-18,24
Pentostatin	-16,48
Zebularine	-13,97
EHNA	-12,06

Table 9. Docking results of the four studied deaminase inhibitors.

We *in vitro* analyzed the activity of the four deaminase inhibitors against ADAT enzyme using RFLP method (177). We analyzed the levels of deamination of *in vitro* transcribed $tRNA^{Ala}_{AGC}$ after 1 hour of incubation at 37ºC with affinity-purified human ADAT upon addition of 1 mM of each small molecule, respectively. None of the four deaminase inhibitors showed inhibition of ADAT-mediated deamination of $tRNA^{Ala}_{AGC}$ (**Figure 54**).

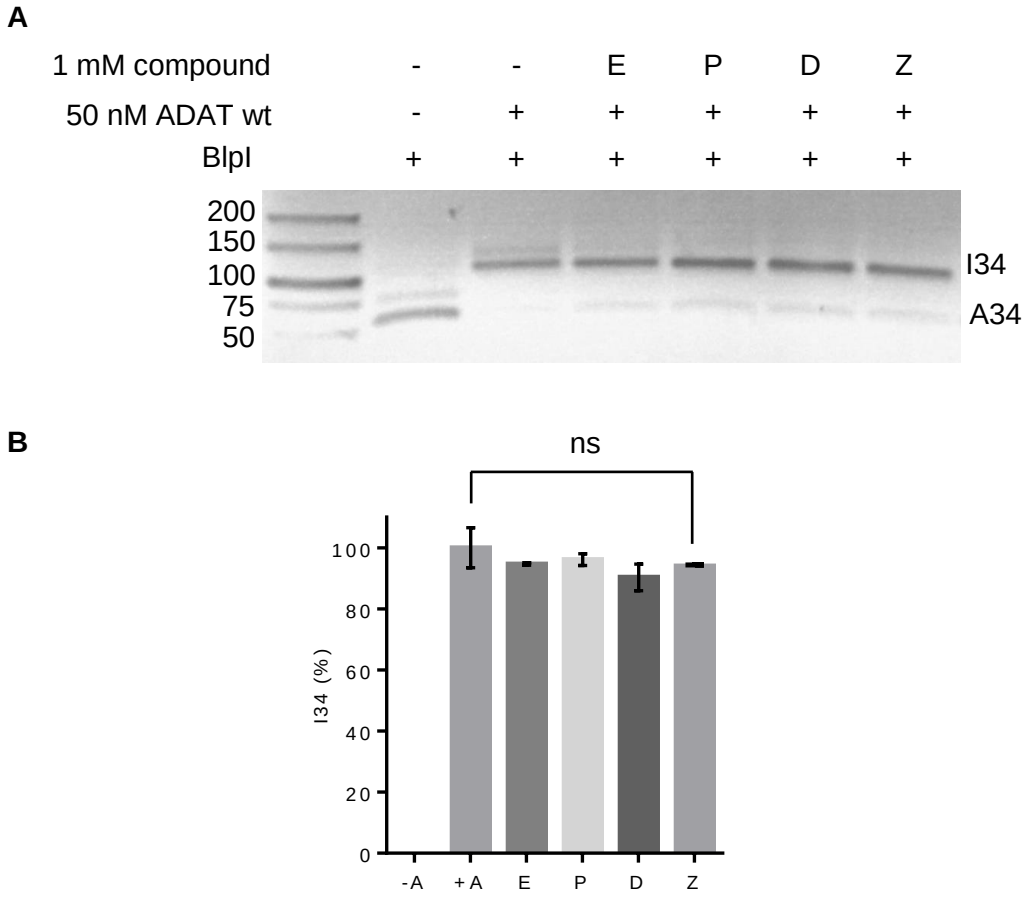


Figure 54. Analysis of the effect of four known deaminase inhibitors on ADAT-mediated deamination of $tRNA^{Ala}_{AGC}$.

A Representative RFLP experiment to assess ADAT-mediated deamination of $tRNA^{Ala}_{AGC}$ upon 1 mM of each compound, respectively. E stands for EHNA hydrochloride, P for Pentostatin, D for 3-deazaadenosine, and Z for Zebularine.

B Levels of I34 modification upon 1 mM of each compound calculated from three independent experiments as in **A** and averaged. Error bars reflect the standard deviation. No statistical differences were observed between different tRNA fragments (P value = 0.0849).

Virtual screening of small molecules targeting ADAT

To identify small molecules that could modulate ADAT activity, we performed a first virtual screening of a chemical library of 200,000 compounds. The homology model of ADAT2 homodimer (PDB ID: 3DH1) was used as molecular docking receptor (**Figure 51**). Since the catalytic center is in ADAT2 subunit and thus our docking cavity involves ADAT2 exclusively (**Figure 55**), the absence of ADAT3 is less likely to alter our analysis.

The crystal structures of TadA from *S. aureus* (PDB ID: 2B3J) (176) and of a cytidine deaminase from *E. coli* (PDB ID: 1CTU) (178) were used as a template for the docking of small molecules to the active site. Both proteins were crystalized with a ligand bound to the catalytic center (nebularine and zebularine, respectively) (**Figure 56**) (176, 178). Two conserved H bonds with the ligands identified in both crystal structures were defined as pharmacophoric restraints for the docking. The first one was with a conserved alanine residue (Ala60), and the second one with the water molecule coordinating the Zn ion in the active site. From the 200,000 compounds in the library, 63,000 were selected. Among the compounds having the two H bond acceptors, we eliminated those in which the H bonds did not have the right distance or orientation. A total of 30,000 compounds were selected.

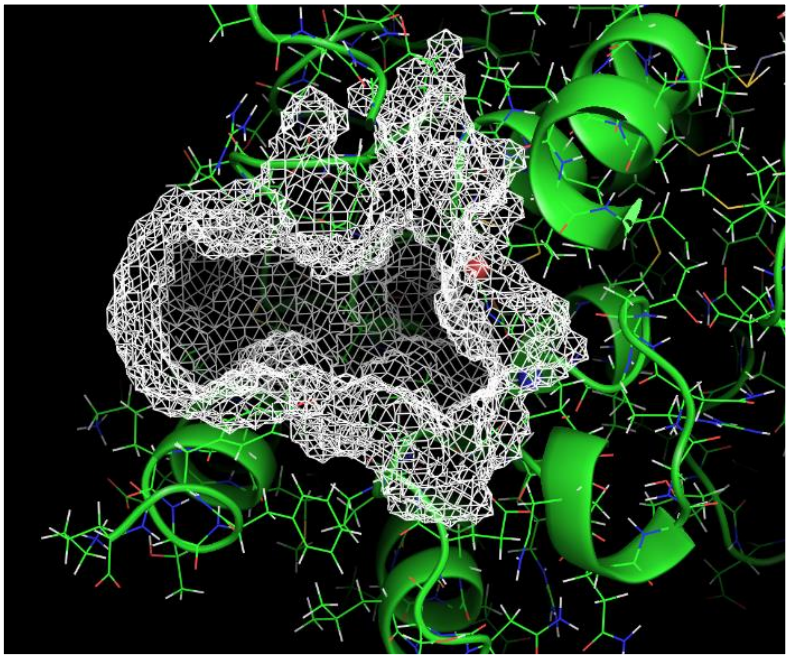


Figure 55. Docking cavity (white) in the active site of ADAT enzyme, in ADAT2 subunit (green).

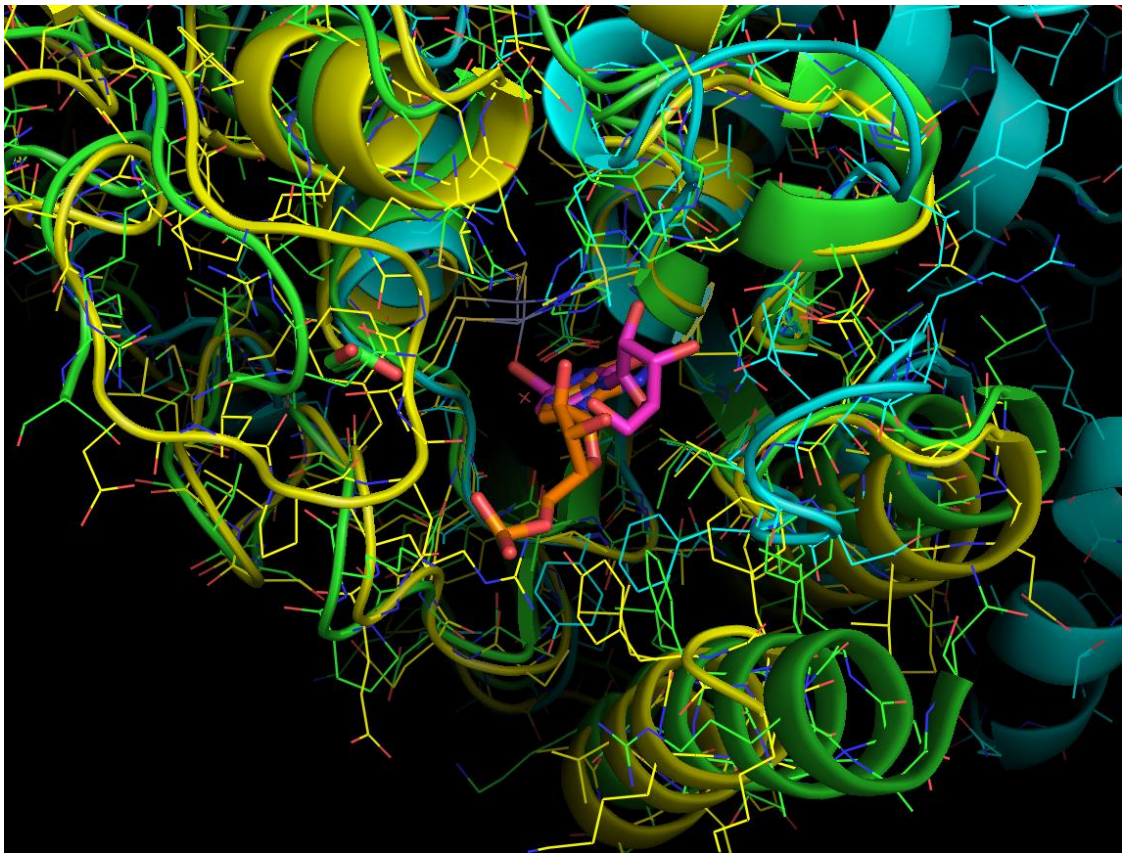


Figure 56. Overlapped crystal structures of the human ADAT2 homodimer (yellow), TadA from *S. aureus* (green), and a cytidine deaminase from *E. coli* (blue). The two ligands bound to TadA and cytidine deaminase are nebularine (orange) and zebularine (pink), respectively.

The 30,000 selected compounds were clustered based on their structural similarity. Using Molecular ACCess System (MACCS) keys, the fingerprint of each compound was determined. A Tanimoto coefficient (Tc) of 85% was used as a threshold to distinguish similar from dissimilar compounds. Compounds with high structural similarity, with Tc >85%, were included in the same cluster. A representative compound with the highest score was selected from each cluster. A total of 8,500 compounds were selected. Within the 8,500 compounds, the ones with a score <-25 (with the putative highest binding affinity to ADAT active site) were selected. Finally, the 14 best candidates were manually selected for the *in vitro* analysis (**Figure 57**).

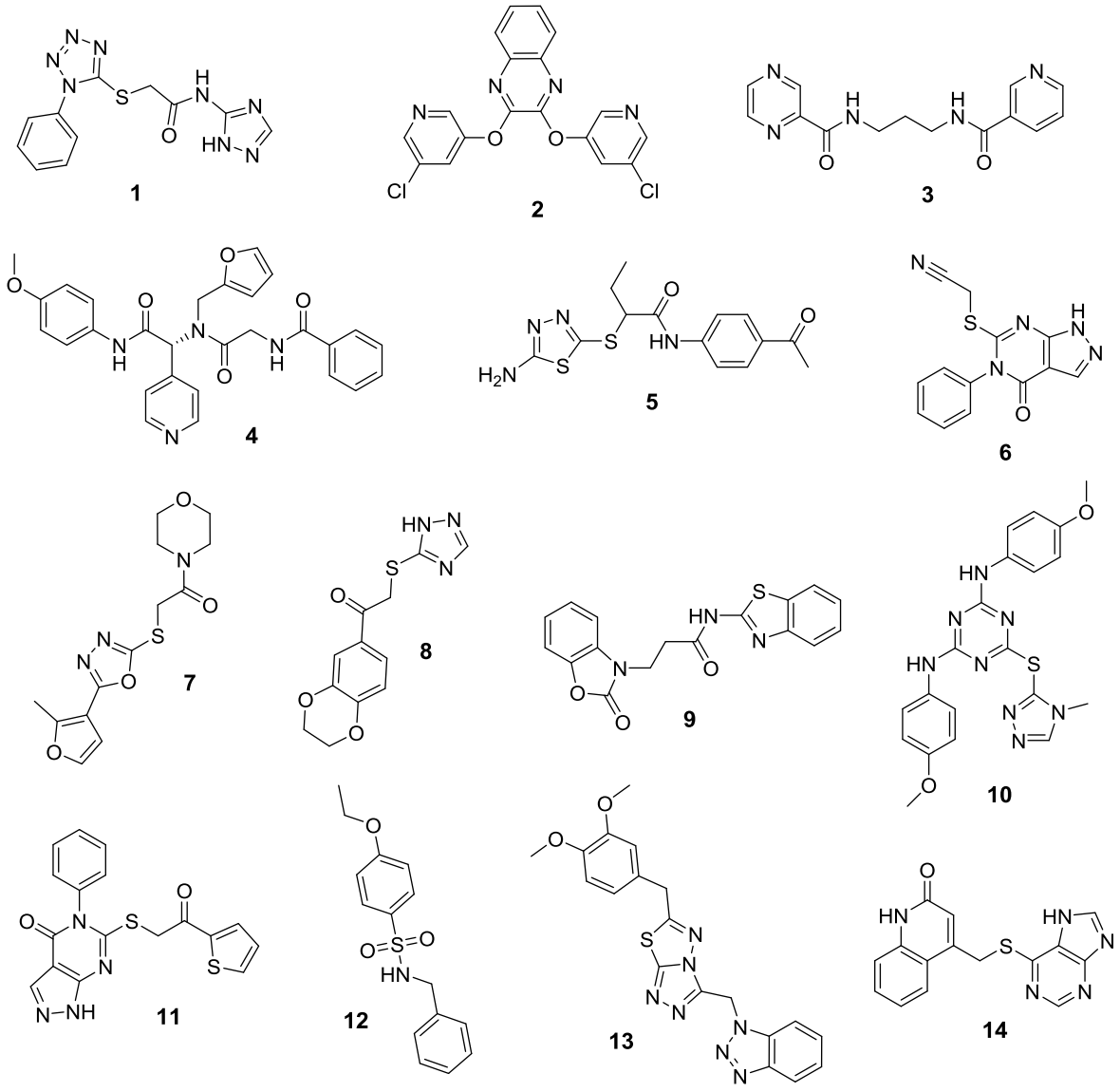


Figure 57. Chemical structures and identification numbers of the fourteen compounds selected in the virtual screening.

In vitro analysis of the virtual screening hits

We *in vitro* analyzed the activity of the fourteen compounds selected in the virtual screening against ADAT enzyme. Using RFLP method, we evaluated the levels of deamination of *in vitro* transcribed $tRNA^{Ala}_{AGC}$ after 1 hour of incubation at 37ºC with affinity-purified human ADAT (50 nM) upon addition of 500 μM of each small molecule, respectively. Six of the compounds showed statistically significant inhibition. However, the levels of inhibition were below 15% in all cases, and none of the candidates was identified as a promising compound (**Figure 58**). Unexpectedly, at a lower ADAT concentration (5 nM), three of the compounds (4, 9, and 10) were able to activate the enzyme at a concentration of 500 μM (**Figure 59**).

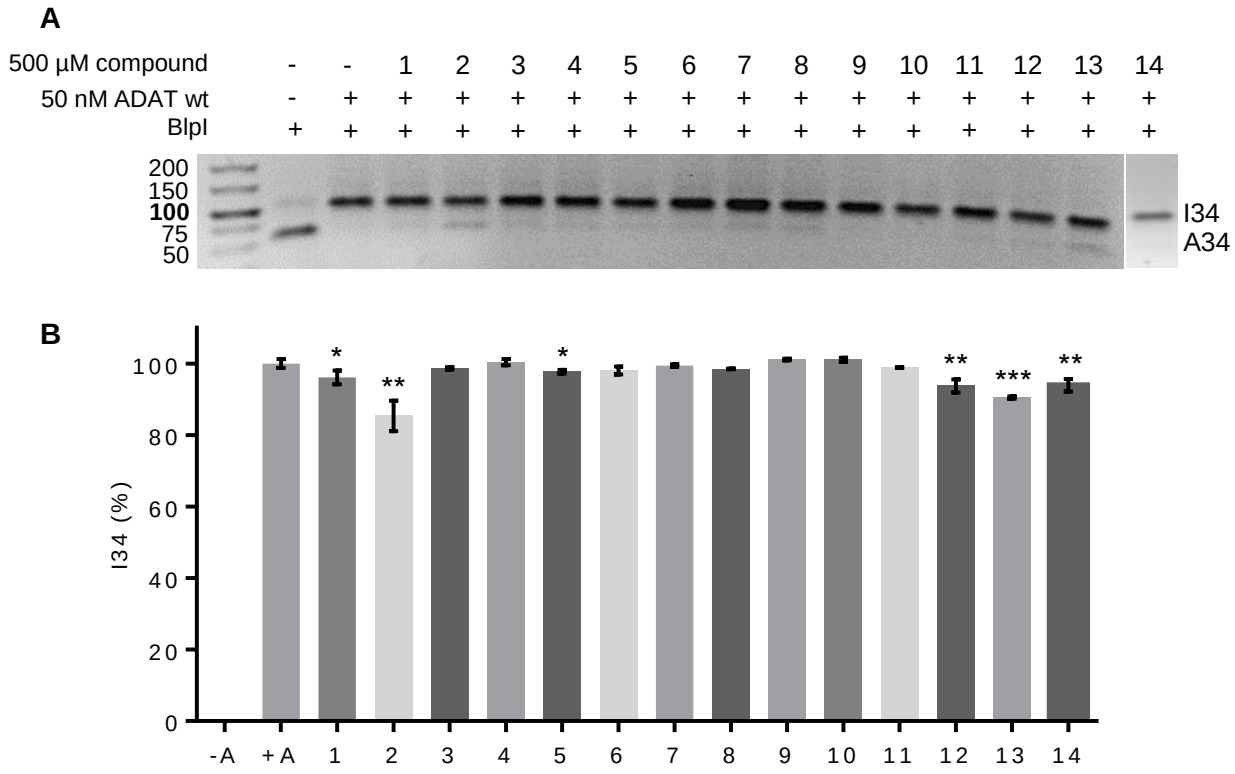


Figure 58. Analysis of the effect of the virtual screening hits on ADAT-mediated deamination of tRNA$^{Ala}_{AGC}$.

