# Characterization of human and bacterial tRNA-specific adenosine deaminases

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

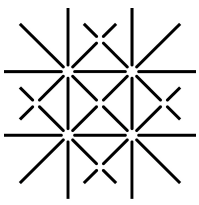
von

**Jeannette Wolf**

aus

Zürich, Bern und Lotzwil

Basel, 2002



Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von Prof. Walter Keller and Prof. Thomas Bickle.

Basel, 4. Juni 2002

Prof. Dr. Andreas D. Zuberbühler
Dekan der Philosophisch-Naturwissenschaftlichen Fakultät

Bruno

und

meinen Eltern

# 1 Summary

Many modified nucleotides are known to be present in tRNAs which have different functions and effects in translation and other cellular processes. Inosine at the wobble position (position 34) in tRNAs has an important role in translation because it increases the base pairing possibilities between mRNA and tRNA.

In bacteria only $tRNA^{Arg2}$ contains $I_{34}$. The deamination at this position is catalyzed by tadA (tRNA-specific adenosine deaminase A) which is the first and so far only bacterial deaminase that acts on RNA. This study presents evidence that bacteria encode only one polypeptide that is sufficient for tRNA:adenosine 34 deaminase activity. tadA is encoded by an essential gene thus underscoring the important role of $I_{34}$ also in bacteria. tadA selectively binds to $tRNA^{Arg2}$, whereas eukaroytic tRNAs are not bound or deaminated by tadA, except when they contain the anticodon loop of $tRNA^{Arg}$. A minisubstrate consisting of the anticodon arm of $tRNA^{Arg2}$ is a substrate for tadA, providing further evidence that important recognition elements are located in this region of the tRNA. Mutational analysis has shown that the sequence UAGC at positions 33 to 36 and a stem-loop structure are sufficient for inosine formation. Thus, the anticodon is the major determinant for tadA acitivity. However, inosine formation is less efficient with the minimal minisubstrate compared to wild-type minisubstrate, suggesting that additional sequences or structures might be required for efficiency.

Recombinant human Tad2 and human Tad3 are sufficient to reconstitute tRNA:adenosine 34 deaminase activity. The complex deaminates eukaryotic tRNAs from humans, *B. mori* and *S. cerevisiae* and also $tRNA^{Arg2}$ from *E. coli*. □-hTad2 antibodies deplete the tRNA:adenosine 34 activity from HeLa cell extracts by depleting both subunits, indicating that they form a stable complex. This is further supported by the result that recombinant hTad2 localizes to the nucleus and transports hTad3 into the nucleus. hTad2 and hTad3 cannot be exchanged *in vitro* with subunits from other organisms, however this is not due to lack of interaction since hTad2 and hTad3 interact with Tad2 and Tad3 proteins from *S. cerevisiae* and *E. coli*. Interestingly, the Arg2 minisubstrate is not deaminated by the human tRNA:adenosine 34 deaminase although this enzyme is known to modify full-length $tRNA^{Arg2}$. This result indicates that tRNA binding is different for prokaryotic and eukaryotic Tad2 and Tad3 proteins.

The analysis of bacterial and human tRNA:adenosine 34 deaminases extended our knowledge on how these enzymes work and how they are evolutionarily related.

# 2 Introduction

| Type | Organism | Transcript (Genome) | Cis-acting elements/ Trans-acting factors | Mechanism | Functional/ physiological consequences |
|---|---|---|---|---|---|
| **Insertion/deletion type editing** | | | | | |
| U insertion/ deletion | kinetoplastids | mRNA(mt) | anchoring sequence/guide RNAs, TUTase, RNA ligase, endo-nuclease, other factors | coupled cleavage/ligation | creation of reading frames |
| C insertion (also U, AA, CU, GU, GC) | *Physarum polycephalum* | mRNAs, rRNAs, tRNAs (mt) | ? | ? | production of functional RNAs |
| G insertion | paramyxovirus | P mRNA (v) | slippery sequence/viral | pseudotemplated transcription | creation of addi-tional downstream reading frames |
| A insertion | Ebola virus | GP mRNA (v) | slippery sequence/ viral polymerase | pseudotemplated transcription | shift from syn-thesis of secreted to structural GP |
| 3' terminal A addition | vertebrates | mRNAs (mt) | flanking tRNA structure/endo-nuclease, TATase | endonuclease, TATase | ? |
| **Conversion type editing** | | | | | |
| C to U | land plants rRNA ? (mt) | mRNAs, tRNAs<br>mRNAs (cp) | flanking sequence (?) | ? | restoration of evolutionary cons. sequences |
| | *Physarum polycephalum* | cox1 mRNAs (mt) | ? | ? | removing frameshifts |
| | marsupials | tRNAs (mt) | ? | ? | correction tRNA identity switch |
| | mammals | mRNAs (n): apoB, NF1 | mooring sequence, efficiency element, AU-rich region/C-deaminase (APOBEC1), auxiliary factors | C-deamination | creation of a stop codon |
| U to C | land plants | mRNAs (mt, cp) | flanking sequence | U-amination | creation of reading frames |
| | mammals | WT1 mRNA,<br>tRNAs (n) | ? | ? | WT1: abolishment of transcriptional repressor activity |
| A to I | mammals | pre-mRNAs (n): GluR, 5-$HT_{2C}$R ADAR2 | dsRNA structure ADAR1, ADAR2 | ds-dependent A-deamination | alteration of receptor determinants |
| | cephalopod (squid) | Kv2 $K^+$ channel mRNA (n) | ? | temperature-dependent | silent amino acid substitutions and changes in channel gating |
| | *Drosophila melanogaster* | $Ca^{2+}$ channel 1 subunit (Dmca1A) 4f-rnp | dADAR (?) | ? | ? |
| | human hepatitis delta virus | antigenomic RNA (v) | dsRNA structure ADAR1, ADAR2? | dsRNA-dependent<br>A-deamination | switch from repli cation to packaging |
| | HIV | TAR element in mRNA (v) | ? | ? | ? |
| | eukaryotes, prokaryotes | anticodon of tRNAs | Tad1p/ADAT1 Tad2p/3p-complex | | enlargement of coding capacities |
| U to A | humans | -galactosidase mRNA (n) | ? | base-replacement (?) | ? |
| C to A, U to G, U to A, A to G | *Acanthamoeba castellani* primitive fungi (chytridiomycetes) | tRNAs (mt) | internal guide | base-replacement | correction of mis-pairing in the acceptor stem |

RNA editing has been defined as co- or post-transcriptional processes that alter the information of primary transcripts other than splicing, capping or 3'-end formation. The term RNA editing describes the mechanisms by which the sequence content of primary transcripts can be changed and can alter the informational content of a mRNA and subsequently also of the protein sequence translated from this mRNA. RNA editing was found in many eukaryotic organisms and variety was also observed for the substrates which include mRNA, tRNA and rRNA. RNA editing can be divided into two types: it either involves the insertion or deletion of nucleotides or their modification.

## 2.1 Insertion-deletion RNA editing

The term RNA editing was first coined by Benne et al. (1986) to describe the insertion of uridines into the cytochrome oxidase subunit II mRNA in kinetoplasts of *Trypanosoma brucei*. In various trypanosomatids 12 of 18 mitochondrial mRNAs undergo U insertions and deletions at precise sites usually within open reading frames (Feagin et al., 1988b; Shaw et al., 1989; Shaw et al., 1988). The extent of editing was shown to vary from the insertion of a few nucleotides to extensive insertion-deletion of uridine residues in which over 50% of the final mRNA product is the result of RNA editing (Feagin et al., 1988a; Shaw et al., 1988). This process creates initiation and termination codons, corrects frameshifts and builds entire reading frames from 'cryptic' genes.

Trypanosomatids have a single mitochondrion (kinetoplast) which contains a network of thousands of DNA minicircles and maxicircles (Shapiro and Englund, 1995; Simpson, 1987). The maxicircles encode ribosomal RNAs and some of the proteins involved in the mitochondrial respiratory chain. Minicircles and also maxicircles encode small RNA molecules (50-70 nt) that were named guide RNAs (gRNAs; Blum et al., 1990). gRNAs contain the sequence information that is needed for both the number of U's and the sites of insertion and deletion (Blum et al., 1990). The 5' region of gRNAs (anchor) is complementary to sequences just downstream of the

**Table I Different types of site-specific editing and their functional consequences (table adapted from Smith et al. (1997) and A. Gerber, thesis, 1999)**

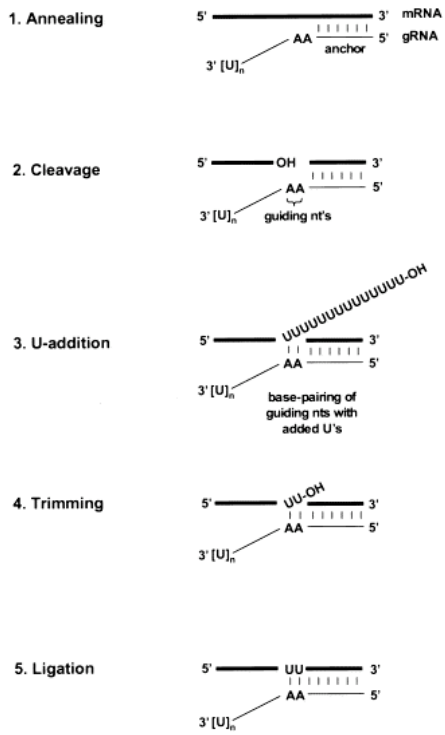
Abbreviations: apoB-apolipoprotein B; cox1-cytochrome c oxidase subunit 1; cp-chloroplast; ds-double stranded; GP-glycoprotein; mt-mitochondria; n-nuclear; TATase-terminal adenylyltransferase; TUTase-terminal uridylyltransferase; v-viral; WT1-Wilms' tumor susceptibility gene 1

editing site in the mRNA and thus provides site-specificity. The central portion of the gRNA contains the editing information of the number of U insertions or deletions and is complementary to the mature edited mRNA. The 3' region of the gRNA has an oligo(U) tail that is added posttranscriptionally. The function of this tail is not known. Editing generally proceeds from 3' to 5' which is due to the creation of new anchor sequences by already edited mRNA sequences (Abraham et al., 1988; Maslov and Simpson, 1992).

Uridine insertion-deletion editing is catalyzed by an enzyme cascade (Fig. 2.1; Blum et al., 1990; Kable et al., 1996; Seiwert et al., 1996; reviewed by Estevez and Simpson, 1999). The gRNA first hybridizes downstream of the first editing site via the anchor and directs an endonuclease to cleave at the first unpaired nucleotide. U's are added by a terminal uridylyl transferase (TUTase) or removed by an exonuclease. The edited mRNA can then basepair with the 5' region of the gRNA and thereby extend the duplex. The two mRNA fragments are joined by a RNA ligase.

**Figure 2.1 Model for U-insertion RNA editing (from Alfonzo et al., 1997).**

The mRNA is represented by a thick line and the cognate gRNA by a thin line. Shown are the different catalytic steps of U-insertion editing. See chapter 2.1 for details about the proteins involved in the catalysis of insertion editing.




The development of *in vitro* editing systems lead to the identification of several proteins involved in insertion-deletion editing. A number of factors were found to bind to gRNAs, among them gBP21 (Koller et al., 1997), TBRGG1 (Vanhamme et al., 1998), RBP16 (Hayman and Read, 1999) and glutamate dehydrogenase (GDH; Bringaud et al., 1997). gBP21 has a high affinity for gRNAs and antibodies against gBP21 depleted *in vitro* editing activity (Lambert et al., 1999). However, a *T. brucei* strain with a knock-out of both gBP21 alleles was still able to

edit and had a normal gRNA pool, but levels of mitochondrial transcripts were reduced (Lambert et al., 1999). Recently, gBP21 was found to accelerate the rate of gRNA-pre-mRNA anchor formation by stimulating base-pairing between gRNA and mRNA (Muller et al., 2001). TBRGG1 belongs to the RGG protein family and was found to bind strongly to synthetic oligo(U) (Vanhamme et al., 1998). The protein co-migrated with *in vitro* editing activity in glycerol gradients, but no direct association with gRNAs or the editing machinery was shown. Because mitochondrial rRNAs are polyuridylated in *Trypanosomes*, TBRGG1 could be involved in rRNA metabolism. RBP16 binds different gRNAs via the oligo(U) tail (Hayman and Read, 1999). Antibodies against RBP16 immunoprecipitated gRNAs and rRNAs, probably due to the oligo(U) tail. As in the case of TBRGG1 this suggested a role in mitochondrial rRNA metabolism. GDH was shown to bind to the oligo(U) tail of gRNAs (Bringaud et al., 1997). This binding probably involves the NADP(H) binding site of the enzyme because binding to gRNAs could be competed with NADP(H). A *T. brucei* GDH knock out strain is viable and showed no differences in abundance of edited mRNAs (Estevez et al., 1999b). GDH also binds UTP and this suggested a regulatory link between mitochondrial metabolism and RNA editing (Bringaud et al., 1997).

Insertion-deletion editing seems to be catalyzed by a large RNP (protein-gRNA complex; reviewed by Madison-Antenucci et al., 2002). In *T. brucei* and *L. tarentolae* three different complexes have been described which sedimented at 10S, 20S and 35-40S (Corell et al., 1996; Peris et al., 1997; Pollard et al., 1992; Rusche et al., 1997). The isolation of different editing complexes could be explained by a dynamic assembly-disassembly of the components that are required for editing (reviewed by Madison-Antenucci et al., 2002). The complex purified by Rusche et al. (1997) was composed of eight major polypeptides, three of which appeared to be RNA ligases. The complex contained the four 'core' activities gRNA-directed endonuclease, TUTase, 3' U-specific exonuclease, RNA ligase and was able to catalyze a complete editing reaction *in vitro*. A monoclonal antibody against the 35-40S complex was used to clone a gene coding for REAP-1 (RNA-editing-associated protein 1; Madison-Antenucci et al., 1998). *In vitro* U-insertion editing was inhibited by this antibody. Madison-Antenucci and Hajduk (2001) showed that REAP-1 is a RNA-binding protein which binds to poly(G) and single-stranded G-rich RNAs. Pre-edited RNAs were shown to be the preferred substrate and this suggested a role for REAP-1 in bringing these RNAs into the editing complex.

In *T. brucei* mitochondrial extracts three different riboendonuclease activities have been described (Piller et al., 1997; Salavati et al., 2002), one of these having the predicted features of a RNA editing riboendonuclease: cleavage of the mRNA immediately 5' of a mRNA-gRNA

duplex and site-specificity for the gRNA-directed editing site. A riboendonuclease termed MAR1 was isolated from *L. tarentolae* (Alfonzo et al., 1998). MAR1 cleaved a pre-edited mRNA in the absence of gRNA and did not distinguish between pre-edited and fully edited substrates, indicating that this protein is not involved in RNA editing. However, additional factors that confer gRNA-dependance could be absent in the purified MAR1 fraction. Cruz-Reyes et al. (1998) showed the existence of differences in adenosine nucleotide requirements for cleavage at U-deletion or U-insertion sites in *T. brucei*. Cleavage at U-deletion sites required high concentrations of ATP and ADP, whereas cleavage at U-insertion sites occured in the absence of these nucleotides and was even inhibited by high concentrations. However, the physiological relevance of this phenomenon remains to be explained. Cleavage at U-insertion and U-deletion sites could also be catalyzed by two different endonuclease with different specificity. After cleavage U residues are added to the 3'-OH of the 5' cleavage fragment by a TUTase or U's are removed by a 3' to 5' riboexonuclease. Recently, a TUTase was purified and the gene cloned from *L. tarentolae* (Aphasizhev et al., 2002). The enzyme was specific for the 3' U-addition. Antibodies against the TUTase co-precipitated a portion of the RNA editing ligases and ~40% of the gRNAs. Inhibition of TUTase by RNAi decreased RNA editing and affected parasite viability. Three 3'-5' exonuclease activities were detected in mitochondrial extract from *L. tarentolae*. One of these enzmyes was purified and was specific for 3'-terminal U's (Aphasizhev and Simpson, 2001). This exonuclease digested RNA in a distributive manner, was specific for single-stranded 3'-oligo(U) and was blocked by a terminal dU. After the removal or addition of U's the edited mRNA fragments are ligated by RNA ligases. The editing complex purified by Rusche et al. (1997) contained three putative RNA ligases whereas Panigrahi et al. (2001b) identified two novel proteins named TbMP52 and TbMP48 by purifying the editing complex. Antibodies against the ligase TbMP52 immunoprecipitated *in vitro* RNA editing activity (Panigrahi et al., 2001a). TbMP52 encodes an essential RNA ligase and knock out experiments showed that mRNAs were not edited in a deletion strain leading to death of the bloodstream form of the parasite (Schnaufer et al., 2001). Huang et al. (2001) showed that TbMP52 is needed to ligate RNAs in U-deletion editing. Probably the second ligase TbMP48 seals RNAs in U-insertion. Finally, the gRNA-mRNA duplex has to unwind to allow the next gRNA to hybridize or to bind to the ribosome. Unwinding may occur due to weak G:U base pairs in the duplex (Maslov and Simpson, 1992) or by a RNA helicase (Missel and Goringer, 1994).

## 2.2 C to U editing

### 2.2.1 Apolipoprotein B editing

The first example of RNA editing discovered in mammals was the tissue-specific editing of apolipoprotein B (apoB) mRNA. Thereby, a single cytidine is converted to uridine which results in the change of a glutamine codon (CAA) to a stop codon (UAA). RNA editing generates two forms of apoB: apoB100 is synthesized in the liver and represents the major protein component of very low density lipoproteins (VLDL) and their maturation products intermediate density lipoproteins (IDL) and low density lipoproteins (LDL). The C-terminus of apoB100 interacts with the LDL receptor and thereby removes LDL from the circulation. LDL transports two-thirds of the plasma cholesterol in humans and high levels of LDL cholesterol is one of the main risk factors for coronary heart disease. apoB48 which is identical to the N-terminal 48% of apoB100 is synthesized in the small intestine and is required for synthesis and secretion of chylomicrons. These particles transport fats from the intestine to the liver and to adipose tissue (reviewed by Anant and Davidson, 2001; Chester et al., 2000; Innerarity et al., 1996; Keegan et al., 2001). apoB100 and apoB48 are both synthesized from the same 14 kb mRNA which encodes 4536 amino acids (Chang et al., 1998). C to U editing at position 6666 is a nuclear event and occurs with high site-specificity on spliced apoB mRNA (Lau et al., 1994a). The establishment of an *in vitro* system was crucial to identify the trans-acting factor and the cis-acting sequence elements around the editing site (Driscoll et al., 1989). cis-acting elements encompass 26 nt consisting of a tripartite motif which includes a 11 nt 'mooring' sequence located 4-5 nt downstream of the edited C, a 'spacer' element of 2-8 nt located between the mooring sequence and the edited C and an enhancer region upstream of the editing site (Backus and Smith, 1992; Davies et al., 1989; Driscoll et al., 1993; Innerarity et al., 1996; Shah et al., 1991; Smith and Sowden, 1996). More distal sequences were also postulated to play a role in apoB editing (Hersberger and Innerarity, 1998). apoB mRNA was suggested to form a stem-loop structure at the editing site with the edited C in the loop (Richardson et al., 1998).

ApoB C to U editing is catalyzed by a protein complex (Fig. 2.2). The catalytic subunit was initially cloned from rat small intestine (Teng et al., 1993) and was named APOBEC1 for apoB mRNA editing catalytic polypeptide 1 (Davidson et al., 1995). APOBEC1 belongs to a family of cytidine deaminases (CDAs) and it contains the motifs that are characteristic of cytidine deaminases. A deaminase domain with three $Zn^{2+}$-binding residues and a glutamic acid that is thought to mediate proton transfer during catalysis is very similar to the deaminase domain of *E. coli* cytidine deaminase (ECCDA; Navaratnam et al., 1995). Two phenylalanine

residues in the active site that are not found in ECCDA are involved in apoB mRNA binding (Anant et al., 1995; Navaratnam et al., 1995). Furthermore, APOBEC1 contains a putative nuclear localization signal (NLS) at the N-terminus, a RNA-binding domain and a leucine-rich domain in the C-terminus which could be involved in dimerization (Lau et al., 1994b). APOBEC1 forms a complex with the auxiliary factor ACF (APOBEC1 complementation factor, also known as ASF, APOBEC1 stimulating factor) that is the minimun requirement for *in vitro* apoB editing (Lellek et al., 2000; Mehta et al., 2000). ACF has three RNA-recognition motifs (RRMs) at the N-terminus and a putative double-stranded RNA binding domain at the C-terminus which might bind the stem that is formed between the mooring sequence and the 3'-efficiency element (Fig. 2.2; Hersberger et al., 1999). ACF binds apoB mRNA and interaction with APOBEC1 might recruit the complex to the editing site (Mehta et al., 2000).

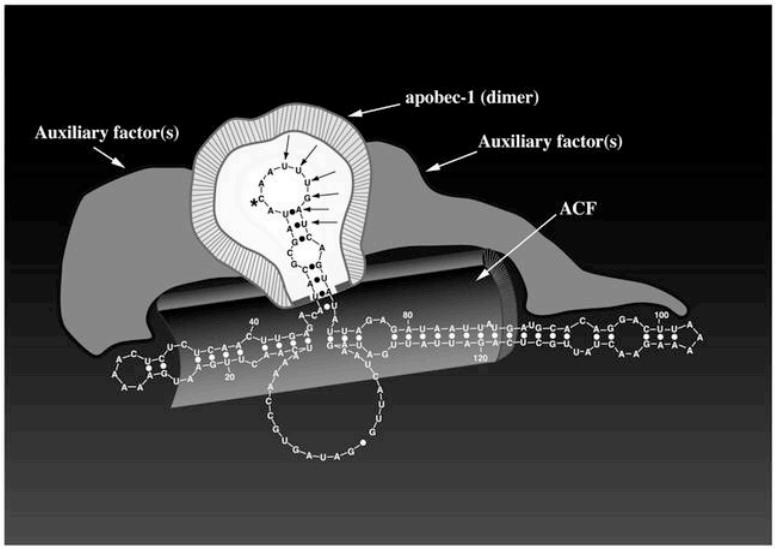



**Figure 2.2 Molecular model of the minimal apoB mRNA editing complex (from Anant and Davidson, 2001).**

The apoB mRNA from nt 6609-6744 was modeled with the mfold program. The model predicts a three branch, stem-loop structure with the edited C (asterisk) within a loop. APOBEC1 binding sites are marked with arrows. APOBEC1 is predicted to form a homodimer (the second APOBEC1 is not shown ), ACF is predicted to bind to the loop containing the edited C as well as the other two loops based on the presence of single- and double-stranded RNA binding domains. Other proteins that might regulate apoB mRNA editing are predicted to bind to the minimal complex and form a holoenzyme.

In humans and rabbits APOBEC1 is expressed exclusively in the small intestine (Teng et al., 1993), whereas in rodents it is also expressed in the liver and in tissues that do not express apoB including spleen, kidney, gonads and brain (Greeve et al., 1993). Due to the broader expression of apoB48 rodents have low LDL levels (Greeve et al., 1993). To determine the

phenotypic effects of apoB mRNA editing, apobec$^{-/-}$ and APOBEC1-overexpressing mice were generated. apobec$^{-/-}$ mice were healthy and had little changes in lipoprotein concentration (Hirano et al., 1996; Morrison et al., 1996; Nakamuta et al., 1996). However, mice are probably not the perfect model system to address the effects of apoB editing because apoB100 levels are naturally low in rodents. APOBEC1-overexpressing mice had reduced LDL concentrations, but had dysplasia and developed hepatocellular carcinomas, probably due to aberrant editing of mRNAs that are normally not edited (Yamanaka et al., 1995; Yamanaka et al., 1997). One such mRNA is Nat1 (novel APOBEC1 target 1) which encodes a eIF4G homologue. Editing of Nat1 mRNA was suggested to interfere with its repressor function which could allow the expression of tumorigenic genes (Yamanaka et al., 1997; reviewed by Chester et al., 2000; Keegan et al., 2001). Furthermore, editing of additional cytidines in the apoB mRNA were found in the APOBEC1-overexpressing mice (Sowden et al., 1998). Homology to the apoB mRNA mooring sequence was found in the human neurofibromatosis type 1 (NF1) mRNA (Mukhopadhyay et al., 2002; Skuse et al., 1996). The appropriately spaced C in this mRNA is indeed deaminated leading to a truncated protein which probably no longer acts as a tumor suppressor.

A human APOBEC1-related protein has been identified which is expressed exclusively in the heart and skeletal muscle (Liao et al., 1999). However, this protein named APOBEC2 cannot edit or bind apoB mRNA and has no cytidine deaminase activity. Jarmuz et al. (2002) identified a cluster of APOBEC1-related genes on chromosome 22. These APOBEC3A to 3G proteins bind zinc, RNA and form homodimers similar to APOBEC1. Expression in a variety of tumor cell lines suggested a role in growth or cell cycle control, however the function of these proteins is not known. Another APOBEC1-related protein is phorbolin-1 which is highly expressed in psoratic lesions. Phorbolin-1 has no cytidine deaminase activity and does not edit apoB mRNA (Madsen et al., 1999). Candidate proteins that might contribute to apoB editing were identified by APOBEC1-binding, RNA affinity and UV cross-linking (Anant et al., 2001b; Greeve et al., 1998; Harris et al., 1993; Lau et al., 1990; Lau et al., 1997; Mehta and Driscoll, 1998; Navaratnam et al., 1993b; Richardson et al., 1998; Schock et al., 1996). One of those is GRY-RBP (glycine-arginine-tyrosine-rich RNA-binding protein) which is 50% homologous to ACF (Blanc et al., 2001). GRY-RBP inhibits apoB mRNA editing by binding ACF thereby inhibiting its interaction with apoB mRNA. The role of the other factors is unclear. They might contribute to specificity or could modulate the editing reaction. Site-specificity might also be achieved by structural elements of the apoB mRNA which could direct the enzymatic machinery to the correct site (reviewed by Davidson, 2002).

### 2.2.2 Activation-induced cytidine deaminase (AID)

AID was identified from germinal center B cells in a subtractive screen for genes involved in class switch recombination (CSR) and somatic hypermutation (SHM), two essential processes for antibody maturation (Muramatsu et al., 1999). AID is homologous to APOBEC1 and has cytidine deaminase activity, but does not edit apoB mRNA *in vitro* (Muramatsu et al., 1999). aid$^{-/-}$ mice did not undergo CSR, had enlarged germinal centers indicative of a hyper-IgM phenotype and were also defective in SHM (Muramatsu et al., 2000). In agreement with these results, 18 patients with the hyper-IgM syndrome HIGM2 all had mutations in the coding sequence of hAID ( Revy et al., 2000; reviewed in Longacre and Storb, 2000). HIGM2 is characterized by defective CSR and SHM and by abnormal germinal centers. Recently, it was shown that AID could induce SHM in hybridoma cells that represent a late stage of B-cell differentiation (Martin et al., 2002). Most of the mutations involved G/C nucleotides, suggesting that AID might be a DNA-specific cytidine deaminase (Martin et al., 2002). Okazaki Im et al. (2002) showed that ectopic expression of AID was sufficient to induce CSR in a synthetic switch substrate in a murine fibroblast cell line. Therefore, all components for CSR except AID are expressed constitutively and ubiquitously. AID is also involved in gene conversion of immunoglobulins, another antibody maturation process (Arakawa et al., 2002). How AID is involved in the three maturation steps remains a mystery. It is unknown whether AID really is a cytidine deaminase acting either on mRNA or DNA. This hypothesis is based solely on its homology with APOBEC1, but so far no substrate has been identified.

## 2.3 A to I editing in pre-mRNAs

Adenosine to inosine deamination has been observed in tRNAs, viral transcripts and mRNAs from numerous eukaryotic species (Bass, 1993; Rueter and Emeson, 1998). The extent of editing varies between transcripts, some RNAs contain only a single editing site whereas in others up to 50% of the As are modified (Rueter and Emeson, 1998). This ,'hypermutation type' of editing has been observed in RNAs encoding the matrix protein of measles virus (Bass et al., 1989; Cattaneo, 1994) and in RNAs of other viruses (Hajjar and Linial, 1995; Kumar and Carmichael, 1997; Murphy et al., 1991; O'Hara et al., 1984). Non-viral examples are the *Drosophila* 4f-rnp gene which encodes a RNA-binding protein of unknown function (Petschek et al., 1996) and transcripts encoding a voltage-dependent potassium channel from the squid *Loligo peali* (Patton

et al., 1997). Because I pase pairs with C during reverse transcription, A to I deaminations are identified as A to G changes when comparing genomic and cDNA sequences. I is read as G by the translational machinery (Basilio et al., 1962), therefore A to I editing in pre-mRNAs can change codons and their amino acid specificity in the mature mRNAs.

Analyses of editing sites in mammals and viruses showed that the substrate RNAs formed double-stranded (ds) structures with bulges and loops (Egebjerg et al., 1994; Higuchi et al., 1993; Rueter et al., 1999). The dsRNA is usually formed between exonic and intronic sequences named ECS (editing site complementary sequence, Fig. 2.4; Burns et al., 1997; Herb et al., 1996; Higuchi et al., 1993; Lomeli et al., 1994; Rueter et al., 1999). Mutations disrupting the base pairing lead to a decrease or complete loss of editing while compensatory mutations that restored base pairing also restored editing (Higuchi et al., 1993; Rueter et al., 1995). The structure of the pre-mRNA rather than its primary sequence was suggested to determine editing sites (Bass, 1997). A 5'-neighbour preference of A, U and C and a disfavour for A close to the 3'-end of a duplex region have been identified (Polson and Bass, 1994).

Pre-mRNAs coding for subunits of the glutamate-gated ion channels (GluRs) expressed in the mammalian brain were the first physiological substrates identified that undergo site-specific A to I editing (reviewed in Seeburg, 1996; Seeburg et al., 1998). The major excitatory neurotransmitter in the central nervous system (CNS) L-glutamate activates the three distinct receptor families NMDA (N-methyl-D-aspartate), AMPA (□-amino-3-hydroxy-5-methyl-4-isoxazolpropionic acid) and kainate receptors. The pentameric receptors are assembled from 18 sequence-related subunits, five of which undergo pre-mRNA editing at up to three sites (Seeburg, 1996; Seeburg et al., 1998). A physiologically important editing event occurs at the Q/R site of the AMPA receptor subunit GluR-B (Higuchi et al., 1993; Sommer et al., 1991). The arginine residue resulting from editing is located in a pore forming domain and is responsible for the reduced permeability to $Ca^{2+}$ ions of AMPA receptors containing this subunit. Another site that is edited in the GluR-B pre-mRNA is the R/G site. Other pre-mRNAs in the CNS that are selectively edited are serotonin receptor $5HT_{2C}$ pre-mRNA (Burns et al., 1997), voltage-gated calcium and sodium channels (Hanrahan et al., 2000; Smith et al., 1996) and a glutamate-gated chloride channel in *Drosophila* (Semenov and Pak, 1999). Serotonin receptors transmit signals via G-proteins and subsequent stimulation of phospholipase C, thereby also regulating synaptic transmission. Pre-mRNAs coding for the serotonin receptor 5-$HT_{2C}$ are edited at four sites. The fully edited receptor has a 10-15 fold reduction in the efficacy of receptor G-protein interaction (Burns et al., 1997). In a screen for novel small non-coding RNAs, a brain-specific C/D box snoRNA was identified that has complementarity to the serotonin receptor 5-$HT_{2C}$ mRNA

covering two of the four editing sites (Cavaille et al., 2000). Interestingly, the mRNA nucleotide potentially targeted for methylation by this snoRNA is one of the four edited adenosines. Therefore, it was suggested that editing and alternative splicing of serotonin receptor 5-$HT_{2C}$ mRNA could be regulated by the binding of the snoRNA. A to I editing in the antigenomic RNA of the Hepatitis Delta Virus (HDV) changes a stop codon (UAG) to a tryptophan codon (UIG), thereby controlling the production of two protein variants, each of which has an important role during the viral life cycle (Polson et al., 1996; Polson et al., 1998). The shorter (unedited) form of the protein is needed for replication whereas the longer (edited) form promotes packaging of the viral genome.

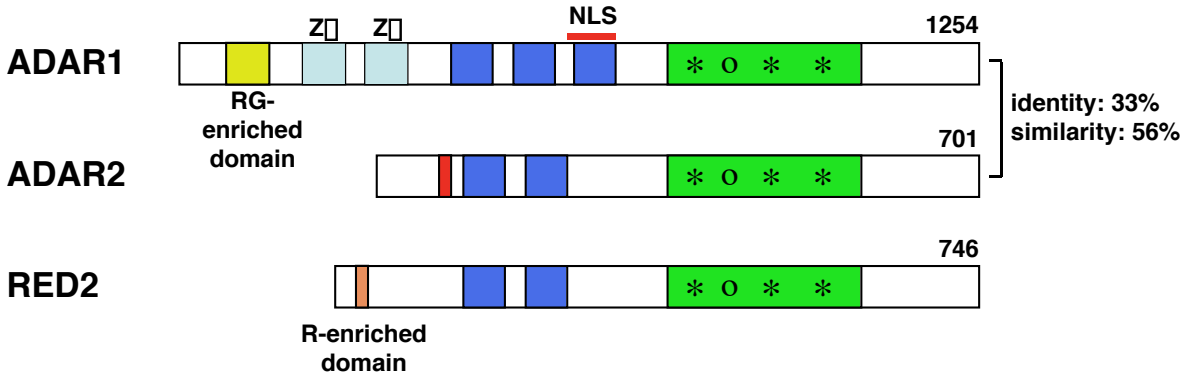



**Figure 2.3 Protein domain structure of mammalian adenosine deaminases acting on RNA (ADARs).**

The enzymes contain a deaminase domain (green) with three Zn2+-chelating residues (asterisks) and a glutamate mediating proton transfer during catalysis (circle), two or three dsRNA binding domains (dsRBDs; blue) and a bipartite nuclear localization signal (NLS; red). The size of the proteins is indicated at the C-terminus by the number of amino acids. The amino acid relationship is shown as percentage identity and similarity (adapted from (Gerber and Keller, 2001).

Pre-mRNA editing is catalyzed by adenosine deaminases acting on RNA (ADARs; Bass et al., 1997). In humans, three members of this family have been identified: ADAR1, ADAR2 and RED2/ADAR3. ADAR1 and ADAR2 are ubiquitously expressed (Wagner et al., 1990) and can convert A to I in long dsRNAs and at specific sites in pre-mRNAs (reviewed in Gerber and Keller, 2001; Keegan et al., 2001; Keller et al., 1999; Maas and Rich, 2000b). Ubiquitous expression goes along with the observation that I is found in mRNAs of all tissues with highest level in brain mRNAs (Paul and Bass, 1998). ADARs do not require cofactors to deaminate A to I. The enzyme binds to the pre-mRNA substrate via its dsRNA binding domains (RBDs, Fig. 2.4) which contribute to the identification of the editing site (Yi-Brunozzi et al., 2001). However, the main deteminant for substrate specificity of ADARs is found within the deaminase domain.

Protein chimaeras between ADAR1 and ADAR2 where the deaminase domain was exchanged showed that this domain has an important role in defining substrate specificity (Wong et al., 2001). Several isoforms of the ADAR1 and ADAR2 have been described which also vary in their editing activity (Gerber et al., 1997; Lai et al., 1997; Liu et al., 1997). ADAR1 and ADAR2 have distinct but overlapping substrate specificity and do not need cofactors in contrast to APOBEC1. ADAR2 edits the Q/R site in GluR-B pre-mRNA, whereas ADAR1 edits the intronic hotspot 1 site. Both enzymes deaminate A to I at the R/G site (reviewed in Keller et al., 1999; Maas and Rich, 2000b). ADAR2 can edit its own pre-mRNA, thereby generating an alternative splice site (Rueter et al., 1999). The edited mRNA includes a 47 nt cassette that changes the reading frame leading to a 82 amino acid protein. Instead translation starts at the second methionine and generates an enzyme that is catalytically active *in vitro*. However, only low amounts of this ADAR2 isoform are detectable, probably due to instability of the protein or inefficient translation initiation. ADAR2 pre-mRNA self-editing could serve as a negative autoregulatory mechanism which prevents ADAR2 overexpression and avoids editing at aberrant sites. RED2/ADAR3 is expressed only in the brain and so far no substrate has been identified (Chen et al., 2000; Melcher et al., 1996). *D. melanogaster* has a single ADAR protein (Palladino et al., 2000a), whereas *C. elegans* encodes two ADARs (Hough et al., 1999). ADARs have a common domain organization: one to three dsRNA binding domains (dsRBDs) are followed by a C-terminal catalytic deaminase domain with three conserved deaminase motifs (Fig. 2.3). The extended N-terminal region of ADAR1 contains two Z-DNA binding domains the function of which is not known (Schade et al., 1999). Recently, the nuclear localizatin signals (NLS) of human and *Xenopus* ADAR1 were mapped. In *Xenopus* ADAR1 a short basic region upstream of the dsRBDs was sufficient for nuclear import whereas the homologous sequence in human ADAR1 does not display NLS activity (Eckmann et al., 2001). In human ADAR1 the NLS overlaps with the third dsRBD, however nuclear import does not depend on RNA binding, showing a dual function for this domain.

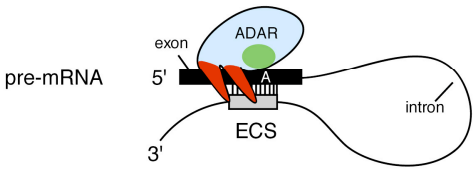



**Figure 2.4 Schematic representation of the substrate recognition by ADARs (from Gerber and Keller, 2001).**

ADARs recognize double-stranded RNA that is formed between the editing site and the ECS (editing site complementary sequence, shown as white box) that is often located in a downstream intron. The enzyme binds to

the pre-mRNA through its double-stranded RNA binding domain (shown in red) and deaminate a specific adenosine to inosine by the deaminase domain (green).

ADAR1 and ADAR2 genes were knocked out in mice and a null mutant of the *Drosophila* ADAR has been characterized. Surprisingly, flies are viable without ADAR and they developed normally (Palladino et al., 2000b). However, the mutant flies showed strong behavioural deficits including locomotion and mating defects and tremors that increase in severity with age. Large lesions in the brain and a disorganized retina were observed, however the flies had an almost normal lifespan. Heterozygous mice for ADAR2 were normal, homozygotes had a normal embryonic development but died from postnatal day P0 to P20 (Higuchi et al., 2000). The adar2$^{-/-}$ phenotype could be rescued with a GluR-B allele that already encoded the edited position at the Q/R site and was thus independent of editing. This experiment showed that the Q/R site is the physiologically most important ADAR2 editing site. ADAR1$^{+/-}$ mice died before embryonic day 14 and had defects in erythropoiesis (Wang et al., 2000). A heterozygous embryonic lethal phenotype is very rare, therefore a dominant-negative effect of the truncated ADAR1 cannot be excluded. In addition, viable ADAR1$^{+/-}$ mice could be generated by another group, but the homozygotes died during embryonic development (reviewed in Keegan et al., 2001). The differences between these results might be due to the different constructs that were used to generate the transgenic mice.

RNA editing generates diversity in proteins that are encoded by a single locus. RNA editing was also suggested to be involved in several diseases either by the lack of or aberrant editing. Sodhi et al. (2001) showed that schizophrenia patients have reduced editing of the serotonin 5-HT$_{2C}$ receptor transcripts. Another study showed elevated levels of editing in people who have committed suicide (Niswender et al., 2001). Underediting of GluR-B mRNA, alterations in serotonin receptor 5-HT$_{2C}$ mRNA editing and alternative splicing were found in malignant human brain tumors (Maas et al., 2001). These results suggested a role for RNA editing in tumor progression and might explain the occurrence of epileptic seizures in association with malignant gliomas. Alterations in the level of editing at the Q/R site in GluR-B pre-mRNA was reported in patients with Alzheimer disease, Huntington disease and schizophrenia (Akbarian et al., 1995). A to G and C to U changes have been found in HIV-1 (human immunodeficiency virus 1) transcripts of chronically infected cells (Bourara et al., 2000), however it is not clear whether these changes are the result of RNA editing or error-prone reverse transcription (Berkhout et al., 2001). More research will be required to elucidate whether RNA editing is indeed involved in all these diseases.