A Representative RFLP experiment to assess ADAT-mediated deamination of tRNA$^{Ala}_{AGC}$ upon 500 µM of each compound, respectively.

B Levels of I34 modification upon 500 µM of each compound calculated from three independent experiments as in **A** and averaged. Error bars reflect the standard deviation. Asterisks indicate statistical significance compared to the control.

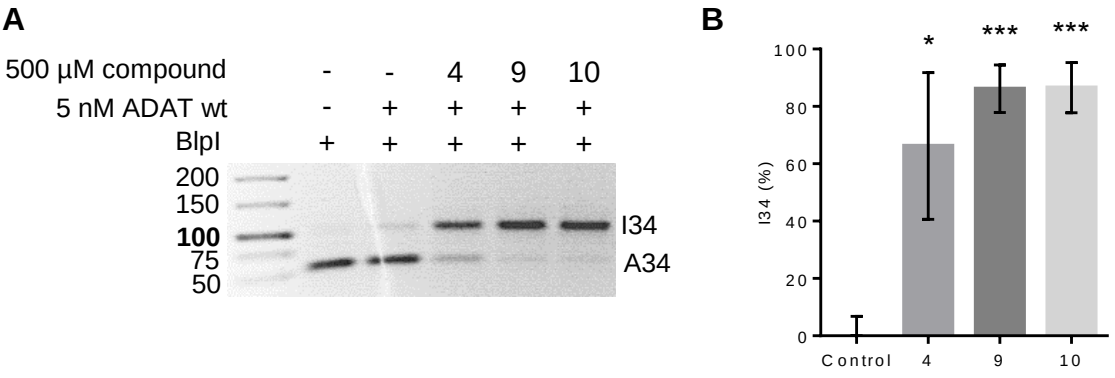


Figure 59. Analysis of the effect of the three activators selected in the virtual screening on ADAT-mediated deamination of tRNA$^{Ala}_{AGC}$.

A Representative RFLP experiment to assess ADAT-mediated deamination of tRNA$^{Ala}_{AGC}$ upon 500 µM of each activator, respectively.

B Levels of I34 modification upon 500 µM of each activator calculated from three independent experiments as in **A** and averaged. Error bars reflect the standard deviation. Asterisks indicate statistical significance compared to the control.

We then studied the effect of the three compounds with the ability to activate the enzyme in the context of V128M mutation. We observed a decreased deamination activity in ADAT V128M compared to the wt enzyme. On this basis, the reported ADAT activators might potentially recover the defective tRNA editing caused by the mutation. The analysis of I34 levels in $tRNA^{Ala}_{AGC}$ upon 500 μM of activator showed a significant activation of the mutant enzyme by compound 10 (**Figure 60**).

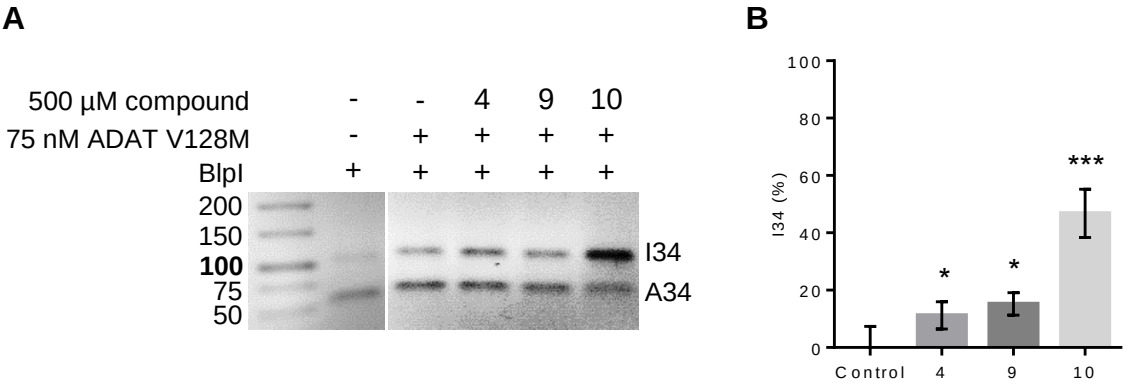


Figure 60. Analysis of the effect of the three activators in ADAT V128M deamination of $tRNA^{Ala}_{AGC}$.

A Representative RFLP experiment to assess ADAT V128M deamination of $tRNA^{Ala}_{AGC}$ upon 500 μM of each activator, respectively.

B Levels of I34 modification upon 500 μM of each activator calculated from three independent experiments as in **A** and averaged. Error bars reflect the standard deviation. Asterisks indicate statistical significance compared to the control.

Second virtual screening of small molecules targeting ADAT

To find better candidates with the ability to modulate ADAT activity we performed a second virtual screening of a 7 million compounds chemical library. The same molecular docking receptor and pharmacophoric restraints defined for the previous virtual screening were used, and the same selection criteria described before was applied. The best 20 compounds were selected for the subsequent *in vitro* analysis (ongoing work) (**Figure 61**).

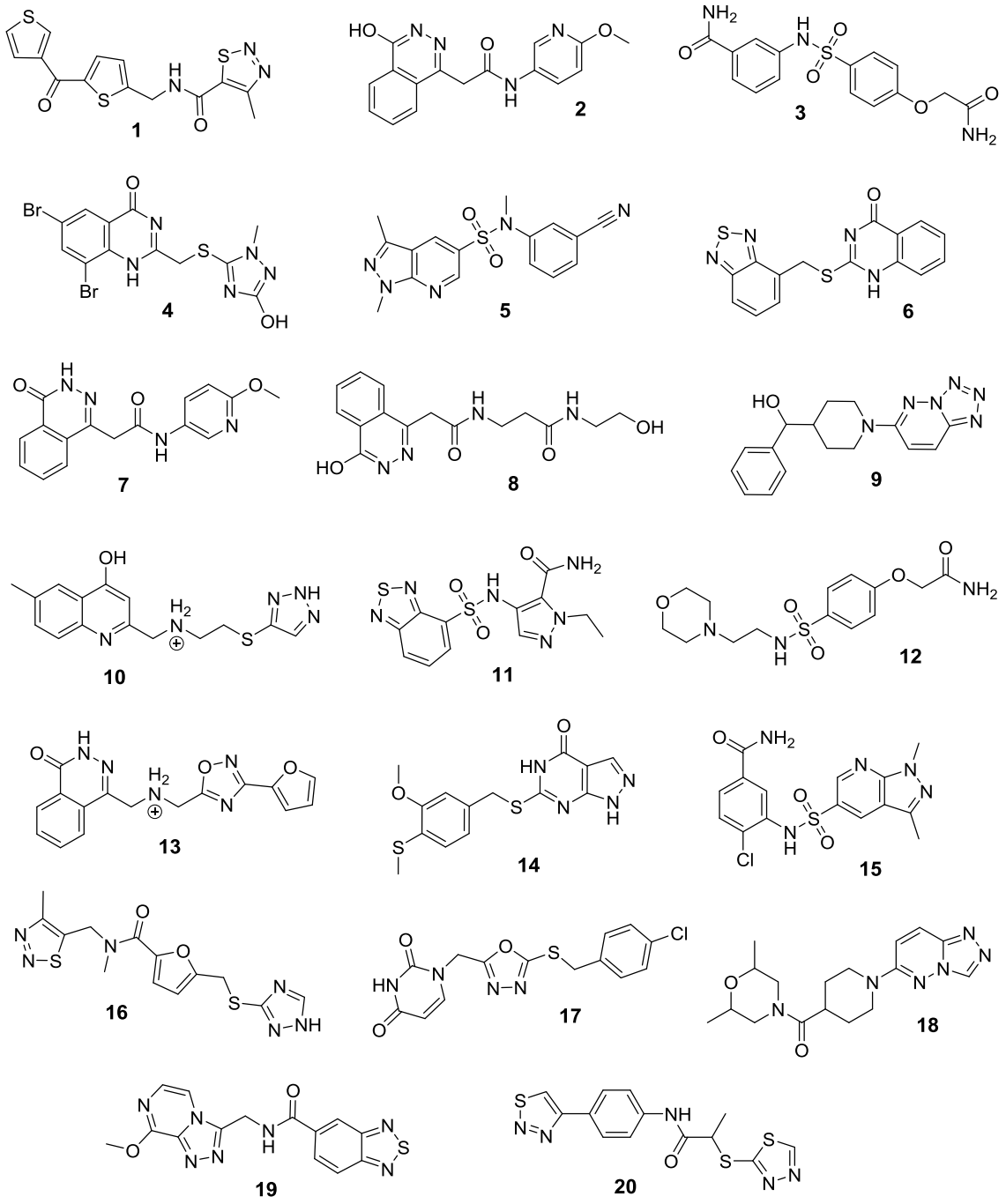


Figure 61. Chemical structures and identification numbers of the twenty compounds selected in the virtual screening.

Discussion

Development of an *in vitro* activity assay to monitor ADAT-mediated deamination

In recent years, tRNA modifications have emerged as a fundamental aspect of tRNA function and regulation and been identified as being highly relevant to human physiology (101, 143). Moreover, while in the past post-transcriptional tRNA modifications were considered rather static, mounting evidence now suggests that they can be dynamic (102, 106). This has prompted the development of new techniques for efficiently detecting, quantifying, and monitoring RNA modifications. The RFLP method for tRNA modifications detection can semiquantitatively detect modifications on specific tRNA bases and in different tRNA species. The method involves four steps: cDNA synthesis, PCR amplification with specific primers, digestion of the obtained amplicon with restriction endonucleases, and detection of cleavage patterns by gel electrophoresis (**Figure 23**).

We applied the RFLP method to evaluate the levels of post-transcriptional tRNA modifications on *in vitro* transcribed tRNAs. We show that, at least for I34 detection, the method can be used in a semiquantitative manner and is sufficiently sensitive to detect the modification even when it is present in only 1% of the tRNA population (**Figure 25**). We monitored A34-to-I34 editing by human ADAT *in vitro* and under different reaction conditions using this approach and showed that this method can be used to determine the optimal parameters for the deamination reaction (**Figure 26**). We later used this method to characterize ADAT substrates and regulation, to study ADAT kinetics and the effect of a mutation in the enzyme linked to intellectual disability (V128M) (18, 19), as well as to evaluate the activity of small molecules targeting ADAT. Although in this work we have used the method to assess the deamination activity *in vitro*, it can be useful to evaluate modifications on endogenous *in vivo* tRNA species (13).

The RFLP method for detecting post-transcriptional tRNA modifications has some limitations. First, it requires the reverse transcription and amplification of the tRNA of interest. This is generally not a problem when using *in vitro* transcribed tRNA, but it might become an issue for *in vivo* tRNAs, i.e. for amplifying tRNAs that may contain other modifications incompatible with RT reactions (145). This may introduce a bias toward reverse transcribing partially modified precursor tRNAs over fully modified mature tRNAs. However, once the RT step is completed the PCR step of the RFLP method will not introduce additional biases as the obtained cDNA template contains standard DNA bases. Finally, RFLP is limited by the need for a restriction enzyme that will differentially recognize the region of interest where the tRNA modifications induce a change in sequence in their transcripts. Nevertheless, we have found useful restriction enzymes to analyze I34 modification in all our tRNAs of interest.

In addition to I34 modification, RFLP method can also be used to analyze other post-transcriptional tRNA modifications that produce a "mutation" upon RT, and with a

restriction enzyme targeting where the modification is. For example, in collaboration with Dr. Hossein Najmabadi research group (in the Genetics Research Center, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran) we used the RFLP method to monitor the formation N2-N2-dimethyl-guanosine (m^2_2G) modification in tRNAs, which causes the incorporation of A or T on the reverse transcribed cDNA (179). The dimethylation is catalyzed by the human enzyme tRNA methyltransferase 1 (TRMT1) (180), and we applied the assay to study the kinetics of the enzyme and the effect of a mutation linked to neurological disorders (181, 182).

On the other hand, the fluorescence-based assay performed at Dr. Tor laboratory at the UCSD is an innovative and promising strategy to quantify ADAT-mediated deamination. Dr. Tor research group developed the first, and only reported to date, emissive adenosine analogs undergoing enzymatic deaminations (14, 16, 154). These fluorescent adenosine analogs showed to be substrates as free nucleosides by adenosine deaminase (ADA) (14, 16) but also as a part of an mRNA molecule by adenosine deaminases acting on RNA (ADARs) (160). On this basis, an emissive adenosine analog at position 34 of a tRNA might also be deaminated by ADAT enzyme. To overcome the possibility that the emissive adenosine analogs were not active ADAT substrates, we designed a second strategy consisting in monitoring I34 formation by measuring changes in the emission of a fluorescent base at position 35 (thG or ^{tz}G). Given that the emission of these molecules highly depends on the chemical environment, the deamination of A34 might cause changes in the emission of the adjacent base at position 35.

An advantage of the fluorescence-based assay compared to the RFLP method is that it does not require processing of the tRNA, thus, deamination levels can be monitored in real time. It is true that both the chemical synthesis of the fluorescent bases and the synthesis of the tRNAs that incorporate the emissive nucleotide analogs are challenging and laborious. Also, establishing the relationship between changes in emission and deamination levels requires optimization. Nevertheless, once the assay conditions have been established, the deamination can be directly quantified in real time. In this work, we prepared four emissive tRNAs containing thA34, ^{tz}A34, thG35, and ^{tz}G35, respectively. However, due to timing constraints, we were unable to finish with the development of the assay.

New insights into substrate recognition and regulation of human ADAT

The changes in substrate specificity between TadA and ADAT, going from one single substrate in TadA ($tRNA^{Arg}$) to eight different tRNAs in eukaryotic ADAT, have been associated to an evolution in substrate recognition. Whereas TadA substrate recognition is focused on the local sequence and structure of the anticodon stem loop (2, 118), the

acquisition of multi-substrate specificity by ADAT possibly required the recognition of additional features of its substrate tRNAs (118, 119). We first investigated whether the primary sequence of ADAT substrate tRNAs might contain distinctive elements that are absent in non-substrate tRNAs. Sequence alignment of a representative set of human tRNAs, including ADAT substrates and non-ADAT substrates, was unable to identify a signature sequence shared between the tRNAs that undergo ADAT-mediated deamination (**Figure 32 and Supplementary Figures S4, S5 and S6**). This indicates that the substrate recognition might rely on elements in the secondary and tertiary structure of the tRNA rather that the nucleotide sequence *per se*.

To explore the dependence of ADAT-mediated deamination on the tRNA structure, we analyzed I34 levels upon different folding conditions. Consistent with the need of a proper tRNA architecture for ADAT activity, the absence of folding of the tRNA resulted in loss of deamination. Interestingly, this loss of activity in unfolded tRNAs could be partially recovered by the addition of Mg (**Figure 33**), which is consistent with the contribution of Mg ions to RNA assembly and tertiary structure stabilization (183). A dependence on Mg ions for deaminase activity was observed in yeast ADAT, which was also explained by its assistance in tRNA folding rather than a need of Mg for the enzymatic reaction (119). The folding method also influenced the deamination levels, as significant differences were observed between fast folding and slow folding (**Figure 33**). Fast-folded tRNA was more efficiently deaminated than slow-folded tRNA, which might be explained because the first method results into a higher proportion of properly folded tRNA. Overall, the tRNA architecture seems to be a key feature for efficient A34-to-I34 conversion by human ADAT, which is in line with the results obtained with other eukaryotic ADATs (from *S. cerevisiae*), which revealed that alterations in the tRNA 3D structure resulted in loss of deamination activity (118, 119). It is also in line with the observation that tRNA modifications located in the anticodon branch and in the D-arm tend to be especially sensitive to perturbations in the tRNA 3D structure, in contrast to modifications in the T-arm and acceptor stem, which are not generally dependent on an intact tRNA architecture (184).

To elucidate the minimum structural features required for ADAT recognition and activity, we designed series of short tRNA variants derived from $tRNA^{Arg}$ and $tRNA^{Ala}$, and checked for ADAT-mediated deamination (**Figure 34**). Upon overnight incubation with ADAT, tRNAs lacking the CCA and acceptor stem showed similar levels of deamination as full-length tRNAs. The full deamination of tRNAs lacking the CCA is consistent with the observation that precursor tRNAs have all the requirements for efficient deamination (3), and suggests that ADAT acts on tRNA precursors after they adopt a canonical tRNA 3D core. We observed statistically significant differences in the levels of deamination between $tRNA^{Ala}$ and $tRNA^{Arg}$-derived fragments lacking the acceptor stem. This might be due to different contributions of the acceptor stem in the binding to ADAT between the two tRNAs, suggesting that the acceptor stem of $tRNA^{Ala}$ would have major involvement.

Supporting this idea, $tRNA^{Ala}$ without acceptor stem showed no deamination after one-hour incubation with ADAT (it was thus less efficiently deaminated than the full-length $tRNA^{Ala}$) (**Supplementary Figures S7**), indicating that the acceptor stem of $tRNA^{Ala}$ might play a role in ADAT substrate recognition. To examine a potential interaction between the acceptor stem of the tRNA and ADAT enzyme, we analyzed whether an acceptor stem analog might be able to inhibit ADAT-mediated deamination (**Figure 36**). In fact, acceptor stem analogs derived from $tRNA^{Pro}$ showed to significantly inhibit the deamination of both $tRNA^{Arg}$ and $tRNA^{Ala}$ to similar extents. Thus, the possibility that the acceptor stem has a role in the recognition of the two tRNAs cannot be excluded.

In contrast to the bacterial homolog TadA, which is able to deaminate a $tRNA^{Arg}$ mini-substrate containing the anticodon stem loop as efficiently as the full-length $tRNA^{Arg}$ (2, 118), equivalent fragments derived from human $tRNA^{Arg}$ and $tRNA^{Ala}$ were not active substrates of the human ADAT enzyme (**Figure 34**). The inability of eukaryotic ADATs to deaminate a tRNA mini-substrate was also observed in *S. cerevisiae* ADAT (118). The loss of deamination of the shortest tRNA fragments by human ADAT is likely to be due to the loss of the 3D core structure. This is consistent with our results showing that unfolded tRNA is not deaminated (**Figure 33**), and the observation that eukaryotic ADATs require a proper tRNA architecture for efficient deamination (118, 119). The absence of significant differences in I34 levels between $tRNA^{Arg}$ and $tRNA^{Ala}$ anticodon stem loop fragments, suggests that the recognition of $tRNA^{Arg}$ by the eukaryotic ADAT enzyme would have evolved as for the rest of the substrates, involving structural features other than the anticodon stem loop (**Figure 34**).