## 2.4 Modifications in tRNAs

### 2.4.1 General aspects

tRNAs are transcribed into long RNA units that are enzymatically trimmed to yield a functional tRNA (Deutscher, 1984). In bacteria, polycistronic as well as monocistronic precursors are present whereas in eukaryotes the majority of the primary transcripts are monocistronic. In both eukaryotes and bacteria, modifications occur at different stages during the processing of the precursor tRNA. In bacteria, several modifications occur in polycistronic precursors. Some modified nucleotides inhibit the RNAse P reaction which generates the mature 5' end of the tRNA suggesting that this cleavage step occurs early in the tRNA maturation process and precedes modification (Schaefer et al., 1973). Thus, on one hand, results show that modification can occur on a molecule of mature size and that some endonucleolytic steps may take place before modifications. On the other hand, when available, precursor molecules may be used as substrate for the modifying enzymes. Therefore, tRNA modification is a function of the processing stage, the concentration of the substrate and the amount and activity of the tRNA-modifying enzyme.

The function of a modification in a tRNA may depend on its position and is not necessarily conserved for a given modified nucleotide. The occurrence of a modification does not necessarily correspond to a necessity in terms of biological functions, because to be modified or not is determined by the specificity of the modifying enzyme and therefore the modification my occur in more tRNAs than are really essential. Modifications in the anticodon arm affect translational efficiency and specificity indirectly by modulating anticodon arm conformation which can affect the tRNA:mRNA complex structure and stability. These modulations can increase translational efficiency and prevent frameshifting. Structural alterations far from the anticodon are known to change the structure of the anticodon and thereby also influence the decoding ability of the tRNA. However, it has been difficult to obtain results demonstrating a function of modified nucleosides outside the anticodon region. Several modified nucleotides are present in the same subset of tRNAs in all organisms, suggesting that their occurrence was an early evolutionary event. Most if not all tRNA-modifying enzymes are position specific and some modifications are catalyzed in several steps requiring several enzymes. Usually, enzymes have evolved to recognize one specific substrate and to disregard other molecules similar to the substrate.

*Escherichia coli* encodes 86 tRNA genes which represent 40 different tRNA species. If the average size of a gene is assumed to be 1 kb, the synthesis of all tRNA-modifying enzymes in *E. coli* requires 1% of the bacterial genome. Assuming the average size of the a tRNA gene in *E. coli* to be 150 nucleotides with mature tRNAs 75-95 nucleotides in length, the genetic information devoted to the synthesis of the primary tRNA transcripts is about 12 kb, which represents about 0.25% of the genome of *E. coli*. Thus, in bacteria, at least four times more genetic information is devoted to the synthesis of the tRNA-modifying enzymes than to the synthesis of their substrates.

Almost 30 different modified nucleotides have been identified in *E. coli* tRNAs which are almost exclusively found in single-stranded regions of the tRNA (Fig. 2.5). All tRNA species contain $\Psi_{55}$, $m^5U_{54}$ and the unmodified $U_{33}$ and modifications at positions 34 and 37 are frequent. Some modified nucleotides ($\Psi$, $s^4U$, D) are found in more than one position of the tRNA and in several cases a certain modification is found in only one tRNA species (Fig. 2.5). The function of a particular modification does not have to be essential for all the tRNA species it is part of, but depends on the position in the tRNA and on the tRNA species (table II). A great variety of modified nucleotides is observed at positions 34 (wobble position) and 37. Modifications at the wobble position can directly affect translation by altering the pattern of hydrogen bond donors and acceptors. Many modifications also affect wobbling by altering nucleoside conformation. Certain of these modifications increase wobbling, while others restrict it. Specific modifications can therefore be either determinants or antideterminants for the translation of specific codons. Position 37 is virtually always a purine and is usually modified. These modified bases may stabilize cognate anticodon:codon interaction, primarily through increased base stacking (Jukes, 1973; Nishimura, 1972). Increased stability of the codon:anticodon pairing is associated with increased translational efficiency and in some cases with enhanced reading frame maintenance. Modifications may also prevent base 37 from pairing with the message (Pieczenik, 1980). In no cytoplasmic tRNA is an unmodified A or U present at position 34 and a pattern can be seen for modifications at this position. tRNAs with a certain modified nucleotide read certain codons and some modifications restrict wobbling whereas others extend wobbling (see chapter 2.4.3).

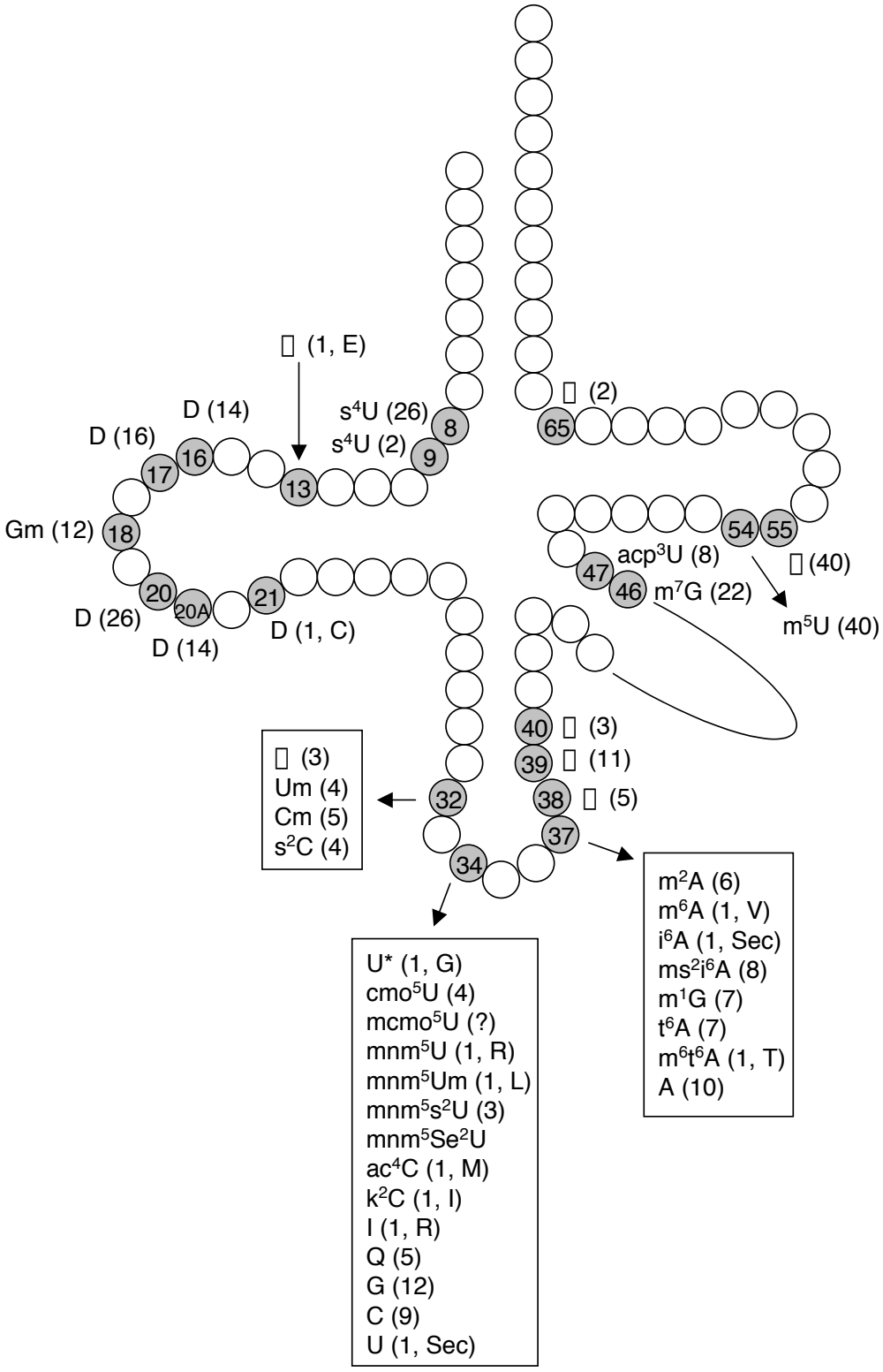



**Figure 2.5 Location of modified nucleotides in tRNAs from *E. coli* (adapted from Björk, 1996).**

Shaded positions are those at which modified nucleotides are present. Figures within parentheses show the number of tRNAs having the indicated modified nucleotide. If this modification is found in only one tRNA, the amino acid specificity of that tRNA is also shown (one-letter code; Sec denotes selenocysteine). U* indicates an unidentified modification.

## Table II Modifications in tRNAs from *E. coli*

Abbreviations to the left of the nucleoside symbol denote modifications of the base, whereas a symbol to the right of the nucleoside symbol denotes modification of the ribose. An index and an exponent indicate the number and the position of the substitution, respectively.

| Modification | Name | Function |
|---|---|---|
| $m^2A$ | 2-methyladenosine | |
| $m^6A$ | $N^6$-methyladenosine | may strengthen base-pairing between pos. 36 of anticodon and first position of codon, may prevent $A_{36}$ from base-pairing other than U |
| $i^6A$ | $N^6$-isopentenyladenosine | |
| $ms^2i^6A$ | 2-methylthio-$N^6$-pentenyladenosine | decodes UNN codons, stabilization of anticodon:codon interactions, effectively compensating for the weak A:U base pair |
| $t^6A$ | $N^6$-threonylcarbamoyladenosine | modification may stabilize U:A base pair at the first codon position, positive determinant for IleRS |
| $m^6t^6A$ | $N^6$-methyl-$N^6$-threonylcarbamoyladenosine | may prevent misreading at the first position with a U:G base pair |
| $s^2C$ | 2-thiocytidine | may increase efficiency of codon:anticodon formation |
| $ac^4C$ | $N^4$-acetylcytidine | reduces reading of AUG codons, decreases misreading of noncognate AUA codons |
| $k^2C$ | lysidine | prevents misacylation, changes base-pairing abilitiy of C to recognize only A |
| Cm | 2'-O-methylcytidine | pos. 32: restricts nucleotide flexibility<br>pos. 34: restricted wobbling with $tRNA^{Met}$ |
| D | dihydrouridine | establishing correct conformation for aminoacylation? |
| Gm | 2'-O-methylguanosine | ? |
| $m^1G$ | 1-methylguanosine | methyl group prohibits base pairing with Watson-Crick geometry, might prevent out-of-phase reading with shifted or expanded anticodon<br>methyl group may increase base stacking |
| $m^7G$ | 7-methylguanosine | ? |
| I | inosine | increase codon:anticodon pairing possibilities |
| Q | queuosine | minor effects on decoding of U and C |
| $s^4U$ | 4-thiouridine | pos. 8: acts as sensor for near-UV light and protects cells from such stress, prevents |

| | | |
|---|---|---|
| | | expression of SOS response and thus reduces mutagenesis |
| Ψ | pseudouridine | different functions depending on the position of Ψ in the tRNA:<br>pos. 32: ?<br>pos. 34, 35: increases translational efficiency by stabilizing anticodon:codon pairing<br>pos. 38-40: increases translational efficiency, especially at pos. 38 |
| Um | 2'-O-methyluridine | ? |
| $cmo^5U$ | uridine 5-oxyacetic acid | enhance wobbling, tRNAs read A, G and U |
| $mcmo^5U$ | uridine 5-oxyacetic acid methyl ester | |
| $mnm^5U$ | 5-methylaminomethyluridine | |
| $mnm^5Um$ | 5-methylaminomethyl-2-O-methyluridine | restricts wobbling, tRNAs read A > G |
| $mnm^5s^2U$ | 5-methylaminomethyl-2-thiouridine | |
| $mnm^5Se^2U$ | 5-methylaminomethyl-2-selenouridine | |
| $acp^3U$ | 3-(3-amino-3-carboxypropyl)-uridine | ? |
| $m^5U$ | ribosylthymine | stabilizes tRNA structure, decreases errors and increases A-site binding |
| U* | unknown modification | ? |

### 2.4.2 Role of modified nucleotides in aminoacylation

Amino acids are attached to tRNAs by aminoacyl-tRNA synthetases. Each of the 20 aminoacyl-tRNA synthetases in a cell must distinguish its own set of isoacceptor tRNAs from the many noncognate tRNAs to catalyze the covalent attachment of the correct amino acid to the 3' end of only these species. Ultimately, the fate of the cell rests on this interaction, as there are no subsequent proof-reading steps in protein synthesis whereby the amino acid is matched against the anticodon to ensure that the proper amino acid is inserted in response to a given codon. The research of the last 25 years on tRNA recognition allows to summarize a few general recognition patterns. Major identity elements for aminoacyl-tRNA synthetases are nucleotides in the anticodon and in the distal part of the acceptor arm. However, the relative importance of these two widely spaced tRNA domains varies greatly among different tRNA species. In *E. coli* $tRNA^{Met}$ the anticodon contains the major recognition elements (Schulman and Pelka, 1988; Stern and Schulman, 1977) whereas in *E. coli* $tRNA^{Ala}$ the acceptor arm alone constitutes the recognition site for alanyl-tRNA synthetase (AlaRS; Hou and Schimmel, 1988; McClain and Foss, 1988). Between these two extremes lie the majority of tRNAs which harbor nucleotides

important for their identities in both these tRNA domains. The delineation of the entire set of identity elements for a given tRNA proved to be difficult because many regions of the tRNA playing no role in recognition could serve to block aminoacylation by competing aminoacyl-tRNA synthetases. Thus, the number of potential identity elements is very large, perhaps most of the nucleotides in a tRNA.

Most modified nucleotides in tRNA may not be essential for the aminoacylation reaction (Björk, 1992). However, nucleotides in the anticodon are important identity determinants for many aminoacyl-tRNA synthetases. *In vitro* transcription of tRNAs allowed a direct comparison between native and unmodified tRNA. Such comparisons of the kinetics of the aminoacylation reactions with cognate and noncognate aminoacyl-tRNA ligases revealed that of 14 different unmodified tRNAs, all except three *E. coli* $tRNA^{Ile1}$, $tRNA^{Glu}$ and $tRNA^{Lys}$ accept the cognate amino acid (Sylvers et al., 1993; Tamura et al., 1992). Thus, modified nucleotides are not a prerequisite for most aminoacylation reactions *in vitro*. However, in all cases but one (*E. coli* $tRNA^{Asp}$) the cognate interactions with the unmodified species have kinetic characteristics that are different from those of the fully modified species. It was shown that modifications at positions 8, 26, 32, 38, 39, 40, 46, 47, 54 and 55 play no significant role in aminoacylation of the tRNA.

Two isoleucine isoacceptors $tRNA^{Ile1}$ and $tRNA^{Ile2}$ (minor species) are present in *E. coli*. These tRNAs contain $G_{34}$ and $k^2C_{34}$ as the wobble nucleotide, respectively (Muramatsu et al., 1988b; Yarus and Barrell, 1971). Although these two nucleotides are quite different, isoleucyl-tRNA synthetase (IleRS) recognizes both species. When $k^2C_{34}$ is replaced with $C_{34,}$ the $tRNA^{Ile2}$ anticodon is changed to the $tRNA^{Met}$ anticodon CAU (Fig. 2.6). Such a mutant tRNA is efficiently misacylated with methionine (Muramatsu et al., 1988a). The $k^2C_{34}$ modification acts thus as a positive identity element for IleRS and as an antideterminant for MetRS preventing deleterious misacylation. *S. cerevisiae* also has two isoleucine acceptors, a major $tRNA^{Ile}$ with the anticodon IAU and a minor $tRNA^{Ile}$ with the anticodon □A□(Pixa et al, 1984; Szweykowska-Kulinska, 1994). Both nucleotides at the wobble position, I and □, contribute significantly to the efficiency of aminoacylation (reviewed in Pallank et al., 1995).

Chemical modification of $mnm^5s^2U_{34}$ present in $tRNA^{Gln}$, $tRNA^{Glu}$ and $tRNA^{Lys}$ significantly reduced amino acid acceptance of these tRNAs (Kern and Lapointe, 1979; Saneyoshi and Nishimura, 1971; Seno et al., 1974). Several studies showed that the $mnm^5$ group, but not the $s^2$ group is a positive determinant for LysRS, GluRS and GlnRS from *E. coli* (reviewed in Björk, 1995). This was confirmed by the crystallographic structure of the GlnRS-$tRNA^{Gln}$ complex which showed a direct interaction between $mnm^5s^2U_{34}$ and the aminoacyl-

tRNA ligase (Rould et al., 1989). Unmodified $tRNA^{Lys}$ which lacked all modifications, among them $mnm^5s^2U_{34}$, had a 140-fold lower lysine acceptance activity compared with native $tRNA^{Lys}$ (Tamura et al., 1992) and is thus one of the few exceptions where aminoacylation depends on modified nucleotides.

### 2.4.3 Modified nucleotides in translation

Cellular physiology is fundamentally dependent on the functions of our translational apparatus and these functions are dependent on modified nucleotides. Modifications at position 34 are frequent, some have obvious effect on the decoding spectrum by altering the array of H-bonding groups. In other cases, modifications alter the relative affinities for codons by altering anticodon loop structure. In addition, message sequences near codons (codon context) can also affect decoding efficiency and accuracy. Aminoacylated tRNAs also undergo conformational changes during ribosomal selection and those changes can depend on tRNA sequences far from the anticodon. Codon choice is highly correlated with the relative abundance of the corresponding tRNAs (Ikemura, 1985). In *S. cerevisiae*, genes expressed at high levels strongly prefer synonymous codons that are transcribed by major isoacceptor tRNAs (Bennetzen and Hall, 1982).

**Position 34 (wobble position)**

Uridine at the wobble position is almost always modified and the nature of the modification is correlated with the base(s) read. Crick (1966) predicted that U should only read A and G. However, tRNAs with unmodified U occur in mitochondria and *Mycoplasma* and those tRNAs read codons ending with all four bases (Andachi et al., 1989; Osawa et al., 1992; Sibler et al., 1986). This might be due to a limited number of tRNA species in these organelles and organisms and a translational system that has been greatly reduced. Codon family boxes that encode two different amino acids have to be read by tRNAs with a modified U at position 34 to avoid translational errors (Fig. 2.6). The A- and G-ending codons of the split codon boxes are frequently read by tRNAs with $xm^5U_{34}$ modifications ($mnm^5U_{34}$, $mcm^5U_{34}$, $mnm^5Um_{34}$, $mnm^5s^2U_{34}$, $mnm^5Se^2U_{34}$, $Um_{34}$). Because these modifications should not read pyrimidine-ending codons, they contribute to restricted wobbling. However, there is evidence that these tRNAs might also be inefficient in reading G-ending codons because in many cases a tRNA with $C_{34}$ is present and $xm^5s^2U_{34}$ was shown to prefer A-ending codons over G-ending codons *in vitro* (Agris et al., 1973; Lustig et al., 1981). Furthermore, *E. coli* $tRNA^{Arg4}$ decodes AGA better than AGG *in*

*vivo* (Spanjaard et al., 1990). Interestingly, lysine and glutamate have single tRNAs to decode their respective codons in *E. coli* (Fig. 2.6). In these cases greater tRNA:mRNA stability of $xm^5U_{34}$:A is correlated with codon usage. Highly expressed gene strongly prefer the A-ending over the G-ending codon (Dalphin et al., 1997). In a minor $tRNA^{Ile}$ from yeast pseudouridine is present at position 34. This tRNA reads the codon AUA and probably does not frequently read the AUG methionine codon. Thus, $\Psi_{34}$ restricts wobbling. Restricted wobbling can also be achieved with unmodified $U_{34}$. It was shown that reading of GGN glycine codons by a tRNA with unmodified $U_{34}$ depended on the presence of $C_{32}$. When $C_{32}$ was mutated, the decoding was strongly restricted to A>G>>U and C (Lustig et al., 1993).

Amino acids that are encoded by family boxes are decoded by at least two tRNAs. One of the tRNAs has $G_{34}$ and should read codons ending in C and U. In eukaryotes, A- and G-ending codons are read by tRNAs that have $xm^5U_{34}$ and $C_{34}$, respectively. Inosine 34-containing tRNAs are common in eukaryotes and read codons ending in C and U and in *S. cerevisiae* also A-ending codons (Percudani, 2001). Bacteria use tRNA with a $xo^5U_{34}$ modification ($cmo^5U_{34}$, $mcmo^5U_{34}$) to read codons within family boxes but not in split boxes (reviewed in Yokoyama and Nishimura, 1995). Because tRNAs with $xo^5U_{34}$ decoded triplets ending in A, G and U (Ishikura et al., 1971), the $xo^5$ adduct facilitates wobbling. By enhanced wobbling, the $xo^5U_{34}$ modifications might increase translational efficiency.

Several tRNAs contain modified $C_{34}$. $Cm_{34}$ allows the reading of G-ending codons, but some tRNAs with this nucleotide can also recognize A-ending codons. However, there are cases for which Cm:A decoding would result in translational errors (Fig. 2.6). In many eukaryotes and archaea $tRNA^{Met}$ has $Cm_{34}$. These tRNAs should only read AUG because AUA is an isoleucine codon. Because the apparent decoding specificity of Cm appears to vary among genetic systems and tRNAs, other unknown features of tRNAs must strongly contribute to specificity. The isoleucine codon group is unique in that it includes three (AUU, AUC and AUA) of the four codons in a family box, but it does not include the AUG methionine codon (Fig. 2.6). In eukaryotes, the isoleucine codons AUC and AUU are read by the inosine-containing tRNA whereas AUA is read by the tRNA with the complementary anticodon UAU. In bacteria, organelles and certain archaea, the isoleucine codons are read by two tRNAs. One contains $G_{34}$ for reading AUU and AUC and the other contains $C_{34}$ modified with lysidine (or a derivative) to read AUA. Apparently, the $k^2C_{34}$ modification switches base pairing specificity from G to A. Isoleucine codons might not be read with tRNAs containing a modified $U_{34}$ because the repertoire of modifications might not include nucleotides that can prevent the decoding of AUG by isoleucine tRNAs.

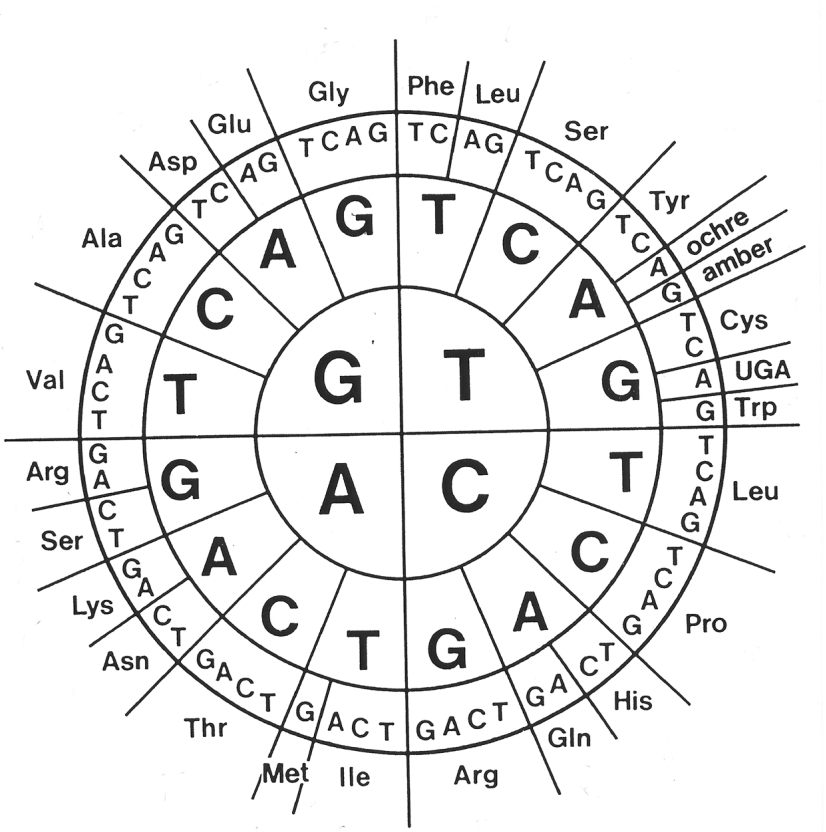



**Figure 2.6 The universal genetic code**

Many bacteria and eukaryotes use tRNAs with $Q_{34}$ to read NAY codons (Sprinzl et al., 1998). This modification does not directly affect the base pairing of G, though it might slightly alter the relative affinities of the base for C and U. An *E. coli* mutant lacking $Q_{34}$ in tRNAs for Tyr, His, Asn and Asp did not show a major growth defect suggesting that codons read with $Q_{34}$ could also be read with $G_{34}$ (Noguchi et al., 1982). Because the NAY codons all occur within split boxes, it is important that these tRNAs do not read A- and G-ending codons (Fig. 2.6). It is not clear whether $Q_{34}$ prevents such misreading.

The elongator $tRNA^{Met}$ is the only tRNA in *E. coli* that contains $ac^4C_{34}$. This suggests a unique function of the acetyl group in this tRNA. Upon removal of the $ac^4$ group by bisulfite, the tRNA gained the ability to misread the isoleucine codon AUA and also became more efficient in recognizing the cognate codon AUG (Stern and Schulman, 1978). Thus, $ac^4$ decreases the efficiency to read both the complementary AUG codon and the noncomplementary AUA codon. $Cm_{34}$ occurs in most eukaryotic and prokaryotic $tRNA^{Trp}$ and UUG-reading $tRNA^{Leu}$ species. *E. coli* $tRNA^{Leu5}$ with the anticodon CmAA recognizes the Leu codons ending in A and G, but not

in U and C. The recognition of A-ending codons is much weaker than that of G-ending codons (reviewed in Yokoyama and Nishimura, 1995). The modified cytidines Cm and $ac^4C$ have the common feature of conformational rigidity and an unusual stability of the C3'-endo form of the ribose.

Inosine at position 34 occurs in eukaryotes and bacteria and a detailed discussion about the function of $I_{34}$ in translation can be found in chapter 5.

**Position 37**

The nucleotide at position 37 is highly correlated with the base at position 36 (Yarus, 1982). The anticodon plus the nearby sequence features (the extended anticodon) contains more coding information than the anticodon alone and can perform more efficiently and accurately at the ribosome. Position 37 is virtually always a purine and is usually modified (Fig. 2.5). tRNAs reading codons starting with U or A have $ms^2i^6A$ or $ms^2io^6A$ and $t^6A$ or $m^6t^6A$ at position 37, respectively. The weak $A_{36}$:U or $U_{36}$:A interactions between position 36 of the anticodon and the first position of the codon might be stabilized by these hypermodifications and prevent first-position errors. If this is true, codons starting with C or G, which results in a strong G:C base pairing, would not require any stabilizing modification at position 37. Indeed, unmodified A or methylated G is found in this position of such tRNA subsets (reviewed in Björk, 1996). Modifications at position 37 may prevent the tRNA from base pairing with the mRNA and also from pairing across the anticodon loop. There is a strong correlation between potential Watson-Crick base pairing within the anticodon loop of primary transcripts and base modifications that would inhibit such pairing (Dao et al., 1994). It has been shown that a suppressor $tRNA^{Trp}$ with $C_{32}$ and $G_{38}$ had an extraordinarily low translational efficiency which could be relieved by mutations at either position that disrupted the putative base pair across the anticodon loop (Yarus et al., 1986). The many roles associated with modifications at position 37 might explain the variety that is observed for modified nucleotides 3' to the anticodon (Fig. 2.5).

The miaA protein catalyzes the first step in the synthesis of $ms^2i^6A_{37}$, the 6-isopentenylation. Since this adduct is required for subsequent modification steps, miaA mutants have completely unmodified $A_{37}$. In general $ms^2i^6A_{37}$-deficient tRNAs were inefficient in translation. Elongation rates were decreased and the average translational cycle was slowed about fourfold *in vivo* (Ericson and Bjork, 1986). Furthermore, lack of $ms^2i^6A_{37}$ also induced an increase in codon context sensitivity (Bouadloun et al., 1986) and increased frameshifting about two-fold (Schwartz and Curran, 1997).

When a guanosine is present at position 37, it is always converted to $m^1G$, suggesting that an unmodified G is detrimental for the tRNA. $m^1G_{37}$ is found in tRNAs that read CUN (leucine), CCN (proline) and CGN (arginine) codons of bacteria. The modification is predicted to have two effects on decoding. First, the methyl adduct at the N1 position of G prohibits base pairing with Watson-Crick geometry. This might prevent out-of-frame reading with a shifted or expanded anticodon (Pieczenik, 1980). Second, because the 1-methyl group increases base hydrophobicity, it may increase base stacking (Nishimura, 1972) which could increase decoding efficiency. Björk et al. (1989) isolated a *S. typhimurium* mutant of the tRNA($m^1G_{37}$)methyltransferase *trmD (trmD3)* that was defective in the methylation of $G_{37}$. This mutant grew 30% more slowly than wild-type cells and the chain elongation rate was reduced. Three proline tRNAs that normally contained $m^1G_{37}$ were found to frequently shift frame when C's were present successively in the message. However, not all codons read by $m^1G_{37}$-containing tRNAs were equally affected by the loss of the methylation. The *leu* attenuator which is sensitive to the rates of translation of four consecutive leucine codons by $m^1G_{37}$-containing tRNAs is not deattenuated by the *trmD3* mutation (Li and Bjork, 1995). This result suggested that the $m^1G$ group affects anticodon arm structure and decoding function, but effect are highly dependent on the broader tRNA sequence and on the codon. The presence of $m^1G_{37}$ affects the cognate anticodon-codon interaction by increasing the aminoacyl-tRNA selection and the polypeptide chain elongation rate.

**Modifications at other positions in the tRNA**

Pseudouridines are found at positions 38, 39 and 40 (Fig. 2.5). Almost half of the tRNAs in *E. coli* contain one or more of these modifications. NMR analyses showed that $\Psi_{39}$ increased the stability of the anticodon arm of *E. coli* $tRNA^{Phe}$ (Davis and Poulter, 1991). Generally, pseudouridines at the 3' side of the anticodon increase translational efficiency, maybe due to stabilization of the anticodon region. Mutations in *hisT* fail to perform pseudouridylation at positions 38-40. The histidine leader mRNA contains seven consecutive histidine codons that were read inefficiently by $tRNA^{His}$ lacking $\Psi_{38}$ and $\Psi_{39}$ which resulted in a derepressed *his* operon (Johnston et al., 1980). The growth rate of the *hisT* mutant was reduced by 30% in minimal medium as was the polypeptide chain elongation (Palmer et al., 1983). Comparison of the effect of $\Psi$ deficiency of two amber suppressors which have $\Psi$ at position 39 alone or at both positions 38 and 39, respectively, suggested that the $\Psi$ at position 38 (in the anticodon loop) exerted in a codon-context independent manner a stronger influence on the activity of the tRNA than $\Psi$ in position 39 (Bossi and Roth, 1980; Hagervall et al., 1990). $\Psi$ has a strong impact on

the activity of the tRNA, and although  in the stem increases the activity, this effect is less than the effect exerted by  in the loop.

Base 32 is always a pyrimidine and it is frequently modified (Fig. 2.5). Structurally, all modifications occuring at this position are predicted to restrict nucleotide flexibility. There is evidence that base identity at position 32 is important for translational efficiency. In *E. coli* $s^2C_{32}$ is present in $tRNA^{Arg2}$. *In vitro* translation of MS2 RNA with cell-free extracts from *E. coli* resulted in the viral synthetase and the coat protein but also in several other polypeptides that were the result of +1 or –1 frameshifting (Baumann et al., 1985). The addition of native $tRNA^{Arg2}$ in 15-fold excess inhibited the –1 frameshifting whereas the addition of $tRNA^{Arg2}(C_{34})$ did not, suggesting that $tRNA^{Arg2}(s^2C_{34})$ efficiently competed with the frameshifting tRNA at the frameshifting site. Mutation of $Cm_{32}$ in an amber suppressor reduced suppression efficiency more than 10-fold (Yarus et al., 1986). Position 32 can also affect the decoding range of tRNAs. An *E. coli* $tRNA^{Gly}$ variant with unmodified $U_{34}$ translated all four GGN codons if base 32 was cytosine (Boren et al., 1993). When $C_{32}$ was changed to $U_{32}$ only GGA was efficiently decoded. The molecular mechanisms for this effect is unknown, but might be related to the microstructure of the anticodon loop.

Little is known about the functions of modified nucleotides outside the anticodon region. One of the reasons might be that their effects on anticodon-codon interaction are minor and that the methods employed to study other effects are not sensitive enough.

The tRNAs of a mutant (*trmA*) lacking $m^5U_{54}$ had a lower stability, increased error levels and decreased A-site binding *in vitro* (Davanloo et al., 1979). Although $m^5U_{54}$ is not essential for cellular growth, the gene (trmA) encoding the tRNA($m^5U_{54}$)mehtyltransferase is essential (Persson et al., 1992). This suggested that trmA has another function which is essential. A *nuvA* mutant lacking $s^4U$ in the tRNA grew normally, but this modification was suggested to be a sensor for near-UV light (Kramer et al., 1988; Thomas and Favre, 1980).

### 2.4.4 A to I editing in tRNAs

A variety of modified nucleotides are present in tRNAs, among them are inosine (I) and $N^1$-methylinosine ($m^1I$) in eukaryotic $tRNA^{Ala}$ (Curran, 1998; Holley et al., 1965). I is found at position 34 (the wobble position of the anticodon) of seven yeast and eight tRNAs of higher eukaryotes, whereas $m^1I$ is present exclusively in eukaryotic $tRNA^{Ala}$ at position 37 (3' adjacent to the anticodon; (Sprinzl et al., 1998). These modifications proceed through hydrolytic

deamination of A to I and subsequent methylation at position 37 (Fig. 2.7 A; Björk et al., 2001; Grosjean et al., 1996b). Based on sequence homology, a *Saccharomyces cerevisiae* open reading frame (ORF) with significant homology to the C-terminal part of ADARs was found. The recombinant protein does not form I in GluR-B pre-mRNAs or extended dsRNA but specifically converts A to I at position 37 in eukaryotic $tRNA^{Ala}$ (Fig. 2.7 C; Gerber et al., 1998). The enzyme was named scADAT1 (*S. cerevisiae* adenosine deaminase acting on tRNA 1) and its gene *TAD1* (tRNA adenosine deaminase 1). The gene is not essential for cell viability and the function of $m^1I_{37}$ is not known. However, in analogy to the function assigned to $m^1G_{37}$ found within several tRNAs from the prokaryote *Salmonella typhimurium*, one possible function of $m^1I_{37}$ (a modification similar to $m^1G_{37}$) could be the prevention of frameshifting and improvement of translation fidelity (Björk et al., 1989; Qian and Björk, 1997). Recently, the human and *Drosophila melanogaster* orthologues of scADAT1 were cloned and the recombinant enzymes were shown to deaminate A to I at position 37 in $tRNA^{Ala}$ (Keegan et al., 2000; Maas et al., 1999). A majority of archaebacterial tRNAs contain $m^1I$ at position 57 (in the T loop). In contrast to the formation of $m^1I_{37}$ in eukaryotes, the first enzymatic step in the *in vitro* formation of $m^1I_{57}$ in archael tRNAs was shown to be a methylation, followed by a deamination (reviewed in Grosjean et al., 1996b). Neither $I_{34}$ nor $m^1I_{37}$ have been identified in archael tRNAs.

In yeast, $A_{34}$ is converted to $I_{34}$ by the tRNA:adenosine 34 deaminase that is comprised of Tad2p and Tad3p (scADAT2/scADAT3) which form a heterodimer (Gerber and Keller, 1999). The C-terminal part of Tad3p is sequence related to Tad2p (Fig. 2.7 B). Interestingly, the two proteins display a deaminase domain that resembles that of the cytidine deaminase (CDA) superfamily, but in contrast to CDAs, Tad2p/Tad3p do not deaminate free cytidine or cytosine *in vitro* (Gerber and Keller, 1999). However, functionally Tad2p/Tad3p belong to the ADAR family. The *TAD2* and *TAD3* genes are both essential for vegetative growth, underscoring the biological importance of this editing event (Gerber and Keller, 1999). Crick (1966) postulated that the first two bases of the codon pair are classical Watson-Crick basepairs. He predicted that the third base of the anticodon (position 34) can basepair in a non-canonical way with different nucleotides in the mRNA ("wobble hypothesis"). $I_{34}$ was predicted to pair with the last position of codons ending in C, U and A (Curran, 1995).

Database searches revealed a number of sequences with homology to Tad2p and Tad3p. Prokaryotic genomes encode only polypeptides that are homologous to Tad2p (see chapter 3). In order to characterize the tRNA:adenosine 34 deaminase from *E. coli* the Tad2 homologue was cloned and the recombinant protein purified (see chapter 3).

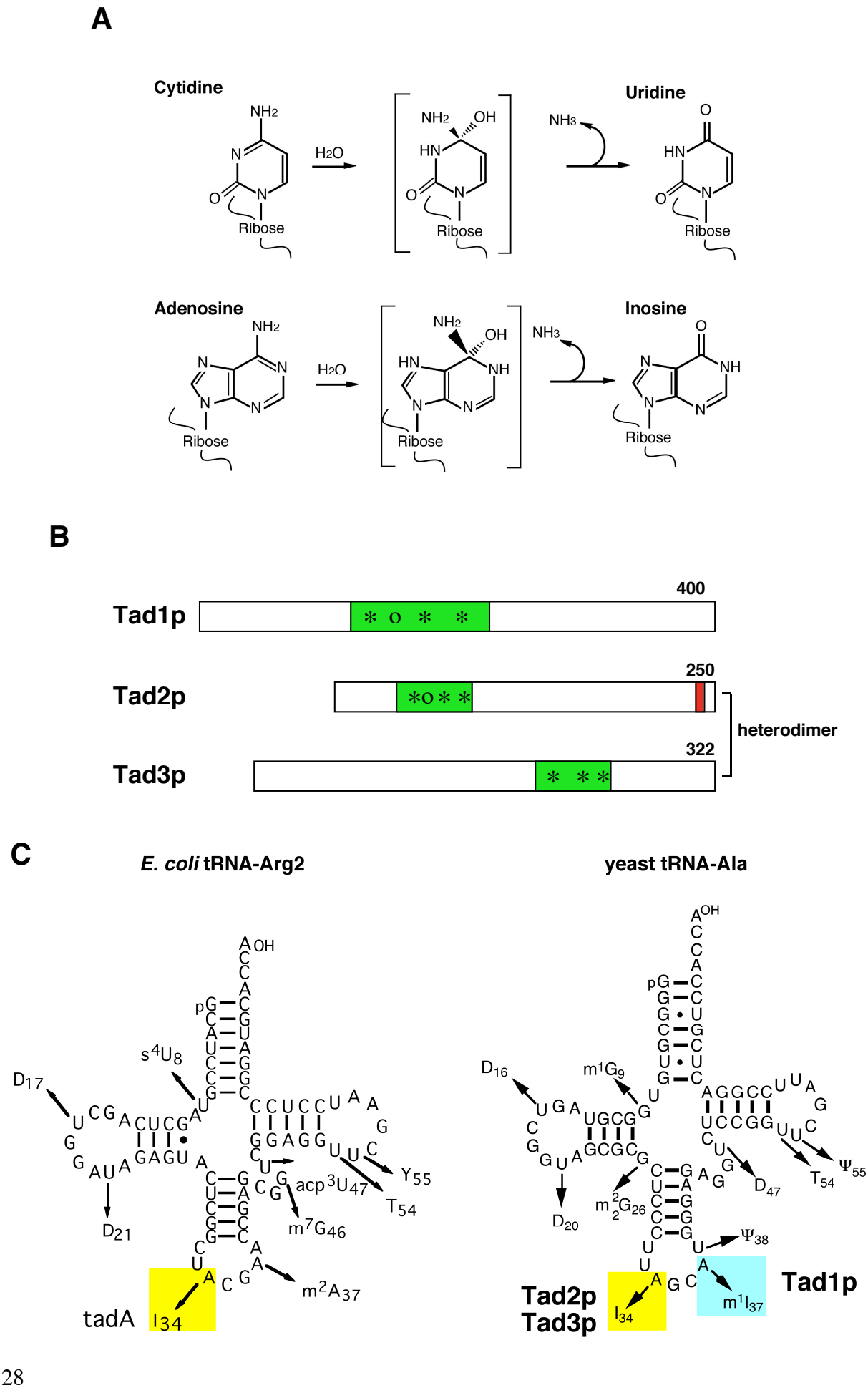

A
Cytidine
Uridine
Adenosine
Inosine
Ribose
H2O
NH3
B
Tad1p
Tad2p
Tad3p
400
250
322
heterodimer
C
E. coli tRNA-Arg2
yeast tRNA-Ala
tadA
Tad2p
Tad3p
Tad1p

**Figure 2.7 Deamination mechanism, protein domain structure of *S. cerevisiae* tRNA-specific adenosine deaminases (Tads) and modifications in $tRNA^{Arg2}$ from *E. coli* and $tRNA^{Ala}$ from *S. cerevisiae*.**

**A)** Reaction mechansim of cytidine and adenosine deamination, respectively (from (Gerber and Keller, 2001). **B)** The proteins contain a deaminase domain (green) with three $Zn^{2+}$-chelating residues (asterisk) and Tad1p and Tad2p contain a glutamate involved in proton transfer during catalysis. A putative nuclear localization signal in Tad2p is shown in red. **C)** Adenosine at position 34 is deaminated in both tRNAs (yellow box) whereas deamination at position 37 by Tad1p and subsequent methylation is present only in eukaryotic $tRNA^{Ala}$ (blue box). Modifications in the two tRNAs are indicated.

## 2.5 Aims of this thesis

In *S. cerevisiae* the heterodimer catalyzing adenosine deamination at the wobble position of tRNAs has been identified (Gerber and Keller, 1999). Although this enzymes has been partially characterized, the isolation of bacterial and human tRNA-specific adenosine deaminases will help to address several questions. (i) Does tRNA editing at the wobble position require conserved proteins in *E. coli*? If yes, can the proteins from different organisms be exchanged? (ii) *E. coli* has only one tRNA containing $I_{34}$. Is the enzyme / are the enzymes catalyzing the reaction essential? If yes, is the formation of $I_{34}$ the essential process or could another reaction be catalyzed by the same enzyme(s)? (iii) What determines substrate specificity and how do tRNA editing enzymes recognize their substrates? What are the differences between eukaryotic and prokaryotic tRNA recognition? (iv) How does $I_{34}$ influence translational efficiency? (v) Where does tRNA editing occur in eukaryotes?

# 3 tadA, an essential tRNA-specific adenosine deaminase from *Escherichia coli*

Jeannette Wolf, André P. Gerber and Walter Keller (2002), *EMBO J.*, **21**, 3841-51

## 3.1 Abstract

We report the characterization of tadA, the first prokaryotic RNA editing enzyme to be identified. *E. coli* tadA displays sequence similarity to the yeast tRNA deaminase subunit Tad2p. Recombinant tadA protein forms homodimers and is sufficient for site-specific inosine formation at the wobble position (position 34) of $tRNA^{Arg2}$, the only tRNA having this modification in prokaryotes. With the exception of yeast $tRNA^{Arg}$ no other eukaryotic tRNA substrates were found to be modified by tadA. However, an artificial yeast $tRNA^{Asp}$ which carries the anticodon loop of yeast $tRNA^{Arg}$ is bound and modified by tadA. Moreover, a $tRNA^{Arg2}$ minisubstrate containing the anticodon stem and loop is sufficient for specific deamination by tadA. We show that nucleotides at positions 33 to 36 are sufficient for inosine formation in mutant Arg2 minisubstrates. The anticodon is thus a major determinant for tadA substrate specificity. Finally, we show that tadA is an essential gene in *E. coli,* underscoring the critical function of inosine at the wobble position in prokaryotes.

## 3.2 Introduction

The nucleotide inosine (I) has been observed in viral transcripts and eukaryotic mRNAs. In all known cases, I is made by the deamination of adenosine, a process termed RNA editing. Because I is read as guanosine by the translational machinery (Basilio et al., 1962), RNA editing can change codon specificity and therefore the amino acid sequence of the encoded protein, resulting in multiple protein products with different biological function from a single mRNA precursor. In mammals for example, the messenger RNA precursors (pre-mRNAs) coding for subunits of glutamate-gated ion channel receptors (GluRs) and the serotonin receptor subunit 5-$HT_{2C}$ are edited (reviewed in Gerber and Keller, 2001; Gott and Emeson, 2000; Keegan et al., 2001; Maas and Rich, 2000b; Seeburg et al., 1998). Editing was also detected in *D. melanogaster*, *C. elegans* and in hepatitis delta virus (Hanrahan et al., 2000; Morse and Bass, 1999; Polson et al., 1996; Semenov and Pak, 1999; Smith et al., 1996). RNA editing therefore represents an important mechanism to increase the genetic diversity in eukaryotes. In all these cases, pre-mRNA editing

requires double-stranded RNA (dsRNA) structures which are formed between exonic sequences encompassing the editing site and downstream intronic sequences.

RNA editing of pre-mRNAs is catalyzed by adenosine deaminases acting on RNA (ADARs; Bass et al., 1997). ADARs have a common modular organization consisting of two or three double-stranded RNA binding domains (dsRBD) and a catalytic deaminase domain containing three $Zn^{2+}$-chelating residues and a proton-shuffling glutamate (reviewed in Bass et al., 1997; Keller et al., 1999; Maas and Rich, 2000b).

Inosine is not only present in mRNAs but also in tRNAs. It was first found in $tRNA^{Ala}$ from yeast (Holley et al., 1965). Eukaryotic $tRNA^{Ala}$ contains I at two positions, at the wobble position of the anticodon (position 34) and as the derivative $N^1$-methylinosine ($m^1I$) at position 37, 3' adjacent to the anticodon (Holley et al., 1965; reviewed in Grosjean et al., 1996b). In eukaryotes, seven to eight tRNAs contain I at position 34 whereas in prokaryotes and plant chloroplasts only $tRNA^{Arg2}$ contains this modification. As in pre-mRNAs, I in tRNAs is the product of hydrolytic deamination of genomically encoded adenosine (Auxilien et al., 1996). $m^1I$ at position 37 is formed in a two-step reaction. First, A is deaminated to I which is further methylated by a methyltransferase (Björk et al., 2001; Grosjean et al., 1996a). The genomes of *S. cerevisiae* and prokaryotes do not encode classical ADAR proteins, but a yeast protein was identified based on sequence homology to ADARs which contains a deaminase domain, but lacks a known RNA-binding motif. This protein catalyzes the deamination of A at position 37 in yeast $tRNA^{Ala}$ and was therefore named adenosine deaminase acting on tRNA 1 and its gene tRNA-specific adenosine deaminase 1 (scADAT1/TAD1; Gerber et al., 1998). Tad1p specifically targets $A_{37}$ in $tRNA^{Ala}$. Mutations affecting the three-dimensional structure of the tRNA or the length of the anticodon loop abolish conversion of A to I *in vitro* (Gerber et al., 1998). Whereas $A_{37}$ is unmodified in a Δ*tad1* strain, modification of $A_{34}$ was unaffected in all tRNAs tested, suggesting that I formation at these two positions is catalyzed by different enzymes. Whereas I in the wobble position ($I_{34}$) is crucial to allow the decoding of three codons by a single tRNA (Crick, 1966), the function of $m^1I_{37}$ is less clear. It was postulated that this modification may prevent translational frameshifts and improve translational fidelity (Björk et al., 1989). ADAT1 proteins have also been cloned from human (Maas et al., 1999), mouse (Maas et al., 2000a) and *D. melanogaster* (Keegan et al., 2000).

The tRNA adenosine deaminase which specifically deaminates $A_{34}$ has been partially purified from yeast extracts (Auxilien et al., 1996) and has been identified by homology to Tad1p (Gerber and Keller, 1999). Activity depended on the correct tRNA structure as well as on the length and the nucleotide sequence of the anticodon loop. The tRNA:$A_{34}$ deaminase is a

heterodimer of sequence-related subunits named scADAT2/Tad2p and scADAT3/Tad3p. Both polypeptides contain a deaminase domain that resembles that of the cytidine deaminase (CDA) superfamily (Gerber and Keller, 1999), although Tad2p/Tad3p deaminates adenosine. Most cytidine/deoxycytidylate deaminases catalyze deaminations of mononucleotides, but there are also a few known cytidine deaminases that act on RNA (CDARs). One of the best studied examples is APOBEC1 which catalyzes the deamination of a cytidine to uridine in the apoB mRNA, resulting in a change of a glutamine codon to a stop codon (reviewed by Chester et al., 2000). Interestingly, the deaminase domain of Tad3p lacks the glutamate which, in analogy to CDAs and CDARs, is thought to be part of the active center. The lack of this glutamate suggests that Tad3p is catalytically inactive and that Tad2p is the catalytic subunit of the complex.

Deamination of tRNA is not limited to eukaryotes but has also been detected in *E. coli* extracts (Auxilien et al., 1996). In *E. coli* the only known A to I conversion in RNA is the deamination of $tRNA^{Arg2}$ at position 34. Based on sequence homology, we identified an *E. coli* protein homologous to Tad2p. Here, we report the identification and characterization of this enzyme which is necessary and sufficient to catalyze this editing reaction. tadA/ecADAT2 shows sequence homology to yeast Tad2p and Tad3p and is encoded by an essential gene, thus underlining the vital function of $I_{34}$. The identification of the first prokaryotic tRNA-specific adenosine deaminase also provides further insight into the evolution of the deaminase family of enzymes.

## 3.3 Results

**tadA acts specifically on $tRNA^{Arg2}$**

Database searches with the *S. cerevisiae* Tad2p sequence and the FASTA program (Pearson and Lipman, 1988) revealed that the *E. coli* open reading frame (ORF) yfhC is 34% identical with Tad2p (Fig. 3.1). To test whether yfhC indeed encodes an adenosine deaminase, the protein was overexpressed in *E. coli* with an N-terminal GST tag and a C-terminal $His_6$ tag and purified to apparent homogeneity (Materials and methods; Fig. 3.2 A and B).

Peak fractions of the final gel filtration column were tested for adenosine deaminase activity on *in vitro* transcribed and uniformally [$\alpha$-$^{33}$P]ATP-labelled $tRNA^{Arg2}$. The recombinant protein indeed converted adenosine to inosine (Fig. 3.2 C). This activity stemmed from the recombinant protein and not from an endogenous enzyme since GST purified from *E. coli* or eluates from $Ni^{2+}$

agarose incubated with extract from empty-vector transformed cells had no activity (results not shown). 2 ng recombinant yfhC protein formed up to 0.5 mol I/mol tRNA (Fig. 3.2 C, fraction 12). No other protein was needed for activity since fraction 12 which lacks the contaminant present in fractions six to eleven deaminated $tRNA^{Arg2}$ efficiently (Fig. 3.2 C, fraction 12). This is in contrast to the yeast enzyme, which acts as a heterodimer composed of Tad2p and Tad3p (Gerber and Keller, 1999).

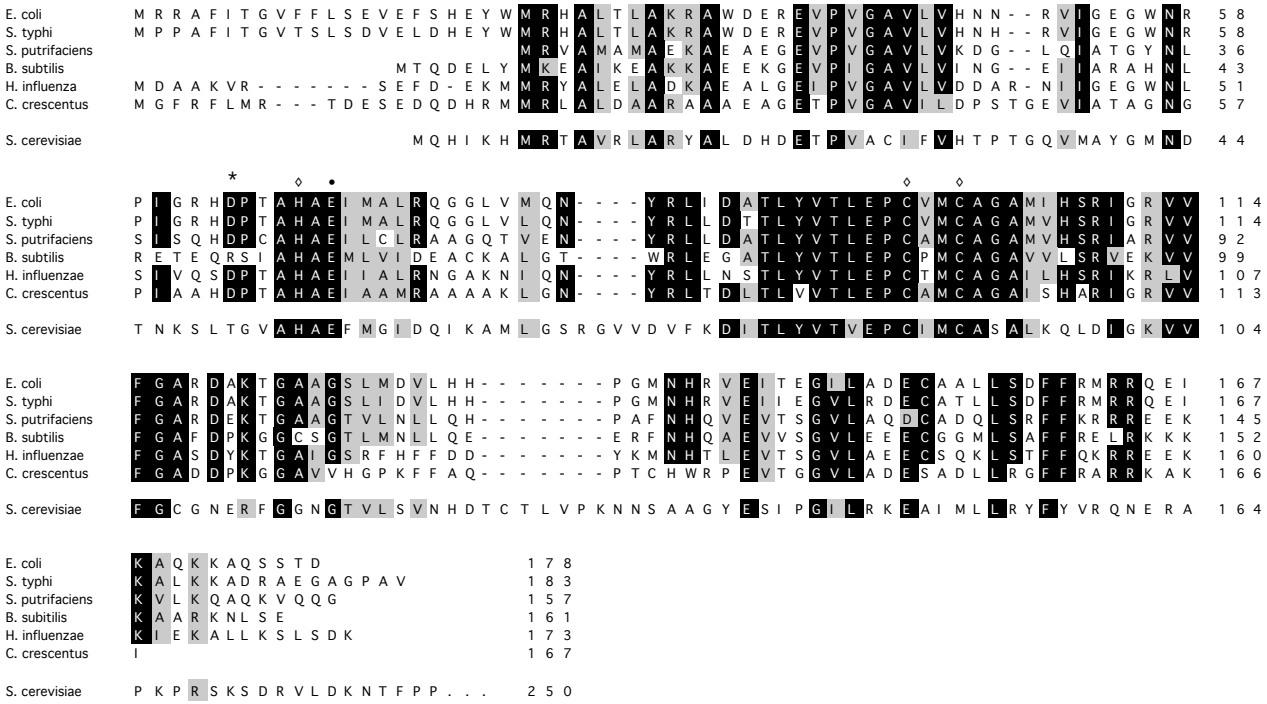



**Figure 3.1 Bacterial genomes encode a protein (tadA) related to Tad2p of *S. cerevisiae***

Multiple sequence alignment of *E. coli* tadA, yeast Tad2p and putative tadA sequences from different bacteria. Residues conserved in more than 83% of the proteins are shaded black, similar amino acids grey. The three putative $Zn^{2+}$-chelating residues (◊) and the glutamate thought to mediate proton transfer (•) are marked. The position of the point mutation (D64E) in the tadA mutant NWL37 is indicated by an asterisk. The alignment was generated with the ClustalW software at the European Bioinformatics Institute (Thompson et al., 1994).

To determine the site in $tRNA^{Arg2}$ that is modified by recombinant yfhC, *in vitro* modified tRNA was amplified by RT-PCR and sequenced. Since I basepairs with C during reverse transcription, A to I deamination changes the sequence from A to G. $tRNA^{Arg2}$ incubated with recombinant yfhC contained a G at position 34, demonstrating that the template-encoded A at this position was deaminated to I. $tRNA^{Arg2}$ incubated in the absence of protein carried A at position 34 and therefore was not modified (Fig. 3.2 D). No other A in $tRNA^{Arg2}$ was modified by the recombinant protein. These results showed that the yfhC gene encodes a protein which is

sufficient to reconstitute A34:tRNA deaminase activity *in vitro*. Therefore, we renamed yfhC to tadA (for tRNA adenosine deaminase A).

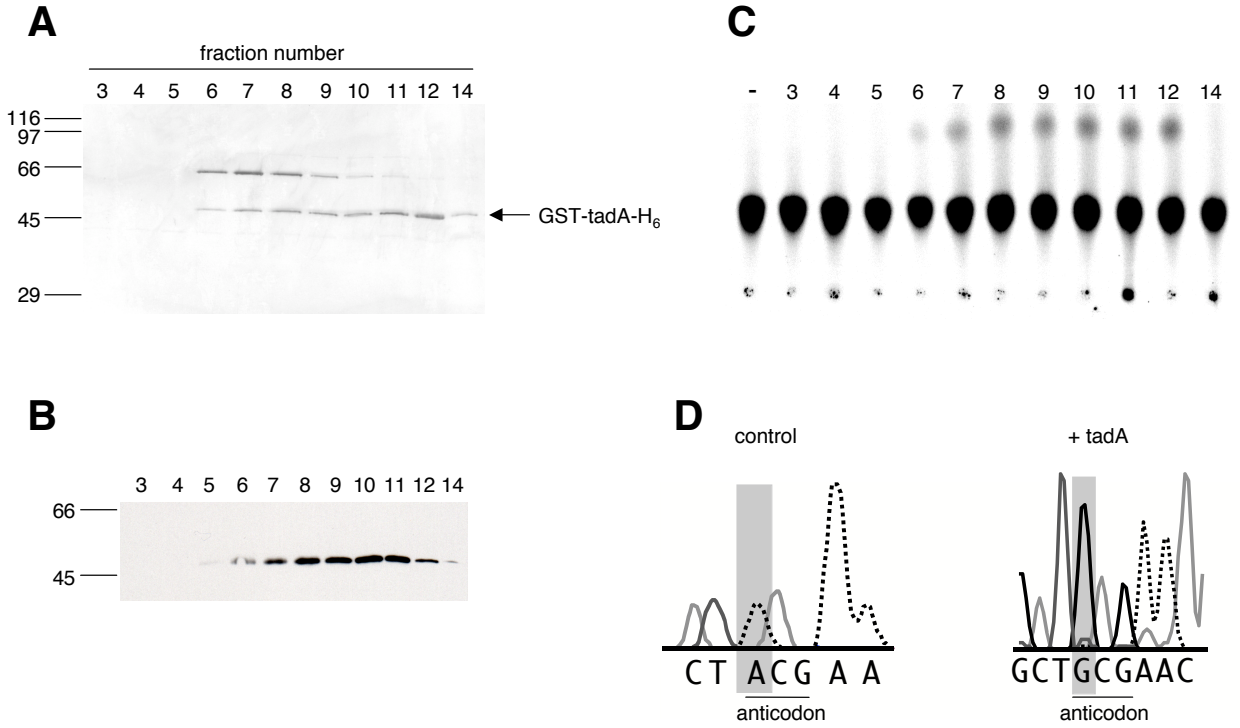



**Figure 3.2 tadA specifically deaminates $A_{34}$ in $tRNA^{Arg2}$**

SDS-polyacrylamide gel stained with silver (A), Western blot (B) and tRNA editing assay (C) of the final gel filtration column of the purification of recombinant tadA.

20 $\mu$l aliquots of each fraction were separated on a 12% SDS-polyacrylamide gel (A) and transferred to a nitrocellulose membrane for detection with antibodies (B). The Western blot was probed with a mouse □-GST monoclonal antibody. Fraction numbers are indicated on the top, the molecular masses of the size standards in kDa on the left.

C) 2 $\mu$l of the fractions indicated at the top were incubated with [$^{33}$P]ATP-labelled $tRNA^{Arg2}$ for 30 min at 37°C. Reactions were treated with P1 nuclease and the products separated by one-dimensional TLC. The positions of AMP, IMP and the origin are indicated on the right. In the reaction indicated with -, $tRNA^{Arg2}$ was incubated with buffer only.

D) Sequence analysis of *in vitro* modified $tRNA^{Arg2}$. The nucleotide sequence surrounding the anticodon is shown. tRNA from reactions shown in (C) was amplified by RT-PCR and sequenced. The nucleotide at position 34 is shaded grey, the anticodon is underlined. Control: $tRNA^{Arg2}$ incubated with buffer only.

The substrate specificity of recombinant tadA was tested by comparing its activity on various tRNA substrates (Fig. 3.3 A). In *E. coli* only one tRNA substrate is known which is deaminated at position 34 ($tRNA^{Arg2}$) in contrast to *S. cerevisiae* where seven different tRNAs are substrates for Tad2p/Tad3p (Auxilien et al., 1996; Gerber and Keller, 1999). Auxilien et al.

(1996) showed that not only tRNA$^{Arg2}$, but also a variant of tRNA$^{Asp}$ from *S. cerevisiae* carrying the anticodon loop of yeast tRNA$^{Arg}$ is deaminated *in vitro* upon incubation with extracts from *E. coli*. The latter observation could be reproduced with recombinant tadA which forms up to 0.8 mol I in this artificial tRNA, suggesting that the anticodon loop carries important recognition signals for the enzyme (Fig. 3. 3A, lane 4). This hypothesis is supported by the result that tadA deaminated yeast tRNA$^{Arg}$ with the same efficiency as *E. coli* tRNA$^{Arg2}$ (Fig. 3.3 A, lanes 10 and 11). These two tRNAs have a conserved anticodon loop sequence (Fig. 3.3 B). However, parts of the anticodon stem and the acceptor stem are also conserved between yeast and *E. coli* tRNA$^{Arg}$ and these nucleotides might also contribute to the recognition of the tRNA by tadA (Fig. 3.3 B).

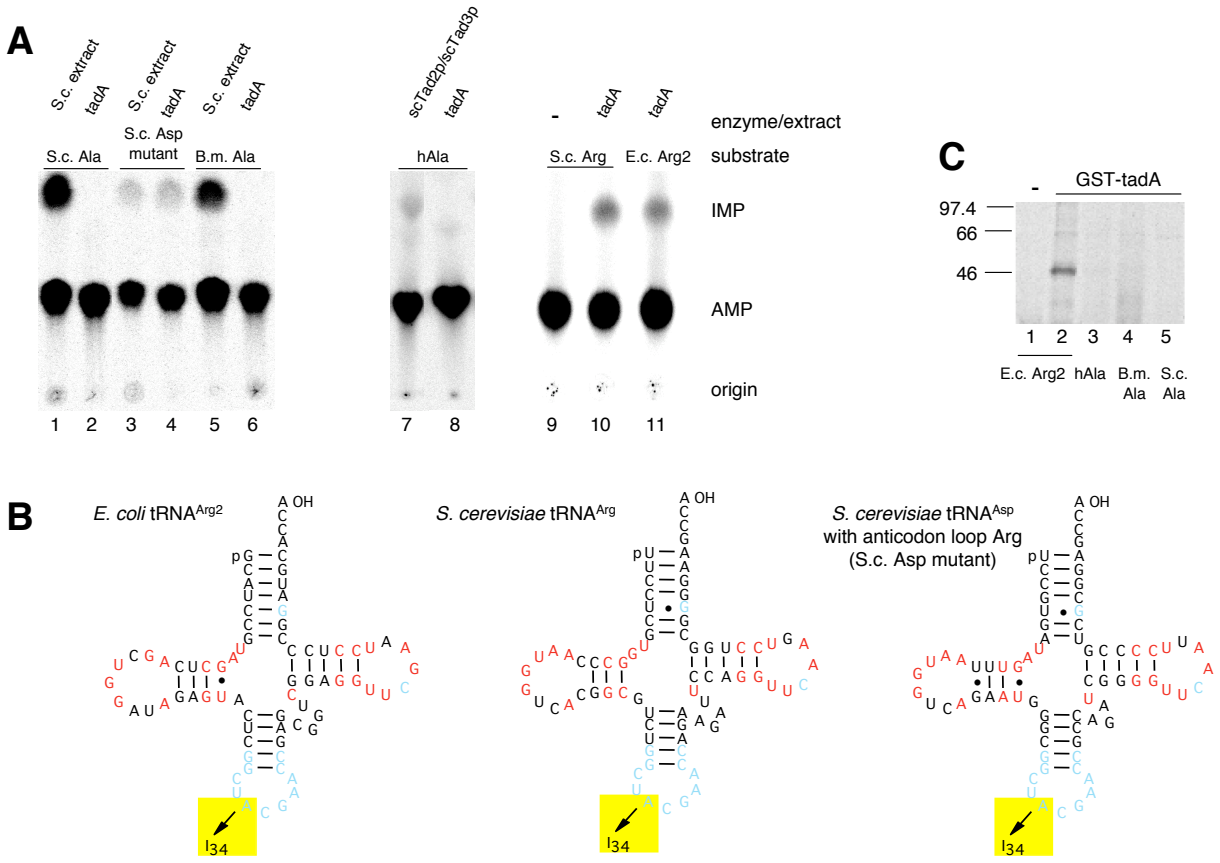



**Figure 3.3 Substrate specificity of tadA**

**A)** tRNA-editing assay with tadA and different tRNAs. tRNA substrates were incubated with either 2 ng recombinant $H_6$-tadA (lanes 2, 4, 6, 8, 10, 11), 20 ng recombinant scTad2p/scTad3p (lane 7) or 40 $\mu$g *S. cerevisiae* total protein (lanes 1, 3, 5). 100 fmol tRNA-Ala from *S. cerevisiae* (lanes 1, 2), a tRNA-Asp mutant from *S. cerevisiae* (lanes 3, 4), tRNA-Ala from *B. mori* (lanes 5, 6), human tRNA-Ala (lanes 7, 8), *S. cerevisiae* tRNA-Arg (lanes 9, 10) and tRNA-Arg2 from *E. coli* (lane 11) were used. All reactions were incubated for 1 h at 30°C and processed as described in the legend to Figure 2 (see also materials and methods).

**B)** Cloverleaf structure of *E. coli* tRNA$^{Arg2}$, *S. cerevisiae* tRNA$^{Arg}$ and *S. cerevisiae* tRNA$^{Asp}$ with the anticodon loop

of $tRNA^{Arg}$. Completely conserved nucleotides and nucleotides conserved as purines or pyrimidines between tRNAs from different species are shown in red (Klingler and Brutlag, 1993). Nucleotides that are conserved in addition between the three tRNAs are depicted in blue.

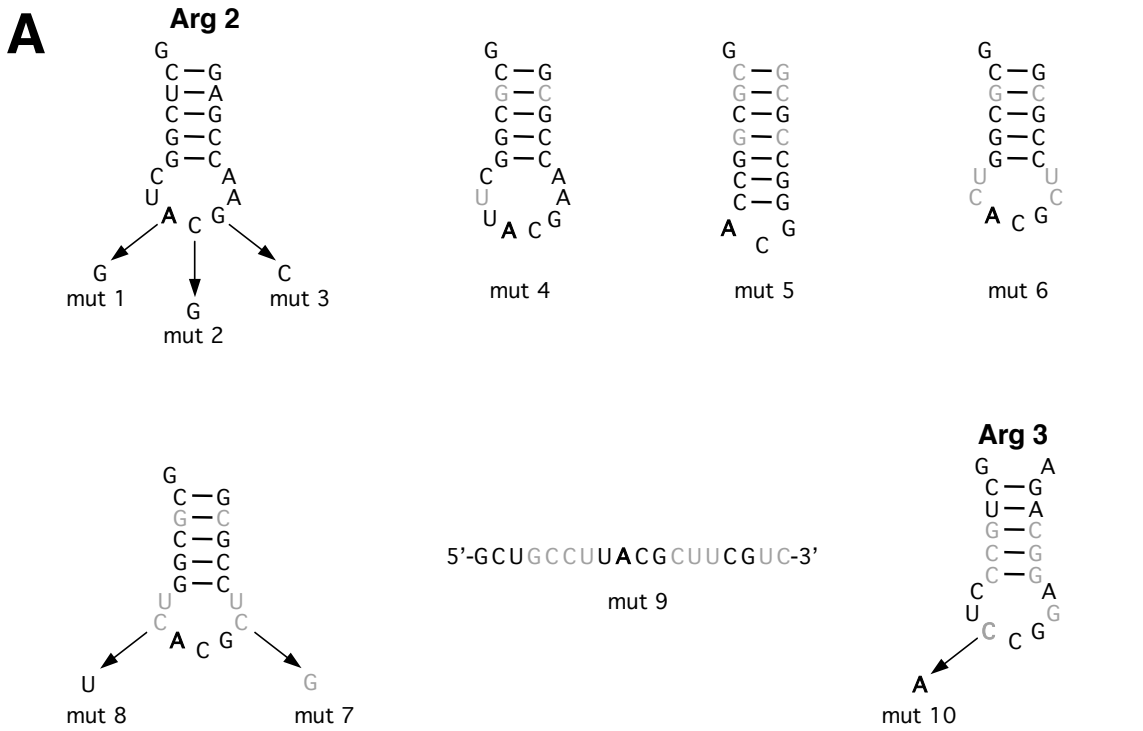
**C)** UV-crosslinking experiments of recombinant tadA and different tRNAs. 400 ng GST-tadA (lanes 2-5) and 100 fmol [$^{33}$P]ATP -labelled tRNA were irradiated and samples treated with RNaseA. Proteins were separated on a 12% SDS polyacrylamide gel and gels exposed to a PhosphorImager screen.

E.c. – *Escherichia coli*, S.c. – *Saccharomyces cerevisiae*, B.m. – *Bombyx mori*, Arg2 – $tRNA^{Arg2}$, hAla – human $tRNA^{Ala}$, B.m. Ala – *B. mori* $tRNA^{Ala}$, S.c. Ala – *S. cerevisiae* $tRNA^{Ala}$.

$tRNA^{Ala}$ from *S. cerevisiae*, *B. mori* and *H. sapiens* are not deaminated by tadA, although they are known to contain $I_{34}$ *in vivo* (Fig. 3.3 A, lanes 2, 6 and 8). Notably, all these tRNAs are substrates for yeast Tad2p/Tad3p (Fig. 3.3 A, lanes 1, 5 and 7; (Auxilien et al., 1996; Gerber and Keller, 1999). In yeast and *B. mori* $tRNA^{Ala}$ positions 34 and 37 are deaminated and 2 mol I/mol tRNA are generated upon incubation with yeast extract (Fig. 3.3 A, lanes 1 and 5). The substrate specificity of tadA could also be shown with UV-crosslinking experiments that indicate binding of the protein to the tRNA substrate. Recombinant GST-tadA could be UV-crosslinked to its natural substrate $tRNA^{Arg2}$ (Fig. 3.3 C, lane 2), but not to $tRNA^{Ala}$ from *H. sapiens*, *B. mori* and *S. cerevisiae* (Fig. 3.3 C, lanes 4, 5 and 6). GST alone could not be crosslinked to any of these substrates (results not shown). In summary, we showed that recombinant tadA specifically deaminates $tRNA^{Arg2}$ at position 34.

To determine whether the anticodon stem and loop is sufficient for deamination by tadA, a $tRNA^{Arg2}$ minisubstrate containing only the anticodon stem and loop was tested *in vitro* (Fig. 3.4 A). This minisubstrate was deaminated as efficiently as full-length $tRNA^{Arg2}$ (Fig. 3.4 B, lane 1 and table III). Quantification revealed that 1.2-1.4 mol I/mol tRNA is formed with WT Arg2 minisubstrate in different experiments, suggesting that an additional A is deaminated. Although the stem-loop of Arg2 serves as a substrate for tadA, there seem to be differences in site-specificity between the full-length $tRNA^{Arg2}$ and mini Arg2. Sequences and structures outside the anticodon arm might be important for deamination at the correct position. Interestingly, when A at the wobble position was mutated to G the minisubstrate was not deaminated, indicating that this nucleotide is essential for tadA recognition (Fig. 3.4 B, lane 2). In order to determine nucleotides that are essential to specify these short RNAs as substrates for tadA, several mutant minisubstrates were generated and tested *in vitro*. No I was formed with minisubstrates that have mutations at position 35 or 36 in the anticodon (Figure 4B, lanes 3 and 4). This showed that nucleotides in the anticodon were key determinants for tadA activity. The importance of the anticodon loop size was investigated with a minisubstrate that has a 8 nt loop. This substrate was still deaminated, but with a lower efficiency compared to the WT RNA (Figure 4B, lane 5 and

table I) whereas a 3 nt loop was not sufficient for tadA activity (Figure 4B, lane 6). A minisubstrate that has a random loop sequence but the correct anticodon was not deaminated, showing that the anticodon is not the only determinant for tadA activity (Figure 4B, lane 7).



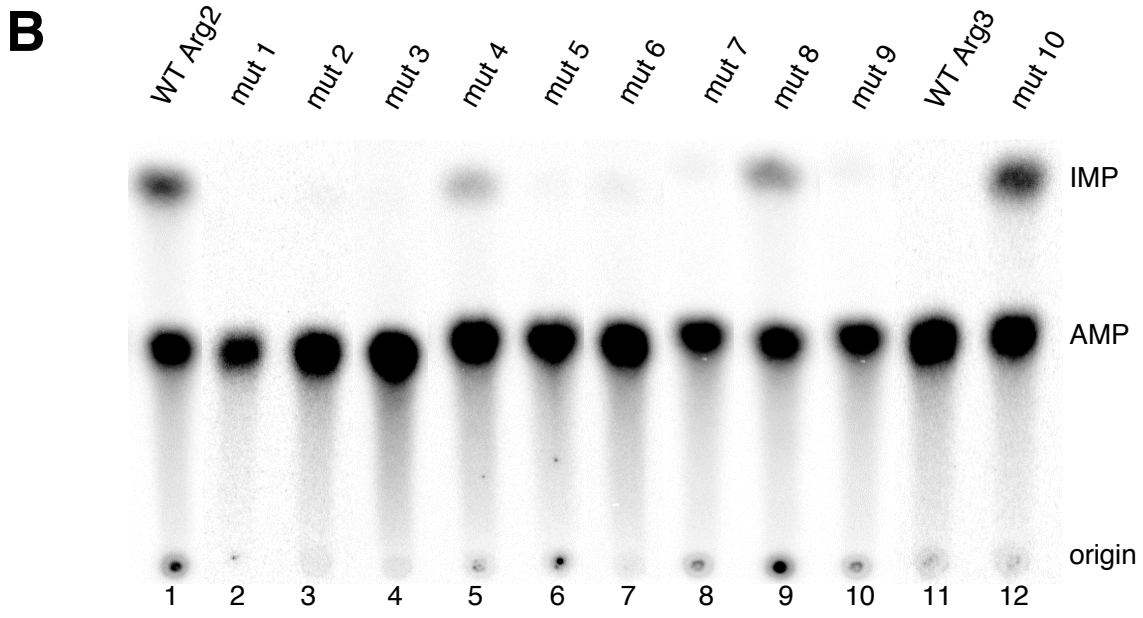



**Figure 3.4 RNA minisubstrates derived from *E. coli* $tRNA^{Arg2}$ and $tRNA^{Arg3}$ are sufficient for deamination by tadA *in vitro***

**A)** Schematic drawing of Arg minisubstrates that were tested *in vitro*. Nucleotides shown in grey are mutated compared to WT Arg2. A at the wobble position of the anticodon is shown in bold. Mutations in each minisubstrate are indicated by arrows.

**B)** Editing assay with minisubstrates and tadA. 25-100 fmol minisubstrate, 100 ng recombinant Flag-tadA-$H_6$ and 500 ng BSA were incubated in a total volume of 25 $\mu$l. WT Arg2 (lane 1), mutant 1 (lane 2), mutant 2 (lane 3), mutant 3 (lane 4), mutant 4 (lane 5), mutant 5 (lane 6), mutant 6 (lane 7), mutant 7 (lane 8), mutant 8 (lane 9), mutant 9 (lane 10), , WT Arg3 (lane 11) and mutant 10 (lane 12) were used. All reactions were incubated for 45 min at room temperature and were processed as described in the legend to Figure 2 (see also material and methods).

**Table III $tRNA^{Arg2}$ minisubstrates**

| substrate | mutations compared to WT | mol I/mol RNA |
|---|---|---|
| Arg2 WT | — | 1.2-1.4 |
| mut 1 | A34G | 0 |
| mut 2 | C35G | 0.10 |
| mut 3 | G36C | 0 |
| mut 4 | 8 nt loop, U28G, A42C | 0.50 |
| mut 5 | 3 nt loop, 7 bp G-C stem | 0 |
| mut 6 | U28G, C32U, U33C, A37C, A38U, A42C | 0 |
| mut 7 | U28G, C32U, U33C, A37G, A38U, A42C | 0 |
| mut 8 | U28G, C32U, A37C, A38U, A42C | 0.26 |
| mut 9 | linear | 0 |
| Arg3 WT | — | 0 |
| mut 10 | C34A | 1.39 |

When position 37 in mutant 6 was changed to G, I formation was not restored (Figure 4B, lane 8). G at position 37 in the WT Arg2 minisubstrate did not affect tadA activity (results not shown). However, reverting position 33 in mutant 6 to the wild-type U residue restored I formation which was, however, not as efficient as with WT minisubstrate (Figure 4B, lane 9 and table I). This suggests that $U_{33}$ is an important determinant for tadA activity, but additional nucleotides probably contribute to efficient deamination. To determine whether a stem-loop structure in addition to the sequence is required for tadA activity, a linear minisubstrate was analyzed (mutant 9, Figure 4A). Although mutant 9 contains the same essential nucleotides for activity as mutant 7, the linear RNA was not deaminated (Figure 4B, lane 10), indicating that a stem-loop structure is essential for tadA activity. As expected, the anticodon arm of *E. coli*

$tRNA^{Arg3}$ was not a substrate for tadA (Figure 4B, lane 11) because this RNA does not contain A at the wobble position (Figure 4A). However, I formation was detected in Arg3 when $C_{34}$ was mutated to A (Figure 4B, lane 12). This result provides further information about key nucleotides for tadA activity. The anticodon stem of Arg3 differs from that of Arg2, indicating that recognition of the stem by tadA is not sequence-specific (Figure 4A). In summary, a minisubstrate with a stem-loop structure is deaminated at the wobble position of the anticodon if the RNA has the sequence UACG within the loop. However, it cannot be excluded that other regions of the tRNA are needed for efficient deamination. The results with the mutant minisubstrates suggest that in addition to the correct sequence in the loop the structure and the loop size but not the stem sequence are important for deamination by tadA.

**A point mutation in tadA leads to an inactive enzyme *in vitro***

Next we asked whether tadA is responsible for inosine formation at position 34 of $tRNA^{Arg2}$ *in vivo* and took advantage of a non-lethal point mutant described previously (Poulsen et al., 1992). The mutant form of tadA carries a glutamic acid instead of an aspartic acid at position 64 (D64E; Fig. 3.1). This amino acid is highly conserved in bacterial, but not in eukaryotic Tad2 proteins.

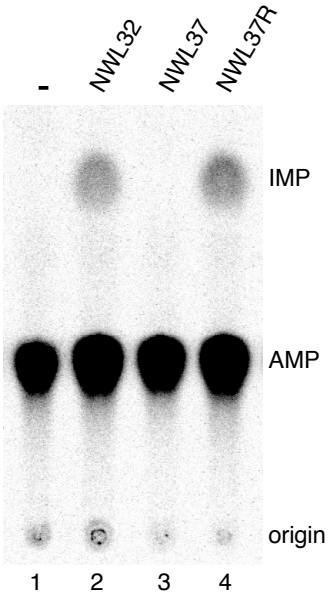



**Figure 3.5 The point mutation D64E abolishes tadA activity *in vitro*.**

Lane 1: no protein, lanes 2 to 4: 40 $\mu$g of total protein of *E. coli* extract. All reactions were incubated 2 h at 37°C and processed as described in the legend to Figure 2 (see also materials and methods).

To determine whether deamination of $A_{34}$ in $tRNA^{Arg2}$ is affected by this mutation *in vivo*, total RNA was isolated from the mutant (NWL37) and the corresponding parental strain (NWL32). After amplification of $tRNA^{Arg2}$ by RT-PCR with specific primers, the cDNAs were subcloned and multiple clones of each strain were sequenced. As expected, all clones derived from the wild-type strain contained G at position 34 (results not shown). Interestingly, 16 out of

17 clones derived from NWL37 also had G at position 34, and only one clone contained A at this position (results not shown). Therefore, this point mutant of tadA is almost fully active *in vivo,* which is not suprising given that the mutant strain showed no growth defect even at 42˚C (results not shown), although tadA is an essential gene.

Even so the mutant protein was functional in cells, its activity might be reduced in an *in vitro* assay system. We therefore prepared extracts from mutant and parental strains and tested equal amounts in a deamination assay. Whereas extracts from wild-type cells showed activity (NWL32; Fig. 3.5, lane 2), no inosine formation could be detected with extracts from the mutant strain (NWL37; Fig. 3.5, lane 3). However, activity could be restored by transforming the mutant cells with a tadA expression construct (NWL37R; Fig. 3.5, lane 4). Thus, the mutation D64E abolishes the enzymatic activity of tadA *in vitro*, although it does not detectably affect its activity in cells.

**tadA cannot substitute for yeast Tad2p *in vivo* and *in vitro***

Because tadA is 34% identical with yeast Tad2p and is functionally conserved, we attempted to replace Tad2p with tadA in a Δ*TAD2* yeast strain (Gerber and Keller, 1999). tadA was cloned into the pGALΔTrp-FLIS6 vector (Gerber et al., 1998) and transformed into a Δ*tad2* yeast strain carrying a copy of *TAD2* on a plasmid with the URA3 marker (pFL38; Gerber and Keller, 1999). Cells were then transferred to plates containing FOA and galactose to force the loss of the URA3-marked *TAD2* plasmid and induce the expression of tadA, respectively. No cells grew showing that tadA cannot substitute for Tad2p *in vivo*.

Next we tested whether tadA could functionally replace Tad2p in an *in vitro* deaminase assay. No activity was detected when tadA preincubated with yeast Tad3p was incubated with *B. mori* $tRNA^{Ala}$, which is a substrate for the Tad2p/Tad3p heterodimer (Fig. 3.6 A, lanes 4 and 2). Thus, tadA can neither replace yeast Tad2p *in vivo* nor *in vitro*.