Potential variations in interactions enzyme-substrate between ADAT and $tRNA^{Arg}$ or $tRNA^{Ala}$ were investigated by monitoring the effect of the abovementioned $tRNA^{Arg}$ fragments on $tRNA^{Ala}$ deamination and the effect of $tRNA^{Ala}$ fragments on $tRNA^{Arg}$ deamination (**Figure 35**). Interestingly, the deamination of $tRNA^{Arg}$ was inhibited by all the $tRNA^{Ala}$-derived fragments, whereas only the longest $tRNA^{Arg}$ fragments inhibited $tRNA^{Ala}$ deamination. These results might indicate a different binding mode between ADAT and the two tRNAs; while the main interaction with $tRNA^{Arg}$ would be in the anticodon arm, in $tRNA^{Ala}$ might be found in the top half of the tRNA. To further study the binding ADAT-tRNA, we designed chimeric tRNAs combining the parts with putatively highest and lowest binding affinity, respectively, and checked for the deamination (**Figure 37**). The tRNA combining the parts with the presumed lowest binding affinity ($tRNA^{Arg}$ with $tRNA^{Ala}$ anticodon arm) showed deamination levels comparable to those for the native substrates $tRNA^{Arg}$ and $tRNA^{Ala}$. By contrast, the $tRNA^{Ala}$ with $tRNA^{Arg}$ anticodon arm (with putatively higher affinity to ADAT) showed extremely low levels of deamination. This might be due to either a very low dissociation rate of the tRNA-ADAT complex (which would not allow the deamination of the rest of tRNAs that are not initially bound to the enzyme), or that combining two parts from different tRNAs, albeit maintaining the cloverleaf secondary

structure (**Supplementary Figure S8**), results in the loss of important interactions for the 3D structure leading to an inactive substrate. To assess this question, are currently performing binding assays to evaluate the affinity between the tRNAs and ADAT. Altogether, despite combining the parts with the putative highest and lowest binding affinity, the deamination levels did not change as expected, suggesting that it is the overall tRNA architecture rather than a compilation of local structural features the main determinant for ADAT recognition.

Based on the observed ability of tRNA fragments to inhibit ADAT activity, we investigated whether human tRNA-derived fragments (tRFs) might be able modulate ADAT activity. Three of the tRFs tested showed inhibition constants in the low μM range (**Figures 38 and 39, and Supplementary Figures S9**). 5'-halves derived from $tRNA^{Ala}$, which showed a considerable inhibition of ADAT deamination of $tRNA^{Arg}$, are generated in response to stress and have been associated to induction of stress granule formation (69), and, together with 5'-halves derived from $tRNA^{Cys}$ (which also showed a significant inhibition of both $tRNA^{Arg}$ and $tRNA^{Ala}$ deamination), have been linked to translational inhibition in human cells (63). The highest inhibitory activity was observed with the 3'-CCA tRF derived from $tRNA^{Ala}$. 3'-CCA tRFs derived from $tRNA^{Ser}_{UGA}$ and $tRNA^{Gly}_{GCC}$ generated by Dicer cleavage have been associated to translational inhibition in human cells via AGO proteins (77, 79). However, the specific function of many of them remains to be determined, and whether they play a role in the regulation of the activity of tRNA modification enzymes has not been reported to date.

The fact that some of the tRFs inhibiting ADAT are generated upon stress suggests that I34 levels might be adjustable and respond to environmental conditions. Nevertheless, I34 modification has previously been shown to be very stable in *S. cerevisiae*, even after exposure to toxic agents, and the tRNAome of cultured human cells has been reported to be saturated in terms of I34 content (102, 123, 185). Even so, we cannot rule out the possibility that, under certain conditions, variations in the levels of I34 might be involved in the regulation of translation of specific genetic programs as a potential adaptive mechanism.

It is known is that specific tRNA modifications in the anticodon loop can regulate the formation of certain tRFs, by protecting the tRNAs from ANG-mediated cleavage. It is the case of NOP2/Sun RNA methyltransferase 2 (NSun2), which modifies many tRNAs to generate 5-mehtylcytidine (m^5C) modification, and whose knockdown was shown to promote stress-induced cleavage of tRNAs (67, 68). Whether I34 modification has a role in tRNA processing or if the generation of tRFs might have an effect on I34 levels in cells remains unknown. Still, the *in vitro* inhibition of ADAT activity reported here might shed light on a new layer of regulation of ADAT-mediated deamination and an additive role of tRFs in the regulation of translation.

Study of ADAT kinetics and structure and effect of a mutation linked to intellectual disability and strabismus

tRNA modifications and the enzymes that catalyze their synthesis are increasingly being linked to human diseases (101, 186). A number of mutations in tRNA modification enzymes have been associated with neurological conditions (18, 19, 96, 97, 181, 187), although the molecular bases for their pathogenesis remain unclear. The modification of adenosine to inosine at position 34 of the anticodon of tRNAs is catalyzed by the heterodimeric ADAT (1, 3). V128M mutation in ADAT3 subunit has been linked to severe neuropathies in patients homozygous for the mutation (18, 19). However, the molecular consequences of the mutation were unknown. We have recently developed a new method for the characterization of ADAT's enzymatic activity (13), and we have applied this method to study the kinetics of wild type human ADAT, and the impact of the ID-causing V128M mutation upon its deamination activity. We have also studied ADAT quaternary structure and the effect of V128M mutation on the enzyme stability.

Size-exclusion chromatography of affinity-purified ADAT revealed that the heterodimeric wt enzyme was found in two states: as a free protein and coeluting with tRNA. The population associated to tRNA eluted a higher elution volume (indicative of smaller particle size), suggesting that the heterodimer might adopt a more compact conformation upon substrate binding (**Figure 40**). To understand the molecular effects of ADAT3-V128M mutation, we tested its impact on the structural stability of the enzyme. *In silico* modeling studies had previously predicted that the mutation might alter a loop found on the surface of ADAT3 subunit, leading to changes in ADAT3 conformation (18). We found that this mutation promotes the proteolytic cleavage of ADAT3 subunit, generating a truncated ADAT3 subunit of a Mw similar to ADAT2. Affinity-purified mutant ADAT after SEC displayed two populations of ADAT-V128M: the full-length heterodimer and a truncated form as a result of the proteolytic cleavage in ADAT3 subunit. The truncated ADAT coeluted with tRNA, suggesting that the cleavage of ADAT3-V128M takes place in the presence of bound tRNA (and not in the wt ADAT). The elution of ADAT-V128M displayed an additional peak at the void volume containing protein aggregates of high Mw, suggestive of ADAT3-V128M aggregates (**Figure 42 and Supplementary Figure S10**). Consistent with the tendency of dissociated ADAT3-V128M to aggregate, it has been previously observed that the expression of soluble ADAT3 in *E. coli* requires co-expression of ADAT2 (114, 188). Structural studies have shown that methionine residues can form non-covalent interactions with aromatic-containing residues such as tryptophan, tyrosine or phenylalanine (189). Thus, the substitution of valine for methionine in ADAT3 surface might result in non-specific interactions with aromatic residues of other ADAT3 subunits leading to the formation of aggregates. Overall, albeit the observed reduced stability of ADAT3-V128M, the mutation does not prevent heterodimerization, as affinity chromatography

against a tag in ADAT3 subunit co-purified the two subunits in stoichiometric amounts (*see Methods section*), and ADAT-V128M eluted from the size-exclusion column mostly as a heterodimer (**Figure 42**).

Using the RFLP method to quantify A34-to-I34 editing (13), we monitored the activity of wt and mutant ADAT enzymes using $tRNA^{Ala}_{AGC}$ as a substrate. We found that the mutation decreases I34 formation and estimated a 6-fold decrease in activity for the mutant enzyme (**Figure 46**). Determination of the respective kinetic constants revealed a 10-fold decrease in the catalytic efficiency for ADAT mutant with respect to ADAT wt, a reduction mainly due to a 5-fold reduction in the k_{cat} value for ADAT V128M (**Table 7**). ADAT wt kinetic values were comparable to those reported for *T. brucei* ADAT and *E. coli* TadA (115, 163). However, we cannot exclude the possibility that the effect of V128M mutation is more severe upon other untested ADAT substrates, such as precursor forms of tRNAs that likely constitute the real substrates of the enzyme (3). Based on the reduction in tRNA editing activity caused by V128M mutation, restoring I34 levels seems a promising therapeutic strategy. Targeting the less active ADAT-V128M with small molecules that activate the enzyme might allow to recover the defective tRNA editing caused by the mutation.

Since ADAT2 is the catalytic subunit of ADAT enzyme, we explored the possibility that ADAT2 might be sufficient for the A34-to-I34 catalysis. *E. coli* expression and affinity-purification of human ADAT2 resulted in the generation of homodimers (**Figure 48**). Consistent with this ability of ADAT2 to homodimerize, human ADAT2 has been crystalized as a homodimer (Structural Genomics Consortium; PDB ID:3DH1), and homodimers of ADAT2 have been reported in *T. brucei* and *S. cerevisiae* (1, 115). These affinity-purified ADAT2 homodimers were unable to *in vitro* deaminate tRNA substrates at an enzyme concentration in the nanomolar range (equivalent to that used for the wt enzyme) (**Figure 49 and Supplementary Figure S11**). Similarly, *T. brucei* ADAT2 homodimers were reported to be capable of binding tRNA and Zn^{2+} ions, but catalytically inactive (115). Nevertheless, increasing ADAT2 concentration to the micromolar range resulted in *in vitro* deamination of tRNAs (**Figure 49 and Supplementary Figure S11**). This suggests that although ADAT2 homodimers have a lower deamination efficiency compared to the heterodimeric ADAT, they are still able to catalyze A-to-I conversion in tRNAs. This is in agreement with ADAT2 being the catalytic subunit and ADAT3 having a structural role (1). Still, the significantly lower catalytic efficiency of ADAT2 compared to the heterodimeric ADAT, indicates that ADAT3 is also important for the catalysis, e.g. contributing to the substrate binding. Thus, alterations in ADAT3 structure or stability are likely to have an impact in ADAT catalytic efficiency.

Given the essential nature of ADAT enzyme (1-5, 114), it is reasonable that only hypomorphic muatations like V128M, which maintain part of the activity, are viable, whereas null mutations would be incompatible with life. The apparently moderate effect of V128M mutation in ADAT activity is also consistent with the fact that only patients homozygous for

the mutation exhibit the syndrome, while heterozygosity does not have notorious effects. The severe neurological conditions in individuals homozygous for the mutation indicates that neurons are more sensitive to depletion of I34 than other cells. It has been proposed that the translation of specific genes might be more dependent on ADAT activity (9, 11). Thus, V128M could be altering the expression of genes that might be important for brain development and function.

We cannot exclude the possibility that the phenotype caused by V128M mutation is not exclusively due to the observed reduction in deaminase activity. Alterations of other non-canonical functions of the enzyme beyond A34-to-I34 editing might be playing a role in the pathology. Several studies in model organisms have shown that ADAT can interact with other components of the translation machinery and participate in processes beyond A34-to-I34 editing. In *A. thaliana*, ADAT3 was found to associate with the DEA(D/H)-box RNA helicase protein (4), whose human homolog is the Eukaryotic Translation Initiation Factor 4A1 (EIF4A1). Rubio *et al.* described an interaction of ADAT with the m^3C methyltransferase TRM140a (METTL6 in humans) in *T. brucei* that was required for the formation of m^3U_{32} in $tRNA^{Thr}$ and to prevent widespread genomic C-to-U deamination by ADAT while present in the nucleus of trypanosomes (190). Recent studies in human cells reported that the mutation altered the cellular localization of ADAT3. ADAT3-V128M was present in both cytosol and nucleus whereas the wt enzyme was restricted to the cytoplasm (*Manuscript under review. Rodríguez-Escribà et al.*). The disruption of the cellular localization of the enzyme might also contribute to altered levels of I34 or induce toxicity through other mechanisms. Taken together with our results, these findings indicate that the pathology related to V128M mutation may be due to a pleiotropic effect caused by the simultaneous impact of the mutation upon the enzyme's activity, stability, and cellular distribution. Further research is needed to understand the role of ADAT in cell physiology and its molecular mechanisms to predict the consequences of a mutation and accelerate the development of therapeutic strategies.

Small molecules targeting ADAT

There are no small molecules targeting ADAT reported to date. Thus, the discovery of either inhibitors (as regulators of the translation of certain genes, e.g. mucins) or activators (that might potentially recover the phenotype caused by V128M mutation) has a noteworthy interest. The first approach to find small molecules targeting ADAT enzyme was to test whether deaminase inhibitors (three ADA inhibitors and one cytidine deaminase inhibitor) might be active against human ADAT. Computational structural analysis predicted low scores for all four candidates (**Figure 53**). In agreement with this, our *in vitro* analysis revealed that none of the compounds was active against ADAT at a concentration of 1 mM (**Figure**

54). Thus, despite the conserved deaminase domains within deaminases, their catalytic centers would have evolved due to substrate specialization and, consequently, they exhibit significant structural differences.

To find better candidates, we performed a virtual screening of a library of chemical compounds. The *in vitro* analysis of the virtual screening hits revealed that three compounds were able to activate ADAT deamination activity at a concentration of 500 μM (**Figure 59**). As we reported that V128M mutation in ADAT enzyme (linked to ID and strabismus (18, 19)) causes a reduction of ADAT deamination activity, ADAT activators are promising candidates to potentially recover the defective tRNA editing caused by the mutation. Interestingly, one of them showed to activate also the mutant enzyme (**Figure 60**).

Finally, as for ADAT inhibitors, although some of the candidates showed mild inhibition of the deamination, we were unable to find potent inhibitors (**Figure 58**). The second virtual screening of a larger library of chemical compounds will increase the chances of finding better candidates. Having a crystal structure of the heterodimeric enzyme bound to a substrate or a small molecule would allow a better prediction of the binding of the candidates and greatly improve the screening of active compounds.

The inhibition of ADAT deamination reported by human tRFs suggests that variations of I34 modification levels might have physiological implications. We propose that ADAT activity might be necessary for the expression of mucins, however, the cellular effects caused by a depletion of I34 need to be further studied. Small molecules with the ability to inhibit the enzyme not only have potential therapeutic applications, but might serve as a tool to further study ADAT function in the cell context.

Conclusions

Development of an *in vitro* activity assay to monitor ADAT-mediated deamination

- We developed an *in vitro* assay to monitor I34 modification based on restriction fragment length polymorphism (RFLP) analyses. We applied this method to characterize human ADAT in terms of activity, substrates, regulation, as well as for drug discovery purposes.

- RFLP method can also be used to analyze other post-transcriptional tRNA modifications that produce the incorporation of a different base on the reverse transcribed cDNA. It is the case of m^2_2G modification catalyzed by TRMT1, which causes a G-to-A or T mutation upon RT.

- Monitoring ADAT-mediated deamination using the first emissive isomorphic and isofunctional nucleotide analogs in the context of a tRNA molecule is a promising strategy to *in vitro* quantify the levels of I34 in real time.

New insights into substrate recognition and regulation of human ADAT

- Human ADAT substrates do not share a common signature sequence for A34 deamination.

- The tRNA architecture is the most important feature for efficient A34-to-I34 conversion by ADAT.

- Both unfolded tRNAs and shorter tRNA variants are poor substrates of ADAT.

- Whereas the bacterial homolog TadA is able to deaminate a $tRNA^{Arg}$ mini-substrate containing the anticodon stem loop as efficiently as the full-length $tRNA^{Arg}$ (2, 118), equivalent fragments derived from human $tRNA^{Arg}$ and $tRNA^{Ala}$ were not active substrates of human ADAT.

- The enzyme-substrate interaction mode varies between different tRNAs, and parts of the tRNA far from the anticodon arm such as the acceptor stem play a role in ADAT binding.

- The recognition of $tRNA^{Arg}$ seems to have evolved as the other emergent substrate tRNAs, and is not restricted to the anticodon. Nevertheless, the deamination of $tRNA^{Arg}$ showed to be less sensitive to alterations in the tRNA structure than $tRNA^{Ala}$.

- The inhibition of ADAT deamination by human tRNA-derived fragments is a potential mechanism of control of the enzyme, which provides new insights into the regulation of ADAT function and may open a door for the development of new strategies to modulate ADAT activity.

Study of ADAT structure and kinetics and effect of a mutation linked to intellectual disability and strabismus

- Human ADAT is a heterodimeric enzyme that binds to tRNAs (from both human and *E. coli*).

- V128M mutation promotes a proteolytic cleavage of ADAT3 subunit, leading to a truncated heterodimer.

- ADAT modifies $tRNA^{Ala}_{AGC}$ in a concentration- and time-dependent manner. Full deamination is achieved at an enzyme concentration of 50 nM after 4 hours of incubation.

- ADAT V128M mutant has reduced tRNA deamination activity, with a 10-fold decrease in catalytic efficiency (k_{cat}/K_m) with respect to ADAT wt.

- The reduced catalytic activity of ADAT V128M may be linked to the observed reduced stability of ADAT3.

- Human ADAT2 forms homodimers, which are not catalytically active at the nanomolar range but deaminate $tRNA^{Ala}_{AGC}$ and $tRNA^{Arg}_{ACG}$ at the micromolar range.

Small molecules targeting ADAT

- The assayed deaminase inhibitors showed no activity against human ADAT enzyme.

- Three of the virtual screening hits activated ADAT enzyme at 500 μM concentration.

- One of the ADAT activators was also active against the V128M mutant enzyme, and might potentially be used to recover the defective deaminase activity caused by the mutation.

Supplementary material

Supplementary Figure S1. Synthetic HsADAT genes optimized for *E. coli* expression.