**tadA forms homodimers**

Because yeast Tad2p functions as a heterodimer with Tad3p and *E. coli* does not encode a Tad3-like protein, we wondered whether tadA might form homodimers instead. This hypothesis was tested with three different approaches. We investigated the interaction of Flag-tadA and *in vitro* translated $^{35}$S-methionine-labelled tadA in a Flag pull-down assay. The tagged tadA and labelled translated tadA were preincubated, bound to anti-Flag agarose, washed and eluted by boiling the matrix in SDS-gel loading buffer. The result is shown in Fig. 3.6 B (lane 2). Approximately 10% of the tadA input was pulled-down by Flag-tadA. In a second experiment, recombinant Flag-

tadA was chromatographed on a gel filtration column which separates proteins according to size and fractions were tested for tadA activity. 50% of tadA eluted at a position corresponding to a dimer and about 50% of tadA eluted as monomer (results not shown). This result suggests that tadA may act as a homodimer. However, elution of tadA monomers indicates that tadA does not interact as strongly with itself as does the yeast Tad2p with Tad3p (Gerber and Keller, 1999). Formation of tadA monomers could occur by separation of tadA dimers on the gel filtration column due to the experimental conditions.

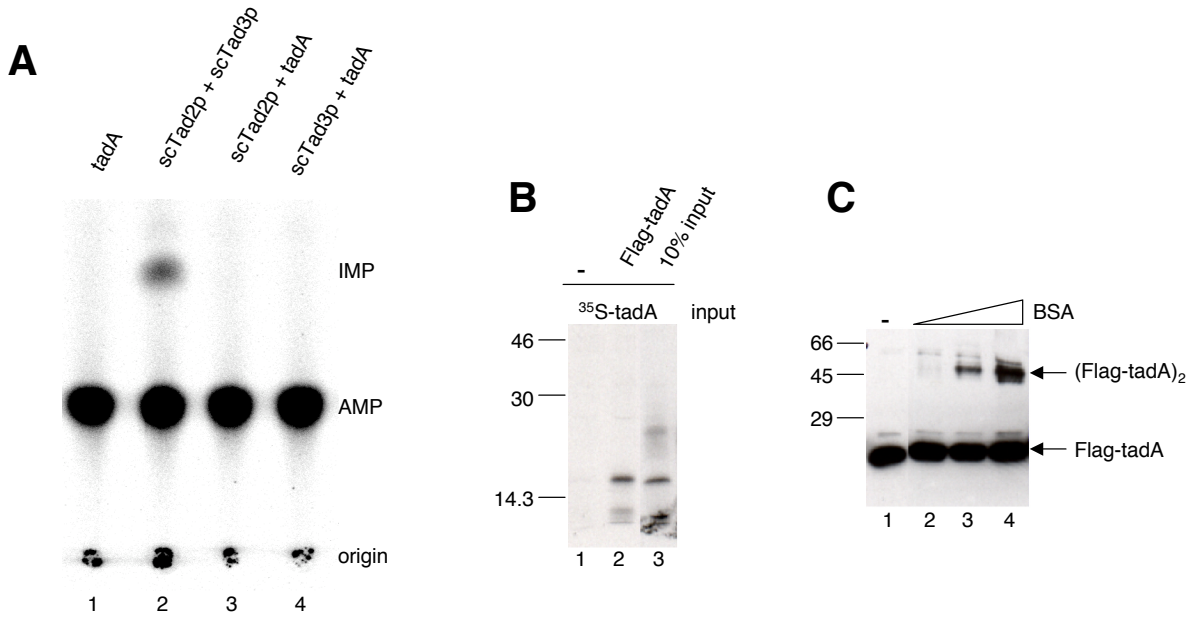



**Figure 3.6 tadA forms homodimers and cannot replace yeast Tad2p *in vitro***

**A)** Recombinant yeast Tad2p, Tad3p and *E. coli* tadA were preincubated as indicated at the top and tested for activity *in vitro* with $tRNA^{Ala}$ from *B. mori*. As a control the yeast Tad2p/Tad3p complex was used (lane 2).

**B)** Flag pull-down assay with *in vitro* translated tadA. $^{35}$S-labelled tadA was incubated with Flag-tagged tadA (lane 2) or buffer (lane 1) and subsequently bound to Flag agarose and eluted. 10% of the input is shown in lane 3. Proteins were separated on a 12% SDS polyacrylamide gel. Molecular masses of the protein standards are indicated in kDa on the left.

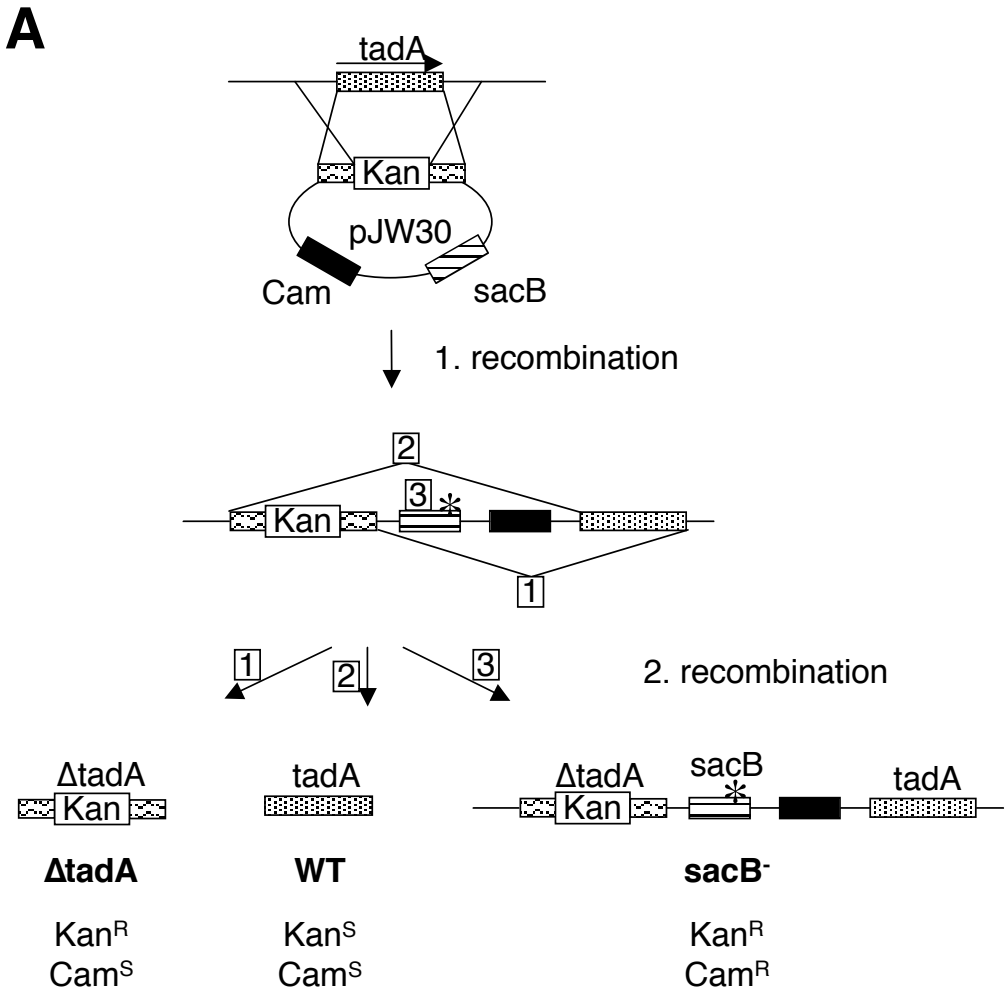
**C)** Chemical crosslinking of recombinant tadA. 300 ng tadA (lanes 2-4) were incubated in a reaction mixture containing tris(2,2'-bipyridyl)dichlororuthenium(II)hexahydrate, APS and BSA (90 ng, 0.9 $\mu$g, 9 $\mu$g). tadA incubated with buffer only is shown in lane 1. Reaction products were separated on a 12% SDS polyacrylamide gel and blotted to a nitrocellulose membrane. tadA was detected with a mouse anti-Flag antibody. Molecular masses of protein standards are indicated in kDa on the left.

To confirm dimerization, we carried out chemical cross-linking experiments and indeed observed a complex corresponding in size to tadA homodimers (Fig. 3.6 C, lanes 2-4). Complex

formation was strongly stimulated by the addition of BSA although a weak cross-link was also detectable in some experiments in the absence of BSA (results not shown). The reason for this stimulation is unclear, however BSA might have a stabilizing effect on the interaction of the highly purified tadA subunits. tadA complex formation was specific and only involved tadA subunits but not BSA. Complexes with BSA would have a higher molecular mass and can therefore easily be distinguished from tadA homodimers by size. Dimerization was not complete and could be observed with about 30% of tadA. This might be due to the experimental procedure where tadA might have been crosslinked under suboptimal dimerization conditions. The addition of *E. coli* total tRNA did not influence dimer formation in these assays, indicating that the observed dimerization did not require substrate tRNA (results not shown). These experiments suggest that tadA can form a homodimer *in vitro*, however it remains to be shown whether tadA acts as a homodimer *in vivo*.

**tadA is essential for cell viability**

tadA has previously been reported to be a non-essential gene in *E. coli* (Poulsen et al., 1992). tadA was deleted by inserting a chloramphenicol acetyltransferase (cat) gene into the SphI restriction site of the ORF which removed the last 17 amino acids of the protein. We reinvestigated whether tadA is essential for viability because recombinant tadA-ΔC, corresponding to the knock-out gene described by Poulsen et al. (1992) was active in a tRNA editing assay (results not shown). The tadA gene was inactivated with the two-step *sacB* counterselection technique (Reece and Phillips, 1995). For this purpose a plasmid-borne tadA gene was produced which was disrupted by an interposon carrying a kanamycin-resistance cassette (pJW30; Fig. 3.7 A). The disrupted tadA cassette was introduced at the chromosomal tadA locus by homologous recombination. Thus, the first recombination event resulted in the introduction of a kanamycin-disrupted tadA gene and plasmid sequences into the genome next to wildtype tadA (Fig. 3.7 A). Subsequent growth of these strains on 5% sucrose selected for excision of the *sacB* gene, which confers sucrose sensitivity, by a second recombination event. Sucrose-resistant clones were then analyzed for kanamycin and plasmid encoded chloramphenicol resistance. If the wild-type copy of tadA can be lost because the gene is not essential, one expects to find similar numbers of WT tadA and Δ*tadA* clones (Fig. 3.7 A). However, all of the 400 clones analyzed were $Kan^S/Cam^S$ and thus wild-type cells, indicating that the disrupted form of tadA was lost selectively and that tadA is an essential gene (Fig. 3.7 B).




**B**

| | Total clones | 1 ΔtadA | 2 WT | 3 sacB$^-$ |
|---|---|---|---|---|
| - | 400 | 0% | 100% | 0% |
| pJW117 | 200 | 47% | 49% | 4% |

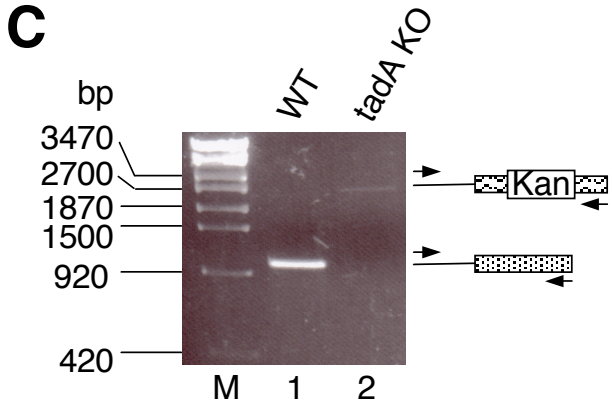



**Figure 3.7 tadA is an essential gene**

**A)** The sacB counterselection procedure. The wildtype chromosomal copy of tadA was exchanged via a two-step homologous recombination reaction with a gene copy disrupted by the insertion of a kanamycin-resistance cassette. After the second recombination step, colonies were analyzed by their antibiotic-resistance profile for the presence or absence of the Kan-cassette in the tadA locus and for mutational inactivation of the *sacB* gene. The mutation in sacB is indicated by an asterisk.

**B)** Summary of the clones obtained by the strategy described in A. The total number of sucrose-resistant colonies

analyzed are indicated in the first column (Total clones). The number of isolated colonies with a disrupted tadA (ΔtadA) gene in the chromosome is shown in the second column, the number of colonies with a restored wildtype tadA locus (WT) after the second recombination step is shown in the third column. Colonies that had acquired sucrose-resistance through *sacB* inactivation are listed in the fourth column (*sacB*$^-$). Plasmids pJW30 and pJW117 carry disrupted and wild-type copies of the tadA ORF, respectively.

**C)** PCR analysis of WT (lane 1) and Δ*tadA* (lane 2) genomic *E. coli* DNA. Primers are indicated by arrows and anneal 500 bp upstream of the tadA ORF and at the 3' end of the ORF, respectively. Sizes of standards (lane M) are indicated in bp on the left.

To corroborate this, we repeated the tadA deletion experiment in the presence of a functional copy of tadA under the control of the arabinose promoter ($P_{BAD}$). In this case, about equal numbers of $Kan^R/Cam^S$ (e.g. Δ*tadA*, 47%) and $Kan^S/Cam^S$ (e.g. WT, 49%) clones were obtained (Fig. 3.7 B). The deletion of tadA was confirmed by PCR analysis of seven of the $Kan^R/Cam^S$ clones (Fig. 3.7 C). PCR analysis of genomic DNA from wild-type cells led to the accumulation of a fragment of 1000 bp (Fig. 3.7 C, lane 1) corresponding to tadA+500 bp upstream sequence. This product was not formed with genomic DNA from a Δ*tadA* strain. Instead, a faint band was observed at ~ 2500 bp, which corresponds to tadA disrupted by the kanamycin cassette and 500 bp upstream sequence (Fig. 3.7 C, lane 2). Note that the primers used are specific for the amplification of the genomic tadA locus whereas the rescue copy of tadA is not amplified. In addition, the deletion was confirmed by Southern blot analysis of three individual clones (results not shown). These results indicate that disruption of the WT chromosomal copy of tadA resulted in viable *E. coli* only when an additional plasmid-born tadA gene was also present. The same results were obtained when these gene disruption experiments were performed with *Caulobacter crescentus* (results not shown). Thus, we conclude that tadA is essential for viability of eubacterial organisms.

## 3.4 Discussion

We describe the cloning and functional characterization of the *E. coli* enzyme tadA/ecADAT2. Based on the following observations we conclude that tadA is an adenosine deaminase that specifically converts $A_{34}$ to $I_{34}$ in bacterial $tRNA^{Arg2}$: (i) Recombinant tadA purified to homogeneity specifically deaminates $tRNA^{Arg2}$ at the wobble position of the anticodon and does not need a cofactor. (ii) The recombinant protein is highly specific for the wobble position of

$tRNA^{Arg2}$. However, eukaryotic tRNAs are deaminated by tadA if they contain the anticodon loop of $tRNA^{Arg}$ from yeast or *E. coli*. (iii) Recombinant tadA forms homodimers. A to I deamination at the wobble position of the anticodon is conserved from bacteria to mammals and the enzymes catalyzing the reaction are also conserved not only in their sequence but probably also in their property to act as dimers. Thus, the tadA homodimers may represent the ancestor of the eukaryotic Tad2p/Tad3p heterodimer. (iv) tadA is encoded by an essential gene. The fact that *E. coli* cells are not viable without tadA underscores the important role of $I_{34}$ in tRNA. There are two major differences between tRNA deamination in *E. coli* and *S. cerevisiae*. First, $tRNA^{Arg2}$ is the only *E. coli* tRNA that has A at position 34 and is therefore the only tRNA target for tadA whereas yeast has seven tRNAs containing $A_{34}$. Second, tadA is sufficient to catalyze the reaction in *E. coli*, whereas in yeast the Tad2p/Tad3p heterodimer is needed to modifiy $A_{34}$ to $I_{34}$ in tRNAs (Gerber and Keller, 1999).

The sequences of tadA and the yeast tRNA deaminase subunit Tad2p were compared to predicted Tad2 proteins from other prokaryotic organisms (Fig. 3.1). Comparing these sequences with those of cytidine deaminases revealed that they share conserved motifs which are typical for cytidine nucleotide deaminases rather than adenosine deaminases (Gerber and Keller, 1999). On the basis of these observations we propose that tadA is the ancestor of the eukaryotic Tad2/Tad3 proteins. Gerber and Keller (1999) suggested that Tad2 and Tad3 are paralogs that appeared by gene duplication after the divergence of prokaryotes and eukaryotes. This hypothesis is now further supported by the identification of tadA which acts as a single polypeptide. Prokaryotic organisms encode only Tad2-like proteins, thus the appearance of Tad3 is a later event which most probably happened after the divergence of prokaryotes and eukaryotes.

I-containing tRNAs are predicted to recognize codons ending in C, U and A. Inosine at the wobble position allows a single tRNA to read three different codons (NNU, NNC, NNA) which therefore must encode the same amino acid (Crick, 1966; Jukes, 1973). tRNAs with inosine at the wobble position are present in all bacteria and eukarya for which representative sets of tRNAs have been characterized (Sprinzl et al., 1998). However, Percudani (2001) suggested that I-containing tRNAs might not contribute significantly to the decoding of NNA codons in higher eukaryotes because there is always a synonymous U-starting anticodon which can read codons ending in A without wobbling. Because prokaryotes have fewer tRNA species, I is also used to decode A. $tRNA^{Arg2}$ is the only tRNA which decodes the CGA arginine codon. CGA is a very rare codon with a frequency of only 0.36 percent in *E. coli* (http://www.kazusa.or.jp/codon/) and might be specifically avoided to encode arginine residues because of its poor translational properties (Curran, 1995). The possibilities of codon-anticodon

basepairing are greatly improved with inosine at the wobble position in tRNAs. Such advantages might have been important for the survival of cells under selective conditions and might have promoted the spreading of this modification and the responsible enzymes because the total number of tRNA genes in an organism can be reduced with tRNAs containing I at position 34.

An *E. coli* strain has been described that has a point mutation in the tadA protein (formerly designated yfhC; Poulsen et al., 1992). This mutation resulted in transient resistance to the overexpression of the membrane protein gef which has a cell-killing function (reviewed in Gerdes et al., 1990). It was speculated that the *gef* genes are involved in plasmid maintenance by killing newborn plasmid-free cells (Gerdes et al., 1986). It is not known how the cells are killed by the proteins of the *gef* gene family or why a mutation in tadA leads to resistance against the gef proteins. We found that extracts from strain NWL32 showed I formation, whereas extracts from the tadA mutant strain NWL37 had no deamination activity. Deamination activity could be rescued by expressing tadA from a plasmid. The absence of activity in NWL37 extracts is probably due to instability of the mutated tadA protein which might be inactivated during extract preparation. tadA might also be stabilized by interactions with other tRNA modifying enzymes and these interactions could be impaired in the mutant protein. This hypothesis is supported by the observation that the mutated amino acid in NWL37 is highly conserved in bacterial tadA sequences and therefore seems to have an important function (Fig. 3.1). NWL37 cells showed no growth defect at 42°C, indicating that the cells did not have a major disadvantage with the mutation. In agreement with this proposal we could not detect a deficiency in tRNA deamination in strain NWL37 *in vivo*. Consequently tadA is active despite the mutation and this explains the presence of modified tRNAs in the mutant strain. This result also explains the wildtype growth phenotype of the mutant strain since a tadA mutant with significantly reduced activity *in vivo* would be expected to be lethal. It is not known why a mutation in the tadA gene makes the cells resistant to gef, but Poulsen et al. (1992) suggest that there is a second mutation in the genome of NWL37 because some *E. coli* strains were only transiently resistant to gef when the mutated tadA was introduced by P1 transduction. This other mutation has not been identified, but could be responsible for the gef resistance.

In yeast seven different tRNAs contain $I_{34}$. The Tad2p/Tad3p complex probably deaminates all these tRNAs, however not all of them have been tested *in vitro* with recombinant proteins. tadA is not able to replace yeast Tad2p neither *in vitro* nor *in vivo*. However, the yeast Tad2p/Tad3p proteins are capable of deaminating $tRNA^{Arg2}$ from *E. coli in vitro* (Auxilien et al., 1996). tadA deaminates yeast $tRNA^{Arg}$ and a hybrid tRNA which consists of the yeast $tRNA^{Asp}$ with the yeast $tRNA^{Arg}$ anticodon loop, indicating that this tRNA region is crucial for tadA to

recognize and deaminate the substrate. The importance of the anticodon loop for recognition by tadA was also shown with $tRNA^{Arg}$ minisubstrates. Mutations in the anticodon of these minisubstrates completely abolished tadA activity, whereas mutations in the stem or at position 37 (3' adjacent to the anticodon) had no effect on I formation. Reduced loop size and increased stem length resulted in less efficient deamination, which nevertheless still ocurred site-specifically. In summary, the results suggest that the stem-loop structure of the anticodon arm and nucleotides in the anticodon and at position 33 are key determinants for tadA activity. Interestingly, it has been shown previously that tRNA modification or processing enzymes need only few determinants in the tRNA sequence or structure to specifically catalyze a reaction. Work with tRNA synthetases has shown that a variety of minisubstrates can be used for aminoacylation (reviewed in Musier-Forsyth and Schimmel, 1993; Schimmel and Alexander, 1998). Synthetic substrates for specific aminoacylation can be as short as a tetraloop with a four base pair stem and a CCA end. For alanyl-tRNA synthetase the critical determinant for aminoacylation of $tRNA^{Ala}$ was shown to be the G3:U70 base pair and this was also true for specific aminoacylation of the minihelix (Francklyn and Schimmel, 1989). Furthermore, it has been shown that RNase P, the enzyme that forms the mature 5' ends of tRNAs by cleaving off the leader sequence, can act on minisubstrates that consist only of a three basepair stem with very short extensions at the 5' and 3' ends (reviewed in Altman and Kirsebom, 1999). Another example is provided by the spliceosomal 15.5 kDa protein. This protein specifically interacts with the 5' stem-loop of U4 snRNA and plays an important role in the late stage of spliceosome assembly (Nottrott et al., 1999). The crystal structure of the 15.5 kDa protein revealed that it forms hydrogen bonds with only five bases of the U4 snRNA loop (Vidovic et al., 2000).

It is not known whether the 3D architecture of the tRNA is important for efficient I formation by tadA. It cannot be excluded that other nucleotides in addition to the anticodon loop of the tRNA are important for tadA binding. Further experiments will be required to show where tadA binds to the tRNA and which nucleotides are contacted by the enzyme. It also remains to be shown which sequences in Tad proteins are needed for tRNA binding since none of the enzymes described so far contains a known RNA binding motif. The high substrate specificity of tadA offers a potential explanation why only one polypeptide is required. A single protein might suffice to modify $tRNA^{Arg2}$ in bacteria whereas the recognition of seven different tRNAs in yeast might require a more elaborate mechanism.

## 3.5 Materials and methods

### Oligonucleotides

| | |
|---|---|
| yfhC.1 | 5'-GGCGGATCCCGCCGCGCTTTTATAACCGG-3' |
| yfhC.2 | 5'-CCCAAGCTTAATCCGTCGAGGATTGCGC-3' |
| yfhC.3 | 5'-CCCACTAGTCATGCGCCGCGCTTTTATAACCGG-3' |
| yfhC.4 | 5'-GGGACTAGTGCATCCGTCGAGGATTGCGC-3' |
| yfhC.5 | 5'-CCCCATATGCGCCGCGCTTTTATAACCGG-3' |
| yfhC.6 | 5'-CCCCATATGCTAGTGATGGTGATGGTGATGATCCGTCGAGGATTGCGC-3' |
| yfhC.7 | 5'-CCCGCTAGCAGAGTAAAGCATAGCGTCCGT-3' |
| yfhC.8 | 5'-CCCCTCGAGTTAATCCGTCGAGGATTGCGC-3' |
| yfhC.9 | 5'-CTCCGGTTATAAAAGCGCGGV-3' |
| Arg1 | 5'-GACACGGTACCACACAACTGGTGCATCCGGGAGGATTCG-3' |
| Arg2 | 5'-GACACGGTACCACACAACTGG-3' |
| Arg3 | 5'-GCATCCGTAGCTCAGCTGG-3' |
| tadA.KO3 | 5'-CCCGCGGCCGCTCCAGAGGAGTATGGTTTGGC-3' |
| tadA.KO4 | 5'-CCCGCGGCCGCGCCGTCGATGCGGTACTGCCGC-3' |
| araC | 5'-TTCATACTCCCGCCATTCAG-3' |

Restriction sites are underlined and nucleotides encoding 6 histidine residues are double underlined.

### Bacterial strains and media

The strains used in this study are described in table IV. *E. coli* strains were grown at 37˚C or 24˚C in LB or 2xYT medium (Sambrook et al., 1989) supplemented with ampicillin (100 $\mu$g/$\mu$l), kanamycin (30 $\mu$g/$\mu$l) or chloramphenicol (20 $\mu$g|$\mu$l) as necessary. *C. crescentus* strains were grown in PYE medium (Ely, 1991), supplemented with nalidixic acid (20 $\mu$g/ml), kanamycin (20 $\mu$g/ml), tetracyclin (2 $\mu$g/ml), spectinomycin (50 $\mu$g/ml) or streptomycin (30 $\mu$g/ml) as needed.

### Table IV Plasmids and strains

| Plasmid | Description | Reference |
|---|---|---|
| pGALΔTrp-Flis6 | Gal promoter, Trp-cassette removed, encodes Flag- and $H_6$-tag | Gerber et al. (1998) |
| pGDV1 | pGEX derivative with TEV cleavage site, no thrombin site | Dichtl and Keller (2001) |

| | | |
|---|---|---|
| pKO3 | *E. coli* gene replacement vector | Link et al. (1997) |
| pRDC15 | pKO3 derivate allowing chromosomal replacement of araBAD genes | Arigoni et al. (1998) |
| pJW24 | tadA in GDV1 via NdeI-XhoI | this work |
| pJW27 | tadA in pQE9 via BamHI-HindIII (Qiagen) | this work |
| pJW28 | tadA in pTrc-FLIS$_6$ via SpeI | this work |
| pJW30 | tadA with 700 nt overhangs, Kan-cassette inserted into tadA | this work |
| pJW106 | tadA in pBS-KS via BamHI/HindIII | this work |
| pJW117 | tadA in pRDC15 via NheI/XhoI | this work |
| pJW118 | tadA in pGALΔTrp-Flis6 via SpeI | this work |

| Strain | Description | Reference |
|---|---|---|
| MG1655 | *E. coli* WT strain | Neidhardt (1996) |
| JW9 | MG1655, P$_{BAD}$-tadA | this work |
| JW33 | MG1655, P$_{BAD}$-tadA, Δ*tadA* | this work |
| NWL32 | *gef-lacZ*, Δ*relF*, Cam$^R$, Kan$^R$ | Poulsen et al. (1992) |
| NWL37 | *gef-lacZ*, Δ*relF*, Cam$^R$, Kan$^R$ | Poulsen et al. (1992) |

**Cloning of the tadA gene**

The ORF yfhC was amplified by PCR from 25 ng *E. coli* genomic DNA with primer pairs yfhC.1/yfhC.2, yfhC.3/yfhC.4, yfhC.5/yfhC.6. The primers anneal at the 5' and the 3' end of the tadA ORF, respectively. The ORF was cloned into pCR2.1 (Invitrogen) and several independent clones were sequenced. tadA was then subcloned in the appropriate restriction sites of either pQE9, pTrc-FLIS$_6$ or pGDV1.

**Expression of tadA in *E. coli* and purification of the recombinant protein**

The procedure is described for the tadA-pGDV1 construct. The tadA-pGDV1 construct was transformed into BL21 cells (Invitrogen). 2 liters of 2xYT medium supplemented with 100 $\mu$g/l ampicillin and 10 $\mu$M $ZnCl_2$ were inoculated with 28 ml of an overnight preculture. Cultures were grown at 37˚C to an $OD_{600}$ of 0.94 and then induced by adding IPTG to a final concentration of 0.4 mM. Incubation was continued at 24˚C for 5 hours. Cells were harvested by centrifugation and the pellet was resuspended in 40 ml lysis buffer (50 mM Tris-HCl pH 8.0, 200 mM KCl, 10% glycerol, 0.02% NP40, 2 mM ß-mercaptoethanol, 0.5 mM phenylmethylsulfonyl fluoride, 0.5 $\mu$g/ml leupeptin, 0.7 $\mu$g/l pepstatin). Cells were lysed by sonication 3x3 min on ice. After centrifugation, the cleared lysate was mixed with $Ni^{2+}$-NTA agarose previously equilibrated with lysis buffer. The mixture was stirred at 4˚C for 45 min and then loaded on the column. The $Ni^{2+}$-agarose was washed with 100 ml wash buffer (lysis buffer + 10 mM

imidazole) and bound proteins were eluted with 250 mM imidazole. Fractions 3-5 were applied to glutathione sepharose (Amersham Pharmacia) equilibrated with buffer A (Tris-HCl pH 8.0, 130 mM KCl, 10% glycerol, 0.02% NP40, 1 mM DTT). Proteins were eluted with buffer B (buffer A + 10 mM reduced glutathione). 50 $\mu$l of fraction 1 were loaded on a gel filtration column (Superdex 200 PC3.2/30, Smart System, Pharmacia) that had been equilibrated with buffer A.

**Construction of an *E. coli* tadA null mutant**

The *E. coli* tadA gene was disrupted by the insertion of an omega cassette (Prentki and Krisch, 1984) via homologous recombination. A 1.9 kb NotI fragment containing 700 nt upstream and downstream flanking regions of the tadA ORF was cloned into pCR2.1 (Invitrogen). The omega cassette was then inserted in the unique EcoNI restriction site in the tadA ORF. pKO3 contains two selection markers for the recombination step: a kanamycin resistance cassette and a *sacB* gene conferring sucrose sensitivity. pJW30 was integrated into the chromosome of *E. coli* wild-type strain MG1655 via homologous recombination. The resulting strain contained a wild-type and a disrupted copy of tadA. To induce the second recombination event, colonies were serially diluted in LB and plated on LB plates containing 5% sucrose. Colonies that were able to grow on sucrose were tested for loss of the plasmid encoded chloramphenicol resistance and for presence of the omega cassette encoding a kanamycin resistance. Clones that were sucrose- and chloramphenicol-resistant had not lost the plasmid and were probably resistant to sucrose due to a mutational inactivation of the *sacB* gene. Kanamycin- and chloramphenicol-sensitive clones had lost the plasmid and recombination left a wild-type tadA in the chromosome. Null mutants of tadA should be Kan-resistant and Cam-sensitive.

For the rescue construct for tadA null mutants, the tadA ORF + 26 nucleotides upstream flanking sequence were amplified by PCR from *E. coli* genomic DNA with primers yfhC.7/yfhC.8. The fragment was cloned into the NheI/XhoI restriction sites of pRDC15 generating construct pJW117. MG1655 cells were transformed with pJW117 and selected for the integration of the plasmid at 43˚C. Integration at the correct locus was verified by PCR with primers araC and yfhC.9. Then the same antibiotic resistance analysis as described above was performed.

***In vitro* transcription of tRNAs and enzyme assays**

*In vitro* transcription of tRNAs was performed essentially as described previously by Gerber et al. (1998). For transcription of minisubstrates T7 primer and substrate primer were mixed in a

1.2:1 ratio in 10 mM Tris pH 8 with a concentration of 20 pmol/$\mu$l substrate primer and 24 pmol/$\mu$l T7 primer. Primers were incubated 3 min at 95˚C and then quickly chilled on ice. 1 $\mu$l primer mix was used for each reaction which was done as described (Gerber et al., 1998). Minisubstrates were purified on a 20% urea gel. tadA activity was assayed in a reaction mixture containing 50 mM Tris-HCl pH 8.0, 25 mM KCl, 2.5 mM $MgSO_4$, 0.1 mM EDTA, 10% glycerol, 2 mM DTT and 100-200 fmol $^{33}$P-labelled tRNA substrate in a 50 $\mu$l reaction as described (Gerber et al., 1998). When purified tadA was used, the reaction was supplemented with 1 $\mu$g BSA. Reactions were incubated at 37˚C for 15-90 min.

**Sequence analysis of tRNAs**

Sequence analysis of edited tRNA was done as described (Gerber et al., 1998). Analysis was performed with *in vitro* edited tRNAs and with total RNA isolated from *E. coli* strains NWL32 and NWL37. Total RNA was isolated with the Total RNA Midi Kit (Qiagen) according to the manufacturer's protocol. Arg1 primer was used for reverse transcription, Arg2 and Arg3 primers were used for the subsequent PCR.

**UV-crosslinking**

10 $\mu$l reactions were carried out with 400 ng of recombinant tadA and 100 fmol of labelled tRNA in 25 mM Tris-HCl pH 8.0, 2.5 mM $MgSO_4$, 0.05 mM EDTA. The reactions were incubated at room temperature for 6 min and then irradiated on ice in a UV Stratalinker at 500 mJ. tRNAs were digested with 500 ng RNaseA for 25 min at 37˚C. Proteins were separated on a denaturing 12% SDS polyacrylamide gel and exposed on a PhosphorImager screen (Molecular Dynamics).

**Chemical crosslinking**

The crosslinking was done according to a modified version of the procedure described by Fancy and Kodadek (1999). Reactions were performed in microtiter plates in a total volume of 20 $\mu$l. Reactions contained 1 x crosslinking buffer (0.1 x PBS, 2.5 mM MgSO4, 0.05 mM EDTA), 300 ng purified recombinant tadA and BSA from 90 ng to 9 $\mu$g. Proteins were dialyzed against 1x PBS since Tris-buffer inhibits the crosslinking reaction. 1 x reaction buffer (15 mM Na-phosphate pH 7.5, 150 mM NaCl, 0.125 mM tris(2,2'-bipyridyl)dichlororuthenium(II)hexahydrate and APS to a final concentration of 2.5 mM were added in the dark just before irradiation. Reactions were irradiated with a flashlight for 30 seconds and immediately afterwards quenched with 5 $\mu$l of 4 x Laemmli loading buffer. Proteins

were denatured at 95˚C for 3 minutes, separated on a 15% SDS polyacrylamide gel and blotted on a nitrocellulose membrane (Schleicher & Schuell).

**Flag pull-down assay**

tadA was *in vitro* transcribed/translated from construct pJW106 with the TNT® Coupled Reticulocyte Lysate System by Promega according to the manufacturer's instructions. 500 µl Flag-matrix slurry was washed twice with buffer TKG25 (50 mM Tris-HCl pH 8, 25 mM KCl, 10% glycerol, 0.1 mM EDTA, 0.01% NP40). The Flag-matrix was then blocked with 10 µg BSA for 30 minutes at 4˚C and afterwards washed 4x with buffer TKG25. 1 µg of Flag-tagged recombinant tadA that has been purified on Ni/NTA agarose, 3 µl of *in vitro* translated tadA and 100 ng BSA in buffer TKG25 in a total volume of 15 µl were incubated at room temperature for 1 hour. 2 x 250 µl of buffer TKG25 were added to the protein mixture and then transferred to a tube containing the Flag-matrix and incubated at 4˚C for 45 minutes. The Flag-matrix was washed with 3 x 1 ml buffer TKG25. 20 µl Laemmli loading buffer were added and the Flag-matrix boiled at 95˚C for 3 minutes. The supernatant was loaded on a 12% SDS-polyacrylamide gel.

**Acknowledgments**

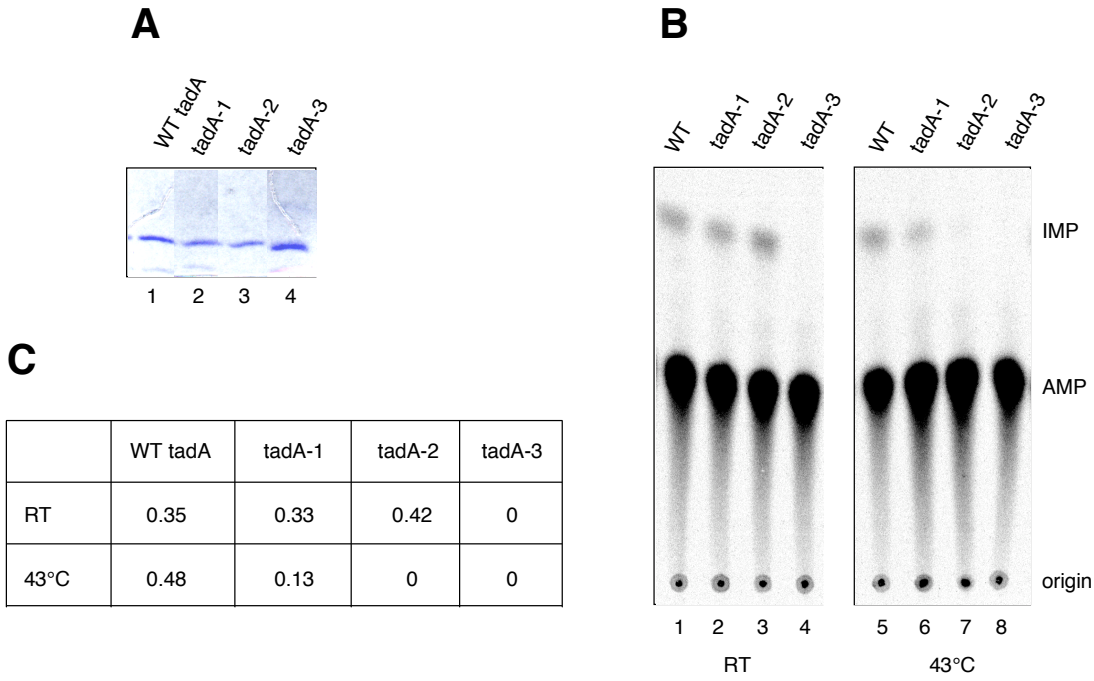
We thank Dr. G. Church (Harvard Medical School, Boston) for providing plasmid pKO3, Dr. H. Loferer (Genome Pharmaceuticals Corporation, Munich) for plasmid pRDC15 and Dr. S. Molin (Technical University of Denmark, Lyngby) for *E. coli* strains NWL32, NWL37 and NWL43. We are grateful to U. Jenal and members of his group for Caulobacter strains and helpful discussions. We thank B. Dichtl, U. Rüegsegger and M. Schaub for critically reading the manuscript. This work was supported by the University of Basel, the Swiss National Science Foundation and the Louis-Jeantet-Foundation for Medicine.

## 3.6 Supplement to chapter 3

### 3.6.1 Isolation of temperature-sensitive tadA mutants

A temperature-sensitive (ts) mutant of the yeast Tad2 protein has been described which contains the two mutations C126S and L153P (Gerber and Keller, 1999). Another allele of Tad2p carrying the mutation E86A was also found to be ts (A.G. and W.K., unpublished results). E86 and L153 are conserved between yeast and *E. coli* Tad2 whereas C126 is within a region of Tad2p that is not present in bacterial sequences (Fig. 3.1). To test whether mutation of these conserved amino acids would also generate ts alleles of tadA, the mutations E96A and L156P were introduced by site-directed mutagenesis either as single or as a double mutations. The resulting alleles were named tadA-1 (L156P), tadA-2 (E96A) and tadA-3 (E96A, L156P). The mutant tadA alleles were expressed in *E. coli* and the recombinant proteins purified on a $Ni^{2+}$/NTA agarose and an □-Flag affinity column. The mutant proteins were expressed equally well as WT tadA (Fig. 3.8 A).

To test tRNA-specific adenosine deaminase activity at high temperature, assay reactions were either incubated at 43˚C or at room temperature as a control. Reactions were also performed with mutant proteins that have been preincubated at 43˚C for 20 min. As can be seen in Fig. 3.8 B WT tadA, tadA-1 and tadA-2 were active at room temperature whereas recombinant tadA-3 did not form I in $tRNA^{Arg2}$ at 22˚C. tadA-3 could either not be folded correctly due to the mutations or the structure of the mutant protein could be weakened and therefore destroyed during purification. At 43˚C WT tadA was slightly more active than at room temperature whereas tadA-1 was reduced in activiy (Fig. 3.8 B, compare lanes 1 and 5, 2 and 6 and C). Interestingly, tadA-2 had no deaminase activity at 43˚C, indicating that the mutation of the conserved E96 leads to a ts allele also in *E. coli*. As expected tadA-3 was also not active at high temperature (Fig. 3.8 B, lane 8). When the reactions were carried out without preincubation of the recombinant protein, I formation could also be detected with tadA-2 (results not shown). This is probably due to the fact that tadA-2 is active at room temperature and that the deamination reaction proceeds very fast. Therefore the time until tadA-2 is inactivated is long enough to allow deamination of the tRNA. tadA-2 deaminated tRNA when the protein was preincubated at 43˚C for 15 min and then tested for activity at room temperature (results not shown). This result showed that tadA-2 is not irreversibly inactivated at high temperature. At 37˚C tadA-1 and tadA-2 are almost as active as WT tadA, even after preincubation of the proteins (results not shown).




C

| | WT tadA | tadA-1 | tadA-2 | tadA-3 |
|---|---|---|---|---|
| RT | 0.35 | 0.33 | 0.42 | 0 |
| 43°C | 0.48 | 0.13 | 0 | 0 |

**Figure 3.8 SDS-polyacrylamide gel analysis of the amount of purified tadA mutants and tRNA-specific adenosine deaminase assay with different tadA mutants.**

**A)** 10 $\mu$l of fraction 2 of each Flag-column were separated on a 12% SDS-PAA gel and stained with Coomassie. The molecular masses of the size standards in kDa are indicated on the left.

**B)** *E. coli* $tRNA^{Arg2}$ was incubated with purified recombinant WT tadA (lanes 1 and 5), tadA-1 (lanes 2 and 6), tadA-2 (lanes 3 and 7) or tadA-3 (lanes 4 and 8) at room temperature or at 43˚C for 30 min. Reactions were processed as described in chapter 3.5.

**C)** Reaction products were quantified with the IPLab Gel Software (Signal Analytics). The numbers are calculated as mol I/mol tRNA.

To test temperature sensitivity *in vivo*, the tadA-1, tadA-2 and tadA-3 alleles were introduced into the *E. coli* genome under the control of the arabinose promoter via homologous recombination (for details see chapter 3.3 and 3.5). For unknown reasons tadA-1 could not be introduced stably into the genome and was not tested *in vivo*. tadA-3 could not rescue the Δ*tadA* strain, indicating that this allele is not functional *in vivo* (results not shown). tadA-2 rescued the Δ*tadA* strain at room temperature (results not shown). However, the $P_{BAD}$-tadA-2 strain grew as efficiently as WT cells at 43˚C on LB plates, indicating that the mutant protein was stabilized within the cells. Interestingly, when tadA-2 was encoded on a multi-copy expression plasmid and was introduced into *E. coli*, the cells had a major growth defect (Fig. 3.9 A). This might be due to the presence of the N-terminal Flag-tag and the C-terminal $His_6$-tag which could prevent correct folding at high temperature. The growth curve of two independent tadA-2 expression

strains was determined. As can be seen in Fig. 3.9 B the mutant strains ceased growth two hours after shift to 44°C whereas the wild-type strain continued growth. It remains to be determined whether $tRNA^{Arg2}$ in the mutant strains has an unmodified position 34.

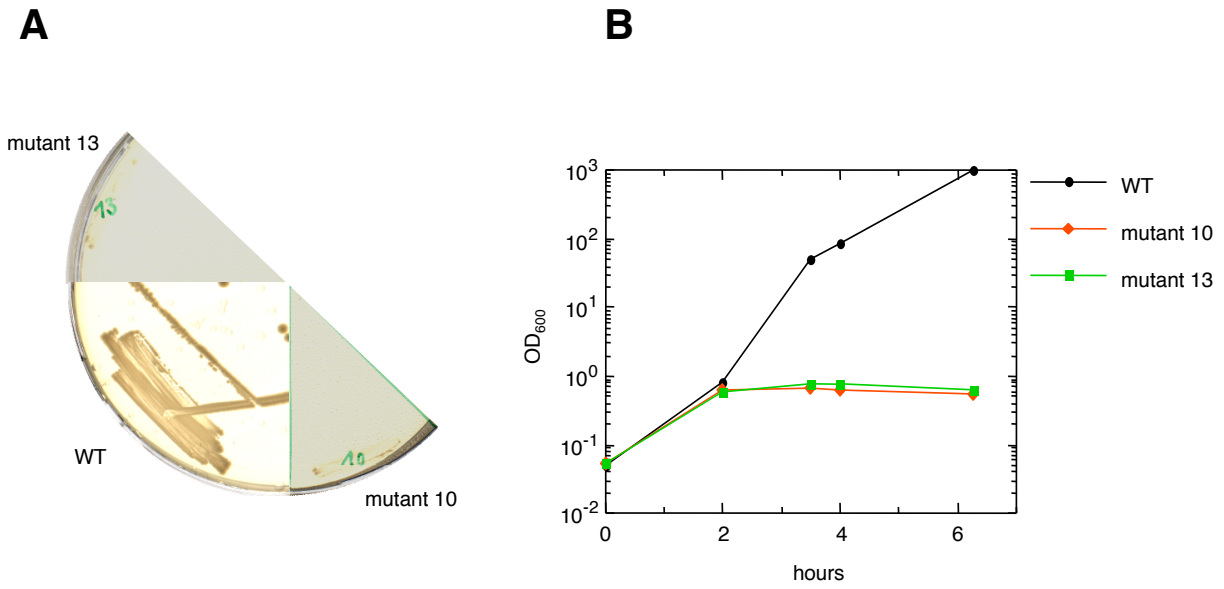



**Figure 3.9 tadA-2 strains have a growth defect at high temperature**

**A)** Growth of WT E. coli and tadA-2/pTrc-FLIS6 cells on LB plates. Plates were incubated overnight at 44°C.
**B)** Growth curve of wild-type and tadA-2 strains after shift to the non-permissive temperature. Cells were incubated at 45°C in a water bath and the optical density was measured at the indicated time points.

It has been predicted but never shown that I-containing tRNAs allow efficient decoding of NNC, NNU and NNA codons (Crick, 1966; Jukes, 1973). It was speculated that I:A base pair formation is inefficient and therefore NNA codons could be specifically avoided (chapter 3.4 and Curran, 1995). Percudani (2001) suggested that I-containing tRNAs do not decode NNA codons in eukaryotes with the exception of yeast tRNAs. However, this hypothesis was so far not confirmed experimentally. Because *S. cerevisiae* has seven I-containing tRNA where the wobble position is deaminated by a complex of two proteins, analysis of translational efficiency would be rather difficult. The identification of tadA offers new possibilities because in *E. coli* one protein deaminates one tRNA substrate. With tadA ts mutants translational efficiency can be determined *in vivo* with a reporter construct encoding several CGA arginine codons followed by a reporter like luciferase which was engineered to contain no CGA, CGC and CGU arginine codons. Upon shift to the non-permissive temperature luciferase activity can be measured in a time-course experiment. At non-permissive temperature newly synthesized $tRNA^{Arg2}$ is not deaminated anymore at the wobble position and with continued growth at high temperature the

already existing modified $tRNA^{Arg2}$ is diluted due to cell division. Unmodified $tRNA^{Arg2}$ is probably not able to base-pair with CGA codons. Base-pairing with CGU should not be affected whereas base-pairing with CGC could be reduced. In summary cells will die because mRNAs encoding for essential proteins are not translated efficiently and processes which they catalyze or where they are involved do not occur efficiently enough. Translational efficiency can also be monitored *in vitro* with the same reporter construct and *E. coli* translation extracts. tadA ts mutants could also be used to monitor overall translation efficiency with pulse-chase experiments. WT and ts strains grown at high temperature would be pulsed with $^{35}$S-methionine and the resulting labelled proteins analyzed by 2-dimensional gel electrophoresis. This would allow to see differences in the translation of specific proteins. Probably highly expressed genes that contain many CGA and CGC codons will be affected most.

**Materials and Methods**

Generation of tadA ts alleles

tadA-1, tadA-2 and tadA–3 were generated by site-directed mutagenesis. For the tadA-1 allele two fragments were amplified by PCR with primer pairs yfhC.3/yfhC.11 and yfhC.4/yfhC.10 introducing the mutations T467C and C468G with numbers referring to the tadA ORF. After gel purification a mix of these two fragments was used as a template for amplification with primers yfhC.3/yfhC.4 for cloning into pTrc-FLIS6 via SpeI. For *in vivo* replacement experiments two fragments were amplified with primer pairs yfhC.7/yfhC.11 and yfhC.8/yfhC.10. A mix of these fragments was used for amplification with primers yfhC.7/yfhC.8 for cloning into pRDC15 via NheI/XhoI. Two fragments for tadA-2 and tadA-3 were amplified with primer pairs yfhC.3/yfhC.11 and yfhC.4/yfhC.13 using either WT tadA (for tadA-2) or tadA-1 (for tadA-3) as template. These reactions introduced the mutations A287C, A288C in tadA-2 and A287C, A288C, T467C, C468G in tadA-3.

| Name | Sequence (5' -> 3') | Restriction site |
|---|---|---|
| yfhC.10 | GGCGTTGCCGAGTGACTTC | - |
| yfhC.11 | CCCCCCTTTCTGCGCTTTAATTTCC | - |
| yfhC.13 | TCACGCTTGCCCCATGTGTAA | - |

Preparation of *E. coli* extracts

20 ml LB supplemented with the appropriate antibiotic were inoculated with either *E. coli* WT or ts strains and grown overnight at 30°C. The next day cells were harvested by centrifugation and

washed twice with buffer Q* (50 mM Tris pH 7.9, 50 mM KCl, 10% glycerol, 5 mM EDTA pH 8, 1 mM DTT, 0.5 mM PMSF, 0.5 $\mu$g/ml leupeptin, 0.7 $\mu$g/ml pepstatin). The pellet was resuspended in 1 ml of buffer Q (same as buffer Q* but with 0.5 mM EDTA) and lysis was performed by sonicating the cells 1 min on ice. The lysate was centrifuged 20 min at 4°C with 14 krpm and the supernatant frozen in liquid nitrogen. The protein concentration of extracts was determined by the Bradford method (Bradford, 1976).

### 3.6.2 tadA chelates $Zn^{2+}$

All deaminases known to date contain $Zn^{2+}$ in their active site (Chang et al., 1991; Moore et al., 1993; Wilson et al., 1991; Yamanaka et al., 1994; Yang et al., 1992). Three amino acids that are characteristic for Zn-chelating enzymes are also found in the deaminase domain of tadA (Fig. 3.1). To test for the presence of $Zn^{2+}$ in tadA the protein was incubated with up to 5 mM o-phenanthroline (OPA) which chelates divalent metal ions and then tested in an assay. Inhibition of I formation could be observed, indicating that tadA contains zinc.

Expression of soluble recombinant tadA was enhanced several fold when 20 mM $ZnCl_2$ was added to the medium. This has been observed previously with yeast Tad2p/Tad3p and also with yeast Tad1p (Gerber et al., 1998; Gerber and Keller, 1999). The amount of active protein was increased significantly when the expression was done in the presence of $ZnCl_2$. The importance of $Zn^{2+}$ was also shown with inactive Tad2p mutants that could be re-activated when the assay reaction was supplemented with $ZnCl_2$ (M.S. and W.K., unpublished results). All these results suggest that Tad proteins bind $Zn^{2+}$ in their active site which is a prerequisite for adenosine deamination.

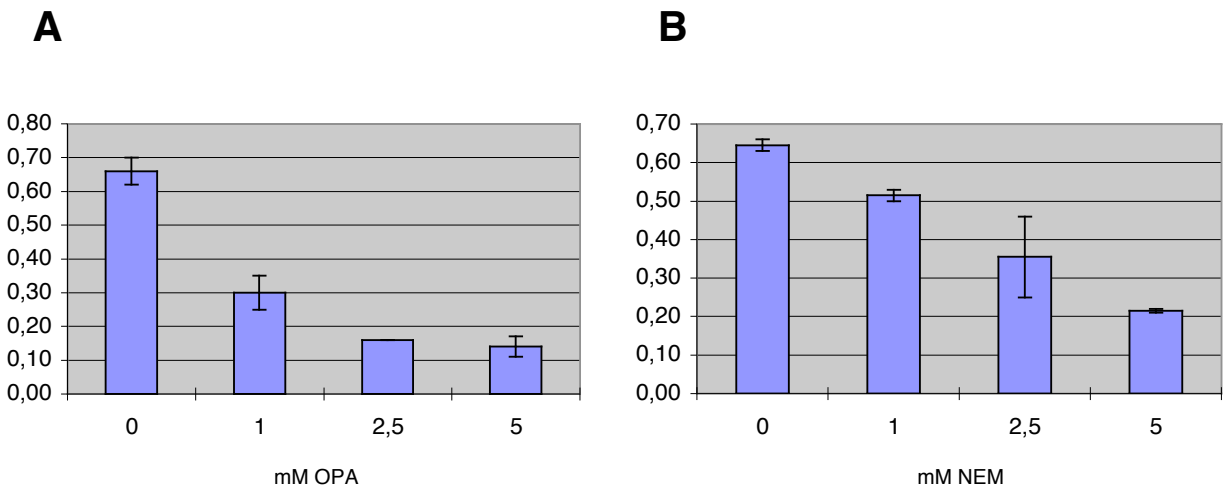


**Figure 3.10 Inhibition of recombinant tadA by OPA (A) and NEM (B).**

Final concentrations of the inhibitors are indicated at the bottom. The amount of I formation is indicated on the left. Reactions were performed for 1 hour at 37°C with 2 ng of Ni-NTA purified $His_6$-tadA.

### 3.6.3 tadA is an essential gene in *Caulobacter crescentus*

To confirm the tadA deletion results obtained in *E. coli*, the tadA homologue in *Caulobacter crescentus* was deleted with the sacB counterselection technique. *Caulobacter* tadA is 45% identical to *E. coli* tadA and is highly conserved in the deaminase domain (Fig. 3.11 boxed sequence). There was no other gene in the *Caulobacter* genome which displayed such a high degree of homology.

```
          ·        20         ·        40         ·        60         ·        80         ·       100         ·       120
  1 MRRAFITGVFFLSEVEFSHEYWMRHALTLAKRAWDEREVPVGAV-L-VHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT
          +    E E      MR AL  A+ A +  E PVGAV L      VI    N PI  HDPTAHAEI A+R     + NYRL D TL VTLEPC MCAGA+ H+RIGRVVFGA D K
  1 ===MGFRFLMRTDESEDQDHRMMRLALDAARAAAEAGETPVGAVILDPSTGEVIATAGNGPIAAHDPTAHAEIAAMRAAAAKLGNYRLTDLTLVVTLEPCAMCAGAISHARIGRVVFGADDPKG
             ·        20         ·        40         ·        60         ·        80         ·       100         ·       120
        ·       140         ·       160         ·
124 AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD      178
    A          P  + R E+T G+LADE A LL  FFR RR+
123 AVVHGPKFFAQPTCHWRPEVTGGVLADESADLLRGFFRARRKAKI==========      167
          ·       140         ·       160
```

**Figure 3.11 Alignment of E. coli and Caulobacter crescentus tadA.**

The *E. coli* tadA amino acid sequence is shown in the upper lane, the *C. crescentus* sequence in the lower lane. The deaminase domain is boxed. Numbers above and below the sequences indicate amino acids.

The sacB counterselection technique is described in detail in chapter 3.3 (see also Fig. 3.7). The results of the deletion are shown in Fig. 3.12. In summary, when tadA was deleted in *C. crescentus* only wild-type and $sacB^-$ clones were obtained. However, Δ*tadA* clones were obtained when a rescue copy of tadA was provided on a plasmid, thus confirming the deletion results with *E. coli*.

| | Total clones | 1 ΔtadA | 2 WT | 3 $sacB^-$ |
|---|---|---|---|---|
| - | 400 | 0% | 40% | 60% |
| pJW72 | 150 | 4% | 42% | 54% |

**Figure 3.12 tadA is essential in *C. crescentus*.**

The total number of sucrose-resistant colonies analyzed are indicated in the first column (Total clones). The number of isolated colonies with a disrupted tadA (ΔtadA) gene in the chromosome is shown in the second column, the number of colonies with a restored wildtype tadA locus (WT) after the second recombination step is shown in the third column. Colonies that had acquired sucrose-resistance through *sacB* inactivation are listed in the fourth column (*sacB*$^-$). Plasmids pJW72 carries a wild-type copy of the *C. crescentus* tadA ORF.

## Methods

### Table V Plasmids, strains and primers Caulobacter tadA

| Name | Description | Reference |
|---|---|---|
| pNPTS138 | $Kan^R$ derivative of pLITMUS28 cloning vector with sacB and oriT | D. Alley |
| pMR20 | RK2-based $Tet^R$ broad host range vector | R. Roberts + C. Moor |
| pJW57 | ±500 bp Caulobacter tadA ORF in pNPTS138 (SpeI, AflII) | this work |
| pJW58 | pJW57 with Spec/Strep-cassette (BamHI) | this work |
| pJW72 | Caulobacter tadA ORF in pMR20 (NcoI, EcoRI) | this work |
| pJW81 | Caulobacter tadA ORF in pBF11 (NdeI, SpeI) | this work |
| NA1000 | Caulobacter WT strain | Evinger and Agabian (1977) |
| JW30 | Caulobacter ΔtadA, pJW72 | this work |

The strains and plasmids used for the deletion of Caulobacter tadA are described in table V. *C. crescentus* strains were grown in PYE medium (Ely, 1991), supplemented with nalidixic acid (20 $\mu$g/ml), kanamycin (20 $\mu$g/ml), tetracyclin (2 $\mu$g/ml), spectinomycin (50 $\mu$g/ml) or streptomycin (30 $\mu$g/ml) as needed.

The Caulobacter tadA gene was disrupted by the insertion of an omega cassette (Prentki and Krisch, 1984) via homologous recombination. A 1 kb SpeI/AflII fragment containing 500 nt upstream and downstream flanking regions of the tadA ORF was cloned into pNPTS138. The omega cassette was then inserted in the unique BamHI restriction site between the upstream and the downstream fragment generating construct pJW58. pNPTS138 contains two selection markers for the recombination step: a kanamycin resistance cassette and a *sacB* gene conferring sucrose sensitivity. pJW58 was integrated into the chromosome of *C. crescentus* wild-type strain NA1000 via conjugation and homologous recombination. The resulting strain contained a wild-type and a disrupted copy of tadA. To induce the second recombination event, colonies were grown overnight at 30˚C in PYE and plated on PYE plates containing 3% sucrose. Colonies that

were able to grow on sucrose were tested for loss of the plasmid encoded kanamycin resistance and for presence of the omega cassette encoding a spectinomycin/streptomycin resistance. Clones that were sucrose- and kanamycin-resistant had not lost the plasmid and were probably resistant to sucrose due to a mutational inactivation of the *sacB* gene. Spectinomycin/streptomycin- and kanamycin-sensitive clones had lost the plasmid and recombination left a wild-type tadA in the chromosome. Null mutants of tadA should be spectinomycin/streptomycin-resistant and kanamycin-sensitive.

For the rescue construct for tadA null mutants, the tadA ORF was amplified by PCR from *C. crescentus* genomic DNA with primers ccTad2.3/ccTad2.2. The fragment was cloned into the NcoI/EcoRI restriction sites of pMR20 generating construct pJW72. pJW72 was inserted in NA1000 via conjugation and correct clones identified by resistance to tetracyclin. A strain containing pJW72 was then used to repeat the same procedure as described above.

# 4 Characterization of hTad2 and hTad3

## 4.1 Introduction

tRNA is the most extensively modified nucleic acid in the cell and modified nucleotides which are derivatives of the four standard nucleosides were found soon after the discovery of the soluble (s)RNA, now called tRNA (Dunn, 1959; Smith and Dunn, 1959). All modifications with the exception of queuosine (Q) are formed post-transcriptionally (reviewed in Björk, 1995; Björk, 1996). To date, more than 80 different modifications have been characterized in tRNA and some of them probably evolved early since they are conserved in Archaea, Bacteria and Eucarya (Sprinzl et al., 1998). Some modified nucleotides are however specific for one of the three phylogenetic domains and are likely to have evolved after the divergence of the three kingdoms (Björk, 1984; Edmonds et al., 1991). Although modified nucleotides are found in various positions in the tRNA, position 34 and 37 contain the largest variety of modifications which influence codon reading patterns in several ways. Base modification contributes to reading frame maintenance by improving the efficiency of the tRNA, influencing the fidelity of translation, improving the recognition by elongation factors or aminoacyl-tRNA synthetases, influencing codon choice, improving the efficiency of translation, decreasing the codon context sensitivity and preventing frameshifting (reviewed in Björk, 1995; Björk, 1996; Curran, 1998). A number of different modifications were predicted to prevent frameshifting (Urbonavicius et al., 2001).

Eukaryotic $tRNA^{Ala}$ contains inosine at two positions: at position 34 (the wobble position of the anticodon) and at position 37 (3' adjacent to the anticodon). Inosine at these positions is generated by hydrolytic deamination of a genomically encoded adenosine (reviewed in Grosjean et al., 1996a). The deamination of $A_{37}$ to $I_{34}$ is catalyzed by Tad1p (tRNA-specific adenosine deaminase 1), an enzyme which is related to the pre-mRNA editing enzymes ADARs (Gerber et al., 1998). In a second step $I_{37}$ is further methylated to $m^1I_{37}$ by a recently identified methyltransferase (Björk et al., 2001). Tad1p contains a deaminase domain with three $Zn^{2+}$-chelating residues and a glutamate involved in catalysis, however it does not contain a known RNA binding domain. Tad1p is encoded by a non-essential gene in *S. cerevisiae* and the function of $m^1I$ is not known (Gerber et al., 1998). It was suggested that this modification prevents translational frameshifts.

In contrast to $m^1I_{37}$, $I_{34}$ is present in seven or eight eukaryotic tRNA and also in $tRNA^{Arg2}$ from *E. coli* and plant chloroplasts. The deamination of $A_{34}$ in yeast is catalyzed by a heterodimer consisting of the two related subunits Tad2p and Tad3p (Gerber and Keller, 1999). Both proteins are encoded by essential genes (Gerber and Keller, 1999). Interestingly, the two proteins display a deaminase domain that resembles that of the cytidine deaminase (CDA)

superfamily (reviewed in Gerber and Keller, 2001). Therefore it was postulated that Tad2p evolved from a prokaryotic cytidine deaminase progenitor and that the appearance of Tad3p was the result of a gene duplication event (Gerber and Keller, 1999). This hypothesis is supported by tadA which is sufficient for bacterial tRNA:adenosine 34 deaminase activity and is likely representing the origin of eukaryotic Tad proteins (Wolf et al., 2002).

Here, we report the characterization of the human tRNA:adenosine 34 deaminase which consists of hTad2 and hTad3. The two proteins are homologous to the previously identified yeast proteins Tad2p/Tad3p. The results of different experiments suggest that $A_{34}$ deamination is a nuclear process and $I_{34}$ might thus contribute to efficient aminoacylation of the tRNA or increase export efficiency.

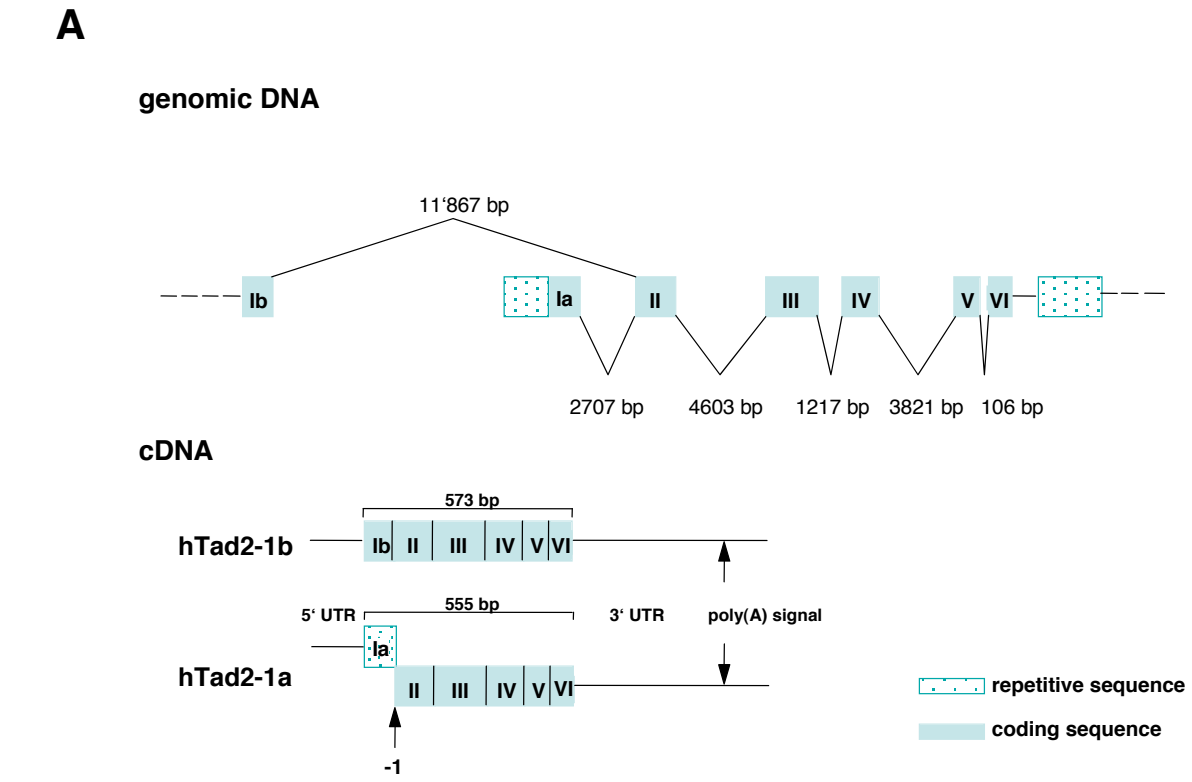
## 4.2 Results

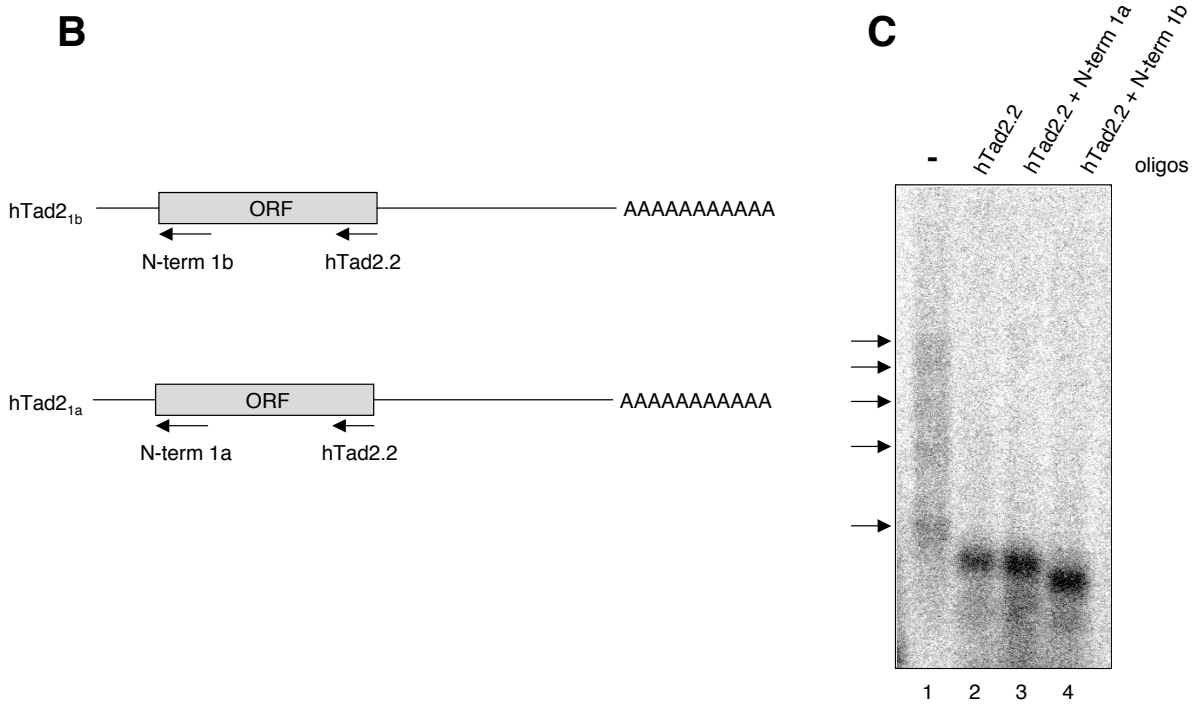
**Two alternatively spliced hTad2 mRNAs are present in HeLa cells**

Database searches with the yeast Tad2 sequence revealed a human EST which encoded part of a homologous protein. The 5' end of this EST T97494 was used to screen a HeLa cDNA library. One of the positive clones identified contained a cDNA of 1514 bp coding for a 184 amino acid protein. The clone had an in frame stop codon upstream of the first methionine, a short (52 bp) 5' UTR and a long (910 bp) 3' UTR. The sequence of the encoded protein was highly conserved in the deaminase domain to yeast Tad2p and contained characteristic residues of deaminases including the three $Zn^{2+}$-chelating amino acids and a glutamate thought to be involved in proton transfer during catalysis. However, the protein was encoded in two different reading frames. Amino acids 26 to 184 were encoded in the –1 frame compared to the first methionine. It was tried to isolate an in frame clone by PCR from HeLa and HEK cDNA, but all the amplified fragments had the same –1 frameshift (results not shown). Therefore an in frame hTad2 clone was generated by insertion of an additional nucleotide at the frameshift site. A second hTad2 clone was isolated that contained an alternative exon 1 and was therefore named $hTad2_{1b}$. hTad2 with the frameshift in its ORF was renamed $hTad2_{1a}$ according to the order in which the two isoforms were identified. $hTad2_{1b}$ and $hTad_{1a}$ were identical except the first 25 aa of $hTad_{1a}$ and 32 aa of $hTad2_{1b}$ (Fig. 4.1 D). However, the $hTad2_{1b}$ protein was encoded in one frame. Comparison of both hTad2 sequences with the genomic sequence revealed that the two different exons 1 were present consecutively on chromosome 6 separated by 9100 bp (Fig. 4.1 A). The

splice site consensus sequence was conserved between exon 1a and exon 2 which could explain alternative splicing at this site and the isolation of $hTad2_{1a}$. hTad2 is encoded in six exons which are separated by introns ranging in size from 106 to 11'867 bp (Fig. 4.1 A). Short repetitive sequence elements were identified upstream of exon 1a and downstream of exon 6. The repetitive sequence in the 5' UTR and first exon belongs to the family of LTR/MaLR elements (long terminal repeats/mammalian apparent LTR-retrotransposons) whereas the sequence in the $hTad2_{1a}$ 3' UTR belongs to the LINE elements (long interspersed nucleotide elements). These elements are inserted into the genome via reverse transcription of RNAs from active elements and are thought to be of viral origin (Deininger et al., 1992). It is not known whether these repetitive sequences have a function in the hTad2 mRNA.

To determine the ration between the $hTad2_{1a}$ and $hTad2_{1b}$ mRNAs *in vivo*, Northern blot analysis was performed with HeLa poly(A)$^+$ RNA. At least five different mRNAs could be detected with a probe against the ORF of hTad2 (Fig. 4.1 C, lane 1). To distinguish between $hTad2_{1a}$ and $hTad2_{1b}$ mRNAs, RNase H cleavage was applied. Poly(A)$^+$ RNA was incubated with oligonucleotide hTad2.2 that annealed at the 3' end of the hTad2 ORF (Fig. 4.1 B). After digestion with RNase H a single band was detected in the Northern (Fig. 4.1 C, lane 2) suggesting that the various mRNAs detected in the undigested sample were due to differences in the 3' UTR. However, the single band in lane 2 could still be a mixture of $hTad2_{1a}$ and $hTad2_{1b}$ transcripts. Therefore, the mRNA was either digested with oligonucleotides hTad2.2/N-term-1a or hTad2.2/N-term-1b. Incubation with the $hTad2_{1a}$-specific primer did not change the band pattern compared to lane 2 (Fig. 4.1 C, lane 3) whereas after incubation with a $hTad2_{1b}$-specific primer a shorter transcript accumulated (Fig. 4.1 C, lane 4). This result showed that *in vivo* $hTad2_{1b}$ is the major form whereas $hTad2_{1a}$ could not be detected with the Northern. However, $hTad2_{1a}$ mRNA can easily be detected by RT-PCR (results not shown).


A
genomic DNA
11‘867 bp
Ib
Ia
II
III
IV
V
VI
2707 bp
4603 bp
1217 bp
3821 bp
106 bp
cDNA
573 bp
hTad2-1b
Ib
II
III
IV
V
VI
5‘ UTR
555 bp
3‘ UTR
poly(A) signal
hTad2-1a
Ia
II
III
IV
V
VI
-1
repetitive sequence
coding sequence


B
hTad2$_{1b}$
ORF
AAAAAAAAAAA
N-term 1b
hTad2.2
hTad2$_{1a}$
ORF
AAAAAAAAAAA
N-term 1a
hTad2.2
C
-
hTad2.2
hTad2.2 + N-term 1a
hTad2.2 + N-term 1b
oligos
1
2
3
4

**D**

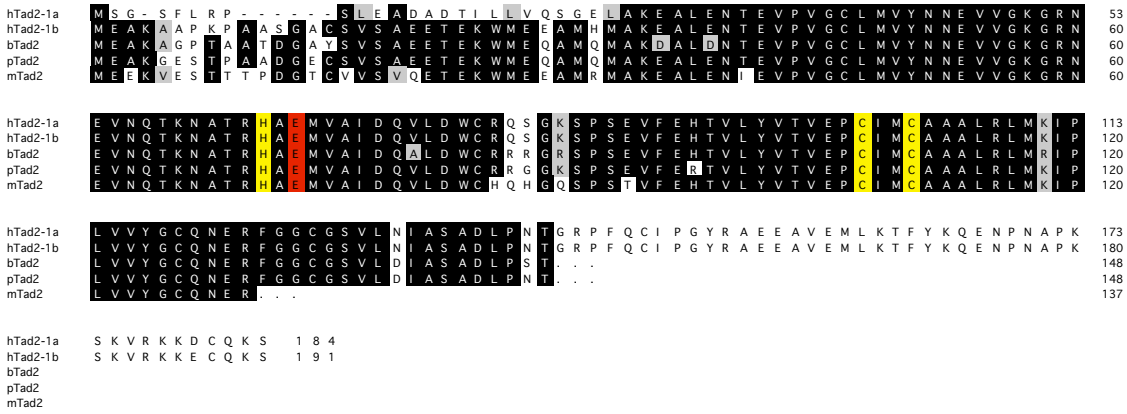



**Figure 4.1 hTad2 mRNA is alternatively spliced**

**A)** Schematic representation of the hTad2 genomic region on chromosome 6. Exons are shown as blue boxes and are not drawn to scale. The sizes of the introns are indicated in base pairs. Dotted boxes correspond to repetitive sequences. The cDNAs corresponding to the two alternatively spliced hTad2 mRNAs are shown with the –1 frameshifting site in $hTad2_{1a}$.

**B)** Schematic representation of the experiment shown in C. The hTad2.2 primer anneals to both hTad2 mRNAs, the primers N-term 1a and N-term 1b are transcript-specific.

**C)** Northern blot analysis of HeLa poly(A)$^+$ RNA that was RNase H cleaved with oligonucleotides as indicated on top of each lane. Poly(A)$^+$ RNA was separated on a 1.5% agarose/formaldehyde gel and blotted to a Hybond N$^+$ membrane for detection with a hTad2 probe. Arrows indicate hTad2 mRNAs in the uncleaved sample.

**D)** Alignment of mammalian Tad2 sequences. Bovine (b), pig (p) and mouse (m) amino acid sequences were predicted from ESTs. Highly conserved amino acids (≥80%) are shown in black, similar residues in grey. Three $Zn^{2+}$-chelating residues are depicted in yellow and the glutamate mediating proton transfer during catalysis in red.

## $hTad2_{1a}$ does not contain a site of programmed translational frameshifting

All the $hTad2_{1a}$ mRNAs that were isolated and those found in the databases had exons 2 to 6 in the –1 frame compared to exon 1. It was therefore investigated whether the $hTad2_{1a}$ mRNA contains a site of programmed translational frameshifting. Although errors that alter the reading frame are extremely rare during translation, some genes have evolved sequences that efficiently induce frameshifting. Frameshifting is conceptually a simple process. At a particular step in the cycle of translational elongation the ribosome shifts its reading frame from the one it initiated translation into a new reading frame. In some genes the frequency of frameshifting is very high, aproaching 100%. These efficient frameshifts are termed programmed frameshifts since conserved structural features of the mRNA template predispose the ribosome toward a shift at a specific site. As a consequence, the protein is not directly encoded in the DNA as a single ORF, but in two overlapping reading frames. Programmed translational frameshifting has been

observed in diverse RNA viruses, *E. coli, B. subtilis,* yeast and mammals (reviewed in Farabaugh, 1996; Gesteland and Atkins, 1996). Mechanisms underlying frameshifting are diverse and in many cases unknown. The strength of the codon-anticodon pairing, low availability of certain tRNAs and secondary RNA structures play a role in +1 and -1 frameshifting. In mammals and *E. coli* frameshifting regulates the levels of specific enzymes. Viruses use frameshifting to translate gag-pro and gag-pro-pol proteins which are encoded in overlapping reading frames (reviewed in Farabaugh, 1996; Gesteland and Atkins, 1996).

Programmed translational frameshifting of $hTad2_{1a}$ was investigated *in vitro* by translation with rabbit reticulocyte lysate in the presence of $^{35}$S-methionine. Two expression constructs encoding $hTad2_{1a}$ either with or without the –1 frameshifting site were *in vitro* translated. Labelled protein was synthesized when the hTad2 ORF did not contain the frameshifting site, however no hTad2 was detected with the –1 frameshift construct (results not shown). To test whether programmed translational frameshifting occured *in vivo*, different $hTad2_{1a}$ expression constructs were transfected into HEK293 cells which have been used previously to study recoding (Grentzmann et al., 1998). The first construct encoded the $hTad2_{1a}$ ORF, either with or without the –1 frameshifting site (Fig. 4.2 A). Extracts of transfected cells were separated on polyacrylamide gels and blotted to nitrocellulose. hTad2 proteins were detected with □-hTad2 antibodies and recombinant hTad2 was distinguished from endogenous hTad2 by size. Tagged hTad2 could readily be detected when the protein was expressed from the corrected clone (Fig. 4.2 A, lane 2), however no recombinant hTad2 was expressed from the frameshifting construct (Fig. 4.2 A, lane 1).