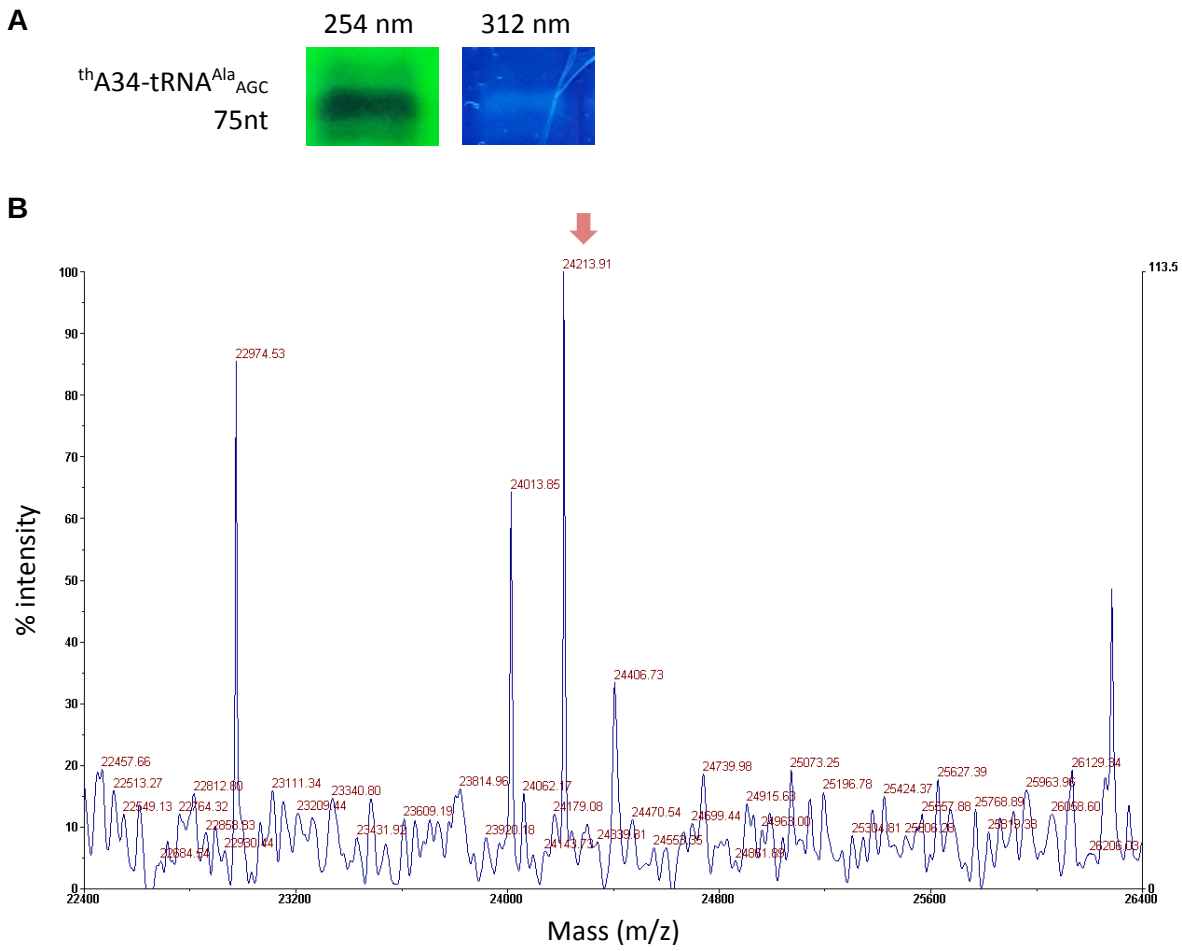
Scheme: HsADAT2 (with stop codon) - RBS (from pETDUET-1, lowercase) - HsADAT3 (without stop codon). Gene sequences were obtained using the amino acid sequences annotated in Uniprot for human ADAT2 (Q7Z6V5-1) and ADAT3 (Q96EY9). The cDNA was optimized for expression in *E. coli*.

ATGGAAGCAAAAGCAGCACCGAAACCGGCAGCAAGCGGTGCATGTAGCGTTAGTGCCGA
AGAAACCGAAAAATGGATGGAAGAGGCAATGCATATGGCAAAAGAAGCACTGGAAAATA
CCGAAGTTCCGGTTGGTTGTCTGATGGTGTATAATAACGAAGTTGTTGGCAAAGGTCGC
AATGAAGTTAATCAGACCAAAAATGCAACCCGTCATGCAGAAATGGTTGCAATTGATCA
GGTTCTGGATTGGTGTCGTCAGAGCGGTAAAAGCCCGAGCGAAGTTTTTGAACATACCG
TTCTGTATGTTACCGTTGAACCGTGTATTATGTGTGCAGCAGCACTGCGTCTGATGAAA
ATTCCGCTGGTTGTTTATGGTTGTCAGAATGAACGTTTTGGTGGTTGTGGTAGCGTTCT
GAATATTGCAAGTGCCGATCTGCCGAATACCGGTCGTCCGTTTCAGTGTATTCCGGGTT
ATCGTGCGGAAGAGGCCGTTGAAATGCTGAAAACCTTCTATAAACAAGAGAATCCGAAT
GCACCGAAAAGCAAAGTGCGTAAAAAAGAATGTCAGAAAAGCTAAtgcttaagtcgaac
agaaagtaatcgtattgtacacggccgcataatcgaaattaatacgactcactataggg
gaattgtgagcggataacaattccccatcttagtatattagttaagtataagaaggaga
tatacatATGGAACCGGCACCGGGTCTGGTTGAACAGCCGAAATGTCTGGAAGCAGGTA
GTCCGGAACCGGAACCTGCACCGTGGCAGGCACTGCCGGTTCTGAGCGAAAAACAGAGC
GGTGATGTTGAACTGGTTCTGGCATATGCAGCACCGGTTCTGGATAAACGTCAGACCAG
CCGTCTGCTGAAAGAAGTTAGCGCACTGCATCCGCTGCCTGCACAGCCGCATCTGAAAC
GTGTTCGTCCGAGCCGTGATGCAGGTTCACCGCATGCACTGGAAATGCTGCTGTGTCTG
GCAGGTCCGGCAAGCGGTCCGCGTAGCCTGGCAGAACTGCTGCCTCGTCCGGCAGTTGA
TCCGCGTGGTCTGGGTCAGCCGTTTCTGGTTCCGGTGCCTGCCCGTCCGCCTCTGACCC
GTGGTCAGTTTGAAGAAGCACGCGCACATTGGCCGACCAGCTTTCATGAAGATAAACAG
GTTACCAGCGCACTGGCAGGCCGTCTGTTTAGCACCCAAGAACGTGCAGCAATGCAGAG
CCATATGGAACGTGCCGTTTGGGCAGCACGTCGTGCAGCAGCACGTGGTCTGCGTGCAG
TTGGTGCCGTTGTTGTTGATCCGGCAAGTGATCGTGTTCTGGCAACCGGTCATGATTGT
AGCTGTGCAGATAACCCGCTGCTGCATGCCGTTATGGTTTGTGTTGATCTGGTTGCACG
CGGTCAGGGTCGTGGCACCTATGATTTTCGTCCGTTTCCGGCATGTAGCTTTGCACCGG
CAGCCGCACCGCAGGCAGTTCGTGCTGGTGCAGTTCGTAAACTGGATGCAGATGAAGAT
GGTCTGCCGTATCTGTGTACCGGTTATGATCTGTATGTTACCCGTGAACCGTGTGCAAT
GTGTGCCATGGCCCTGGTTCATGCACGTATTCTGCGTGTTTTTTATGGTGCACCGAGTC
CGGATGGTGCGCTGGGCACCCGTTTTCGTATTCATGCGCGTCCGGATCTGAATCATCGT
TTTCAGGTTTTTCGTGGTGTTCTGGAAGAACAGTGTCGTTGGCTGGATCCGGATACC

Supplementary Figure S2. Representative SDS-PAGE (A) and MALDI (B) analyses of thA34-tRNA$^{Ala}_{AGC}$.

A SDS-PAGE was visualized at 254 nm, where nucleic acids are detected, and at 312 nm, where only fluorescent molecules are detected. The band observed at 254 nm and 312 nm coincided, confirming thus that the tRNA was emissive.

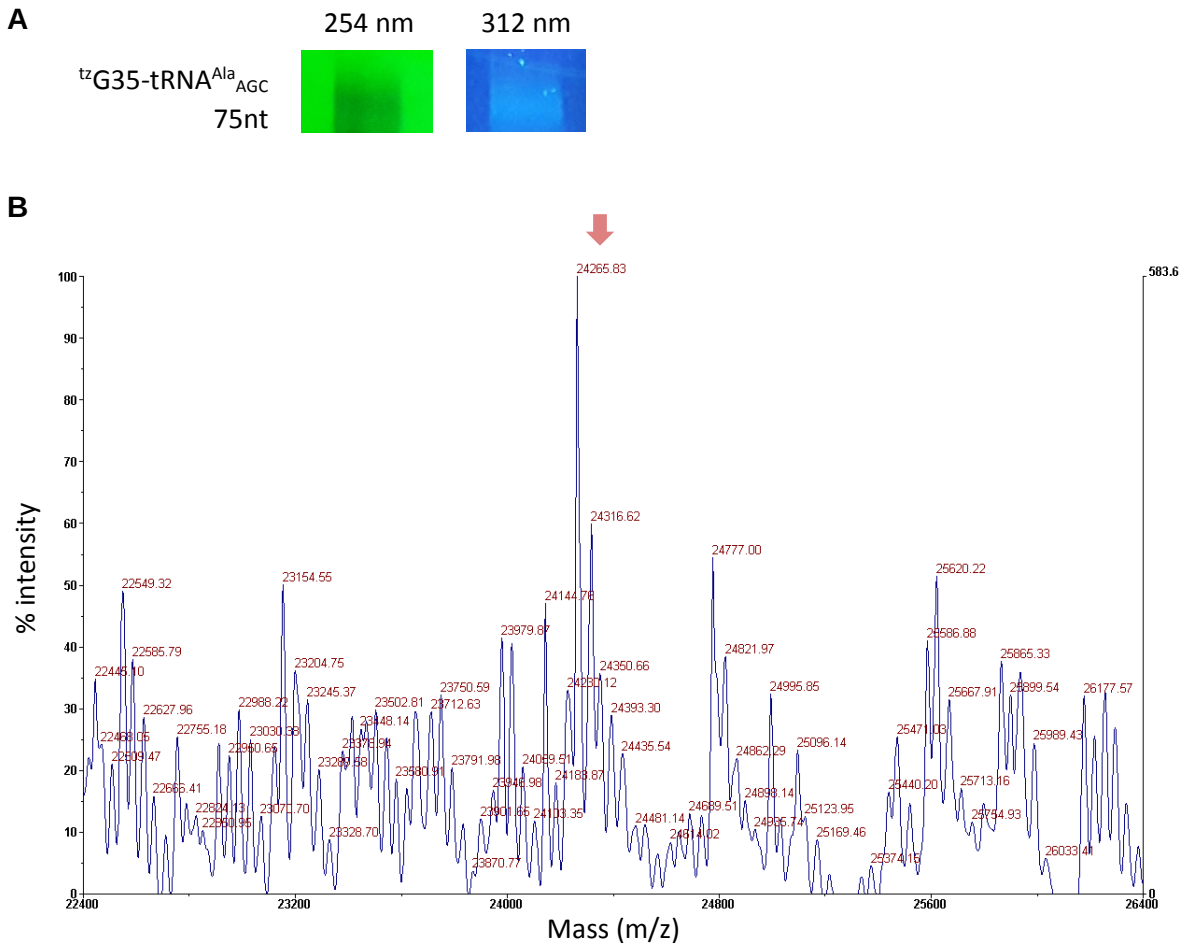
B The main peak of the MALDI spectra (red arrow) coincided with the expected molecular weight for thA34-tRNA$^{Ala}_{AGC}$ (23,984.66 g/mol).



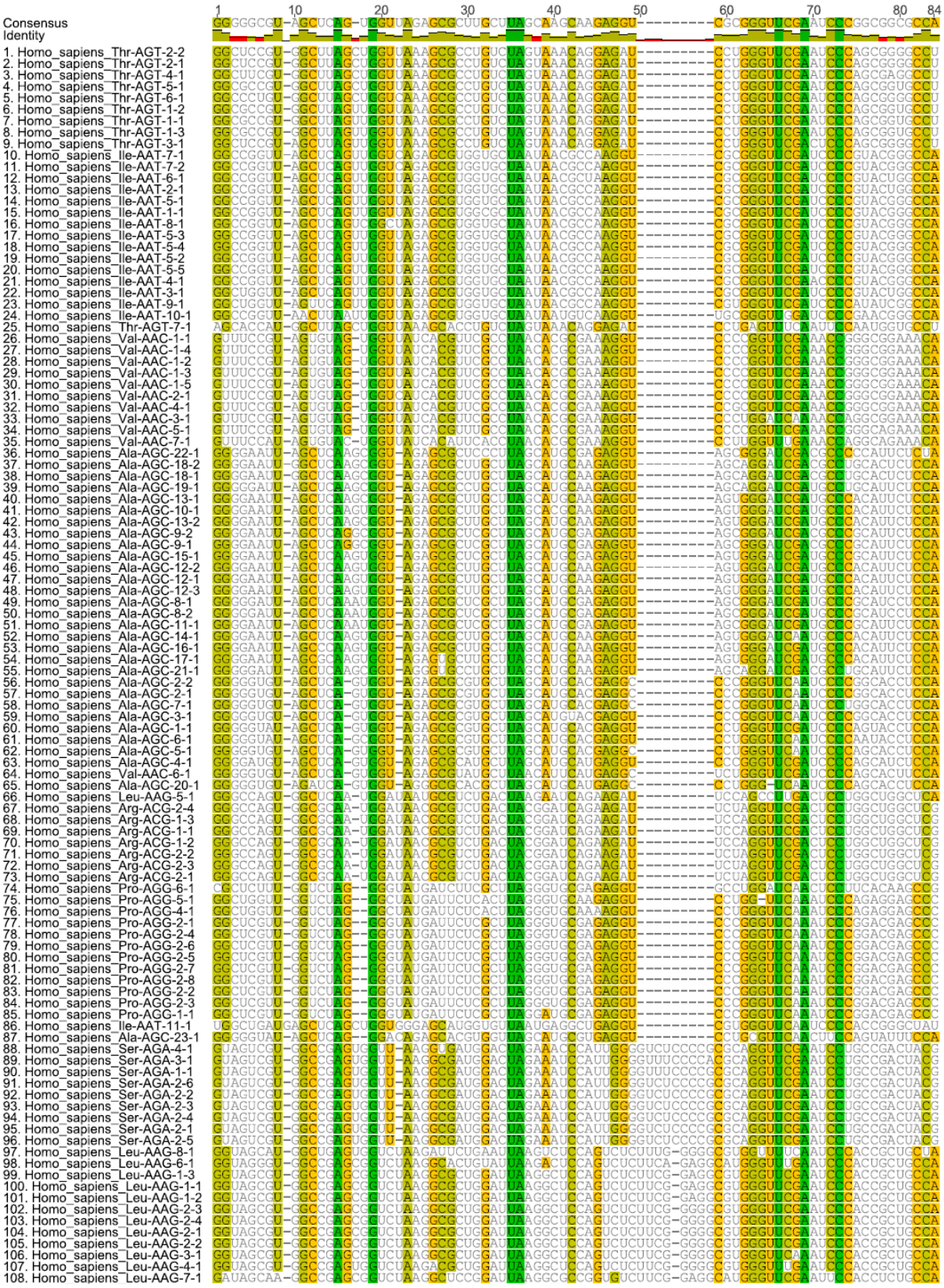
Supplementary Figure S3. Representative SDS-PAGE (A) and MALDI (B) analyses of ^{tz}G35-tRNA$^{Ala}_{AGC}$.

A SDS-PAGE was visualized at 254 nm, where nucleic acids are detected, and at 312 nm, where only fluorescent molecules are detected. The band observed at 254 nm and 312 nm coincided, confirming thus that the tRNA was emissive.

B The main peak of the MALDI spectra (red arrow) coincided with the expected molecular weight for ^{tz}G35-tRNA$^{Ala}_{AGC}$ (23,985.64 g/mol).

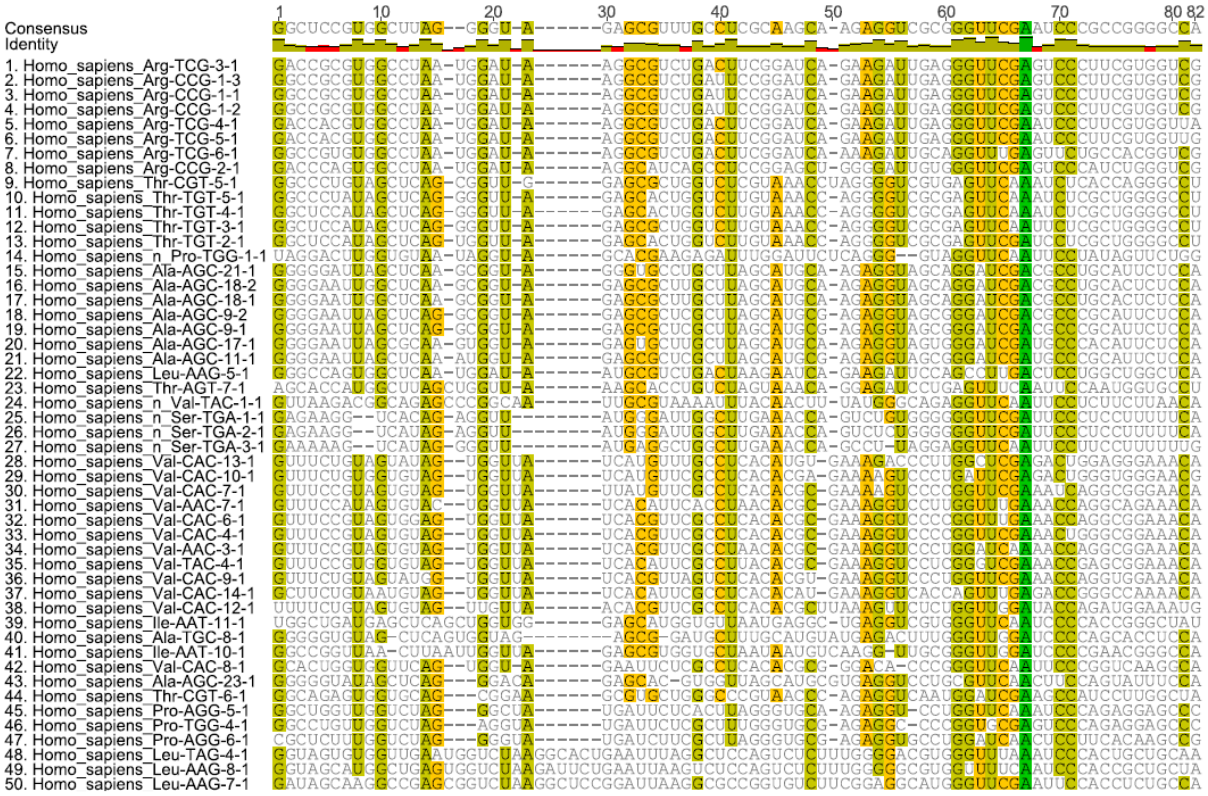


Supplementary Figure S4. Sequence alignment of all human ADAT substrate tRNAs.