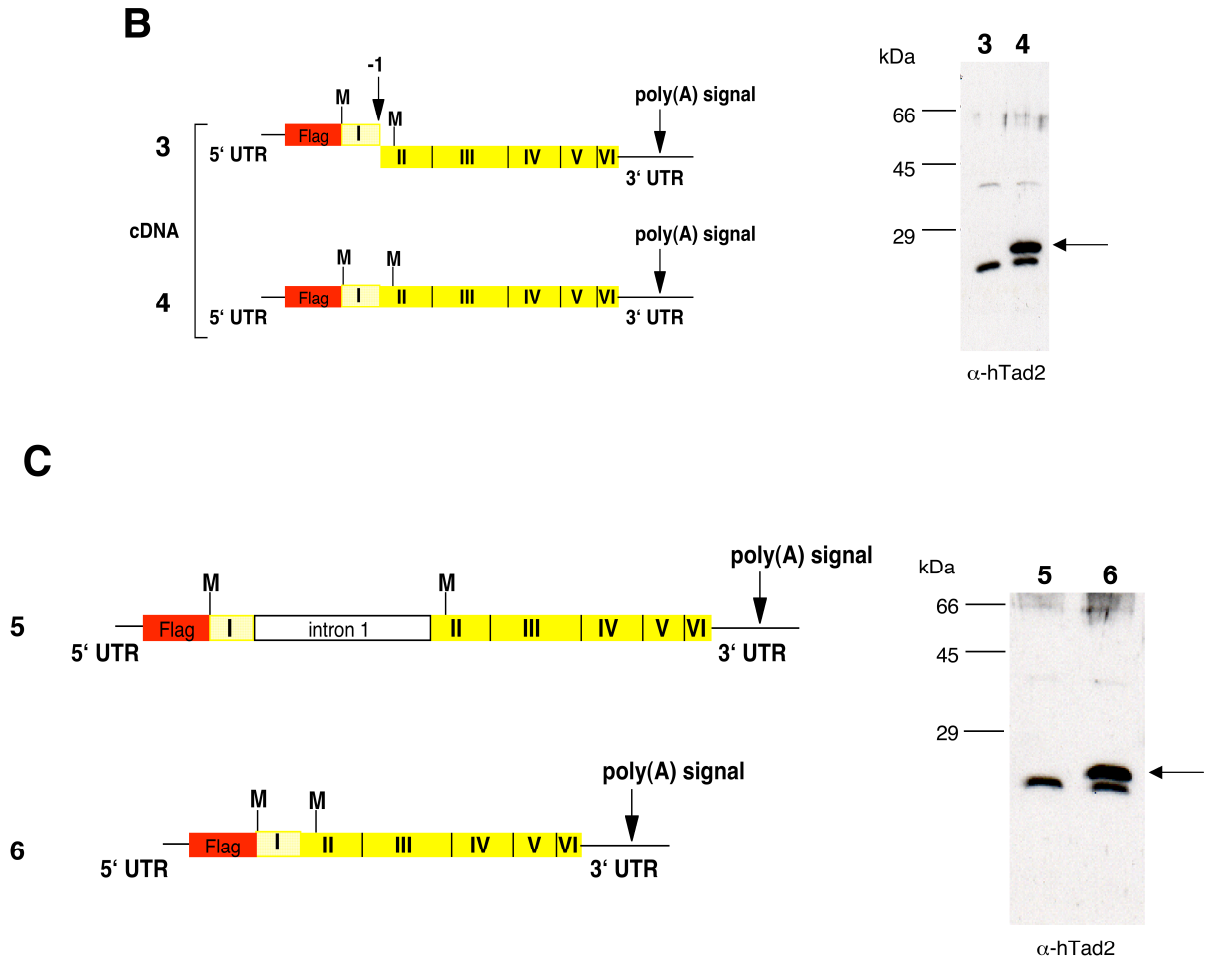



**Figure 4.2 hTad2$_{1a}$ does not contain a –1 frameshifting site.**

**A)** Schematic drawing of the constructs used for transfection and Western blot analysis of extracts from transfected HEK293 cells. The hTad2$_{1a}$ ORF was cloned either in frame (lane 2) or with exons 2 to 6 in the –1 frame (lane 1). The recombinant protein had a N-terminal Flag-tag and a C-terminal His$_6$-tag and was detected with a □-hTad2 antibody. The arrow indicates the position of recombinant hTad2. The molecular masses of the size standards in kDa is indicated on the left.

**B)** Same experiment as described in A) except that the hTad2$_{1a}$ cDNA encoding a Flag-tag at the 5’ end of the ORF either with (lane 3) or without (lane 4) the -1 frameshifting site was transfected.

**C)** HEK293 cells were transfected with a hTad2$_{1a}$ cDNA containing intron 1 (lane 5) and as a control the hTad2$_{1a}$ cDNA without frameshift was used (lane 6).

It has been shown previously that secondary structures like pseudoknots stimulated translational frameshifting by pausing the ribosome and thereby increasing the likeliness of alternative decoding ( ten Dam et al., 1990; reviewed in Farabaugh, 1996; Gesteland and Atkins, 1996). Such structures could be formed between sequences around the frameshifting site and sequences further upstream or downstream, e.g. in the UTR. Therefore, a hTad2 expression construct

containing the 5' and 3' UTR sequences was transfected into HEK293 cells (Fig. 4.2 B). Also with these constructs expression of recombinant hTad2 could only be observed with the corrected clone (Fig. 4.2 B, lanes 3 and 4). Because the frameshift site was present exactly at the exon-exon junction, it was tested whether splicing at this site could determine framehifting. Intron 1 was therefore reintroduced into the $hTad2_{1a}$ cDNA and transfected into HEK293 cells. Expression of recombinant hTad2 could not be observed with this construct (Fig. 4.2 C, lane 1).

Frameshifting was also investigated with dual-luciferase reporter constructs that encoded the firefly luciferase gene the expression of which was coupled to short upstream frameshfting sequences (pluc; Grentzmann et al., 1998). For normalization of the firefly luciferase activity the activity of a second luciferase gene from *Renilla reniformis* was measured. Firefly luciferase activity could only be measured when frameshifting occurred and was thus linearily coupled to the frameshifting efficiency. The hTad2 coding region around the frameshifting site (33 nt upstream sequence and 100 nt downstream sequence) was cloned into pluc and transfected into HEK293 cells. Extracts of transfected cells were tested for both luciferase activities, however no significant frameshifting was observed with hTad2 (results not shown). Although frameshifting could not be detected in all these experiments it cannot be excluded that hTad2 is synthesized from the $hTad2_{1a}$ mRNA. Frameshifting might require different conditions or could be regulated or $hTad2_{1a}$ could be translated by a different mechanism.

**Cloning of hTad3 and reconstitution of human tRNA:adenosine 34 deaminase activity with recombinant proteins**

hTad3 sequences were identified by homology to yeast Tad3p. The hTad3 ORF was amplified from EST AI367911 and the recombinant protein expressed in *E. coli* CodonPlus cells. These cells encode additional copies of rare *E. coli* tRNAs Arg(AGA and AGG) and Pro(CCC). The hTad3 ORF has a G/C content of 73% and 21 CCC proline codons which is a very rare codon in *E. coli*. hTad3 expression was increased several fold in BL21 CodonPlus cells compared to BL21 cells (results not shown). hTad3 was purified on a $Ni^{2+}$/NTA column (Fig. 4.3 A). To reconstitute human tRNA:adenosine 34 deaminase activity *in vitro* recombinant hTad2 was expressed in *E. coli* from the corrected $hTad2_{1a}$ and from the $hTad2_{1b}$ clone with an N-terminal Flag-tag and a C-terminal $His_6$-tag. The proteins were purified on a $Ni^{2+}$/NTA and an □-Flag-column (Fig. 4.3 A) and tested for adenosine deaminase activity with hTad3 on *in vitro* transcribed *B. mori* $tRNA^{Ala}$ (A37G) tRNA. Neither hTad2 alone nor hTad3 deaminated A to I (Fig. 4.3 A, lanes 1 and 2). However, when the proteins were mixed, they converted A to I up to 0.8 mol I/mol tRNA (Fig. 4.3 A, lane 3). Thus, hTad2 and hTad3 are sufficient to reconstitute

tRNA:adenosine 34 deaminase activity. Other eukaryotic-specific cofactors are not needed since both recombinant proteins were expressed in *E. coli*. hTad2 and hTad3 are therefore the sequence and also functional homologues of the previously described yeast proteins Tad2p/Tad3p (Gerber and Keller, 1999). hTad2 and hTad3 could also be co-expressed in *E. coli* with a N-terminal Flag-tag (hTad2) and a N-terminal $His_6$-tag (hTad3). The complex was purified on a $Ni^{2+}$/NTA column and an□-Flag column (Fig. 4.3 B). Both complexes, containing either $hTad2_{1a}$ or $hTad2_{1b}$, were active *in vitro* (Fig. 4.3 B).

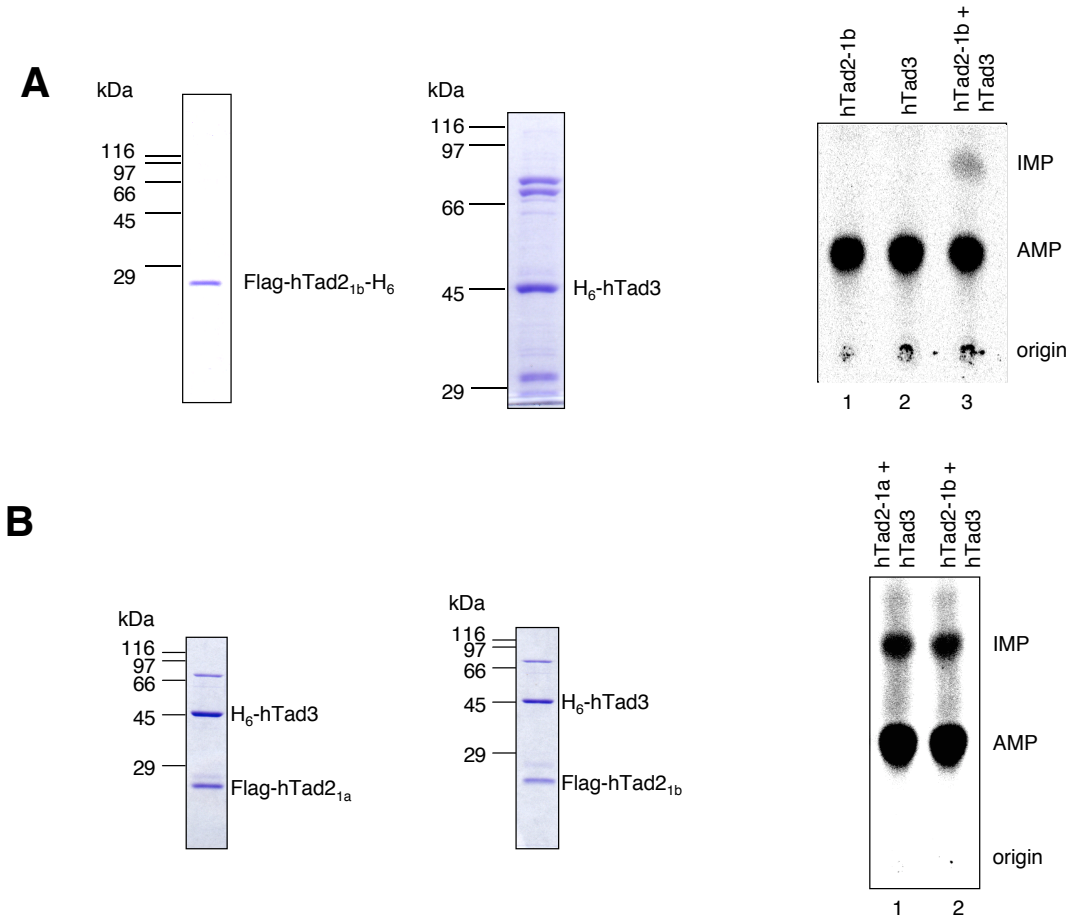



**Figure 4.3 hTad2 and hTad3 are required for inosine formation at position 34.**

**A)** Coomassie stained gels of the final purification step of $hTad2_{1b}$ and hTad3. Approximately 20 ng of recombinant $hTad2_{1b}$ (lane 1), hTad3 (lane 2) or 20 ng of each protein were tested for tRNA:adenosine 34 deaminase activity with *B. mori* $tRNA^{Ala}$ (A37G). The reactions were incubated at 37°C for 1 hour.

**B)** Coomassie stained polyacrylamide gel of the purification of the hTad2/hTad3 complex. Peak fractions of the final Flag-column were separated on 15% polyacrylamide gels. The $hTad2_{1a}$/hTad3 complex (lane 1) and the $hTad2_{1b}$/hTad3 complex (lane 2) were tested for tRNA deaminase activity with *B. mori* $tRNA^{Ala}$ (A37G) for 1 hour at 37°C.

**Depletion of tRNA:adenosine 34 deaminase activity from HeLa extracts**

Polyclonal antibodies were raised against aa 29-184 of hTad2 (ΔN-hTad2). These antibodies recognized the antigen (Fig. 4.4 A, lane 2), recombinant hTad2 (Fig. 4.4 A, lane 1) and similar amounts of a single protein in HeLa cell cytoplasmic and nuclear extracts (Fig. 4.4 A, lanes 3 and 4). ΔN-hTad2 was shorter than the protein detected in HeLa cell extracts (Fig. 4.4 A, compare lanes 2, 3 and 4), showing that the N-terminally deleted form did not correspond to the endogenous hTad2. hTad2 could be depleted from HeLa cell extract whereas depletion with preimmun hTad2 serum did not change the amount of hTad2 (Fig. 4.4 B, lanes 2 and 3). The depleted extracts did not form I in *B. mori* $tRNA^{Ala}$ (A37G) (Fig. 4.4 C, lane 3) whereas extracts depleted with the preimmun α-hTad2 serum had the same tRNA adenosine deaminase activity as untreated HeLa cell cytoplasmic extract (Fig. 4.4 C, compare lanes 1 and 2). Addition of recombinant hTad2 to depleted extracts did not restore activity (Fig. 4.4 C, lane 4). However, activity could be restored by adding hTad2 and hTad3 (Fig. 4.4 C, lane 6), indicating that the α-hTad2 serum depleted the hTad2/hTad3 complex. This is not surprising because both proteins are required for tRNA adenosine deaminase activity (Fig. 4.3 A) and hTad2 and hTad3 form a heterodimer which is supported by the purification of the recombinant complex from *E. coli* lysate (Fig. 4.3 B).

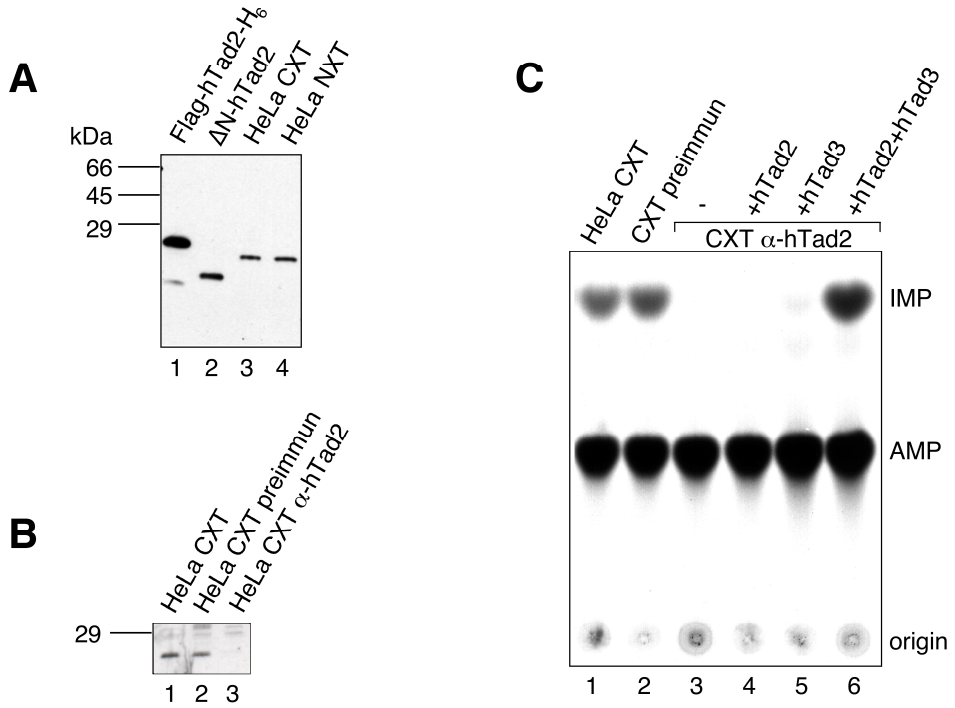



**Figure 4.4 α-hTad2 antibodies deplete the hTad2/hTad3 complex from HeLa cell extracts.**

**A)** Western blot analysis of Flag-hTad2-$His_6$ (lane 1), ΔN-hTad2 (lane 2), HeLa cell cytoplasmic (lane 3) and nuclear extract (lane 4). The proteins were separated on a 15% polyacrylamide gel and blotted to a nitrocellulose membrane for detection with α-hTad2 antibodies.

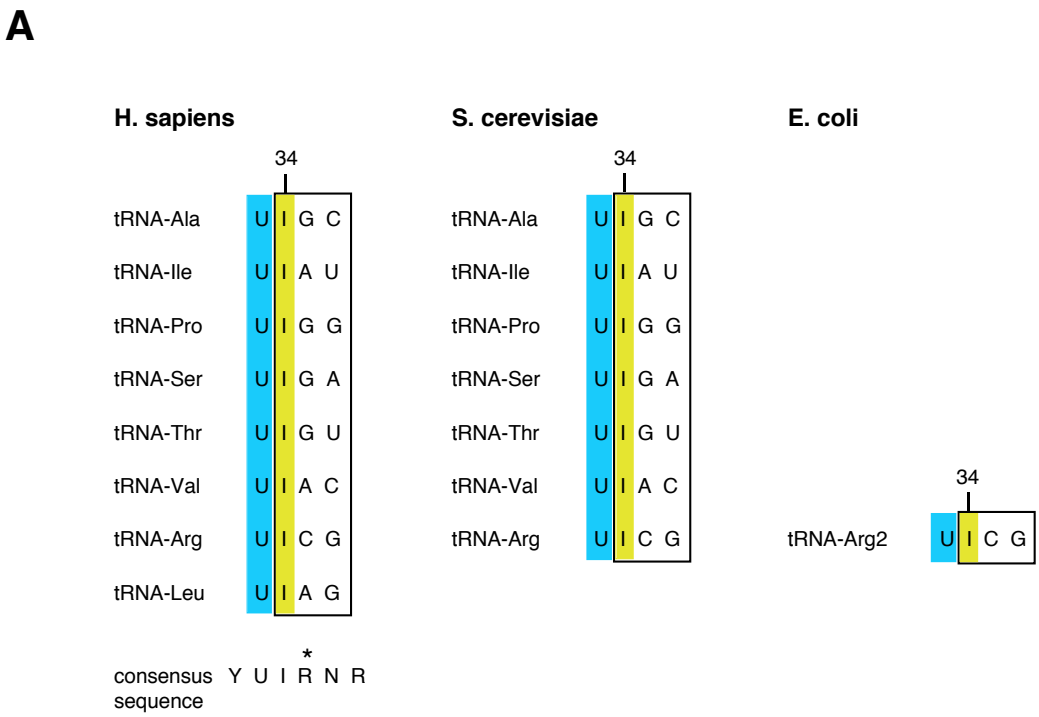
**B)** Western blot analysis of HeLa cell extracts depleted with -hTad2 antibodies. Untreated HeLa cell cytoplasmic extract (lane 1), extract depleted with preimmun hTad2 serum (lane 2) and extract depleted with -hTad2 antibodies were tested for the presence of hTad2.
**C)** tRNA-specific adenosine deaminase assay with depleted extracts. *B. mori* $tRNA^{Ala}$ (A37G) was incubated with 10 $\mu$g total protein HeLa cell cytoplasmic extract (lane 1), extract depleted with preimmun serum (lane 2) or extract depleted with -hTad2 serum (lane 3). To reconstitute tRNA deaminase activity depleted extracts were supplemented with recombinant hTad2 (lane 4), hTad3 (lane 5) or hTad2 and hTad3 (lane 6).

## Substrate specificity of hTad2/hTad3

The substrate specificity of recombinant hTad2/hTad3 was tested with various eukaryotic and prokaryotic tRNA substrates. Eight human tRNAs are known to contain $I_{34}$ (Fig. 4.5 A). Site-specificity of hTad2/hTad3 was tested with a mutant $tRNA^{Ala}$ from *B. mori* that contained G instead of A at position 34. This tRNA was not deaminated, indicating that adenosine deamination is specific for position 34 (Fig. 4.5 B, lane 2). Yeast $tRNA^{Ala}$ and human $tRNA^{Ala}$ were a substrate for hTad2/hTad3 (Fig. 4.5 B, lane 3 and results not shown). It has been shown previously that $tRNA^{Arg2}$ from *E. coli* is a substrate for yeast Tad2p/Tad3p (Auxilien et al., 1996). The result that hTad2/hTad3 also deaminated $tRNA^{Arg2}$ (Fig. 4.5 B, lane 4) confirmed that eukaryotic tRNA:adenosine 34 deaminases recognize prokaryotic substrates. In summary, eukaryotic tRNAs that contain A at the wobble position of the anticodon were deaminated however the efficiency varied and was highest with human and *B. mori* tRNAs.

Because hTad2 and hTad3 are functionally conserved with yeast Tad2p/Tad3p (Gerber and Keller, 1999) and also with *E. coli* tadA (Wolf et al., 2002), it was tried to exchange subunits of the different complexes *in vitro*. Recombinant proteins were pre-incubated and tested for activity with $tRNA^{Ala}$ (A37G) from *B. mori*. In summary, the yeast proteins could not replace the human proteins (Fig. 4.5 B, lanes 5 and 6) although the yeast enzyme is known to deaminate *B. mori* $tRNA^{Ala}$ (Fig. 4.5 B, lane 8; Gerber and Keller, 1999). tadA could also not replace hTad2 *in vitro* (Fig. 4.5 B, lane 7). Interestingly, hTad2 and hTad3 could be replaced *in vitro* with *Drosophila* Tad2 and Tad3, respectively (M.Schaub, pers. communication). To test replacement of Tad2p with hTad2 *in vivo*, a pGalΔTrp-FLIS6-hTad2 expression construct was transformed into the Δ*tad2* yeast strain which carried a copy of *TAD2* on a plasmid with the URA3 marker (Gerber and Keller, 1999). Transformed cells were transferred to plates containing FOA and galactose which forced the loss of the URA3-plasmid and induced hTad2 expression. Cells did not grow on these plates, showing that hTad2 cannot replace Tad2p *in vivo* (results not shown). The absence of activity *in vitro* with hybrid complexes could be due to several reasons. Either the two mixed subunits did not interact or they did interact but could not recognize the substrate

tRNA or they did not form a binding site for the tRNA. With Flag pull-down experiments it was shown that the different subunits could interact. As a positive control hTad2 was shown to interact with hTad3 (Fig. 4.5 C, lane 4). The human Tad2/Tad3 proteins also interacted with the yeast Tad2/Tad3 proteins (M.S. and W.K., unpublished results) showing that a hybrid complex could form but not deaminate tRNA. hTad2 also interacted with tadA (Fig. 4.5 C, lane 8) and with itself (Fig. 4.5 C, lane 6). None of the labelled proteins bound to Flag-agarose (Fig. 4.5 C, lanes 5, 7 and 9). Interestingly, it has been shown previously that tadA can form homodimers (Wolf et al., 2002). Because tadA is likely to be the ancestor of the eukaryotic Tad2/Tad3 proteins, all eukaryotic Tad2 proteins could have the ability to form homodimers. With the appearance of Tad3, which is suggested to be a duplication of Tad2, the complex was a heterodimer, however Tad2 was still able to dimerize. This is also supported by the result that hTad2 interacted with tadA. Thus, the different subunits did interact but could not deaminate tRNA.



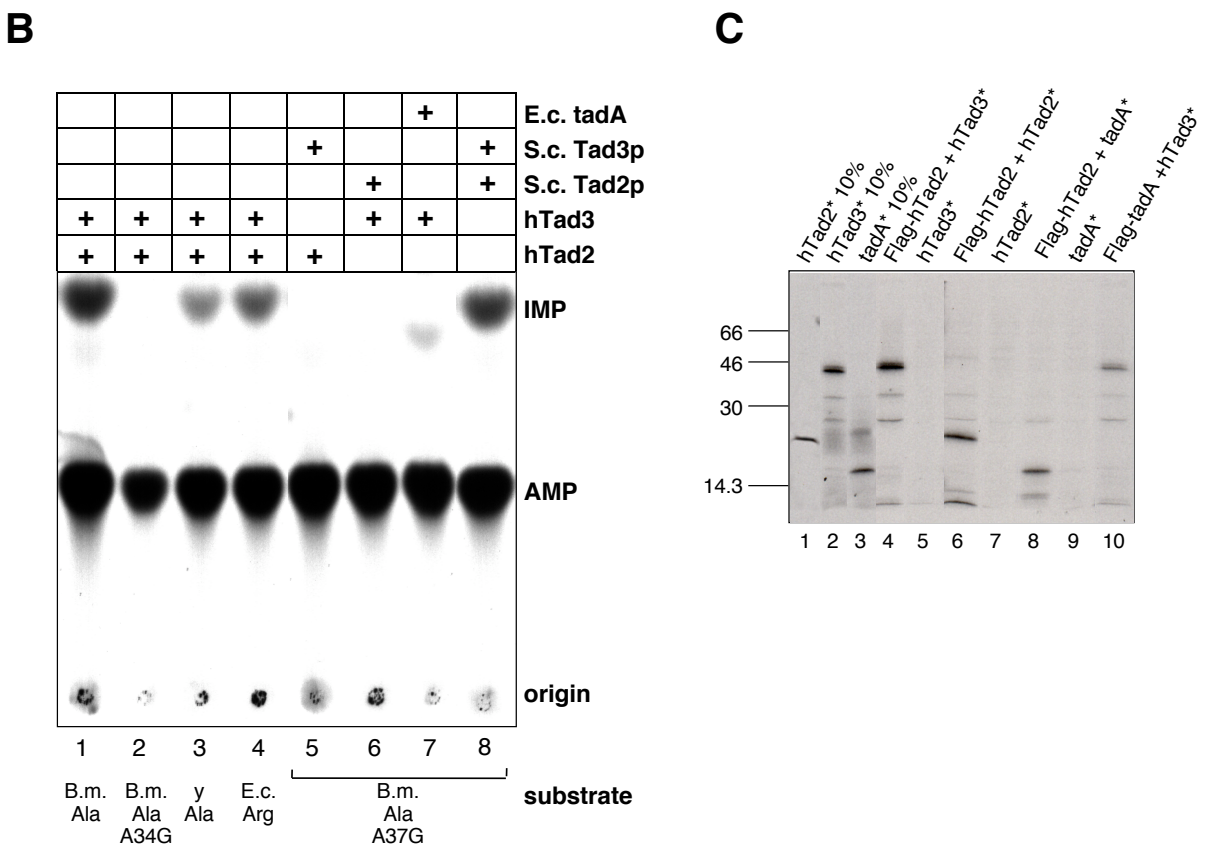



**Figure 4.5 hTad2 and hTad3 interact with other Tad2 proteins, but these complexes are noct active *in vitro*.**

**A)** tRNAs from H. sapiens, *S. cerevisiae* and *E. coli* that contain $I_{34}$. The anticodon sequence is boxed and position 34 is shown in yellow. The invariable $U_{33}$ is boxed in blue. The consensus sequence of the human tRNAs is shown at the bottom. The asterisk at position 35 indicates that there is a purine in all tRNAs except $tRNA^{Arg}$.

**B)** hTad2/hTad3 is specific for position 34 and none of the subunits can be exchanged with Tad2 or Tad3 proteins from other organisms. hTad2/hTad3 was tested for deaminase activity with WT $tRNA^{Ala}$ from *B. mori* (lane 1), *B. mori* $tRNA^{Ala}$ (A34G) (lane 2), *S. cerevisiae* $tRNA^{Ala}$ (lane 3) or *E. coli* $tRNA^{Arg2}$. *B. mori* $tRNA^{Ala}$ (A37G) was used as a substrate to test adenosine deaminase activity of hTad2/yeast Tad3p (lane 5), hTad3/yeast Tad2p (lane 6), hTad3/*E. coli* tadA (lane 7) or yeast Tad2p/yeast Tad3p (lane 8) as a positive control.

**C)** Flag pull-down assay with human and *E. coli* Tad2 and Tad3 proteins. 10% of the labelled protein input is shown in lanes 1, 2 and 3. Asterisks indicate $^{35}S$-labelled proteins. Flag-hTad2 was tested for interaction with hTad3 (lane 4), with hTad2 (lane 6) or with tadA (lane 8). Flag-tadA interaction with hTad3 is shown in lane 10. Mock incubation of labelled proteins with the Flag agarose was tested for hTad3 (lane 5), hTad2 (lane 7) and tadA (lane 9).

## hTad2 contains a NLS and transports hTad3 into the nucleus

A putative NLS was identified in yeast Tad3p (Gerber and Keller, 1999). However, it has not been shown experimentally that this motif really functions as a NLS. Therefore, $hTad2_{1a}$ and $hTad2_{1b}$ were cloned into an expression plasmid adding either a N-terminal red fluorescent protein tag (RFP) or a green fluorescent protein (GFP) tag. A N-terminal GFP-tag was added to

hTad3. Expression and localization of the tagged proteins in HeLa monolayer cells was monitored by confocal microscopy. GFP-hTad2$_{1a}$ was observed mainly in the nucleus and was excluded from the nucleolus (Fig. 4.6 A). A minor fraction of the recombinant protein localized to the cytoplasm. Interestingly, GFP-hTad2$_{1b}$ was detected exclusively in the nucleus and was also excluded from the nucleolus (Fig. 4.6 A). Sequences which determine the different localization of the hTad2 isoforms are most probably within the first exon since this is the only difference between these two proteins (Fig. 4.1 D). GFP-hTad3 was detected in the cytoplasm and was excluded from the nucleus (Fig. 4.6 A). The result that hTad2 and hTad3 localized to different compartments when expressed independently raised the question how the complex could form. The absence of co-localization could be due to the expression of only one subunit of the complex. Therefore, hTad2 and hTad3 were co-expressed from two plasmids with different fluorescent protein tags. For unknown reasons, RFP-hTad2$_{1a}$ could not be co-expressed with hTad3. When co-expressed with hTad2$_{1b}$ hTad3 could also be detected in the nucleus (Fig. 4.6 B), suggesting that hTad3 is transported into the nucleus by forming a complex with hTad2. This result further supports the idea that hTad2 and hTad3 form a tight complex. This was already shown with the Flag pull-down experiments and the co-purification of the recombinant complex.

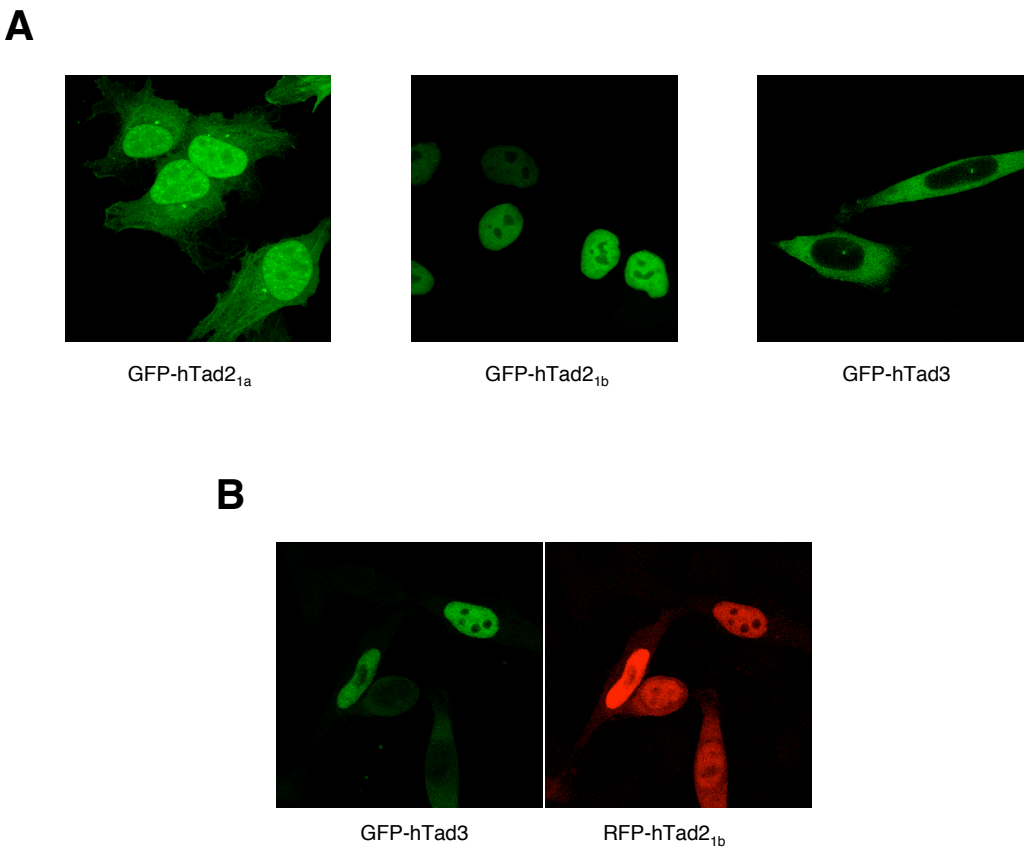


**Figure 4.6 hTad2 localizes to the nucleus and transports hTad3 into the nucleus.**

**A)** HeLa cells were transfected with GFP-tagged $hTad2_{1a}$, $hTad2_{1b}$ or hTad3 as indicated. Cells were analyzed two days after transfection by confocal microscopy.

**B)** HeLa cells were co-transfected with GFP-hTad3 and RFP-$hTad2_{1b}$ and analyzed by confocal microscopy two days after transfection.

## 4.3 Discussion

This chapter describes the identification and characterization of human Tad2 and Tad3. These two proteins are sufficient to reconstitute human tRNA:adenosine 34 deaminase activity *in vitro*. The results in this chapter raise several questions and could help to answer questions that are still open after the identification of the yeast Tad2p/Tad3p complex and *E. coli* tadA.

Two alternatively spliced hTad2 mRNAs were detected by RT-PCR. The two recombinant proteins expressed from these cDNAs differed only in their N-terminal sequence. $hTad2_{1a}$ mRNA encoded the protein in two different reading frames, thus raising the possibility of programmed translational frameshifting. However, frameshifting could not be detected with *in vitro* translation and cell culture expression systems. There are two possible explanations for this result. First, the $hTad2_{1a}$ mRNA does not contain a site for programmed –1 frameshifting. Second, $hTad2_{1a}$ contains a site for programmed frameshifting which could however not be detected with the methods that were used. If frameshifting occurs in $hTad2_{1a}$ mRNA, it could be a mechanism to regulate the expression and also adenosine deamination in tRNAs. It has been shown previously that frameshifting can be regulated. Mammalian ornithine decarboxylase antizyme requires +1 recoding. Antizyme plays a key role in polyamine biosynthesis by binding to ornithine decarboxylase (ODC) which is the key enzyme in polyamine synthesis. This binding targets ODC for degradation by proteasomes. Expression of antizyme is autoregulatory because the required frameshift event is modulated by polyamines. High concentrations of polyamines lead to greater frameshifting, giving more antizyme. Antizyme causes faster turnover of ODC and synthesis of fewer polyamines, thereby completing the autoregulatory circuit ( Matsufuji et al., 1995; Rom and Kahana, 1994; reviewed in Gesteland and Atkins, 1996). A similar mechanism could also apply for hTad2 expression. However, because $hTad2_{1a}$ mRNA was underrepresented compared to $hTad2_{1b}$ mRNA as shown by Northern blot, it is likely that the protein detected in HeLa cell extracts corresponds to $hTad2_{1b}$. It is still possible that a minor fraction of hTad2 is expressed from the $hTad2_{1a}$ mRNA. However, this protein cannot be

distinguished from $hTad2_{1b}$ with the □-hTad2 antibodies. It was tried to isolate hTad2 from HeLa cell extracts with a hTad2 antibody affinity column in order to sequence the N-terminus of the protein. The amount of the isolated hTad2 was not sufficient for sequencing and a major contaminant which had the same size as hTad2 complicated analysis. It is therefore still unclear whether $hTad2_{1a}$ is expressed. Because $hTad2_{1a}$ and $hTad2_{1b}$ have slightly different pI values, HeLa extracts could be tested for $hTad2_{1a}$ by 2-dimensional gel electrophoresis and Western blotting. In case $hTad2_{1b}$ is the only form of the protein, the question arises why the $hTad2_{1a}$ mRNA exists or why the pre-mRNA is spliced between exon 1a and exon 2. In addition, what happens to $hTad2_{1a}$ mRNA in the cells? Is it degraded by the nonsense-mediated decay (NMD) pathway? Pre-mRNAs that contain premature stop codons are eliminated by NMD thereby ensuring that these pre-mRNAs are removed as templates for translation. Translation of $hTad2_{1a}$ without frameshifting would lead to a 47 amino acid protein which could potentially be harmful to the cells. NMD could prevent such an aberrant protein, but then the question remains why splicing generates $hTad2_{1a}$. The splice site consensus sequence is conserved between exon 1a and 2 and a putative branch site was identified. These elements might be sufficient for splicing at this site, but might generate a mRNA that is useless for the cell.

It was shown in chapter 4.2 that hTad2/hTad3 deaminated tRNAs from humans, *B. mori*, *S. cerevisiae* and also from *E. coli*. In eukaryotes seven or eight different tRNAs contain $I_{34}$, therefore the anticodon sequence cannot be a determinant for deamination. An *E. coli* $tRNA^{Arg2}$ minisubstrate consisting of the anticodon arm was not deaminated by hTad2/hTad3. Recognition of the tRNA by eukaryotic Tad2/Tad3 requires therefore sequences and structures outside the anticodon arm. This is not surprising given the fact that the eukaryotic complex has to recognize more tRNAs than *E. coli* tadA. What determines a eukaryotic tRNA as a substrate? Every eukaryotic tRNA that contains A at the wobble position is deaminated, it is therefore possible that Tad2/Tad3 recognizes only the overall tRNA structure and if this structure contains $A_{34}$ it is deaminated. Substrate tRNAs would therefore not be selected actively, but passively by exclusion of the tRNAs that do not contain $A_{34}$. This hypothesis could be tested with an eukaryotic tRNA that is mutated to contain $A_{34}$. Glycin is the only eukaryotic amino acid that is encoded by a family box the codons of which are not read by an $I_{34}$-containing tRNA. One of the three $tRNA^{Gly}$ could be mutated and tested for deamination *in vitro* with recombinant hTad2/hTad3. If indeed the overall tRNA structure is the determinant for Tad2/Tad3, it is also possible that adenosines in tRNA-like structures in mRNAs are deaminated. Tad2/Tad3 could therefore also be involved in pre-mRNA editing and deaminate mRNAs that are not recognized by ADARs. The identification of new editing substrates was a major goal of editing research

during the last years. For ADAR2 it has been shown that the most important editing site is the Q/R site in GluR-B. The other sites targeted by ADAR2 are either not essential or they are edited by ADAR1 due to overlapping substrate specificity. It is possible that hTad2/hTad3 might also contribute to redundancy in pre-mRNA editing by recognizing tRNA-like structures. New editing substrates could be identified by comparing genomic DNA and cDNA sequences or by searching for tRNA-like structures in pre-mRNAs.

## 4.4 Materials and methods

**Table VI hTad2 and hTad3 primers**

| Name | Sequence (5' -> 3') | restriction site (underlined) |
|---|---|---|
| hTad2.1 | GGCGGATCCCTCGAAAATACTGAAGTTCC | BamHI |
| hTad2.2 | CCCAAGCTTCAAGATTTCTGACAATCCTTTTTCCG | HindIII |
| hTad2.5 | CCCGGATCCATGAGCGGCAGCTTCCTGAGG | BamHI |
| hTad2.7 | CCCACTAGTCATGAGTGGAAGCTTCCTGAGG | SpeI |
| hTad2.8 | CCGACTAGTGATTTCTGACAATCCTTTTTCCG | SpeI |
| hTad2.16 | CCCACTAGTCATGGAGGCGAAGGCGGCA | SpeI |
| hTad2.19 | CCCCTCGAGCCATGAGTGGAAGCTTCCTG | XhoI |
| hTad2.20 | CCCCTCGAGCCATGGAGGCGAAGGCGGCA | XhoI |
| 92.SalI | CCCGTCGACGGCTGATACCATACTTCTTGT | SalI |
| 225.SacI | GGCGGAGCTCGGTTTGGTTAACTTCATTTCT | SacI |
| | | |
| hTad3.5 | CCCGGATCCATGATCCTCTGCTCCCGTCTCTGT | BamHI |
| hTad3.8 | TATAAGCTTCAGGCTTGTGCGGAAGAAATCGAG | HindIII |
| hTad3.9 | CCCCTCGAGCCATGATCCTCTGCTCCCGTCTCT | XhoI |

**Table VII Plasmids and strains hTad2 and hTad3**

| Name | Description | Reference |
|---|---|---|
| E. coli BL21 | | Stratagene |

| | | |
|---|---|---|
| codon plus | | |
| pCEP-Pu | expression vector for cell culture | Kohfeldt et al. (1997) |
| pEGFP | cell culture expression vector with GFP tag | Clontech |
| dsRed2 | cell culture expression vector with RFP tag | Clontech |
| p2luc | expression and *in vitro* translation vector to study programmed frameshifting | Grentzmann et al. (1998) |
| pET28 | *E. coli* expression vector | Novagen |
| pJW7 | $hTad2_{1a}$/pTrc-FLIS6 (SpeI) | this study |
| pJW9 | ΔN-$hTad2_{1a}$/pQE9 (BamHI, HindIII) | this study |
| pJW12 | hTad3/pQE9 (BamHI, HindIII) | this study |
| pJW13 | $hTad2_{1a}$ frameshift cDNA/pBS | this study |
| pJW71 | $hTad2_{1a}$ frameshift/pCEP-Pu (KpnI, HindIII) | this study |
| pJW75 | $hTad2_{1a}$ no frameshift/pCEP-Pu (KpnI, HindIII) | this study |
| pJW93 | $hTad2_{1a}$ cDNA 92-225 frameshift/p2luc (SalI, SacI) | this study |
| pJW94 | $hTad2_{1a}$ cDNA 92-225 no frameshift/p2luc | this study |
| pJW114 | $hTad2_{1a}$ cDNA frameshift/pCEP-Pu (NotI, XhoI) | this study |
| pJW115 | $hTad2_{1a}$ cDNA no frameshift/pCEP-Pu (NotI, XhoI) | this study |
| pJW116 | $hTad2_{1a}$ cDNA + intron 1/pCEP-Pu (NotI, XhoI) | this study |
| pJW120 | $hTad2_{1b}$/pTrc-FLIS6 (SpeI) | this study |
| pJW122 | hTad3/pET28 (BamHI, HindIII) | this study |
| pJW133 | $hTad2_{1a}$/pTrc-Flag (SpeI, HindIII) | this study |
| pJW134 | $hTad2_{1b}$/pTrc-Flag (SpeI, HindIII) | this study |
| pJW135 | $hTad2_{1b}$/pEGFP (XhoI, HindIII) | this study |
| pJW136 | hTad3/pEGFP (XhoI, HindIII) | this study |
| pJW139 | $hTad2_{1a}$/pEGFP (XhoI, HindIII) | this study |
| pJW159 | $hTad2_{1a}$/dsRed2 (XhoI, HindIII) | this study |
| pJW160 | $hTad2_{1b}$/dsRed2 (XhoI, HindIII) | this study |

**Cloning of $hTad2_{1a}$**

A HeLa cDNA ZAP II library (Stratagene) was screened with a probe encompassing nucleotides 1-352 of Genbank EST T97494. Approximately $10^6$ clones were screened in duplicates. Filters were hybridized overnight in 5xSSC, 20 mM $NaH_2PO_4$, 1% SDS, 50% formamide, 100 $\mu$g/ml salmon sperm DNA, $2x10^7$ [$\alpha$-$^{32}$P]dATP-labelled hTad2. The following wash steps were done: 2 x 15 min at room temperature in 2xSSC, 0.1% SDS, 2 x 20 min at 65°C in 0.1xSSC, 0.1% SDS,

40 min at 68˚C in 0.1xSSC, 0.1% SDS. 3 positive clones were identified and the corresponding plaques recovered. These clones were re-screened twice and afterwards sequenced. The insert of clone 2 (1514 bp) was used to amplify the hTad2 ORF for expression with primer pairs hTad2.5/hTad2.2, hTad2.7/hTad2.8 and cloned into the BamHI/HindIII and SpeI sites of pQE9 (Qiagen) and pTrc-FLIS6, respectively.