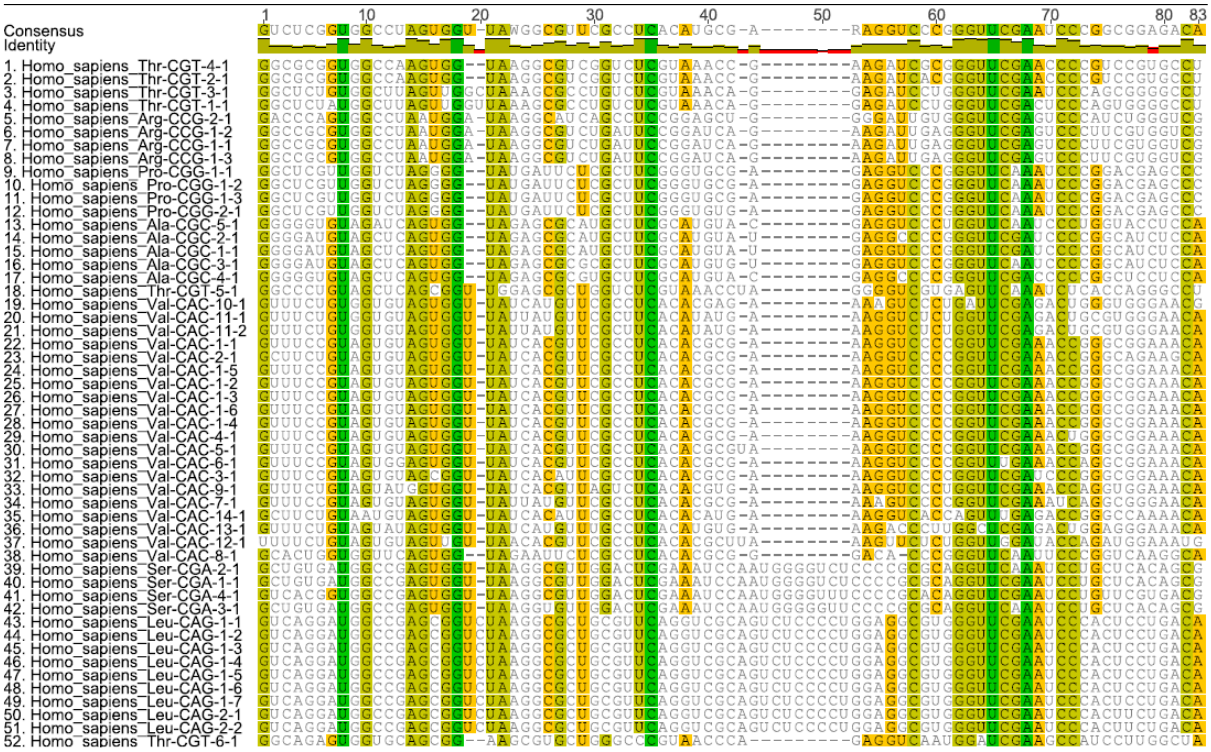
Supplementary Figure S5. Sequence alignment a selection of 50 representative tRNAs, including ADAT substrates and non-ADAT substrates (all of them encoding for amino acids T, A, P, S, L, I, V, R).

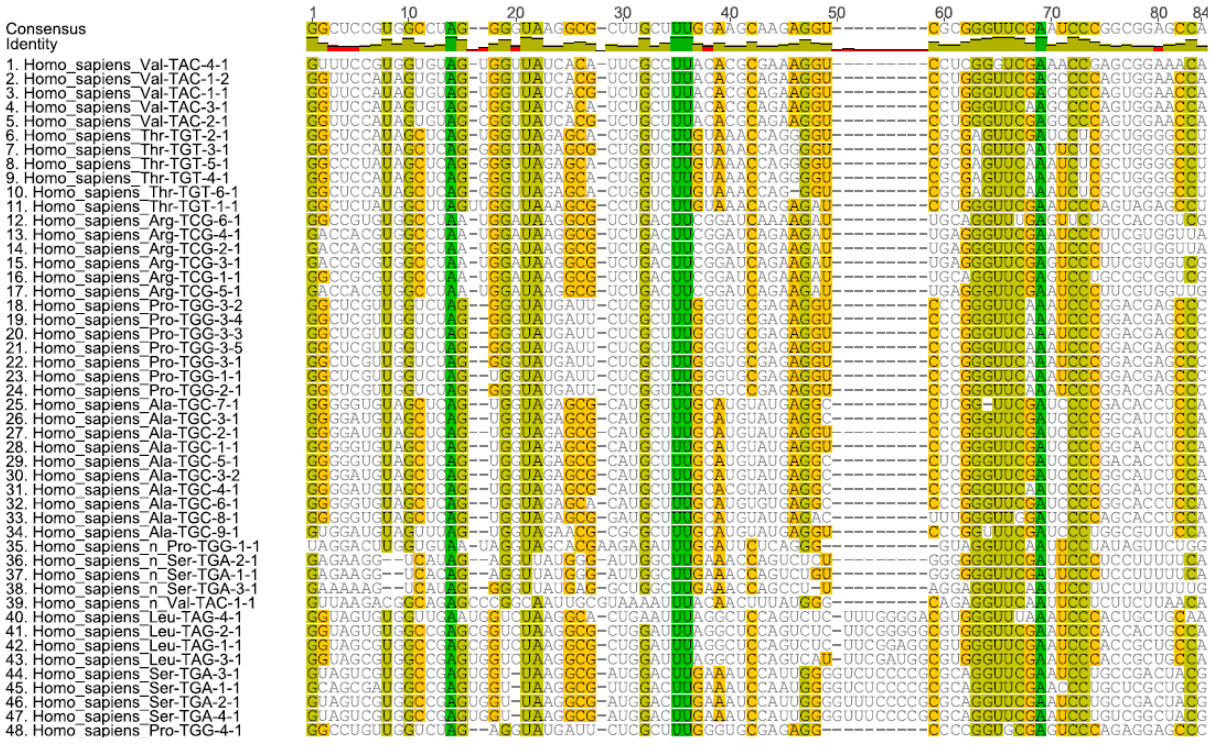
The selected tRNAs containing either A34, C34, or U34 (no G34-tRNAs encoding for T, A, P, S, L, I, V, R are present in human). T-coffee was used to select the 50 tRNAs with the highest sequence variability from a total of 180.



Supplementary Figure S6. Sequence alignment of non-ADAT substrate tRNAs (all of them encoding for amino acids T, A, P, S, L, I, V, R).

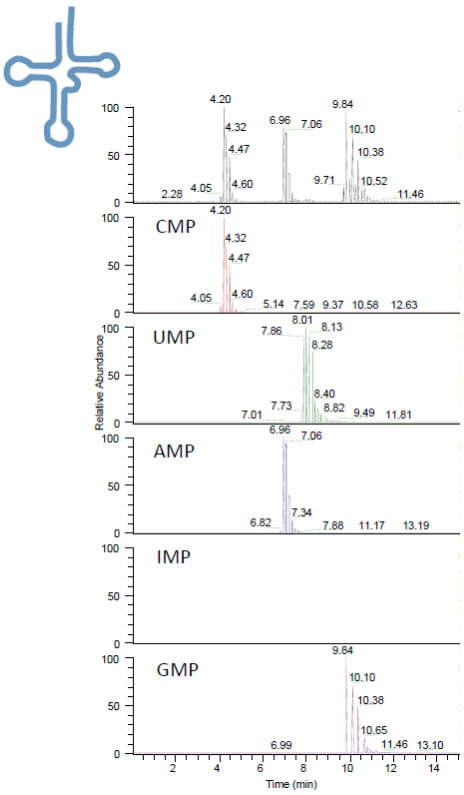
Aligned sequences of non-ADAT substrate tRNAs with C34 (upper panel) and U34 (lower panel) (G34 tRNAs encoding for T, A, P, S, L, I, V, R are not found).

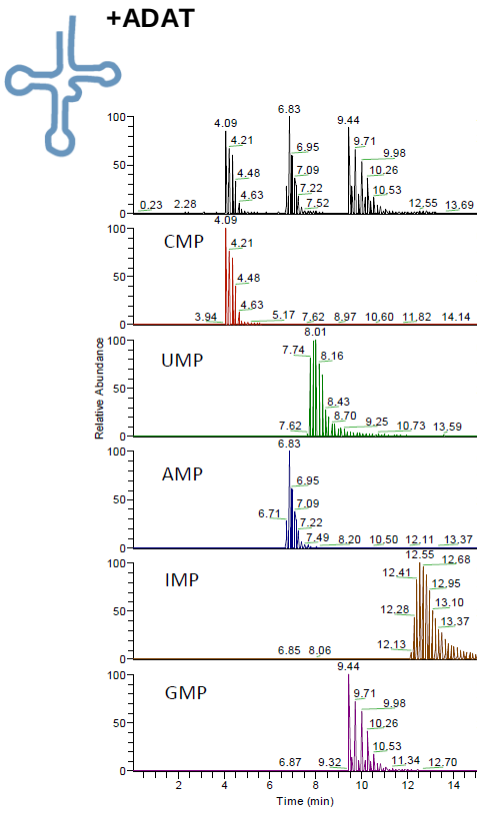


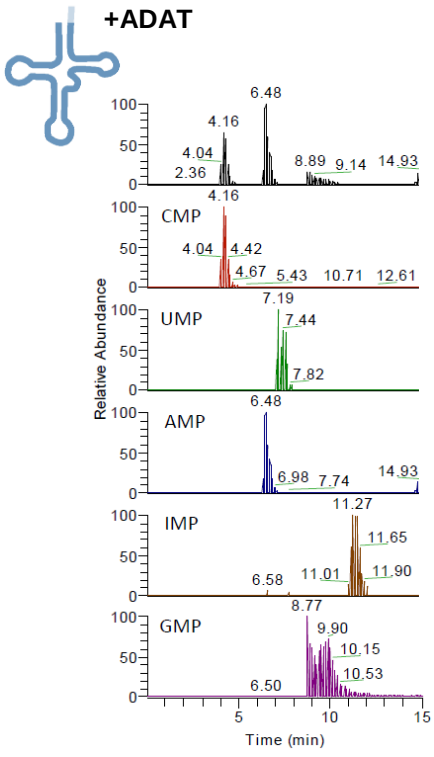


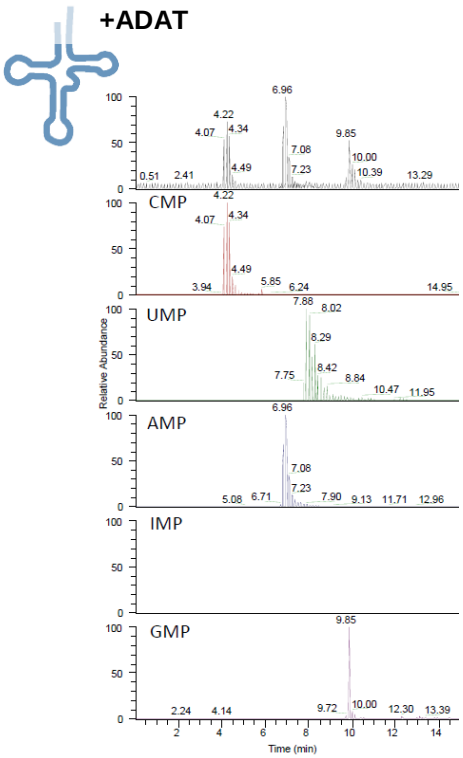
Supplementary Figure S7. Analysis of ADAT-mediated deamination of $tRNA^{Ala}_{AGC}$ and tRNA mini-substrates derived from $tRNA^{Ala}_{AGC}$ after 1 hour incubation at 37ºC.

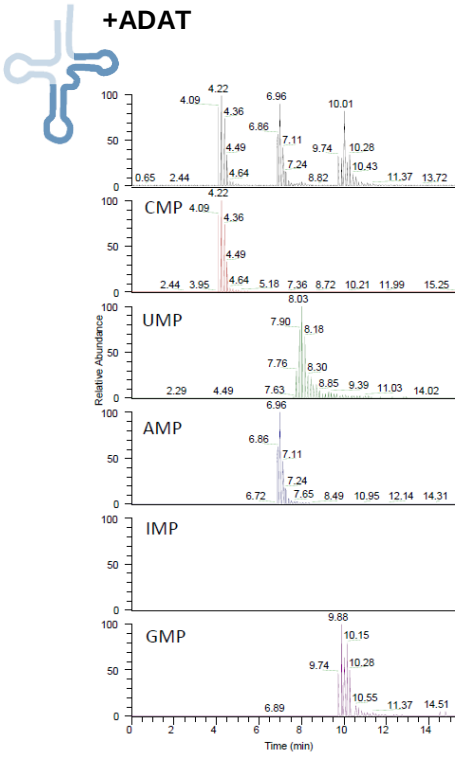
UPLC-MS analysis of the tRNA fragments after incubation with 800 nM ADAT for 1 hour at 37ºC. IMP was only detected in the full-length $tRNA^{Ala}_{AGC}$ and $tRNA^{Ala}_{AGC}$ lacking the CCA.

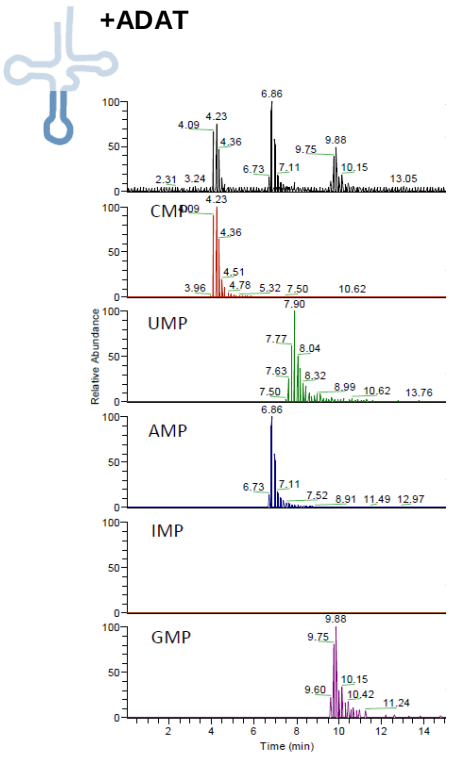




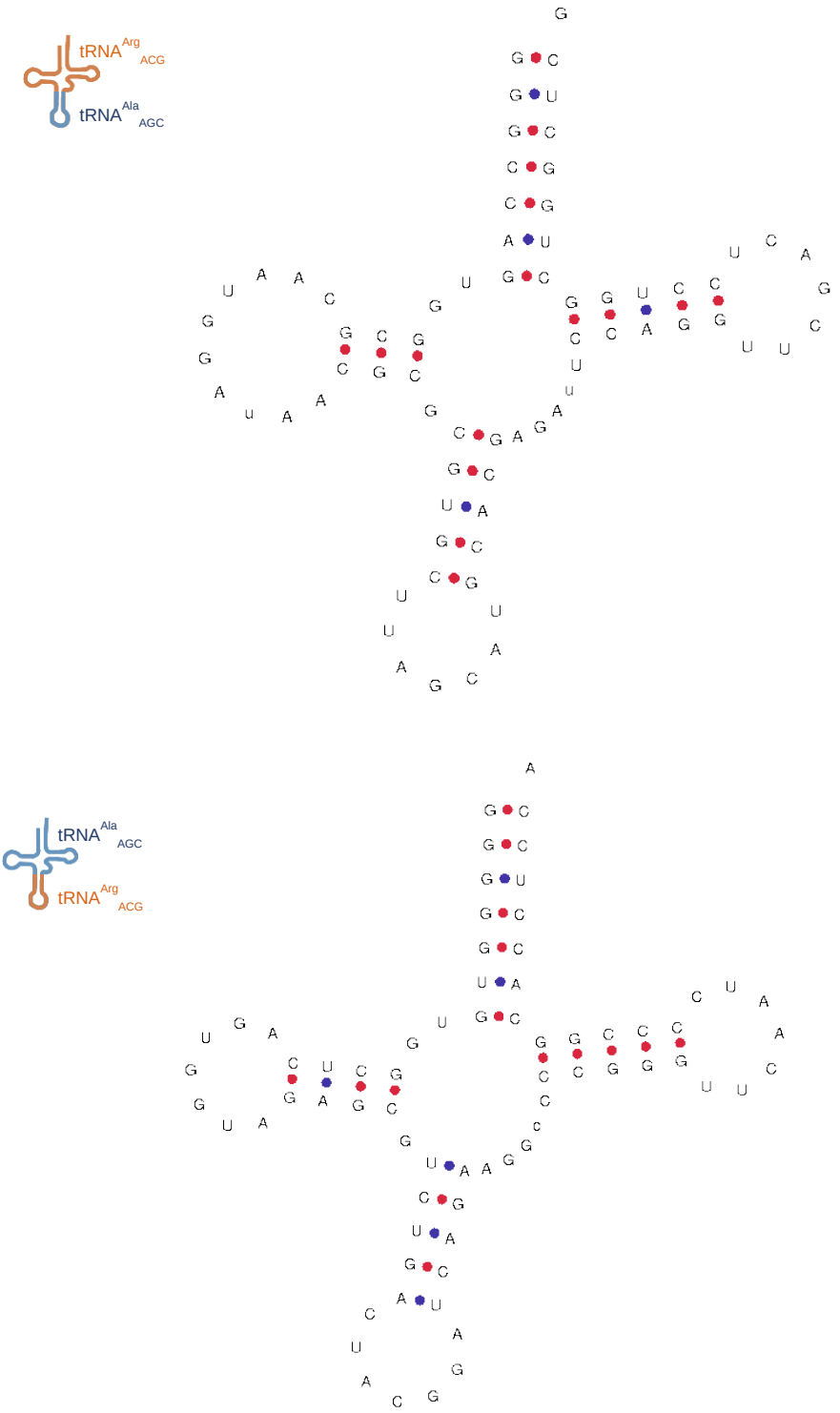





+ADAT
CMP
UMP
AMP
IMP
GMP
Relative Abundance
Time (min)


+ADAT
CMP
UMP
AMP
IMP
GMP
Relative Abundance
Time (min)

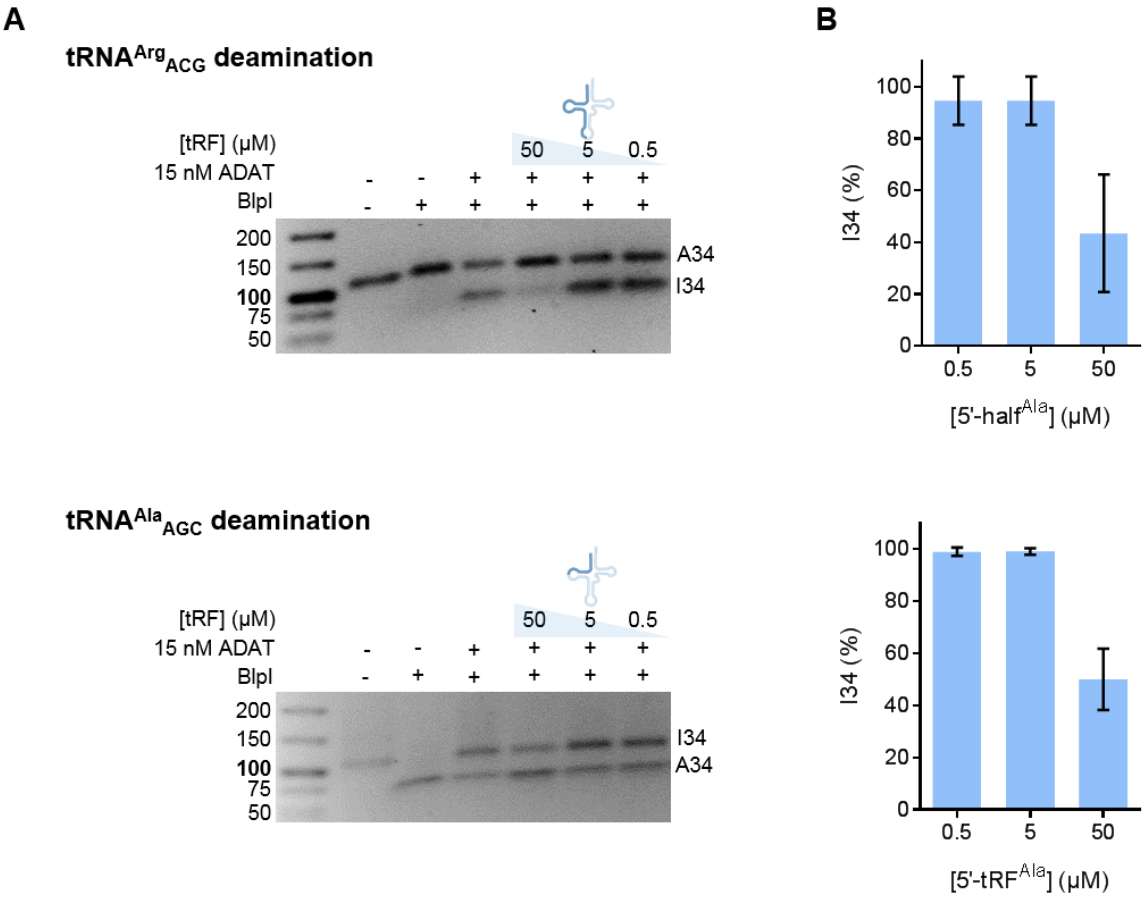
Supplementary Figure S8. Prediction of the seconday structure of the chimeric tRNAs. Both chimeric tRNAs display the typical cloverleaf secondary structure.