**Cloning of hTad2$_{1b}$**

The hTad2-1b ORF was amplified by PCR from HeLa cDNA with primers hTad2.16/hTad2.8 and subsequently cloned into the SpeI restriction site of pTrc-FLIS6.

**Cloning of hTad3**

The ORF of hTad3 with a 93 nt extension at the 3' end was amplified by PCR from Genbank EST AI367911 with the Expand Long Template kit (Boehringer) and primers hTad3.5/hTad3.8. The PCR mix contained 0.38 mM deaza-dGTP and 0.125 mM dGTP to facilitate amplification of G/C-rich sequences. The PCR product was cloned into the BamHI/HindIII restriction sites of pQE9 (Qiagen) and transformed into BL21-CodonPlus™ *E. coli* cells (Stratagene).

**GFP and RFP constructs**

The ORFs of hTad2$_{1a}$, hTad2$_{1b}$ and hTad3 were amplified by PCR with primer pairs hTad2.19/hTad2.2, hTad2.20/hTad2.2 and hTad3.9/hTad3.8, respectively, and cloned into the XhoI/HindIII sites of pEGFP-C1 (Clontech) adding a N-terminal GFP. For RFP constructs, inserts from the pEGFP-C1 constructs were subcloned into the XhoI/HindIII sites of pDsRed2-C1 (Clontech) adding a N-terminal red fluorescent protein (RFP).

**hTad2 constructs for frameshifting experiments**

hTad2 from nt 40-173 with and without the –1 frameshifting site was amplified with primers 92.SalI/225.SacI and cloned into the SalI/SacI site of pluc (Grentzmann et al., 1998) generating constructs pJW93 and pJW94. hTad2$_{1a}$ ORF with and without the –1 frameshifting site was subcloned into pCEP-Pu (Reference) via KpnI/HindIII restriction sites generating constructs pJW71 and JW75, respectively. hTad2 cDNA from the full-length hTad2$_{1a}$ library clone was subcloned into pCEP-Pu via NotI/XhoI generating constructs pJW114 and pJW115. hTad2 cDNA + intron 1 was generated by digesting a partial genomic hTad2 clone with HindIII, followed by a partial digest with AccI. The correct fragment of ca. 3000 bp was subcloned into

the HindIII/AccI sites of hTad2 cDNA/pBS. The hTad2 cDNA + intron 1 was then cloned into the NotI/XhoI sites of pCEP-Pu generating construct pJW116.

**Expression and purification of recombinant proteins**

All proteins were expressed in BL21 *E. coli* cells and 2xYT medium. 1l cultures were inoculated at an $OD_{600}$ of 0.05-0.1 and grown at 37˚C. Induction was done at $OD_{600}$ 0.7-1.4 with 0.4 mM isopropyl-ß-D-thiogalactopyranoside (IPTG) and growth was continued for 2-5h at 28˚C. Cells were harvested and frozen at –70˚C. The next day cells were resuspended in buffer A (50 mM Tris pH 7.9, 150 mM KCl, 10% glycerol, 0.02% NP40, 2 mM ß-mercaptoethanol, 0.5 mM phenylmethylsulfonyl fluoride, 0.7 $\mu$g/ml pepstatin, 0.4 $\mu$g/ml leupeptin) and sonicated on ice. After centrifugation lysates were mixed with 1 ml of Ni2+/NTA agarose (Qiagen) in buffer A, stirred on ice 45 min and poured into a column. The column was washed with buffer A containing 10 mM imidazole, buffer A with 300 mM KCl, buffer a with 25 mM KCl and the recombinant protein eluted with buffer A containing 250 mM imidazole. Aliquots of each fraction were analyzed by electrophoresis on a SDS-PAA gel and stained with Coomassie Brilliant Blue R-250 (BioRad). Ni-fractions containing recombinant protein were further purified on a anti M2-Flag affinity matrix (Kodak) and eluted with buffer B (50 mM Tris pH 7.9, 25 mM KCl, 10% glycerol, 0.1 mM DTT, 100 $\mu$g/ml Flag peptide (Kodak). Flag-tagged proteins were tested for activity with a tRNA-specific adenosine deaminase assay.

**Generation of polyclonal antibodies against hTad2**

The hTad2 ORF from nt 85-555 was amplified with primers hTad2.1/hTad2.2 and cloned into the BamHI/HindIII restriction sites of the *E. coli* expression vector pQE9. Expression from this constructs yielded a amino-terminal $His_6$-tagged fusion protein. The recombinant protein was expressed in BL21 *E. coli* cells. The culture (1l) was induced with 0.4 mM IPTG at $OD_{600}$ 1.34 and grown further for 2h at 37˚C. Cells were sonicated and the recombinant protein purified under denaturing conditions on a $Ni^{2+}$/NTA affinity column according to the manufacturer's instructions (Qiagen). Fraction 1 was further purified on a large 15% denaturing polyacrylamide gel, stained with $CuCl_2$ and $H_6$-hTad2 cut out. The protein was eluted with 1x Tris-Glycin. A rabbit was immunized 5x with 100 $\mu$g hTad2.

**Depletion of HeLa extracts**

ProteinA-sepharose beads were equilibrated in IPP buffer (10 mM Tris pH 7.9, 150 mM NaCl, 0.1% NP40). 20 $\mu$l packed beads were mixed with 20 $\mu$l □-hTad2 serum (terminal bleeding) or

hTad2 pre-immun serum in a total volume of 500 $\mu$l IPP and were rotated overnight at 4°C. Beads were washed with 3x1 ml IPP buffer, 70 $\mu$l HeLa cytoplasmic extract was added and rotated 2h at 4°C. Extracts were then tested for tRNA adenosine deaminase activity.

**Northern analysis**

Total RNA was isolated from HeLa cells with the RNeasy Kit (Qiagen) according to the manufacturer's instructions. Poly(A)$^+$ RNA was subsequently isolated with an oligo(dT) Sepharose column. For RNase H digestion 8 $\mu$g poly(A)$^+$ RNA was mixed with different oligos as indicated in H-buffer (100 mM KCl, 50 mM Tris pH 7.5, 5 mM MgCl2, 1 mM DTT) and incubated at 70°C for 15 min. The reaction mix was cooled to 37°C and RNase H and RNAguard added. Digestion was done at 37°C for 1 h. The RNA was then phenol/chloroform extracted and separated on a 1.5% agarose-formaldehyde gel. The gel was washed 10 min in water, 20 min in 75 mM NaOH, 2x15 min in 0.5 M Tris pH 7.9, 1.5 M NaCl, 20 min in 10x SSC. The RNA was blotted to nitrocellulose with 10x SSC and crosslinked with a UV Stratalinker (Stratagene). hTad2 mRNA was detected with a hTad2 ORF probe.

**Flag pull-down assay**

The Flag pull-down assays were done as described (chapter 3.5).

**Transfection of HeLa monolayer cells**

HeLa monolayer cells were cultured in Joklik's modified Eagle medium (Sigma) supplemented with 10% newborn calf serum at 37°C in 5% $CO_2$. For transfections $1x10^5$-$3x10^5$ cells were seeded on glass cover slips in 6-well plates. 1.5 $\mu$g plasmid DNA was transfected with PolyFect Transfection Reagent (Qiagen) according to the manufacturer's instructions. Cells were grown for 1-2 more days and prepared for confocal microscopy by fixation in 3.7% paraformaldehyde for 15 min, saturation in 50 mM $NH_4Cl$ for 15 min and permeabilization in 0.2% Triton for 15 min. Double-stranded nucleid acids were stained with TOTO-3 (Molecular Probes) diluted 1:1000 in PBS for 2 min. The cover slips were placed upside down on 5 $\mu$l Vecta Shield (Vector Laboratories) on a microscope slide

# 5 Conclusions and outlook

Since almost 40 years inosine is known to be present in tRNAs. However, the enzymes which catalyze the deamination of the genomically encoded adenosine were identified only recently. This work describes the characterization of *E. coli* tadA which catalyzes the formation of $I_{34}$ in $tRNA^{Arg2}$ and of hTad2/hTad3 which catalyze $I_{34}$ formation in eight human tRNAs. The cloning of Tad proteins allows to study the molecular mechanism of tRNA editing in detail. The identification of tadA and hTad2/hTad3 could help to answer several questions: (i) How do tadA and Tad proteins in general bind the tRNA substrate? (ii) Which amino acids in tadA are needed for tRNA binding and which for catalysis? (iii) What is the role of $I_{34}$? (iv) Is $I_{34}$ formation linked to other processes in the cell, e.g. tRNA export or aminoacylation? (v) Where does tRNA editing occur?

Neither Tad1p nor Tad2p/Tad3p or tadA contain a known RNA binding motif. It has been observed previously that the capacity to bind RNA does not go along with the presence of a known RNA binding motif (e.g. 3' end processing factors, aminoacyl-tRNA synthetases). This suggests that there have to be RNA binding motifs that remain to be identified. Obviously, tadA and hTad2/hTad3 have to bind tRNA to catalyze adenosine deamination. All Tad2 proteins found so far are small proteins with a molecular weight between 20 and 28 kDa. The small size should facilitate the identification of amino acids that are critical for tRNA binding. Such amino acids could be identified by site-specific mutation and the generation and analysis of tadA mutants would certainly give more insight into tRNA recognition and binding. Kinetic parameters of mutant and wildtype tadA could be determined and tadA mutants could also be analyzed *in vivo*. Mutation of highly conserved residues in Tad2p or Tad3p resulted in loss of deaminase activity or the ability to form a heterodimer (A.G., M.S. and W.K., unpublished results). C- and N-terminally deleted Tad2p and Tad3p also had no deaminase activity (M.S., A.G. and W.K., unpublished results). *In vivo* the mutants were lethal except for one mutation. All these mutant proteins were tested for tRNA binding by UV-crosslinking. However, bandshift experiments should also be done to investigate binding since this method allows to observe differences in binding efficiency. Establishing the conditions for bandshifts with tadA (and other Tad proteins) is difficult because these proteins do not bind stably to their substrates but release the tRNA after catalysis. Efficiency of bandshifts could be improved with a tRNA substrate that has a chemically modified nucleotide at position 34 which would trap the enzyme on the substrate. For ADAR2 it has been shown that methylation of the 2'-hydroxyl group at the edited

adenosine significantly decreased the rate of the deamination reaction (Yi-Brunozzi et al., 1999). It might also be possible that tRNA binding of Tads cannot be attributed to a short amino acid sequence but rather involves sequences throughout the protein. Different parts of tadA might bind to different regions of the tRNA. This would be one explanation why no known RNA binding motif is present. It has been shown for human ADAR1 that the NLS overlaps with the third dsRNA binding domain which thus has a dual function. Likewise, the tRNA binding domain of tadA might overlap with other domains, e.g. the deaminase domain and therefore it might be difficult to assign certain functions to certain motifs. For the yeast Tad2p/Tad3p complex it has been proposed that Tad2p catalyzes the deamination and Tad3p is responsible for binding to the tRNA. Because *E. coli* only encodes a Tad2-like protein tRNA binding might be different, also because tadA recognizes only one tRNA and not seven different tRNAs like the yeast complex.

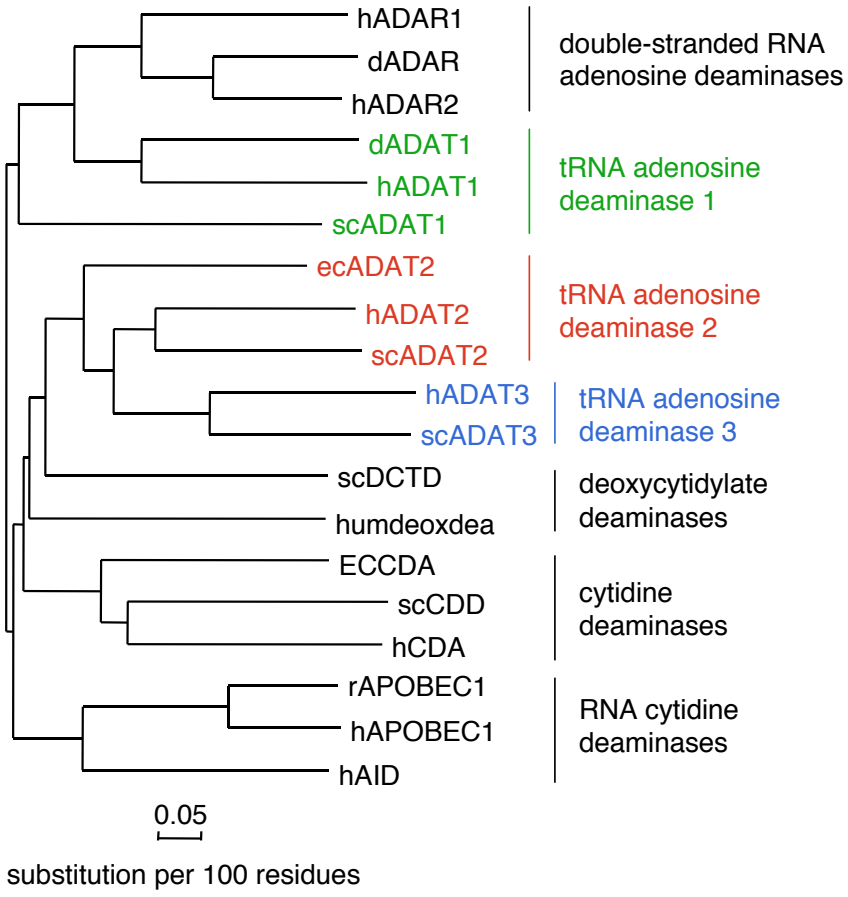



**Figure 5.1 Phylogenetic tree of deaminase domains from adenosine and cytidine deaminases.**

The alignment was generated with the neighbour joining method. Amino acid sequences were used (with numbering starting with the first methionine) of hADAR1 (908-1039), dADAR (326-452), hADAR2 (392-519), dADAT1 (76-194), hADAT1 (85-302), scADAT1 (99-226), ecADAT2 (66-105), hADAT2 (69-114), scADAT2 (52-95),

hADAT3 (237-314), scADAT3 (214-261), scDCTC (231-267), humdeoxdea (82-117), ECCDA (102-136), scCDD (60-103), hCDA (63-106), rAPOBEC1 (59-100), hAPOBEC1 (59-100), hAID (54-94).

It has been shown in chapter 4.2 that Tad proteins from different organisms can interact. This interaction could have its origin in bacteria where tadA is likely to act as a homodimer. Bacterial tadA is proposed to be the ancestor of the eukaryotic Tad2p (Fig. 5.1; Gerber and Keller, 1999; Gerber and Keller, 2001). Tad3 is suggested to have appeared through a genome duplication in eukaryotes and its appearance probably enabled the eukaryotic tRNA:adenosine 34 deaminase to modify additional substrates. Since Tad2p can interact with itself it can also interact with the related Tad3p and even with Tad2 proteins from human and *E. coli*. It was proposed that a Tad2p-like enzyme might have evolved to Tad1p by establishing an additional deaminase motif and changing site-specificity to position 37. Tad1p might then have evolved to the metazoan ADARs by acquiring dsRNA binding domains (Fig. 5.1; Gerber et al., 1998; Gerber and Keller, 1999). All Tads and ADARs are thought to be related and have a conserved deaminase domain. Tad2 also interacts with Tad1, possibly because the two proteins have the same origin and are related. This is supported by the presence of a conserved deaminase domain in both proteins. Furthermore, *D. melanogaster* ADAR can also dimerize (Mary O'Connell, pers. communication). *E. coli* cytidine deaminase and APOBEC1 form homodimers (Betts et al., 1994; Lau et al., 1994b), so dimerization seems to be a general feature of cytidine deaminases and adenosine deaminase that act on RNA and supports the idea that all these enzymes are related (Fig. 5.1).

Why do all these enzmyes form homodimers? Is dimerization needed for catalysis? It has been shown that APOBEC1 forms homodimers in which the active site residues in the deaminase domain are involved in RNA binding. A structural model proposes that the deaminase domain of one subunit targets C at the editing site, whereas the same domain in the other subunit binds to a downstream U in the apoB mRNA (Navaratnam et al., 1998). However, despite extensive mutagenesis, a specific U has not been identified as an essential requirement for apoB editing. Probably, APOBEC1 binds to one of several downstream uridines and is thereby searching for a editing site in a U or A/U-rich context. Mutational analysis confirmed that a homodimeric structure of APOBEC1 is essential for its RNA binding and editing functions (Navaratnam et al., 1998). Mutations that abolished dimerization also eliminated RNA binding and RNA editing. Thus, for APOBEC1 and ECCDA is has clearly been shown that dimerization is a prerequisite for activity and this could also prove correct for tadA. Such a result would further support the hypothesis that tadA evolved from the *E. coli* cytidine deaminase that acts on mononucleotides. It is likely that in eukaryotic tRNA:adenosine 34 deaminases Tad2 is the

catalytic subunit of the complex and that binding and catalysis has been separated into two different subunits. There are two possible explanations for ADAR dimerization to be important. First, as for APOBEC1 dimerization of ADARs could also help the enzyme to find the target A at the editing site and thus contribute to site specificity. Second, ADAR dimers could either be the active or the inactive form of the enzyme. If dimerization inhibits ADARs, this could be a mechanism to regulate pre-mRNA editing. In case of endogenous ADAR overexpression hyperediting could be prevented by dimerization.

A minimal tRNA substrate for tadA has been identified using $tRNA^{Arg2}$ mutant minisubstrates (chapter 3.3 and 3.4). tadA requires a stem-loop structure, the invariant $U_{33}$ and the correct anticodon sequence. Although deamination could be detected with such a substrate, the efficiency was low compared to wild-type tRNA. A possible explanations for this result is that tadA might require other tRNA structures or sequences than the anticodon arm for efficient deamination. tadA could maybe not bind as efficiently to the minisubstrate as to full-length tRNA or tadA could not efficiently dimerize on the minisubstrate. An interesting question is also where site-specificity in the minisubstrates comes from. The minimal requirements $U_{33}$-$A_{34}$-$C_{35}$-$G_{36}$ and stem-loop structure are probably also present in other RNAs. It is therefore likely that additional tRNA sequences contribute to specificity which is supported by the result that more than 1 mol inosine/mol WT minisubstrate is detected with tadA. Interestingly, hTad2/hTad3 is not able to deaminate the arginine minisubstrate although this complex is known to deaminate full-length $tRNA^{Arg2}$ (chapter 4.3). This result further supports the hypothesis that tRNA recognition by Tads is different in bacteria and eukaryotes. It is very likely that hTad2/hTad3 need additional tRNA sequences and/or structures to deaminate $A_{34}$ and this might explain why the eukaryotic complex recognizes seven or eight different tRNAs. Nucleotides in human tRNAs that are important for Tad2/Tad3 binding could be identified with footprint experiments. Potential binding sites could then be verified by mutating specific nucleotides in the tRNA. However, the best way to answer all these questions would be the crystal structure of a tRNA-Tad protein complex. Such a structure would reveal the contacts between the protein and the substrate and could to help to identify determinants for substrate- and site-specificity. Yeast Tad2p/Tad3p could be successfully overexpressed in *E. coli* and purified as a complex, however so far no crystals were obtained. Also for a co-crystal position 34 needs to be modified to trap the enzyme on the substrate.

What is the function of $I_{34}$ in tRNAs? One obvious function is alternative base pairing of these tRNAs with several codons as suggested by Crick (1966). In *E. coli* $tRNA^{Arg2}$ decodes the codons

CGC, CGU and CGA. The three base pairs with I have different structures. I:C is a Watson-Crick base pair and I:U has the same geometry as a G:U pair. I:A is the only purine: purine pair and has therefore a different conformation. Crick (1966) suggested a „long wobble" conformation for the I:A pair and also predicted that A:G and A:A would not form during translation. However, yeast mitochondria $tRNA^{Arg}$ (ACG) and *Mycoplasma* $tRNA^{Thr}$ (AGU) read all four codons of their respective codon family boxes with unmodified $A_{34}$. If $A_{34}$ can base pair with A, C, G and U at the wobble position of the codon, why then is A modified to I in tRNAs that read codons of a four codon family box? By structural modelling it was confirmed that A can base pair with all four nucleotides, however a significant propeller twist is required to read C, G and A which will weaken the anticodon:codon complex (Lim, 1995). Interestingly, the most severe problem occurs after translocation of the tRNA:mRNA complex to the P site of the ribosome. The 6-amino group of $A_{34}$ in all A:N pairs is predicted to interfere with reading the next codon in the A site (Lim, 1995). Thus, the deamination of $A_{34}$ relieves this steric hindrance. However, this potential interference has not been shown experimentally. The A:I base pair is predicted to be inefficient (Curran, 1995), the arginine codon CGA which is read by the I-containing $tRNA^{Arg2}$ is therefore avoided in *E. coli* genes and especially in highly expressed genes.

It is assumed that also *S. cerevisiae* uses I-containing tRNAs to read NNA codons. However, the A-ending codons in four codon family boxes have two cognate isoacceptors, one with inosine and one with a modified uridine which is capable of reading the NNA codon without wobbling. Functional redundancy is avoided in decoding systems and this is supported by the fact that the gene copy number of a certain tRNA is correlated with the abundance of the respective codon and thus the translational need. This means that in the absence of such need due to redundancy the gene number should be zero and this specific tRNA would not be present. Genetic evidence against A:I base pairing in *Schizosaccharomyces pombe* was provided by Munz et al. (1981). Mutation of the two tRNAs reading the serine codon UCA resulted in lethality, suggesting that the remaining I-containing tRNA cannot read this codon. However, the cells might not have survived because the UCA codon is abundant in yeast and translation was not efficient enough with the I-containing tRNA. It was predicted that in *S. cerevisiae* only the arginine codon CGA is read with I:A wobbling due to the higher number of tRNA genes compared to bacteria (Percudani, 2001). This prediction could be tested with yeast *tad2* or *tad3* mutants in which the $tRNA^{Arg}$ is overexpressed from a plasmid. $tRNA^{Arg}$ containing adenosine at position 34 would read codons ending in U and C. Arginine codons ending in A will neither be read by the $I_{34}$- nor by the $C_{34}$-containing tRNA. Therefore, an artificial $tRNA^{Arg}$ with $U_{34}$ could

be expressed in yeast cells. Such an artificial tRNA may rescue the t*ad2*/*tad3* mutants if $tRNA^{Arg}$ is indeed the only tRNA that forms an I:A base pair. However, this experiment includes several critical steps. First, it is not clear whether the artificial $tRNA^{Arg}$ would be recognized by arginyl-tRNA synthetase and aminoacylated. Second, reading of arginine codons ending in C with unmodified $tRNA^{Arg}$ containing $A_{34}$ could not be efficient enough for the cells to survive. Efficiency depends largely on the abundance of this codon. The first base of the anticodon is not necessarily required to form an „standard" base pair with the third position of the codon when the other two base pairs between the anticodon and the codon are G:C base pairs (Lagerkvist, 1978). Because this is the case for the CGN arginine codons, the A:C base pair is thought to be formed. *C. elegans* has twice as many tRNA genes as yeast. Interestingly, four additional species with different anticodons were found in the *C. elegans* genome which are complementary to the four codons in yeast that are read through I:A and U:G wobbling. Thus, it is very likely that I:A base pairing does not occur in higher eukaryotes and the higher number of tRNA genes in these organisms can be explained by a more restricted use of wobbling. It has been shown previously that a tRNA modification is not essential for all tRNAs it is part of. 1-methyladenosine at position 58 ($m^1A_{58}$) is found in 18 yeast tRNAs and its formation is catalyzed by Gcd10p and Gcd14p. The lethality of strains lacking Gcd10p could be complemented by genes encoding $tRNA^{Met}_i$, indicating that modification of this tRNA is the only essential role of Gcd10p in yeast cells (Anderson et al., 1998).

Conditional tadA mutants would be very helpful in elucidating functions of $I_{34}$ *in vivo*. A temperature-sensitive tadA mutant (tadA-2) was generated by mutation of a single amino acid that is conserved in all Tad2 proteins and was previously shown to result in a ts phenotype in yeast (M.S., A.G. and W.K., unpublished results). In *E. coli* the ts phenotype was more pronounced when the mutant protein contained a Flag- and a $His_6$-tag. The cells did not grow at high temperature, however it remains to be determined whether $tRNA^{Arg2}$ was unmodified in this strain. As discussed in chapter 3.6.1 tadA ts mutants are an excellent tool to study the effect of $I_{34}$ *in vivo*. The expression of a reporter construct that is coupled to a series of arginine codons could be analyzed. *E. coli* as a model organism is suited much better for these experiments than yeast because in *tad2*/*tad3* mutants seven tRNAs lack $I_{34}$ and the resulting effects are thus more difficult to analyze. It would be especially difficult to analyze the loss of $I_{34}$ on translation with *tad2*/*tad3* mutants because most likely any reporter gene used (e.g. luciferase, CAT, ß-galactosidase) would not be translated in this mutant since expression of the reporter also depends on I-containing tRNAs. In *E. coli* tadA mutants the reporter gene could be mutated such that it does not encode arginine residues read by $tRNA^{Arg2}$.

It is not known whether the eight human tRNAs that are deaminated by hTad2/hTad3 contain $I_{34}$ at all time. It could be possible that $I_{34}$ formation is regulated developmentally, is organ-specific or could be changed by metabolic stress. Hypomodified tRNAs might translate less efficiently and this in turn might regulate gene expression. Hypomodification could also lead to increased frameshifting or read-through and could thereby affect gene expression. Modified nucleotides in tRNAs not only affect the translational efficiency but also the codon context sensitivity and reading frame maintenance. tRNA modifying enzymes could also have other functions in the cells. The modified nucleotides could act as sensors for environmental stress or developmental signals. It is also not known whether hTad2/hTad3 could deaminate adenosines in tRNA-like structures that are present in pre-mRNAs. tRNAs were integrated into the genome through reverse transcription and are present in protein-coding genes and in short interspersed repetitive elements (SINEs; Brosius, 1999).

Modified nucleotides at position 34 improve the decoding capacity of tRNAs, but decrease the codon sensitivity. However, certain modifications also limit the decoding capacity of the tRNA mainly when the codons read by these tRNAs encode two different amino acids. Modification thus directs base pairing of tRNAs with mRNAs.

Could $I_{34}$ formation be coupled to other processes in the cell, e.g. tRNA export or aminoacylation? tRNAs undergo several maturation steps in the nucleus and are then exported to the cytoplasm by the tRNA-specific export receptor exportin-t (Arts et al., 1998a; Kutay et al., 1998). Mature tRNAs from HeLa cells bound exportin-t 10 times better than *in vitro* transcribed tRNA lacking all modifications (Kutay et al., 1998). It was suggested that modifications and the 3' CCA end of the tRNA contribute to binding specificity. tRNAs were only exported when they had a mature 5' and 3' end, intron-containing tRNAs were exported in the presence of excess exportin-t (Arts et al., 1998b; Lund and Dahlberg, 1998). Interestingly, *in vitro* transcribed tRNAs were exported to the cytoplasm when injected into the nucleus of *Xenopus* oocytes. This suggests that modifications in the tRNAs are not essential for tRNA export, but might increase specificity and efficiency.

For a long time aminoacylation was believed to be restricted to the cytoplasm. However, uncharged tRNAs became aminoacylated upon injection into the nucleus of *Xenopus* oocytes, thus providing clear evidence for nuclear aminoacylation (Lund and Dahlberg, 1998). In *S. cerevisiae* 15 cytoplasmic tRNA synthetases contain a classical nuclear localization signal (Schimmel and Wang, 1999). Nuclear aminoacylation was suggested to be required for efficient tRNA export and might provide a final check after synthesis and processing of the tRNA are

complete. Most modified nucleotides in tRNA are probably not essential for aminoacylation (Björk, 1992). *In vitro* transcribed tRNAs could be aminoacylated with the cognate amino acid in many cases. However, the kinetic characteristics of the reaction were different. Modified nucleotides might affect recognition by the cognate aminoacyl-tRNA ligase either directly or indirectly through conformational changes (Hall et al., 1989). The anticodon of $tRNA^{Arg2}$ is a recognition motif for *E. coli* Arg-tRNA synthetase (Schulman and Pelka, 1989). Since the $tRNA^{Arg}$ isoacceptors have different nucleotides at the wobble position (C, I and $mnm^5U$), this position might not contribute significantly to recognition by the synthetase. An unmodified $tRNA^{Arg}$ with $A_{34}$ or $C_{34}$ instead of $I_{34}$ has the same aminoacylation kinetics as mature $tRNA^{Arg2}$ (Tamura et al., 1992). Thus, $I_{34}$ and $mnm^5U_{34}$ are not essential for aminoacylation of $tRNA^{Arg2}$. Yeast has two $tRNA^{Ile}$, a major tRNA with the anticodon IAU and a minor tRNA with the anticodon □A□. $I_{34}$ and □$_{34}$ contribute significantly to the efficiency of aminoacylation (Senger et al., 1997). Thus, in *S. cerevisiae* $I_{34}$ is critical for correct aminoacylation and a lack of TAD2/TAD3 not only affects translation by the decoding properties of $A_{34}$, but also by the absence of aminoacylation of certain tRNAs.

Where does $I_{34}$ formation occur in human cells? Localization experiments with GFP-tagged hTad2 and hTad3 revealed that hTad2 localizes to the nucleus and transports hTad3 into the nucleus. hTad3 alone is detected in the cytoplasm. A putative NLS was identified in the hTad2 sequence which should be verified by mutation. GFP-hTad2 with a mutated NLS should not localize to the nucleus and co-expression with hTad3 should also result in cytoplasmic accumulation of both proteins. The nuclear localization of the hTad2/hTad3 complex suggests that $I_{34}$ formation occurs in the nucleus and it is likely that exported tRNAs all contain $I_{34}$. By Western blot analysis hTad2 was also detected in the cytoplasm (chapter 4.3). This result is in contrast to the localization experiments with GPF-tagged proteins, however GFP-hTad2 might accumulate exclusively in the nucleus due to the overexpression with the CMV promoter. The endogenous hTad2 might also be found in the cytoplasm. This hypothesis could be verified by immunofluorescence with the □-hTad2 antibodies. If hTad2 indeed is present in both compartments, then the question arises where adenosine deamination occurs. To find an answer to this question *in vitro* transcribed ATP-labelled tRNA could be injected into *Xenopus* oocytes. Two hours later the oocytes are dissected into nuclear and cytoplasmic fraction and the distribution of the tRNA could be analyzed by separating the fractions on denaturing polyacrylamide gels. The tRNAs could also be analyzed for the presence of inosine by thin layer chromatography and autoradiography. If I is detectable in the nuclear tRNA population, tRNA

deamination is also a nuclear process. tRNA deaminase activity was already detected *in vitro* with HeLa cell nuclear extracts which is thus consistent with the presence of hTad2 in the nucleus as determined by Western blot. Since tRNAs seem to be deaminated already in the nucleus, it might be possible that this modification contributes to tRNA quality control before export to the cytoplasm. Although tRNAs are exported without any modified nucleotides, certain modifications could improve the export efficiency, e.g. by increasing the binding efficiency of the tRNA to the export receptor exportin-t. This would also be in agreement with the suggestion that tRNA processing, modification and export are coupled. Due to interactions of the different enzymes that are involved in these processes, such a coupling would „channel“ the tRNA from its site of synthesis to its site of action thereby increasing the processing efficiency.

In summary, the identification and characterization of eukaryotic and prokaryotic adenosine deaminases that act on tRNA allows to investigate different aspects of these enzymes. There are conserved features, but clearly there are also differences, e.g. in tRNA binding. Future research will provide new interesting insights into the formation of inosine in tRNAs.

# 6 List of Figures

# 7 List of Tables

# 8 References


Abraham, J.M., Feagin, J.E. and Stuart, K. (1988) Characterization of cytochrome c oxidase III transcripts that are edited only in the 3' region. *Cell*, **55**, 267-72.

Agris, P.F., Soll, D. and Seno, T. (1973) Biological function of 2-thiouridine in Escherichia coli glutamic acid transfer ribonucleic acid. *Biochemistry*, **12**, 4331-7.

Akbarian, S., Smith, M.A. and Jones, E.G. (1995) Editing for an AMPA receptor subunit RNA in prefrontal cortex and striatum in Alzheimer's disease, Huntington's disease and schizophrenia. *Brain Res*, **699**, 297-304.

Alfonzo, J.D., Thiemann, O. and Simpson, L. (1997) The mechanism of U insertion/deletion RNA editing in kinetoplastid mitochondria. *Nucleic Acids Res*, **25**, 3751-9.

Alfonzo, J.D., Thiemann, O.H. and Simpson, L. (1998) Purification and characterization of MAR1. A mitochondrial associated ribonuclease from Leishmania tarentolae. *J Biol Chem*, **273**, 30003-11.

Altman, S. and Kirsebom, L. (1999) Ribonuclease P. In Gesteland, R.F., Cech, T.R. and Atkins, J.F. (eds.), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Anant, S. and Davidson, N.O. (2001) Molecular mechanisms of apolipoprotein B mRNA editing. *Curr Opin Lipidol*, **12**, 159-65.

Anant, S., Henderson, J.O., Mukhopadhyay, D., Navaratnam, N., Kennedy, S., Min, J. and Davidson, N.O. (2001b) Novel role for RNA-binding protein CUGBP2 in mammalian RNA editing. CUGBP2 modulates C to U editing of apolipoprotein B mRNA by interacting with apobec-1 and ACF, the apobec-1 complementation factor. *J Biol Chem*, **276**, 47338-51.

Anant, S., MacGinnitie, A.J. and Davidson, N.O. (1995) apobec-1, the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme, is a novel RNA-binding protein. *J Biol Chem*, **270**, 14762-14767.

Andachi, Y., Yamao, F., Muto, A. and Osawa, S. (1989) Codon recognition patterns as deduced from sequences of the complete set of transfer RNA species in Mycoplasma capricolum. Resemblance to mitochondria. *J Mol Biol*, **209**, 37-54.

Anderson, J., Phan, L., Cuesta, R., Carlson, B.A., Pak, M., Asano, K., Bjork, G.R., Tamame, M. and Hinnebusch, A.G. (1998) The essential Gcd10p-Gcd14p nuclear complex is required for 1- methyladenosine modification and maturation of initiator methionyl-tRNA. *Genes Dev*, **12**, 3650-62.

Aphasizhev, R., Sbicego, S., Peris, M., Jang, S.H., Aphasizheva, I., Simpson, A.M., Rivlin, A. and Simpson, L. (2002) Trypanosome Mitochondrial 3' Terminal Uridylyl Transferase (TUTase). The Key Enzyme in U-Insertion/Deletion RNA Editing. *Cell*, **108**, 637-48.

Aphasizhev, R. and Simpson, L. (2001) Isolation and characterization of a U-specific 3'-5'-exonuclease from mitochondria of Leishmania tarentolae. *J Biol Chem*, **276**, 21280-4.

Arakawa, H., Hauschild, J. and Buerstedde, J.M. (2002) Requirement of the Activation-Induced Deaminase (AID) Gene for Immunoglobulin Gene Conversion. *Science*, **295**, 1301-6.

Arigoni, A., Talabot, F., Peitsch, M., Edgerton, M.D., Meldrum, e., Allet, E., Fish, R., Jamotte, T., Curchod, M.-L. and Loferer, H. (1998) A genome-based approach for the identification of essential bacterial genes. *Nat Biotechnol.*, **16**, 851-856.

Arts, G.J., Fornerod, M. and Mattaj, I.W. (1998a) Identification of a nuclear export receptor for tRNA. *Curr Biol*, **8**, 305-14.

Arts, G.J., Kuersten, S., Romby, P., Ehresmann, B. and Mattaj, I.W. (1998b) The role of exportin-t in selective nuclear export of mature tRNAs. *EMBO J*, **17**, 7430-41.

Auxilien, S., Crain, P.F., Trewyn, R.W. and Grosjean, H. (1996) Mechanism, specificity and general properties of the yeast enzyme catalysing the formation of inosine 34 in the anticodon of transfer RNA. *J Mol Biol*, **262**, 437-458.

Backus, J.W. and Smith, H.C. (1992) Three distinct RNA sequence elements are required for efficient apolipoprotein B (apoB) RNA editing in vitro. *Nucleic Acids Res*, **20**, 6007-14.

Basilio, C., Wahba, A.J., Lengyel, P., Speyer, J.F. and Ochoa, S. (1962) Synthetic polynucleotides and the amino acid code. *Proc Natl Acad Sci USA*, **48**, 613-616.

Bass, B.L. (1993) New uses for old players in the RNA world. In Gesteland, R.F. and Atkins, J.F. (eds.), *The RNA world*. Cold Spring Harbor Laboratory Press, Plainview, NY, pp. 383-418.

Bass, B.L. (1997) RNA editing and hypermutation by adenosine deamination. *Trends Biochem. Sci.*, **22**, 157-162.

Bass, B.L., Nishikura, K., Keller, W., Seeburg, P.H., Emeson, R.B., O'Connell, M.A., Samuel, C.E. and Herbert, A. (1997) A standardized nomenclature for adenosine deaminases that act on RNA. *RNA*, **3**, 947-949.

Bass, B.L., Weintraub, H., Cattaneo, R. and Billeter, M.A. (1989) Biased hypermutation of viral RNA genomes could be due to unwinding/modification of double-stranded RNA. *Cell*, **56**, 331.

Baumann, U., Fischer, W. and Sprinzl, M. (1985) Analysis of modification-dependent structural alterations in the anticodon loop of Escherichia coli tRNAArg and their effects on the translation of MS2 RNA. *Eur J Biochem*, **152**, 645-9.

Benne, R., Van den Burg, J., Brakenhoff, J., Sloof, P., Van Boom, J.H. and Tromp, M.C. (1986) Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell*, **46**, 819-826.

Bennetzen, J.L. and Hall, B.D. (1982) Codon selection in yeast. *J Biol Chem*, **257**, 3026-31.

Berkhout, B., Das, A.T. and Beerens, N. (2001) HIV-1 RNA editing, hypermutation, and error-prone reverse transcription. *Science*, **292**, 7.

Betts, L., Xiang, S., Short, S.A., Wolfenden, R. and Carter, C.W., Jr. (1994) Cytidine deaminase. The 2.3 A crystal structure of an enzyme: transition-state analog complex. *J Mol Biol*, **235**, 635-656.