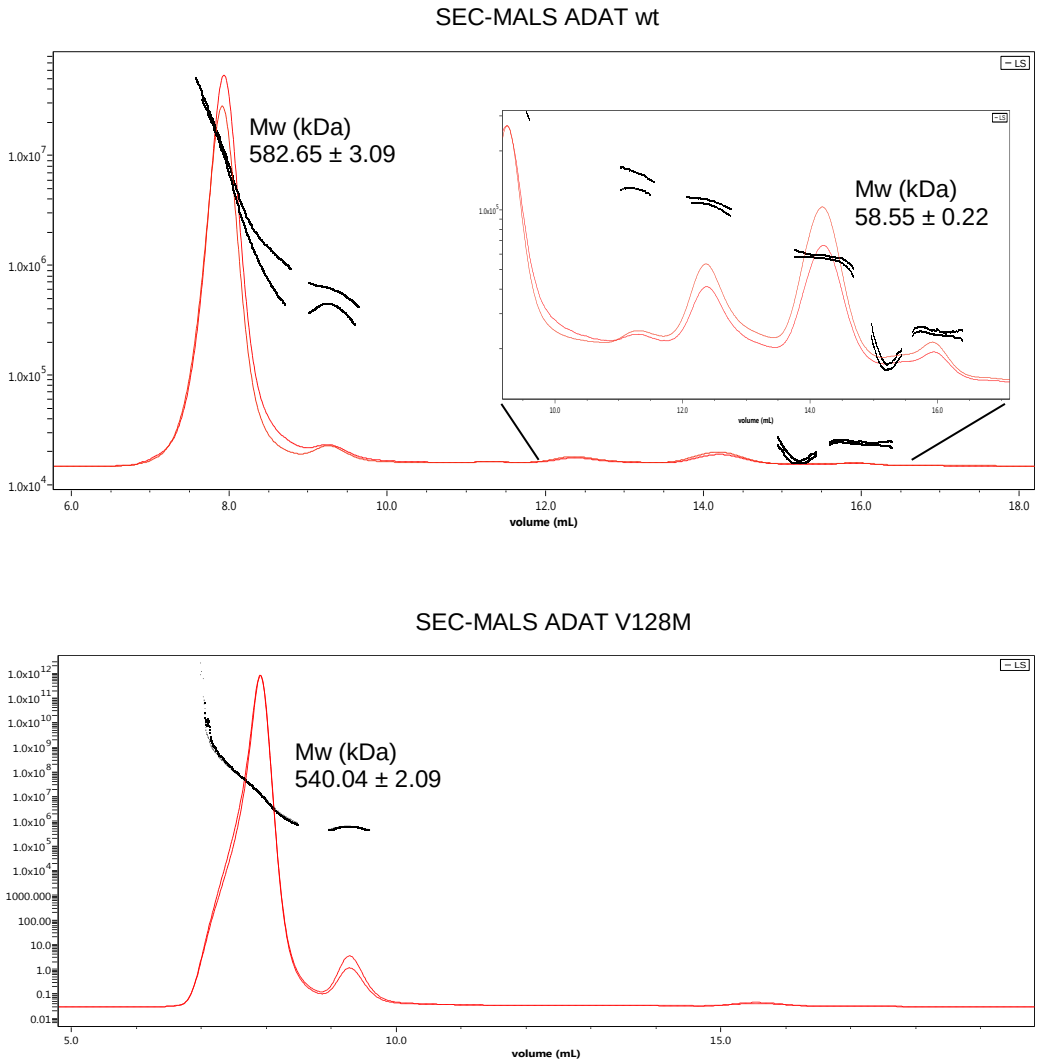
Supplementary Figure S9. Dose-response curves for 5'-half and 5'-tRF derived from $tRNA^{Ala}_{AGC}$ inhibitory effect on ADAT-mediated deamination of $tRNA^{Arg}_{ACG}$ and $tRNA^{Ala}_{AGC}$.

A Representative RFLP experiments to assess the inhibitory effect of 5'-half derived from $tRNA^{Ala}_{AGC}$ and 5'-tRF derived from $tRNA^{Ala}_{AGC}$ on ADAT-mediated deamination of $tRNA^{Arg}_{ACG}$ (upper panel) and $tRNA^{Ala}_{AGC}$ (lower panel).

B Percentage of I34 modification upon varying concentrations of the tRFs from three independent experiments and averaged. Error bars reflect the standard deviation.

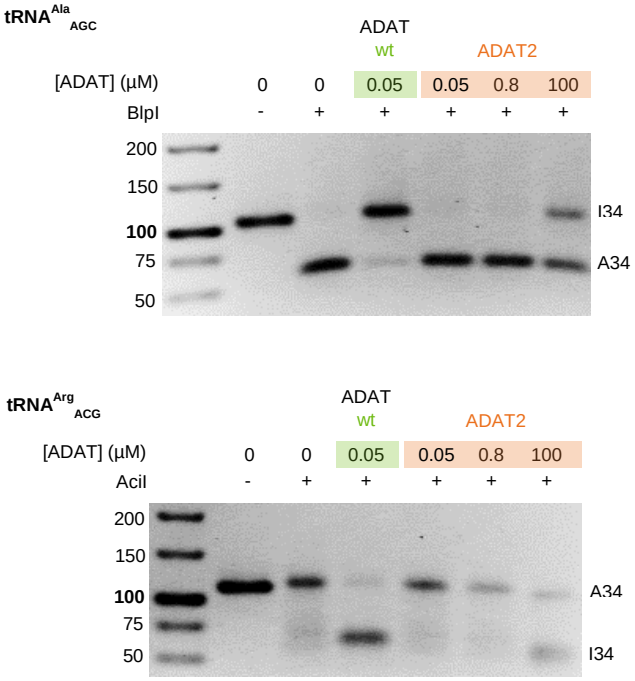


Supplementary Figure S10. SEC-MALS analyses for ADAT wt (upper panel) and ADAT V128M (lower panel). Elution volume (mL) is represented on the x-axis and Molar mass (g/mol) is represented on the y-axis.



Supplementary Figure S11. Analysis of the deamination of $tRNA^{Ala}_{AGC}$ and $tRNA^{Arg}_{ACG}$ by human ADAT2 homodimer (RFLP method).

RFLP experiment to assess the deamination activity of ADAT2 homodimer on $tRNA^{Ala}_{AGC}$ (upper panel) and $tRNA^{Arg}_{ACG}$ (lower panel). Inosine formation was monitored at the indicated enzyme concentrations after 4 hours of incubation. 50 nM ADAT wt was used as a positive control. BlpI cuts A34-$tRNA^{Ala}_{AGC}$-derived PCR amplicons. AciI cuts I34-$tRNA^{Arg}_{ACG}$-derived PCR amplicons.



Acknowledgments

Gràcies al meu director de tesi, Lluís Ribas, per haver confiat en mi i haver-me assessorat en el que ha estat la meva principal font de mal de caps durant els darrers quatre anys.

A Toni Riera, el meu inicial co-director de tesi, amb la recança d'haver-la acabada sense arribar a passar una temporada al teu laboratori.

A Elena Escubedo, per acceptar ser la meva tutora.

Als membres del meu *Thesis Advisory Committee*, Modesto Orozco, Oriol Gallego i Maria Dolors Pujol, pel seguiment que han fet del meu treball durant aquests quatre anys. I també el meu agraïment a les seves aportacions individuals: a Modesto Orozco, per ser el meu supervisor de la *lab rotation*; a Oriol Gallego, per la seva col·laboració en el projecte d'ADAT i per ensenyar-me la seva tècnica PICT (Protein interactions from Imaging of Complexes after Translocation); i, de manera molt especial, a Dolors Pujol, per guiar-me i donar-me sempre els millors consells des que vaig començar Farmàcia.

A Pepita Badia, per ser la coordinadora del Programa de doctorat de Biotecnologia i respondre incansablement a tots els meus dubtes.

A Xavier Barril i a Sergi Ruiz, per les anàlisis computacionals per trobar moduladors d'ADAT.

A la Mass Spectrometry Facility de l'IRB, en especial a Marta Vilaseca i a Mireia Díaz.

A Roman Bonet i a Joan Pous, pel seu suport tècnic amb el SEC-MALS i DLS.

A Jorge García, por optimizar la expresión y purificación de ADAT, ha sido un placer colaborar contigo.

A tots els membres del Ribas lab. Sobretot a Maria Castellví, a Carla Fernández i a Adrianos Skaros, per ser els millors *minions* del món. Em sento molt afortunada que hagueu format part de la meva tesi i no tinc cap dubte que hem format un equip immillorable. A Marta, a Noelia i a Adrian, per la seva col·laboració amb el projecte d'ADAT. A Enric, perquè encara que no ho reconeguem les teves bromes fan bastanta gràcia. Sedi from Iran, it was a pleasure working with you, but even more eating your delicious Iranian food. A Núria, tot i que ens canviessis per Suïssa no ens hem oblidat de tu i encara repetim les teves frases cèlebres. A Maria Carretero, per ser l'artista del lab, la pintura que em vas fer segueix sent el meu quadre preferit. Als meus pilars del laboratori, Alba, Albert i Àlbert. A l'Alba, tot i ser de les darreres incorporacions al lab, és com si ens coneguéssim de tota la vida; gràcies per mimar-me sempre i per *ses nostres berenetes* al banc. A l'Albert, per ser el millor company de tesi que podia haver imaginat, perquè sense les nostres converses terapèutiques probablement hagués hagut de recórrer al psicòleg, i perquè ningú no sap tant de restaurants *top* com tu. A l'Àlbert, per cuidar-me sempre, per fer-me veure que els problemes són menys problemes, per fer-me riure encara que no vulgui i per ser la meva part preferida de la tesi.

Thanks to my laboratory and family in San Diego. Thanks Dr. Tor for accepting me in your group and being the best boss. Thanks Yao for being the best supervisor and my BFF (best foodie friend). And thanks to the rest of the members of the lab, especially Paul, Alex, Keivin, Patricia, François, Jackie, and Andrea. Also thanks to my lovely housemates, Ayat and Melisa, and to my partner in crime Becca.

A la meva família de l'IRB. A l'Anna Casas, des del PhD Retreat a Nijmegen vaig saber que no em separaria de tu. Als meus preferits de can Giralt, Cris, Martí i Salva. Cris i Martí, tots sabem que les festes que no acaben amb un esmorzar o un audio des del llit no es consideren festes. Salva, el nostre mestre cerveser i creador de la Salvafia, des de l'inici del doctorat sempre has estat la meva referència de farmacèutic top. A Leyre, por los viajes, las comidas détox, las fiestas de tranquis y nuestra luna de miel en Nusa Lembongan. A l'Anna Bellmunt, pels cursos de Bio-business compartits, perquè estem fetes per ser CEOs i perquè la nostra startup és imminent. A tota la resta de companys de l'IRB, Susi, Gemma, Laura, Ernest, Craig, Jürgen, Jordi, Adrian, Ricardo, Anna, Maria, Gianmarco, Marina, Clara, Irene, Ela, Busra, Judith, Anna, Berta, Rosa, Sandra, i un llarg etcètera, no oblidaré els sopars, festes i viatges compartits.

A la meva Farmília, per ser els millors amics que podria demanar. I molt especialment a l'inseparable equip Namaste, Ona, Anna i Emma, perquè encara que estiguem en la distància és com si no ens haguem separat mai. També a la Sara, per les quedades a la biblioteca, tu estudiant FIR i jo escrivint tesi.

Als meus pares. Tot agraïment es queda curt. Per ser els principals responsables que m'embranqués en aquesta aventura, per ajudar-me i animar-me sempre, i perquè és gràcies a vosaltres que he arribat fins on he arribat.

I gràcies a tots aquells que en algun moment, en més o menys mesura, han format part del meu doctorat.

References

1. Gerber AP, Keller W. An adenosine deaminase that generates inosine at the wobble position of tRNAs. Science. 1999;286(5442):1146-9. Epub 1999/11/05.

2. Wolf J, Gerber AP, Keller W. tadA, an essential tRNA-specific adenosine deaminase from Escherichia coli. EMBO J. 2002;21(14):3841-51. Epub 2002/07/12.

3. Torres AG, Pineyro D, Rodriguez-Escriba M, Camacho N, Reina O, Saint-Leger A, et al. Inosine modifications in human tRNAs are incorporated at the precursor tRNA level. Nucleic Acids Res. 2015;43(10):5145-57. Epub 2015/04/29.

4. 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. Epub 2014/10/16.

5. Tsutsumi S, Sugiura R, Ma Y, Tokuoka 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 Biol Chem. 2007;282(46):33459-65. Epub 2007/09/19.

6. Crick FH. Codon--anticodon pairing: the wobble hypothesis. Journal of molecular biology. 1966;19(2):548-55. Epub 1966/08/01.

7. Torres AG, Pineyro D, Filonava L, Stracker TH, Batlle E, Ribas de Pouplana L. A-to-I editing on tRNAs: biochemical, biological and evolutionary implications. FEBS Lett. 2014;588(23):4279-86. Epub 2014/09/30.

8. Phizicky EM, Hopper AK. tRNA biology charges to the front. Genes & development. 2010;24(17):1832-60. Epub 2010/09/03.

9. Novoa EM, Pavon-Eternod M, Pan T, Ribas de Pouplana L. A role for tRNA modifications in genome structure and codon usage. Cell. 2012;149(1):202-13. Epub 2012/04/03.

10. Schaub M, Keller W. RNA editing by adenosine deaminases generates RNA and protein diversity. Biochimie. 2002;84(8):791-803. Epub 2002/11/30.

11. Rafels-Ybern A, Attolini CS, Ribas de Pouplana L. Distribution of ADAT-Dependent Codons in the Human Transcriptome. Int J Mol Sci. 2015;16(8):17303-14. Epub 2015/08/01.

12. Rafels-Ybern A, Torres AG, Grau-Bove X, Ruiz-Trillo I, de Pouplana LR. Codon adaptation to tRNAs with Inosine modification at position 34 is widespread among Eukaryotes and present in two Bacterial phyla. RNA biology. 2017:0. Epub 2017/09/08.

13. Wulff TF, Arguello RJ, Molina Jordan M, Roura Frigole H, Hauquier G, Filonava L, et al. Detection of a Subset of Posttranscriptional Transfer RNA Modifications in Vivo with a Restriction Fragment Length Polymorphism-Based Method. Biochemistry. 2017;56(31):4029-38. Epub 2017/07/14.

14. Sinkeldam RW, McCoy LS, Shin D, Tor Y. Enzymatic interconversion of isomorphic fluorescent nucleosides: adenosine deaminase transforms an adenosine analogue into an inosine analogue. Angew Chem Int Ed Engl. 2013;52(52):14026-30. Epub 2013/11/30.

15. McCoy LS, Shin D, Tor Y. Isomorphic emissive GTP surrogate facilitates initiation and elongation of in vitro transcription reactions. Journal of the American Chemical Society. 2014;136(43):15176-84. Epub 2014/09/26.

16. Rovira AR, Fin A, Tor Y. Chemical Mutagenesis of an Emissive RNA Alphabet. J Am Chem Soc. 2015;137(46):14602-5. Epub 2015/11/03.

17. Juhling F, Morl M, Hartmann RK, Sprinzl M, Stadler PF, Putz J. tRNAdb 2009: compilation of tRNA sequences and tRNA genes. Nucleic acids research. 2009;37(Database issue):D159-62. Epub 2008/10/30.

18. Alazami AM, Hijazi H, Al-Dosari MS, Shaheen R, Hashem A, Aldahmesh MA, et al. Mutation in ADAT3, encoding adenosine deaminase acting on transfer RNA, causes intellectual disability and strabismus. Journal of medical genetics. 2013;50(7):425-30. Epub 2013/04/27.

19. El-Hattab AW, Saleh MA, Hashem A, Al-Owain M, Asmari AA, Rabei H, et al. ADAT3-related intellectual disability: Further delineation of the phenotype. American journal of medical genetics Part A. 2016;170A(5):1142-7. Epub 2016/02/05.

20. Bianconi E, Piovesan A, Facchin F, Beraudi A, Casadei R, Frabetti F, et al. An estimation of the number of cells in the human body. Annals of human biology. 2013;40(6):463-71. Epub 2013/07/09.

21. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. Cell. 2015;161(5):1202-14. Epub 2015/05/23.

22. Crick F. Central dogma of molecular biology. Nature. 1970;227(5258):561-3. Epub 1970/08/08.

23. Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature. 1953;171(4356):737-8. Epub 1953/04/25.

24. Berg JM, Tymoczko JL, Stryer L, Stryer L. Biochemistry. 5th ed. New York: W.H. Freeman; 2002.

25. Woese CR. Order in the genetic code. Proceedings of the National Academy of Sciences of the United States of America. 1965;54(1):71-5. Epub 1965/07/01.

26. Alff-Steinberger C. The genetic code and error transmission. Proceedings of the National Academy of Sciences of the United States of America. 1969;64(2):584-91. Epub 1969/10/01.

27. Ibba M, Soll D. Aminoacyl-tRNA synthesis. Annual review of biochemistry. 2000;69:617-50. Epub 2000/08/31.

28. Laporte D, Huot JL, Bader G, Enkler L, Senger B, Becker HD. Exploring the evolutionary diversity and assembly modes of multi-aminoacyl-tRNA synthetase complexes: lessons from unicellular organisms. FEBS Lett. 2014;588(23):4268-78. Epub 2014/10/16.

29. Alberts B. Molecular biology of the cell. 4th ed. New York: Garland Science; 2002. xxxiv, 1548 p. p.

30. Kirchner S, Ignatova Z. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. Nature reviews Genetics. 2015;16(2):98-112. Epub 2014/12/24.

31. Holley RW, Apgar J, Everett GA, Madison JT, Marquisee M, Merrill SH, et al. Structure of a Ribonucleic Acid. Science. 1965;147(3664):1462-5. Epub 1965/03/19.

32. Giege R, Puglisi JD, Florentz C. tRNA structure and aminoacylation efficiency. Progress in nucleic acid research and molecular biology. 1993;45:129-206. Epub 1993/01/01.

33. Giege R, Juhling F, Putz J, Stadler P, Sauter C, Florentz C. Structure of transfer RNAs: similarity and variability. Wiley interdisciplinary reviews RNA. 2012;3(1):37-61. Epub 2011/10/01.

34. Lodish HF. Molecular cell biology. 4th ed. New York: W.H. Freeman; 2000. xxxvi, 1084, G-17, I-36 p. p.

35. Chan PP, Lowe TM. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. Nucleic acids research. 2009;37(Database issue):D93-7. Epub 2008/11/06.

36. Iben JR, Maraia RJ. tRNA gene copy number variation in humans. Gene. 2014;536(2):376-84. Epub 2013/12/18.

37. Parisien M, Wang X, Pan T. Diversity of human tRNA genes from the 1000-genomes project. RNA biology. 2013;10(12):1853-67. Epub 2014/01/23.

38. Plotkin JB, Kudla G. Synonymous but not the same: the causes and consequences of codon bias. Nature reviews Genetics. 2011;12(1):32-42. Epub 2010/11/26.

39. Tuller T, Carmi A, Vestsigian K, Navon S, Dorfan Y, Zaborske J, et al. An evolutionarily conserved mechanism for controlling the efficiency of protein translation. Cell. 2010;141(2):344-54. Epub 2010/04/21.

40. Dittmar KA, Goodenbour JM, Pan T. Tissue-specific differences in human transfer RNA expression. PLoS genetics. 2006;2(12):e221. Epub 2006/12/30.

41. Novoa EM, Ribas de Pouplana L. Speeding with control: codon usage, tRNAs, and ribosomes. Trends in genetics : TIG. 2012;28(11):574-81. Epub 2012/08/28.

42. Gingold H, Tehler D, Christoffersen NR, Nielsen MM, Asmar F, Kooistra SM, et al. A dual program for translation regulation in cellular proliferation and differentiation. Cell. 2014;158(6):1281-92. Epub 2014/09/13.

43. Fredrick K, Ibba M. How the sequence of a gene can tune its translation. Cell. 2010;141(2):227-9. Epub 2010/04/21.

44. Zhang G, Ignatova Z. Folding at the birth of the nascent chain: coordinating translation with co-translational folding. Current opinion in structural biology. 2011;21(1):25-31. Epub 2010/11/30.

45. Raina M, Ibba M. tRNAs as regulators of biological processes. Front Genet. 2014;5:171. Epub 2014/06/27.

46. Sheppard K, Yuan J, Hohn MJ, Jester B, Devine KM, Soll D. From one amino acid to another: tRNA-dependent amino acid biosynthesis. Nucleic Acids Res. 2008;36(6):1813-25. Epub 2008/02/07.

47. Jahn D, Verkamp E, Soll D. Glutamyl-transfer RNA: a precursor of heme and chlorophyll biosynthesis. Trends Biochem Sci. 1992;17(6):215-8. Epub 1992/06/01.

48. Shepherd J, Ibba M. Direction of aminoacylated transfer RNAs into antibiotic synthesis and peptidoglycan-mediated antibiotic resistance. FEBS Lett. 2013;587(18):2895-904. Epub 2013/08/03.

49. Miller SB, Yildiz FZ, Lo JA, Wang B, D'Souza VM. A structure-based mechanism for tRNA and retroviral RNA remodelling during primer annealing. Nature. 2014;515(7528):591-5. Epub 2014/09/12.

50. Ruggero K, Guffanti A, Corradin A, Sharma VK, De Bellis G, Corti G, et al. Small noncoding RNAs in cells transformed by human T-cell leukemia virus type 1: a role for a tRNA fragment as a primer for reverse transcriptase. J Virol. 2014;88(7):3612-22. Epub 2014/01/10.

51. Mogk A, Schmidt R, Bukau B. The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. Trends Cell Biol. 2007;17(4):165-72. Epub 2007/02/20.

52. Bachmair A, Finley D, Varshavsky A. In vivo half-life of a protein is a function of its amino-terminal residue. Science. 1986;234(4773):179-86. Epub 1986/10/10.

53. Karakozova M, Kozak M, Wong CC, Bailey AO, Yates JR, 3rd, Mogilner A, et al. Arginylation of beta-actin regulates actin cytoskeleton and cell motility. Science. 2006;313(5784):192-6. Epub 2006/06/24.

54. Raab JR, Chiu J, Zhu J, Katzman S, Kurukuti S, Wade PA, et al. Human tRNA genes function as chromatin insulators. EMBO J. 2012;31(2):330-50. Epub 2011/11/17.

55. Mei Y, Yong J, Liu H, Shi Y, Meinkoth J, Dreyfuss G, et al. tRNA binds to cytochrome c and inhibits caspase activation. Molecular cell. 2010;37(5):668-78. Epub 2010/03/17.

56. Hou YM, Yang X. Regulation of cell death by transfer RNA. Antioxid Redox Signal. 2013;19(6):583-94. Epub 2013/01/29.

57. Marshall L, White RJ. Non-coding RNA production by RNA polymerase III is implicated in cancer. Nature reviews Cancer. 2008;8(12):911-4. Epub 2008/11/07.

58. Winter AG, Sourvinos G, Allison SJ, Tosh K, Scott PH, Spandidos DA, et al. RNA polymerase III transcription factor TFIIIC2 is overexpressed in ovarian tumors. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(23):12619-24. Epub 2000/11/01.

59. Daly NL, Arvanitis DA, Fairley JA, Gomez-Roman N, Morton JP, Graham SV, et al. Deregulation of RNA polymerase III transcription in cervical epithelium in response to high-risk human papillomavirus. Oncogene. 2005;24(5):880-8. Epub 2004/12/14.

60. Zhou Y, Goodenbour JM, Godley LA, Wickrema A, Pan T. High levels of tRNA abundance and alteration of tRNA charging by bortezomib in multiple myeloma. Biochemical and biophysical research communications. 2009;385(2):160-4. Epub 2009/05/20.

61. Anderson P, Ivanov P. tRNA fragments in human health and disease. FEBS letters. 2014;588(23):4297-304. Epub 2014/09/16.

62. Pliatsika V, Loher P, Telonis AG, Rigoutsos I. MINTbase: a framework for the interactive exploration of mitochondrial and nuclear tRNA fragments. Bioinformatics. 2016;32(16):2481-9. Epub 2016/05/07.

63. Ivanov P, Emara MM, Villen J, Gygi SP, Anderson P. Angiogenin-induced tRNA fragments inhibit translation initiation. Molecular cell. 2011;43(4):613-23. Epub 2011/08/23.

64. Thompson DM, Parker R. Stressing out over tRNA cleavage. Cell. 2009;138(2):215-9. Epub 2009/07/28.

65. Thompson DM, Lu C, Green PJ, Parker R. tRNA cleavage is a conserved response to oxidative stress in eukaryotes. RNA. 2008;14(10):2095-103. Epub 2008/08/23.

66. Czech A, Wende S, Morl M, Pan T, Ignatova Z. Reversible and rapid transfer-RNA deactivation as a mechanism of translational repression in stress. PLoS genetics. 2013;9(8):e1003767. Epub 2013/09/07.

67. Schaefer M, Pollex T, Hanna K, Tuorto F, Meusburger M, Helm M, et al. RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. Genes Dev. 2010;24(15):1590-5. Epub 2010/08/04.

68. Blanco S, Dietmann S, Flores JV, Hussain S, Kutter C, Humphreys P, et al. Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders. The EMBO journal. 2014;33(18):2020-39. Epub 2014/07/27.

69. Emara MM, Ivanov P, Hickman T, Dawra N, Tisdale S, Kedersha N, et al. Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. The Journal of biological chemistry. 2010;285(14):10959-68. Epub 2010/02/05.

70. Saikia M, Krokowski D, Guan BJ, Ivanov P, Parisien M, Hu GF, et al. Genome-wide identification and quantitative analysis of cleaved tRNA fragments induced by cellular stress. The Journal of biological chemistry. 2012;287(51):42708-25. Epub 2012/10/23.

71. Saikia M, Jobava R, Parisien M, Putnam A, Krokowski D, Gao XH, et al. Angiogenin-cleaved tRNA halves interact with cytochrome c, protecting cells from apoptosis during osmotic stress. Molecular and cellular biology. 2014;34(13):2450-63. Epub 2014/04/23.

72. Zou H, Li Y, Liu X, Wang X. An APAF-1.cytochrome c multimeric complex is a functional apoptosome that activates procaspase-9. The Journal of biological chemistry. 1999;274(17):11549-56. Epub 1999/04/17.

73. Cole C, Sobala A, Lu C, Thatcher SR, Bowman A, Brown JW, et al. Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. RNA. 2009;15(12):2147-60. Epub 2009/10/24.

74. Li Z, Ender C, Meister G, Moore PS, Chang Y, John B. Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. Nucleic acids research. 2012;40(14):6787-99. Epub 2012/04/12.

75. Sobala A, Hutvagner G. Small RNAs derived from the 5' end of tRNA can inhibit protein translation in human cells. RNA Biol. 2013;10(4):553-63. Epub 2013/04/09.

76. Shigematsu M, Kirino Y. tRNA-Derived Short Non-coding RNA as Interacting Partners of Argonaute Proteins. Gene regulation and systems biology. 2015;9:27-33. Epub 2015/09/25.

77. Haussecker D, Huang Y, Lau A, Parameswaran P, Fire AZ, Kay MA. Human tRNA-derived small RNAs in the global regulation of RNA silencing. RNA. 2010;16(4):673-95. Epub 2010/02/26.

78. Burroughs AM, Ando Y, de Hoon MJ, Tomaru Y, Suzuki H, Hayashizaki Y, et al. Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. RNA Biol. 2011;8(1):158-77. Epub 2011/02/02.

79. Maute RL, Schneider C, Sumazin P, Holmes A, Califano A, Basso K, et al. tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. Proc Natl Acad Sci U S A. 2013;110(4):1404-9. Epub 2013/01/09.

80. Dhahbi JM. 5' tRNA Halves: The Next Generation of Immune Signaling Molecules. Frontiers in immunology. 2015;6:74. Epub 2015/03/10.

81. Mishima E, Inoue C, Saigusa D, Inoue R, Ito K, Suzuki Y, et al. Conformational change in transfer RNA is an early indicator of acute cellular damage. Journal of the American Society of Nephrology : JASN. 2014;25(10):2316-26. Epub 2014/05/17.

82. Pavon-Eternod M, Gomes S, Geslain R, Dai Q, Rosner MR, Pan T. tRNA over-expression in breast cancer and functional consequences. Nucleic acids research. 2009;37(21):7268-80. Epub 2009/09/29.

83. Lee YS, Shibata Y, Malhotra A, Dutta A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). Genes & development. 2009;23(22):2639-49. Epub 2009/11/26.

84. Li S, Hu GF. Emerging role of angiogenin in stress response and cell survival under adverse conditions. Journal of cellular physiology. 2012;227(7):2822-6. Epub 2011/10/25.

85. Gao X, Xu Z. Mechanisms of action of angiogenin. Acta biochimica et biophysica Sinica. 2008;40(7):619-24. Epub 2008/07/08.

86. Goodarzi H, Liu X, Nguyen HC, Zhang S, Fish L, Tavazoie SF. Endogenous tRNA-Derived Fragments Suppress Breast Cancer Progression via YBX1 Displacement. Cell. 2015;161(4):790-802. Epub 2015/05/11.

87. Wang Q, Lee I, Ren J, Ajay SS, Lee YS, Bao X. Identification and functional characterization of tRNA-derived RNA fragments (tRFs) in respiratory syncytial virus infection. Molecular therapy : the journal of the American Society of Gene Therapy. 2013;21(2):368-79. Epub 2012/11/28.

88. Greenway MJ, Andersen PM, Russ C, Ennis S, Cashman S, Donaghy C, et al. ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. Nature genetics. 2006;38(4):411-3. Epub 2006/02/28.

89. van Es MA, Schelhaas HJ, van Vught PW, Ticozzi N, Andersen PM, Groen EJ, et al. Angiogenin variants in Parkinson disease and amyotrophic lateral sclerosis. Annals of neurology. 2011;70(6):964-73. Epub 2011/12/23.

90. Kieran D, Sebastia J, Greenway MJ, King MA, Connaughton D, Concannon CG, et al. Control of motoneuron survival by angiogenin. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2008;28(52):14056-61. Epub 2008/12/26.

91. Sebastia J, Kieran D, Breen B, King MA, Netteland DF, Joyce D, et al. Angiogenin protects motoneurons against hypoxic injury. Cell death and differentiation. 2009;16(9):1238-47. Epub 2009/05/16.

92. Wu D, Yu W, Kishikawa H, Folkerth RD, Iafrate AJ, Shen Y, et al. Angiogenin loss-of-function mutations in amyotrophic lateral sclerosis. Annals of neurology. 2007;62(6):609-17. Epub 2007/09/22.

93. Ivanov P, O'Day E, Emara MM, Wagner G, Lieberman J, Anderson P. G-quadruplex structures contribute to the neuroprotective effects of angiogenin-induced tRNA fragments. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(51):18201-6. Epub 2014/11/19.

94. Robberecht W, Philips T. The changing scene of amyotrophic lateral sclerosis. Nat Rev Neurosci. 2013;14(4):248-64. Epub 2013/03/07.

95. Bendotti C, Marino M, Cheroni C, Fontana E, Crippa V, Poletti A, et al. Dysfunction of constitutive and inducible ubiquitin-proteasome system in amyotrophic lateral sclerosis: implication for protein aggregation and immune response. Prog Neurobiol. 2012;97(2):101-26. Epub 2011/10/29.

96. Abbasi-Moheb L, Mertel S, Gonsior M, Nouri-Vahid L, Kahrizi K, Cirak S, et al. Mutations in NSUN2 cause autosomal-recessive intellectual disability. American journal of human genetics. 2012;90(5):847-55. Epub 2012/05/01.

97. Khan MA, Rafiq MA, Noor A, Hussain S, Flores JV, Rupp V, et al. Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. American journal of human genetics. 2012;90(5):856-63. Epub 2012/05/01.

98. Martinez FJ, Lee JH, Lee JE, Blanco S, Nickerson E, Gabriel S, et al. Whole exome sequencing identifies a splicing mutation in NSUN2 as a cause of a Dubowitz-like syndrome. J Med Genet. 2012;49(6):380-5. Epub 2012/05/12.

99. Jackman JE, Alfonzo JD. Transfer RNA modifications: nature's combinatorial chemistry playground. Wiley interdisciplinary reviews RNA. 2013;4(1):35-48. Epub 2012/11/10.

100. Piñeyro DT, A. G.; Ribas de Pouplana, L. Biogenesis and Evolution of Functional tRNAs. In: Ane Sesma TvdH, editor. Fungal RNA Biology. Switzerland: Springer International Publishing Switzerland; 2014. p. 233-67.

101. Torres AG, Batlle E, Ribas de Pouplana L. Role of tRNA modifications in human diseases. Trends Mol Med. 2014;20(6):306-14. Epub 2014/03/04.

102. Chan CT, Dyavaiah M, DeMott MS, Taghizadeh K, Dedon PC, Begley TJ. A quantitative systems approach reveals dynamic control of tRNA modifications during cellular stress. PLoS Genet. 2010;6(12):e1001247. Epub 2010/12/29.

103. Patil A, Chan CT, Dyavaiah M, Rooney JP, Dedon PC, Begley TJ. Translational infidelity-induced protein stress results from a deficiency in Trm9-catalyzed tRNA modifications. RNA biology. 2012;9(7):990-1001. Epub 2012/07/27.

104. Begley U, Dyavaiah M, Patil A, Rooney JP, DiRenzo D, Young CM, et al. Trm9-catalyzed tRNA modifications link translation to the DNA damage response. Mol Cell. 2007;28(5):860-70. Epub 2007/12/18.

105. Kalhor HR, Clarke S. Novel methyltransferase for modified uridine residues at the wobble position of tRNA. Mol Cell Biol. 2003;23(24):9283-92. Epub 2003/12/04.

106. Chan CT, Pang YL, Deng W, Babu IR, Dyavaiah M, Begley TJ, et al. Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins. Nature communications. 2012;3:937. Epub 2012/07/05.

107. Bilbille Y, Vendeix FA, Guenther R, Malkiewicz A, Ariza X, Vilarrasa J, et al. The structure of the human tRNALys3 anticodon bound to the HIV genome is stabilized by modified nucleosides and adjacent mismatch base pairs. Nucleic Acids Res. 2009;37(10):3342-53. Epub 2009/03/28.

108. Cabello-Villegas J, Winkler ME, Nikonowicz EP. Solution conformations of unmodified and A(37)N(6)-dimethylallyl modified anticodon stem-loops of Escherichia coli tRNA(Phe). J Mol Biol. 2002;319(5):1015-34. Epub 2002/06/25.

109. Vendeix FA, Murphy FVt, Cantara WA, Leszczynska G, Gustilo EM, Sproat B, et al. Human tRNA(Lys3)(UUU) is pre-structured by natural modifications for cognate and wobble codon binding through keto-enol tautomerism. J Mol Biol. 2012;416(4):467-85. Epub 2012/01/10.

110. Noma A, Suzuki T. Ribonucleome analysis identified enzyme genes responsible for wybutosine synthesis. Nucleic Acids Symp Ser (Oxf). 2006(50):65-6. Epub 2006/12/08.

111. de Crecy-Lagard V, Brochier-Armanet C, Urbonavicius J, Fernandez B, Phillips G, Lyons B, et al. Biosynthesis of wyosine derivatives in tRNA: an ancient and highly diverse pathway in Archaea. Mol Biol Evol. 2010;27(9):2062-77. Epub 2010/04/13.

112. Bjork GR, Jacobsson K, Nilsson K, Johansson MJ, Bystrom AS, Persson OP. A primordial tRNA modification required for the evolution of life? The EMBO journal. 2001;20(1-2):231-9. Epub 2001/02/28.

113. Agris PF, Vendeix FA, Graham WD. tRNA's wobble decoding of the genome: 40 years of modification. J Mol Biol. 2007;366(1):1-13. Epub 2006/12/26.

114. Rubio MA, Pastar I, Gaston KW, Ragone FL, Janzen CJ, Cross GA, et al. An adenosine-to-inosine tRNA-editing enzyme that can perform C-to-U deamination of DNA. Proc Natl Acad Sci U S A. 2007;104(19):7821-6. Epub 2007/05/08.

115. Spears JL, Rubio MA, Gaston KW, Wywial E, Strikoudis A, Bujnicki JM, et al. A single zinc ion is sufficient for an active Trypanosoma brucei tRNA editing deaminase. J Biol Chem. 2011;286(23):20366-74. Epub 2011/04/22.

116. Grosjean H, de Crecy-Lagard V, Marck C. Deciphering synonymous codons in the three domains of life: co-evolution with specific tRNA modification enzymes. FEBS Lett. 2010;584(2):252-64. Epub 2009/11/26.

117. Chan PP, Lowe TM. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. Nucleic acids research. 2016;44(D1):D184-9. Epub 2015/12/18.

118. Elias Y, Huang RH. Biochemical and structural studies of A-to-I editing by tRNA:A34 deaminases at the wobble position of transfer RNA. Biochemistry. 2005;44(36):12057-65. Epub 2005/09/07.

119. Auxilien S, Crain PF, Trewyn RW, Grosjean H. Mechanism, specificity and general properties of the yeast enzyme catalysing the formation of inosine 34 in the anticodon of transfer RNA. J Mol Biol. 1996;262(4):437-58. Epub 1996/10/04.

120. French BT, Trewyn RW. Modification of the anticodon wobble position of tRNA(Ala) in vitro does not require 5' or 3' processing. Gene. 1990;96(2):301-4. Epub 1990/12/15.

121. Hopper AK, Phizicky EM. tRNA transfers to the limelight. Genes & development. 2003;17(2):162-80. Epub 2003/01/21.

122. Ragone FL, Spears JL, Wohlgamuth-Benedum JM, Kreel N, Papavasiliou FN, Alfonzo JD. The C-terminal end of the Trypanosoma brucei editing deaminase plays a critical role in tRNA binding. RNA. 2011;17(7):1296-306. Epub 2011/05/24.

123. Saint-Leger A, Bello C, Dans PD, Torres AG, Novoa EM, Camacho N, et al. Saturation of recognition elements blocks evolution of new tRNA identities. Science advances. 2016;2(4):e1501860. Epub 2016/07/08.

124. Rafels-Ybern A, Torres AG, Grau-Bove X, Ruiz-Trillo I, Ribas de Pouplana L. Codon adaptation to tRNAs with Inosine modification at position 34 is widespread among Eukaryotes and present in two Bacterial phyla. RNA Biol. 2017:1-8. Epub 2017/09/08.

125. Pavlov MY, Watts RE, Tan Z, Cornish VW, Ehrenberg M, Forster AC. Slow peptide bond formation by proline and other N-alkylamino acids in translation. Proc Natl Acad Sci U S A. 2009;106(1):50-4. Epub 2008/12/24.

126. Guimaraes JC, Rocha M, Arkin AP. Transcript level and sequence determinants of protein abundance and noise in Escherichia coli. Nucleic acids research. 2014;42(8):4791-9. Epub 2014/02/11.

127. Ude S, Lassak J, Starosta AL, Kraxenberger T, Wilson DN, Jung K. Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. Science. 2013;339(6115):82-5. Epub 2012/12/15.

128. Doerfel LK, Wohlgemuth I, Kothe C, Peske F, Urlaub H, Rodnina MV. EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. Science. 2013;339(6115):85-8. Epub 2012/12/15.

129. Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature. 2011;478(7367):57-63. Epub 2011/09/23.

130. Salehi Chaleshtori AR, Miyake N, Ahmadvand M, Bashti O, Matsumoto N, Noruzinia M. A novel 8-bp duplication in ADAT3 causes mild intellectual disability. Hum Genome Var. 2018;5:7. Epub 2018/05/26.

131. Grosjean H, Constantinesco F, Foiret D, Benachenhou N. A novel enzymatic pathway leading to 1-methylinosine modification in Haloferax volcanii tRNA. Nucleic Acids Res. 1995;23(21):4312-9. Epub 1995/11/11.

132. Gerber A, Grosjean H, Melcher T, Keller W. Tad1p, a yeast tRNA-specific adenosine deaminase, is related to the mammalian pre-mRNA editing enzymes ADAR1 and ADAR2. The EMBO journal. 1998;17(16):4780-9. Epub 1998/08/26.

133. Zhou W, Karcher D, Bock R. Importance of adenosine-to-inosine editing adjacent to the anticodon in an Arabidopsis alanine tRNA under environmental stress. Nucleic acids research. 2013;41(5):3362-72. Epub 2013/01/29.

134. Macbeth MR, Schubert HL, Vandemark AP, Lingam AT, Hill CP, Bass BL. Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. Science. 2005;309(5740):1534-9. Epub 2005/09/06.

135. Becker HF, Corda Y, Mathews MB, Fourrey JL, Grosjean H. Inosine and N1-methylinosine within a synthetic oligomer mimicking the anticodon loop of human tRNA(Ala) are major epitopes for anti-PL-12 myositis autoantibodies. RNA. 1999;5(7):865-75. Epub 1999/07/20.

136. Bass BL, Nishikura K, Keller W, Seeburg PH, Emeson RB, O'Connell MA, et al. A standardized nomenclature for adenosine deaminases that act on RNA. RNA. 1997;3(9):947-9. Epub 1997/09/18.

137. Keegan LP, Leroy A, Sproul D, O'Connell MA. Adenosine deaminases acting on RNA (ADARs): RNA-editing enzymes. Genome biology. 2004;5(2):209. Epub 2004/02/05.

138. Basilio C, Wahba AJ, Lengyel P, Speyer JF, Ochoa S. Synthetic polynucleotides and the amino acid code. V. Proceedings of the National Academy of Sciences of the United States of America. 1962;48:613-6. Epub 1962/04/15.

139. Zinshteyn B, Nishikura K. Adenosine-to-inosine RNA editing. Wiley interdisciplinary reviews Systems biology and medicine. 2009;1(2):202-9. Epub 2010/09/14.

140. Keegan LP, Gallo A, O'Connell MA. The many roles of an RNA editor. Nat Rev Genet. 2001;2(11):869-78. Epub 2001/11/21.

141. Gerber AP, Keller W. RNA editing by base deamination: more enzymes, more targets, new mysteries. Trends Biochem Sci. 2001;26(6):376-84. Epub 2001/06/19.

142. Maas S, Rich A. Changing genetic information through RNA editing. BioEssays : news and reviews in molecular, cellular and developmental biology. 2000;22(9):790-802. Epub 2000/08/17.

143. Torres AG, Ribas de Pouplana L. Transfer RNA Modifications: From Biological Functions to Biomedical Applications. . Modified Nucleic Acids in Biology and Medicine (Jurga, S, Erdmann, V A, Barciszewski, J, Eds). Cham, Switzerland: Springer International Publishing; 2016. p. 1-26.

144. Grosjean H, Droogmans L, Roovers M, Keith G. Detection of enzymatic activity of transfer RNA modification enzymes using radiolabeled tRNA substrates. Methods in enzymology. 2007;425:55-101. Epub 2007/08/04.

145. Motorin Y, Muller S, Behm-Ansmant I, Branlant C. Identification of modified residues in RNAs by reverse transcription-based methods. Methods in enzymology. 2007;425:21-53. Epub 2007/08/04.

146. Lamichhane TN, Arimbasseri AG, Rijal K, Iben JR, Wei FY, Tomizawa K, et al. Lack of tRNA-i6A modification causes mitochondrial-like metabolic deficiency in S. pombe by limiting activity of cytosolic tRNATyr, not mito-tRNA. RNA. 2016;22(4):583-96. Epub 2016/02/10.

147. Igloi GL, Kossel H. Affinity electrophoresis for monitoring terminal phosphorylation and the presence of queuosine in RNA. Application of polyacrylamide containing a covalently bound boronic acid. Nucleic acids research. 1985;13(19):6881-98. Epub 1985/10/11.

148. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American journal of human genetics. 1980;32(3):314-31. Epub 1980/05/01.

149. Grover A, Sharma PC. Development and use of molecular markers: past and present. Critical reviews in biotechnology. 2016;36(2):290-302. Epub 2014/11/29.

150. Gu T, Buaas FW, Simons AK, Ackert-Bicknell CL, Braun RE, Hibbs MA. Canonical A-to-I and C-to-U RNA editing is enriched at 3'UTRs and microRNA target sites in multiple mouse tissues. PloS one. 2012;7(3):e33720. Epub 2012/03/27.

151. Kawahara Y, Ito K, Sun H, Kanazawa I, Kwak S. Low editing efficiency of GluR2 mRNA is associated with a low relative abundance of ADAR2 mRNA in white matter of normal human brain. The European journal of neuroscience. 2003;18(1):23-33. Epub 2003/07/16.

152. Palladino MJ, Keegan LP, O'Connell MA, Reenan RA. A-to-I pre-mRNA editing in Drosophila is primarily involved in adult nervous system function and integrity. Cell. 2000;102(4):437-49. Epub 2000/08/31.

153. Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC. Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. Nucleic acids research. 1987;15(21):8783-98. Epub 1987/11/11.

154. Shin D, Sinkeldam RW, Tor Y. Emissive RNA alphabet. Journal of the American Chemical Society. 2011;133(38):14912-5. Epub 2011/08/27.

155. ULC CCG. Molecular Operating Environment (MOE) 2013.

156. Schmidtke P, Bidon-Chanal A, Luque FJ, Barril X. MDpocket: open-source cavity detection and characterization on molecular dynamics trajectories. Bioinformatics. 2011;27(23):3276-85. Epub 2011/10/05.

157. Ruiz-Carmona S, Alvarez-Garcia D, Foloppe N, Garmendia-Doval AB, Juhos S, Schmidtke P, et al. rDock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. PLoS computational biology. 2014;10(4):e1003571. Epub 2014/04/12.

158. Schrödinger L. Schrödinger Release 2018-1: LigPrep. 2018.

159. Rubio MA, Ragone FL, Gaston KW, Ibba M, Alfonzo 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(1):115-20. Epub 2005/11/05.

160. Mizrahi RA, Shin D, Sinkeldam RW, Phelps KJ, Fin A, Tantillo DJ, et al. A Fluorescent Adenosine Analogue as a Substrate for an A-to-I RNA Editing Enzyme. Angew Chem Int Ed Engl. 2015;54(30):8713-6. Epub 2015/06/23.

161. Su D, Chan CT, Gu C, Lim KS, Chionh YH, McBee ME, et al. Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry. Nat Protoc. 2014;9(4):828-41. Epub 2014/03/15.

162. Pliatsika V, Loher P, Magee R, Telonis AG, Londin E, Shigematsu M, et al. MINTbase v2.0: a comprehensive database for tRNA-derived fragments that includes nuclear and mitochondrial fragments from all The Cancer Genome Atlas projects. Nucleic Acids Res. 2018;46(D1):D152-D9. Epub 2017/12/01.

163. Kim J, Malashkevich V, Roday S, Lisbin M, Schramm VL, Almo SC. Structural and kinetic characterization of Escherichia coli TadA, the wobble-specific tRNA deaminase. Biochemistry. 2006;45(20):6407-16. Epub 2006/05/17.

164. Rose MC, Voynow JA. Respiratory tract mucin genes and mucin glycoproteins in health and disease. Physiological reviews. 2006;86(1):245-78. Epub 2005/12/24.

165. Kreda SM, Davis CW, Rose MC. CFTR, mucins, and mucus obstruction in cystic fibrosis. Cold Spring Harbor perspectives in medicine. 2012;2(9):a009589. Epub 2012/09/07.

166. Kufe DW. Mucins in cancer: function, prognosis and therapy. Nature reviews Cancer. 2009;9(12):874-85. Epub 2009/11/26.

167. Nakajima Y, Kanno T, Nagaya T, Kuribayashi K, Nakano T, Gotoh A, et al. Adenosine deaminase inhibitor EHNA exhibits a potent anticancer effect against malignant pleural mesothelioma. Cell Physiol Biochem. 2015;35(1):51-60. Epub 2014/12/31.

168. Cristalli G, Franchetti P, Grifantini M, Vittori S, Lupidi G, Riva F, et al. Adenosine deaminase inhibitors. Synthesis and biological activity of deaza analogues of erythro-9-(2-hydroxy-3-nonyl)adenine. J Med Chem. 1988;31(2):390-3. Epub 1988/02/01.

169. Johnston JB. Mechanism of action of pentostatin and cladribine in hairy cell leukemia. Leuk Lymphoma. 2011;52 Suppl 2:43-5. Epub 2011/04/06.

170. Kane BJ, Kuhn JG, Roush MK. Pentostatin: an adenosine deaminase inhibitor for the treatment of hairy cell leukemia. Ann Pharmacother. 1992;26(7-8):939-47. Epub 1992/07/01.

171. Lupidi G, Riva F, Cristalli G, Grifantini M. Inhibition of adenosine deaminase by deaza derivatives of adenosine and purine riboside. Ital J Biochem. 1982;31(6):396-403. Epub 1982/11/01.

172. Jurgensen CH, Huber BE, Zimmerman TP, Wolberg G. 3-deazaadenosine inhibits leukocyte adhesion and ICAM-1 biosynthesis in tumor necrosis factor-stimulated human endothelial cells. J Immunol. 1990;144(2):653-61. Epub 1990/01/15.

173. Lemaire M, Momparler LF, Raynal NJ, Bernstein ML, Momparler RL. Inhibition of cytidine deaminase by zebularine enhances the antineoplastic action of 5-aza-2'-deoxycytidine. Cancer Chemother Pharmacol. 2009;63(3):411-6. Epub 2008/04/10.

174. Marquez VE, Kelley JA, Agbaria R, Ben-Kasus T, Cheng JC, Yoo CB, et al. Zebularine: a unique molecule for an epigenetically based strategy in cancer chemotherapy. Ann N Y Acad Sci. 2005;1058:246-54. Epub 2006/01/06.

175. Takemura Y, Satoh M, Hatanaka K, Kubota S. Zebularine exerts its antiproliferative activity through S phase delay and cell death in human malignant mesothelioma cells. Biosci Biotechnol Biochem. 2018:1-6. Epub 2018/04/25.

176. Losey HC, Ruthenburg AJ, Verdine GL. Crystal structure of Staphylococcus aureus tRNA adenosine deaminase TadA in complex with RNA. Nature structural & molecular biology. 2006;13(2):153-9. Epub 2006/01/18.

177. Wulff TF, Arguello RJ, Molina Jordan M, Roura Frigole H, Hauquier G, Filonava L, et al. Detection of a Subset of Posttranscriptional Transfer RNA Modifications in vivo With a Restriction Fragment Length Polymorphism-Based Method. Biochemistry. 2017. Epub 2017/07/14.

178. Xiang S, Short SA, Wolfenden R, Carter CW, Jr. Transition-state selectivity for a single hydroxyl group during catalysis by cytidine deaminase. Biochemistry. 1995;34(14):4516-23. Epub 1995/04/11.

179. Arimbasseri AG, Blewett NH, Iben JR, Lamichhane TN, Cherkasova V, Hafner M, et al. RNA Polymerase III Output Is Functionally Linked to tRNA Dimethyl-G26 Modification. PLoS genetics. 2015;11(12):e1005671. Epub 2016/01/01.

180. Liu J, Straby KB. The human tRNA(m(2)(2)G(26))dimethyltransferase: functional expression and characterization of a cloned hTRM1 gene. Nucleic Acids Res. 2000;28(18):3445-51. Epub 2000/09/13.

181. Davarniya B, Hu H, Kahrizi K, Musante L, Fattahi Z, Hosseini M, et al. The Role of a Novel TRMT1 Gene Mutation and Rare GRM1 Gene Defect in Intellectual Disability in Two Azeri Families. PloS one. 2015;10(8):e0129631. Epub 2015/08/27.

182. Dewe JM, Fuller BL, Lentini JM, Kellner SM, Fu D. TRMT1-Catalyzed tRNA Modifications Are Required for Redox Homeostasis To Ensure Proper Cellular Proliferation and Oxidative Stress Survival. Mol Cell Biol. 2017;37(21). Epub 2017/08/09.

183. Serebrov V, Clarke RJ, Gross HJ, Kisselev L. Mg2+-induced tRNA folding. Biochemistry. 2001;40(22):6688-98. Epub 2001/06/09.

184. Grosjean H, Edqvist J, Straby KB, Giege R. Enzymatic formation of modified nucleosides in tRNA: dependence on tRNA architecture. Journal of molecular biology. 1996;255(1):67-85. Epub 1996/01/12.

185. Dedon PC, Begley TJ. A system of RNA modifications and biased codon use controls cellular stress response at the level of translation. Chemical research in toxicology. 2014;27(3):330-7. Epub 2014/01/16.

186. Torres AG, Ribas de Pouplana L. Transfer RNA Modifications: From Biological Functions to Biomedical Applications. In: Jurga S, Erdmann VA, Barciszewski J, editors. Modified Nucleic Acids in Biology and Medicine. Switzerland: Springer International Publishing; 2016. p. 1-26.

187. Freude K, Hoffmann K, Jensen LR, Delatycki MB, des Portes V, Moser B, et al. Mutations in the FTSJ1 gene coding for a novel S-adenosylmethionine-binding protein cause nonsyndromic X-linked mental retardation. American journal of human genetics. 2004;75(2):305-9. Epub 2004/05/27.

188. Delker RK, Zhou Y, Strikoudis A, Stebbins CE, Papavasiliou FN. Solubility-based genetic screen identifies RING finger protein 126 as an E3 ligase for activation-induced cytidine deaminase. Proc Natl Acad Sci U S A. 2013;110(3):1029-34. Epub 2013/01/02.

189. Zauhar RJ, Colbert CL, Morgan RS, Welsh WJ. Evidence for a strong sulfur-aromatic interaction derived from crystallographic data. Biopolymers. 2000;53(3):233-48. Epub 2000/02/19.

190. Rubio MA, Gaston KW, McKenney KM, Fleming IM, Paris Z, Limbach PA, et al. Editing and methylation at a single site by functionally interdependent activities. Nature. 2017;542(7642):494-7. Epub 2017/02/24.

A Helena.

Vull viure intensament, l'avui em crida;
l'instant és la mesura del present.
Gaudir cada segon tota la vida
sens res que em privi d'apreciar el moment.

Vull fer que la raó posi la mida
als somnis que em voltegen per la ment,
i un bàlsam nou a la vella ferida,
i un sol radiant en el meu firmament.

La vida ens ve guarnida de sorpreses,
on joies i dissorts vénen ensems,
tot alternant carícies i escomeses,

donant-nos lliris o bé crisantems.
La mar és plana, o plena d'agreses,
meva és la barca i jo moc els rems.

Jaume Roura (Tercer Maset)