Björk, G.R. (1984) Transfer RNA modification in different organisms. *Chemica Scripta*, **26B**, 91-95.

Björk, G.R. (1992) The role of modified nucleosides in tRNA interactions. In Hatfield, L., Lee, B.J. and Pirtle, R.M. (eds.), *Transfer RNA in Protein Synthesis*. CRC Press, Boca Raton, Fla, pp. 23D-85D.

Björk, G.R. (1995) Biosynthesis and function of modified nucleosides. In Söll, D. and RajBhandary, U.L. (eds.), *tRNA: Structure, Biosynthesis and Function.* American Society Microbiology Press, Washington D.C., pp. 165-205.

Björk, G.R. (1996) Stable RNA modification. In Neidhardt, R.C.e.a. (ed.) *Escherichia coli and Salmonella: Cellular and Molecular Biology, 2nd edition.* ASM Press, Washington, D.C., pp. 861-886.

Björk, G.R., Jacobsson, K., Nilsson, K., Johansson, M.J.O., Byström, A.S. and Persson, O.P. (2001) A primordial tRNA modifiaction required for the evolution of life? *EMBO J.*, **20**, 231-239.

Björk, G.R., Wikström, P.M. and Byström, A.S. (1989) Prevention of translational frameshifting by the modified nucleoside 1-methylguanosine. *Science*, **244**, 986-990.

Blanc, V., Navaratnam, N., Henderson, J.O., Anant, S., Kennedy, S., Jarmuz, A., Scott, J. and Davidson, N.O. (2001) Identification of GRY-RBP as an apolipoprotein B RNA-binding protein that interacts with both apobec-1 and apobec-1 complementation factor to modulate C to U editing. *J Biol Chem*, **276**, 10272-83.

Blum, B., Bakalara, N. and Simpson, L. (1990) A model for RNA editing in kinetoplastid mitochondria: "Guide" RNA molecules transcribed from maxicircle DNA provide the edited information. *Cell*, **60**, 189-198.

Boren, T., Elias, P., Samuelsson, T., Claesson, C., Barciszewska, M., Gehrke, C.W., Kuo, K.C. and Lustig, F. (1993) Undiscriminating codon reading with adenosine in the wobble position. *J Mol Biol*, **230**, 739-49.

Bossi, L. and Roth, J.R. (1980) The influence of codon context on genetic code translation. *Nature*, **286**, 123-127.

Bouadloun, F., Srichaiyo, T., Isaksson, L.A. and Bjork, G.R. (1986) Influence of modification next to the anticodon in tRNA on codon context sensitivity of translational suppression and accuracy. *J Bacteriol*, **166**, 1022-7.

Bourara, K., Litvak, S. and Araya, A. (2000) Generation of G-to-A and C-to-U changes in HIV-1 transcripts by RNA editing. *Science*, **289**, 1564-6.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, **72**, 248-254.

Bringaud, F., Stripecke, R., Frech, G.C., Freedland, S., Turck, C., Byrne, E.M. and Simpson, L. (1997) Mitochondrial glutamate dehydrogenase from Leishmania tarentolae is a guide RNA-binding protein. *Mol Cell Biol*, **17**, 3915-23.

Brosius, J. (1999) Transmutation of tRNA over time. *Nat. Genet.*, **22**, 8-9.

Burns, C.M., Chu, H., Rueter, S.M., Hutchinson, L.K., Canton, H., Sanders-Bush, E. and Emeson, R.B. (1997) Regulation of serotonin-2C receptor G-protein coupling by RNA editing. *Nature*, **387**, 303-308.

Cattaneo, R. (1994) Biased (A-I) hypermutation of animal RNA virus genomes. *Current Opinion in Genetics and Development*, **4**, 895-900.

Cavaille, J., Buiting, K., Kiefmann, M., Lalande, M., Brannan, C.I., Horsthemke, B., Bachellerie, J.P., Brosius, J. and Huttenhofer, A. (2000) Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc Natl Acad Sci U S A*, **97**, 14311-6.

Chang, B.H.-J., Lau, P.P. and Chan, L. (1998) Apolipoprotein B mRNA editing. In Grosjean, H. and Benne (eds.), *Modification and editing of RNA*. American Society Microbiology Press, Washington DC., pp. 325-342.

Chang, Z., Nygaard, P., Chinault, A.C. and Kellems, R.E. (1991) Deduced amino acid sequence of *Escherichia coli* adenosine deaminase reveals evolutionarily conserved amino acids residues: Implications for catalytic function. *Biochemistry*, **30**, 2273-2280.

Chen, C.X., Cho, D.S., Wang, Q., Lai, F., Carter, K.C. and Nishikura, K. (2000) A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *Rna*, **6**, 755-767.

Chester, A., Scott, J., Anant, S. and Navaratnam, N. (2000) RNA editing: cytidine to uridine conversion in apolipoprotein B mRNA. *Biochim Biophys Acta*, **1494**, 1-13.

Corell, R.A., Read, L.K., Riley, G.R., Nellissery, J.K., Allen, T.E., Kable, M.L., Wachal, M.D., Seiwert, S.D., Myler, P.J. and Stuart, K.D. (1996) Complexes from Trypanosoma brucei that exhibit deletion editing and other editing-associated properties. *Mol Cell Biol*, **16**, 1410-8.

Crick, F.H.C. (1966) Codon-Anticodon Pairing: the wobble hypothesis. *J Mol Biol*, **19**, 548-555.

Cruz-Reyes, J., Rusche, L.N., Piller, K.J. and Sollner-Webb, B. (1998) T. brucei RNA editing: adenosine nucleotides inversely affect U- deletion and U-insertion reactions at mRNA cleavage. *Mol Cell*, **1**, 401-9.

Curran, J.F. (1995) Decoding with the A:I wobble pair is inefficient. *Nucleic Acids Res*, **23**, 683-8.

Curran, J.F. (1998) Modified nucleosides in translation. In Grosjean, H. and Benne, R. (eds.), *Modification and editing of RNA*. American Society Microbiology Press, Washington DC., pp. 493-516.

Dalphin, M.E., Brown, C.M., Stockwell, P.A. and Tate, W.P. (1997) The translational signal database, TransTerm: more organisms, complete genomes. *Nucleic Acids Res*, **25**, 246-7.

Dao, V., Guenther, R., Malkiewicz, A., Nawrot, B., Sochacka, E., Kraszewski, A., Jankowska, J., Everett, K. and Agris, P.F. (1994) Ribosome binding of DNA analogs of tRNA requires base modifications and supports the "extended anticodon". *Proc Natl Acad Sci U S A*, **91**, 2125-9.

Davanloo, P., Sprinzl, M., Watanabe, K., Albani, M. and Kersten, H. (1979) Role of ribothymidine in the thermal stability of transfer RNA as monitored by proton magnetic resonance. *Nucleic Acids Res*, **6**, 1571-81.

Davidson, N.O. (2002) The challenge of target sequence specificity in C-->U RNA editing. *J Clin Invest*, **109**, 291-4.

Davidson, N.O., Innerarity, T.L., Scott, J., Smith, H., Driscoll, D.M., Teng, B. and Chan, L. (1995) Proposed nomenclature for the catalytic subunit of the mammalian apolipoprotein B mRNA editing enzyme: APOBEC-1. *RNA*, **1**, 3.

Davies, M.S., Wallis, S.C., Driscoll, D.M., Wynne, J.K., Williams, G.W., Powell, L.M. and Scott, J. (1989) Sequence requirements for apolipoprotein B RNA editing in transfected rat hepatoma cells. *J Biol Chem*, **264**, 13395-8.

Davis, D.R. and Poulter, C.D. (1991) 1H-15N NMR studies of Escherichia coli tRNA(Phe) from hisT mutants: a structural role for pseudouridine. *Biochemistry*, **30**, 4223-31.

Deininger, P.L., Batzer, M.A., Hutchison, C.A., 3rd and Edgell, M.H. (1992) Master genes in mammalian repetitive DNA amplification. *Trends Genet*, **8**, 307-11.

Deutscher, M.P. (1984) Processing of tRNA in prokaryotes and eukaryotes. *CRC Crit Rev Biochem*, **17**, 45-71.

Dichtl, B. and Keller, W. (2001) Recognition of polyadenylation sites in yeast pre-mRNAs by cleavage and polyadenylation factor. *EMBO*, **20**, 3197-3209.

Driscoll, D., Wynne, J., Wallis, S. and Scott, J. (1989) An *in vitro* system for the editing of apolipoprotein B mRNA. *Cell*, **58**, 519-525.

Driscoll, D.M., Lakhe-Reddy, S., Oleksa, L.M. and Martinez, D. (1993) Induction of RNA editing at heterologous sites by sequences in apolipoprotein B mRNA. *Mol. Cell. Biol.*, **13**, 7288-7294.

Dunn, D.B. (1959) Additional components in ribonucleic acid of rat-liver fractions. *Biochim. Biophys. Acta*, **34**, 286-288.

Eckmann, C.R., Neunteufl, A., Pfaffstetter, L. and Jantsch, M.F. (2001) The human but not the Xenopus RNA-editing enzyme ADAR1 has an atypical nuclear localization signal and displays the characteristics of a shuttling protein. *Mol Biol Cell*, **12**, 1911-24.

Edmonds, C.G., Crain, P.F., Gupta, R., Hashizume, T., Hocart, C.H., Kowalak, J.A., Pomerantz, S.C., Stetter, K.O. and McCloskey, J.A. (1991) Posttranscriptional modification of tRNA in thermophilic archaea (Archaebacteria). *J Bacteriol*, **173**, 3138-48.

Egebjerg, J., Kukekov, V. and Heinemann, S.F. (1994) Intron sequence directs RNA editing of the glutamate receptor subunit GluR2 coding sequence. *Proc Natl Acad Sci USA*, **91**, 10270-10274.

Ely, B. (1991) Genetics of Caulobacter crescentus. *Methods Enzymol*, **204**, 372-84.

Ericson, J.U. and Bjork, G.R. (1986) Pleiotropic effects induced by modification deficiency next to the anticodon of tRNA from Salmonella typhimurium LT2. *J Bacteriol*, **166**, 1013-21.

Estevez, A.M., Kierszenbaum, F., Wirtz, E., Bringaud, F., Grunstein, J. and Simpson, L. (1999b) Knockout of the glutamate dehydrogenase gene in bloodstream Trypanosoma brucei in culture has no effect on editing of mitochondrial mRNAs. *Mol Biochem Parasitol*, **100**, 5-17.

Estevez, A.M. and Simpson, L. (1999) Uridine insertion/deletion RNA editing in trypanosome mitochondria--a review. *Gene*, **240**, 247-60. bin/cas/tree/store/gene/cas_sub/browse/browse.cgi?year=1999&volume=240&issue=2&aid=12146.

Evinger, M. and Agabian, N. (1977) Envelope-associated nucleoid from Caulobacter crescentus stalked and swarmer cells. *J Bacteriol*, **132**, 294-301.

Fancy, D.A. and Kodadek, T. (1999) Chemistry for the analysis of protein-protein interactions: rapid and efficient cross-linking triggered by long wavelength light. *Proc Natl Acad Sci U S A*, **96**, 6020-4.

Farabaugh, P.J. (1996) Programmed translational frameshifting. *Annu. Rev. Genet.*, **30**, 507-528.

Feagin, J.E., Abraham, J.M. and Stuart, K. (1988a) Extensive editing of the cytochrome c oxidase III transcript in Trypanosoma brucei. *Cell*, **53**, 413-22.

Feagin, J.E., Shaw, J.M., Simpson, L. and Stuart, K. (1988b) Creation of AUG initiation codons by addition of uridines within cytochrome b transcripts of kinetoplastids. *Proc Natl Acad Sci U S A*, **85**, 539-43.

Francklyn, C. and Schimmel, P. (1989) Aminoacylation of RNA minihelices with alanine. *Nature*, **337**, 478-81.

Gerber, A., Grosjean, H., Melcher, T. and Keller, W. (1998) Tad1p, a yeast tRNA-specific adenosine deaminase, is related to the mammalian pre-mRNA editing enzymes ADAR1 and ADAR2. *EMBO J.*, **17**, 4780-4789.

Gerber, A., O'Connell, M.A. and Keller, W. (1997) Two forms of human double-stranded RNA-specific editase 1 (hRED1) generated by the insertion of an Alu cassette. *RNA*, **3**, 453-463.

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

Gerber, A.P. and Keller, W. (2001) RNA editing by base deamination: more enzymes, more targets, new mysteries. *Trends Biochem Sci*, **26**, 376-84.

Gerdes, K., Poulsen, L.K., Thisted, T., Nielsen, A.K., Martinussen, J. and Andreasen, P.H. (1990) The hok killer gene family in gram-negative bacteria. *New Biol*, **2**, 946-956.

Gerdes, K., Rasmussen, P.B. and Molin, S. (1986) Unique type of plasmid maintenance: Postsegregational killing of plasmid-free cells. *Proc Natl Acad Sci*, **83**, 3116-3120.

Gesteland, R.F. and Atkins, J.F. (1996) Recoding: Dynamic reprogramming of translation. *Annu. Rev. Biochem.*, **65**, 741-768.

Gott, J.M. and Emeson, R.B. (2000) Functions and Mechanisms of RNA Editing. *Annu Rev Genet*, **34**, 499-531.

Greeve, J., Altkemper, I., Dieterich, J.H., Greten, H. and Windler, E. (1993) Apolipoprotein B mRNA editing in 12 different mammalian species: hepatic expression is reflected in low concentrations of apoB- containing plasma lipoproteins. *J Lipid Res*, **34**, 1367-83.

Greeve, J., Lellek, H., Rautenberg, P. and Greten, H. (1998) Inhibition of the apolipoprotein B mRNA editing enzyme-complex by hnRNP C1 protein and 40S hnRNP complexes. *Biol Chem*, **379**, 1063-73.

Grentzmann, G., Ingram, J.A., Kelly, P.J., Gesteland, R.F. and Atkins, J.F. (1998) A dual-luciferase reporter system for studying recoding signals. *Rna*, **4**, 479-86.

Grosjean, H., Auxilien, S., Constantinesco, F., Simon, C., Corda, Y., Becker, H.F., Foiret, D., Morin, A., Fournier, M. and Fourrey, J.L. (1996a) Enzymatic formation of adenosine to inosine and to $N^1$-methylinosine in transfer RNAs: A review. *Biochimie*, **78**, 488-501.

Grosjean, H., Edqvist, J., Stråby, K.B. and Giegé, R. (1996b) Enzymatic formation of modified nucleosides in tRNA: dependence on tRNA structure. *J Mol Biol*, **255**, 67-85.

Hagervall, T.G., Ericson, J.U., Esberg, K.B., Li, J.N. and Bjork, G.R. (1990) Role of tRNA modification in translational fidelity. *Biochim Biophys Acta*, **1050**, 263-6.

Hajjar, A.M. and Linial, M.L. (1995) Modification of retroviral RNA by double-stranded RNA adenosine deaminase. *J. of Virology*, **69**, 5878-5882.

Hall, K.B., Sampson, J.R., Uhlenbeck, O.C. and Redfield, A.G. (1989) Structure of an unmodified tRNA molecule. *Biochemistry*, **28**, 5794-801.

Hanrahan, C.J., Palladino, M.J., Ganetzky, B. and Reenan, R.A. (2000) RNA editing of the Drosophila para Na(+) channel transcript. Evolutionary conservation and developmental regulation. *Genetics*, **155**, 1149-60.

Harris, S.G., Sabio, I., Mayer, E., Steinberg, M.F., Backus, J.W., Sparks, J.D., Sparks, C.E. and Smith, H.C. (1993) Extract-specific heterogeneity in high-order complexes containing apolipoprotein B mRNA editing activity and RNA-binding proteins. *J Biol Chem*, **268**, 7382-7392.

Hayman, M.L. and Read, L.K. (1999) Trypanosoma brucei RBP16 is a mitochondrial Y-box family protein with guide RNA binding activity. *J Biol Chem*, **274**, 12067-74.

Herb, A., Higuchi, M., Sprengel, R. and Seeburg, P.H. (1996) Q/R site editing in kainate receptor GluR5 and GluR6 pre-mRNAs requires distant intronic sequences. *Proc Natl Acad Sci USA*, **93**, 1875-1880.

Hersberger, M. and Innerarity, T.L. (1998) Two efficiency elements flanking the editing site of cytidine 6666 in the apolipoprotein B mRNA support mooring-dependent editing. *J Biol Chem*, **273**, 9435-42.

Hersberger, M., Patarroyo-White, S., Arnold, K.S. and Innerarity, T.L. (1999) Phylogenetic analysis of the apolipoprotein B mRNA-editing region. Evidence for a secondary structure between the mooring sequence and the 3' efficiency element. *J Biol Chem*, **274**, 34590-7.

Higuchi, M., Maas, S., Single, F.N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R. and Seeburg, P.H. (2000) Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature*, **406**, 78-81.

Higuchi, M., Single, F.N., Köhler, M., Sommer, B., Sprengel, R. and Seeburg, P.H. (1993) RNA editing of AMPA receptor subunit GluR-B: A base-paired intron-exon structure determines position and efficiency. *Cell*, **75**, 1361-1370.

Hirano, K., Young, S.G., Farese, R.V., Jr., Ng, J., Sande, E., Warburton, C., Powell-Braxton, L.M. and Davidson, N.O. (1996) Targeted disruption of the mouse apobec-1 gene abolishes apolipoprotein B mRNA editing and eliminates apolipoprotein B48. *J Biol Chem*, **271**, 9887-90.

Holley, R.W., Everett, G.A., Madison, J.T. and Zamir, A. (1965) Nucleotide sequences in yeast alanine transfer RNA. *J Biol Chem*, **240**, 2122-2127.

Hou, Y.M. and Schimmel, P. (1988) A simple structural feature is a major determinant of the identity of a transfer RNA. *Nature*, **333**, 140-5.

Hough, R.F., Lingam, A.T. and Bass, B.L. (1999) Caenorhabditis elegans mRNAs that encode a protein similar to ADARs derive from an operon containing six genes. *Nucleic Acids Res*, **27**, 3424-32.

Huang, C.E., Cruz-Reyes, J., Zhelonkina, A.G., O'Hearn, S., Wirtz, E. and Sollner-Webb, B. (2001) Roles for ligases in the RNA editing complex of Trypanosoma brucei: band IV is needed for U-deletion and RNA repair. *EMBO J*, **20**, 4694-703.

Ikemura, T. (1985) Codon usage and tRNA content in unicellular and multicellular organisms. *Mol Biol Evol*, **2**, 13-34.

Innerarity, T.L., Borén, J., Yamanaka, S. and Olofsson, S.-O. (1996) Biosynthesis of apolipoprotein B48-containing lipoproteins. *J. Biol. Chem*, **271**, 2353-2356.

Ishikura, H., Yamada, Y. and Nishimura, S. (1971) Structure of serine tRNA from Escherichia coli. I. Purification of serine tRNA's with different codon responses. *Biochim Biophys Acta*, **228**, 471-81.

Jarmuz, A., Chester, A., Bayliss, J., Gisbourne, J., Dunham, I., Scott, J. and Navaratnam, N. (2002) An Anthropoid-Specific Locus of Orphan C to U RNA-Editing Enzymes on Chromosome 22. *Genomics*, **79**, 285-96.

Johnston, H.M., Barnes, W.M., Chumley, F.G., Bossi, L. and Roth, J.R. (1980) Model for regulation of the histidine operon of Salmonella. *Proc Natl Acad Sci U S A*, **77**, 508-12.

Jukes, T.H. (1973) Possibilities for the evolution of the genetic code from a preceding form. *Nature*, **246**, 22-6.

Kable, M.L., Seiwert, S.D., Heidmann, S. and Stuart, K.D. (1996) RNA editing: a mechanism for gRNA-specified uridylate insertion into precursor mRNA. *Science*, **273**, 1189-1195.

Keegan, L.P., Gallo, A. and O'Connell, M.A. (2001) The many roles of an RNA editor. *Nature reviews*, **2**, 869-879.

Keegan, L.P., Gerber, A.P., Brindle, J., Leemans, R., Gallo, A., Keller, W. and O'Connell, M.A. (2000) The properties of a tRNA-specific adenosine deaminase from Drosophila melanogaster support an evolutionary link between pre-mRNA editing and tRNA modification. *Mol Cell Biol*, **20**, 825-33.

Keller, W., Wolf, J. and Gerber, A. (1999) Editing of messenger RNA precursors and of tRNAs by adenosine to inosine conversion. *FEBS Lett.*, **452**, 71-76.

Kern, D. and Lapointe, J. (1979) Glutamyl transfer ribonucleic acid synthetase of Escherichia coli. Effect of alteration of the 5-(methylaminomethyl)-2-thiouridine in the anticodon of glutamic acid transfer ribonucleic acid on the catalytic mechanism. *Biochemistry*, **18**, 5819-26.

Klingler, T.M. and Brutlag, D.L. (1993) Detection of correlations in tRNA sequences with structural implications. *Proc Int Conf Intell Syst Mol Biol*, **1**, 225-33.

Kohfeldt, E., Maurer, P., Vannahme, C. and Timpl, R. (1997) Properties of the extracellular calcium binding module of the proteoglycan testican. *FEBS Lett*, **414**, 557-61.

Koller, J., Muller, U.F., Schmid, B., Missel, A., Kruft, V., Stuart, K. and Goringer, H.U. (1997) Trypanosoma brucei gBP21. An arginine-rich mitochondrial protein that binds to guide RNA with high affinity. *J Biol Chem*, **272**, 3749-57.

Kramer, G.F., Baker, J.C. and Ames, B.N. (1988) Near-UV stress in Salmonella typhimurium: 4-thiouridine in tRNA, ppGpp, and ApppGpp as components of an adaptive response. *J Bacteriol*, **170**, 2344-51.

Kumar, M. and Carmichael, G.C. (1997) Nuclear antisense RNA induces extensive adenosine modifications and nuclear retention of target transcripts. *Proc Natl Acad Sci USA*, **94**, 3542-3547.

Kutay, U., Lipowsky, G., Izaurralde, E., Bischoff, F.R., Schwarzmaier, P., Hartmann, E. and Gorlich, D. (1998) Identification of a tRNA-specific nuclear export receptor. *Mol Cell*, **1**, 359-69.

Lagerkvist, U. (1978) "Two out of three": an alternative method for codon reading. *Proc Natl Acad Sci U S A*, **75**, 1759-62.

Lai, F., Chen, C.-X., Carter, K.C. and Nishikura, K. (1997) Editing of glutamate receptor B subunit ion channel RNAs by four alternatively spliced DRADA2 double-stranded RNA adenosine deaminases. *Mol. Cell. Biol.*, **17**, 2413-2424.

Lambert, L., Muller, U.F., Souza, A.E. and Goringer, H.U. (1999) The involvement of gRNA-binding protein gBP21 in RNA editing-an in vitro and in vivo analysis. *Nucleic Acids Res*, **27**, 1429-36.

Lau, P.P., Chen, S.H., Wang, J.C. and Chan, L. (1990) A 40 kilodalton rat liver nuclear protein binds specifically to apolipoprotein B mRNA around the RNA editing site. *Nucleic Acids Res.*, **18**, 6781-6786.

Lau, P.P., Xiong, W., Zhu, H.-J., Chen, J.-H. and Chan, L. (1994a) Apolipoprotein B mRNA editing is an intranuclear event that occurs posttranscriptionally coincident with splicing and polyadenylation. *Proc Natl Acad Sci USA*, 8522-8526.

Lau, P.P., Zhu, H.-J., Baldini, A., Charnsangavej, C. and Chan, L. (1994b) Dimeric structure of a human apolipoprotein B mRNA editing protein and cloning and chromosomal localization of its gene. *Proc Natl Acad Sci USA*, **91**, 8522-8526.

Lau, P.P., Zhu, H.J., Nakamuta, M. and Chan, L. (1997) Cloning of an Apobec-1-binding protein that also interacts with apolipoproteinB mRNA and evidence for its involvement in RNA editing. *J Biol Chem*, **272**, 1452-1455.

Lellek, H., Kirsten, R., Diehl, I., Apostel, F., Buck, F. and Greeve, J. (2000) Purification and molecular cloning of a novel essential component of the apolipoprotein B mRNA editing enzyme-complex. *J Biol Chem*, **275**, 19848-56.

Li, J.N. and Bjork, G.R. (1995) 1-Methylguanosine deficiency of tRNA influences cognate codon interaction and metabolism in Salmonella typhimurium. *J Bacteriol*, **177**, 6593-600.

Liao, W., Hong, S.H., Chan, B.H., Rudolph, F.B., Clark, S.C. and Chan, L. (1999) APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. *Biochem Biophys Res Commun*, **260**, 398-404.

Lim, V.I. (1995) Analysis of action of the wobble adenine on codon reading within the ribosome. *J Mol Biol*, **252**, 277-82.

Link, A.J., Phillips, D. and Church, G.M. (1997) Methods for generating precise deletions and insertions in the genome of wild-type Escherichia coli: application to open reading frame characterization. *J Bacteriol*, **179**, 6228-37.

Liu, Y., George, C.X., Patterson, J.B. and Samuel, C. (1997) Functionally distinct double-stranded RNA-binding domains associated with alternative splice site variants of the interferon-inducible double-stranded RNA-specific adenosine deaminase. *J Biol Chem*, **272**, 4419-4428.

Lomeli, H., Mosbacher, J., Melcher, T., Höger, T., Geiger, J.R., Kuner, T., Monyer, H., Higuchi, M., Bach, A. and Seeburg, P.H. (1994) Control of kinetic properties of AMPA receptor channels by nuclear RNA editing. *Science*, **266**, 1709-1713.

Longacre, A. and Storb, U. (2000) A novel cytidine deaminase affects antibody diversity. *Cell*, **102**, 541-4.

Lund, E. and Dahlberg, J.E. (1998) Proofreading and aminoacylation of tRNAs before export from the nucleus. *Science*, **282**, 2082-5.

Lustig, F., Boren, T., Claesson, C., Simonsson, C., Barciszewska, M. and Lagerkvist, U. (1993) The nucleotide in position 32 of the tRNA anticodon loop determines ability of anticodon UCC to discriminate among glycine codons. *Proc Natl Acad Sci U S A*, **90**, 3343-7.

Lustig, F., Elias, P., Axberg, T., Samuelsson, T., Tittawella, I. and Lagerkvist, U. (1981) Codon reading and translational error. Reading of the glutamine and lysine codons during protein synthesis in vitro. *J Biol Chem*, **256**, 2635-43.

Maas, S., Gerber, A.P. and Rich, A. (1999) Identification and characterization of a human tRNA-specific adenosine deaminase related to the ADAR family of pre-mRNA editing enzymes. *Proc Natl Acad Sci USA*, **96**, 8895-8900.

Maas, S., Kim, Y.G. and Rich, A. (2000a) Sequence, genomic organization and functional expression of the murine tRNA-specific adenosine deaminase ADAT1. *Gene*, **243**, 59-66.

Maas, S., Patt, S., Schrey, M. and Rich, A. (2001) Underediting of glutamate receptor GluR-B mRNA in malignant gliomas. *Proc Natl Acad Sci U S A*, **98**, 14687-92.

Maas, S. and Rich, A. (2000b) Changing genetic information through RNA editing. *Bioessays*, **22**, 790-802.

Madison-Antenucci, S., Grams, J. and Hajduk, S.L. (2002) Editing machines. The complexities of trypanosome RNA editing. *Cell*, **108**, 435-8.

Madison-Antenucci, S. and Hajduk, S.L. (2001) RNA editing-associated protein 1 is an RNA binding protein with specificity for preedited mRNA. *Mol Cell*, **7**, 879-86.

Madison-Antenucci, S., Sabatini, R.S., Pollard, V.W. and Hajduk, S.L. (1998) Kinetoplastid RNA-editing-associated protein 1 (REAP-1): a novel editing complex protein with repetitive domains. *EMBO J*, **17**, 6368-76.

Madsen, P., Anant, S., Rasmussen, H.H., Gromov, P., Vorum, H., Dumanski, J.P., Tommerup, N., Collins, J.E., Wright, C.L., Dunham, I., MacGinnitie, A.J., Davidson, N.O. and Celis, J.E. (1999) Psoriasis upregulated phorbolin-1 shares structural but not functional similarity to the mRNA-editing protein apobec-1. *J Invest Dermatol*, **113**, 162-9.

Martin, A., Bardwell, P.D., Woo, C.J., Fan, M., Shulman, M.J. and Scharff, M.D. (2002) Activation-induced cytidine deaminase turns on somatic hypermutation in hybridomas. *Nature*, **415**, 802-6.

Maslov, D.A. and Simpson, L. (1992) The polarity of editing within a multiple gRNA-mediated domain is due to formation of anchors for upstream gRNAs by downstream editing. *Cell*, **70**, 459-67.

Matsufuji, S., Matsufuji, T., Miyazaki, Y., Murakami, Y., Atkins, J.F., Gesteland, R.F. and Hayashi, S. (1995) Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase antizyme. *Cell*, **80**, 51-60.

McClain, W.H. and Foss, K. (1988) Changing the identity of a tRNA by introducing a G-U wobble pair near the 3' acceptor end. *Science*, **240**, 793-6.

Mehta, A. and Driscoll, D.M. (1998) A sequence-specific RNA-binding protein complements apobec-1 To edit apolipoprotein B mRNA. *Mol Cell Biol*, **18**, 4426-32.

Mehta, A., Kinter, M.T., Sherman, N.E. and Driscoll, D.M. (2000) Molecular cloning of apobec-1 complementation factor, a novel RNA-binding protein involved in the editing of apolipoprotein B mRNA. *Mol Cell Biol*, **20**, 1846-54.

Melcher, T., Maas, S., Herb, A., Sprengel, R., Higuchi, M. and Seeburg, P.H. (1996) RED2, a brain specific member of the RNA-specific adenosine deaminase family. *J Biol Chem*, **271**, 31795-31798.

Missel, A. and Goringer, H.U. (1994) Trypanosoma brucei mitochondria contain RNA helicase activity. *Nucleic Acids Res*, **22**, 4050-6.

Moore, J.T., Silversmith, R.E., Maley, G.F. and Maley, F. (1993) T4-phage deoxycytidylate deaminase is a metalloprotein containing two zinc atoms per subunit. *J Biol Chem*, **268**, 2288-2291.

Morrison, J.R., Paszty, C., Stevens, M.E., Hughes, S.D., Forte, T., Scott, J. and Rubin, E.M. (1996) Apolipoprotein B RNA editing enzyme-deficient mice are viable despite alterations in lipoprotein metabolism. *Proc Natl Acad Sci U S A*, **93**, 7154-9.

Morse, D.P. and Bass, B.L. (1999) Long RNA hairpins that contain inosine are present in Caenorhabditis elegans poly(A)+ RNA. *Proc Natl Acad Sci U S A*, **96**, 6048-6053.

Mukhopadhyay, D., Anant, S., Lee, R.M., Kennedy, S., Viskochil, D. and Davidson, N.O. (2002) C-->U editing of neurofibromatosis 1 mRNA occurs in tumors that express both the type II transcript and apobec-1, the catalytic subunit of the apolipoprotein B mRNA-editing enzyme. *Am J Hum Genet*, **70**, 38-50.

Muller, U.F., Lambert, L. and Goringer, H.U. (2001) Annealing of RNA editing substrates facilitated by guide RNA-binding protein gBP21. *EMBO J*, **20**, 1394-404.

Munz, P., Leupold, U., Agris, P. and Kohli, J. (1981) In vivo decoding rules in Schizosaccharomyces pombe are at variance with in vitro data. *Nature*, **294**, 187-8.

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y. and Honjo, T. (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*, **102**, 553-63.

Muramatsu, M., Sankaranand, V.S., Anant, S., Sugai, M., Kinoshita, K., Davidson, N.O. and Honjo, T. (1999) Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J Biol Chem*, **274**, 18470-6.

Muramatsu, T., Nishikawa, K., Nemoto, F., Kuchino, Y., Nishimura, S., Miyazawa, T. and Yokoyama, S. (1988a) Codon and amino-acid specificities of a transfer RNA are both converted by a single post-transcriptional modification. *Nature*, **336**, 179-81.

Muramatsu, T., Yokoyama, S., Horie, N., Matsuda, A., Ueda, T., Yamaizumi, Z., Kuchino, Y., Nishimura, S. and Miyazawa, T. (1988b) A novel lysine-substituted nucleoside in the first position of the anticodon of minor isoleucine tRNA from Escherichia coli. *J Biol Chem*, **263**, 9261-7.

Murphy, D.G., Dimock, K. and Kang, C.Y. (1991) Numerous transitions in human parainfluenza virus 3 RNA recovered from persistently infected cells. *Virology*, **181**, 760-763.

Musier-Forsyth, K. and Schimmel, P. (1993) Aminoacylation of RNA oligonucleotides: minimalist structures and origin of specificity. *Faseb J*, **7**, 282-9.

Nakamuta, M., Chang, B.H., Zsigmond, E., Kobayashi, K., Lei, H., Ishida, B.Y., Oka, K., Li, E. and Chan, L. (1996) Complete phenotypic characterization of apobec-1 knockout mice with a wild-type genetic background and a human apolipoprotein B transgenic background, and restoration of apolipoprotein B mRNA editing by somatic gene transfer of Apobec-1. *J Biol Chem*, **271**, 25981-8.

Navaratnam, N., Bhattacharya, S., Fujino, T., Patel, D., Jarmuz, A.L. and Scott, J. (1995) Evolutionary origins of apoB mRNA editing: catalysis by a cytidine deaminase that has acquired a novel RNA-binding motif at its active site. *Cell*, **81**, 187-195.

Navaratnam, N., Fujino, T., Bayliss, J., Jarmuz, A., How, A., Richardson, N., Somasekaram, A., Bhattacharya, S., Carter, C. and Scott, J. (1998) *Echerichia coli* cytidine deaminase provides a molecular model for apoB RNA editing and a mechanism for RNA substrate recognition. *J Mol Biol*, **275**, 695-714.

Navaratnam, N., Shah, R., Patel, D., Fay, V. and Scott, J. (1993b) Apolipoprotein B mRNA editing is associated with UV crosslinking of proteins to the editing site. *Proc Natl Acad Sci U S A*, **90**, 222-6.

Neidhardt, F.C., *Editor in Chief.* (1996) *Escherichia coli and Salmonella, Cellular and Molecular Biology*. ASM Press, Washington, D.C.

Nishimura, S. (1972) Minor components in transfer RNA: their characterization, location, and function. *Prog Nucleic Acid Res Mol Biol*, **12**, 49-85.

Niswender, C.M., Herrick-Davis, K., Dilley, G.E., Meltzer, H.Y., Overholser, J.C., Stockmeier, C.A., Emeson, R.B. and Sanders-Bush, E. (2001) RNA editing of the human serotonin 5-HT2C receptor. alterations in suicide and implications for serotonergic pharmacotherapy. *Neuropsychopharmacology*, **24**, 478-91.

Noguchi, S., Nishimura, Y., Hirota, Y. and Nishimura, S. (1982) Isolation and characterization of an Escherichia coli mutant lacking tRNA-guanine transglycosylase. Function and biosynthesis of queuosine in tRNA. *J Biol Chem*, **257**, 6544-50.

Nottrott, S., Hartmuth, K., Fabrizio, P., Urlaub, H., Vidovic, I., Ficner, R. and Luhrmann, R. (1999) Functional interaction of a novel 15.5kD [U4/U6.U5] tri-snRNP protein with the 5' stem-loop of U4 snRNA. *EMBO J*, **18**, 6119-33.

O'Hara, P.J., Nihol, S.R., Horodyski, F.M. and Holland, J.J. (1984) Vesicular stomatitis virus defective interfering particles can contain extensive genomic sequence rearrangements and base substitutions. *Cell*, **36**, 915-924.

Okazaki Im, I., Kinoshita, K., Muramatsu, M., Yoshikawa, K. and Honjo, T. (2002) The AID enzyme induces class switch recombination in fibroblasts. *Nature*, **416**, 340-345.

Osawa, S., Jukes, T.H., Watanabe, K. and Muto, A. (1992) Recent evidence for evolution of the genetic code. *Microbiol Rev*, **56**, 229-64.

Palladino, M.J., Keegan, L.P., O'Connell, M.A. and Reenan, R.A. (2000a) dADAR, a Drosophila double-stranded RNA-specific adenosine deaminase is highly developmentally regulated and is itself a target for RNA editing. *Rna*, **6**, 1004-18.

Palladino, M.J., Keegan, L.P., O'Connell, M.A. and Reenan, R.A. (2000b) A-to-I pre-mRNA editing in Drosophila is primarily involved in adult nervous system function and integrity. *Cell*, **102**, 437-49.

Pallank, L., Pak, M. and Schulman, L.H. (1995) tRNA discrimination in aminoacylation. In Söll, D. and RajBhandary, U. (eds.), *tRNA: Structure, biosynthesis and function*. American Society for Microbiology, Washinton, D.C., pp. 371-394.

Palmer, D.T., Blum, P.H. and Artz, S.W. (1983) Effects of the hisT mutation of Salmonella typhimurium on translation elongation rate. *J Bacteriol*, **153**, 357-63.

Panigrahi, A.K., Gygi, S.P., Ernst, N.L., Igo, R.P., Jr., Palazzo, S.S., Schnaufer, A., Weston, D.S., Carmean, N., Salavati, R., Aebersold, R. and Stuart, K.D. (2001b) Association of two novel proteins, TbMP52 and TbMP48, with the Trypanosoma brucei RNA editing complex. *Mol Cell Biol*, **21**, 380-9.

Panigrahi, A.K., Schnaufer, A., Carmean, N., Igo, R.P., Jr., Gygi, S.P., Ernst, N.L., Palazzo, S.S., Weston, D.S., Aebersold, R., Salavati, R. and Stuart, K.D. (2001a) Four related proteins of the Trypanosoma brucei RNA editing complex. *Mol Cell Biol*, **21**, 6833-40.

Patton, D.E., Silvia, T. and Bezanilla, F. (1997) RNA editing generates a diverse array of transcripts encoding squid kv2 $K^+$ channels with altered functional properties. *Neuron*, **19**, 711-722.

Paul, M.S. and Bass, B.L. (1998) Inosine exists in mRNA at tissue-specific levels and is most abundant in brain mRNA. *EMBO J.*, **17**, 1120-1127.

Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A*, **85**, 2444-8.

Percudani, R. (2001) Restricted wobble rules for eukaryotic genomes. *TRENDS in Genetics*, **17**, 133-135.

Peris, M., Simpson, A.M., Grunstein, J., Liliental, J.E., Frech, G.C. and Simpson, L. (1997) Native gel analysis of ribonucleoprotein complexes from a Leishmania tarentolae mitochondrial extract. *Mol Biochem Parasitol*, **85**, 9-24.

Persson, B.C., Gustafsson, C., Berg, D.E. and Bjork, G.R. (1992) The gene for a tRNA modifying enzyme, m5U54-methyltransferase, is essential for viability in Escherichia coli. *Proc Natl Acad Sci U S A*, **89**, 3995-8.

Petschek, J.P., Mermer, M.J., Scheckelhoff, M.R., Simone, A.A. and Vaughn, J.C. (1996) RNA editing in *Drosophila 4f-rnp* gene nuclear transcripts by multiple A-to-G conversions. *J Mol Biol*, **259**, 885-890.

Pieczenik, G. (1980) Predicting coding function from nucleotide sequence or survival of "fitness" of tRNA. *Proc Natl Acad Sci U S A*, **77**, 3539-43.

Piller, K.J., Rusche, L.N., Cruz-Reyes, J. and Sollner-Webb, B. (1997) Resolution of the RNA editing gRNA-directed endonuclease from two other endonucleases of Trypanosoma brucei mitochondria. *RNA*, **3**, 279-90.

Pixa, G., Dirheimer, G. and Keith, G. (1984) Sequence of tRNA Ile IAU from brewer's yeast. *Biochem Biophys Res Commun*, **119**, 905-12.

Pollard, V.W., Harris, M.E. and Hajduk, S.L. (1992) Native mRNA editing complexes from Trypanosoma brucei mitochondria. *EMBO J*, **11**, 4429-38.

Polson, A.G. and Bass, B.L. (1994) Preferential selection of adenosines for modification by double-stranded RNA adenosine deaminase. *EMBO*, **13**, 5701-5711.

Polson, A.G., Bass, B.L. and Casey, J.L. (1996) RNA editing of hepatitis delta virus antigenome by dsRNA-adenosine deaminase. *Nature*, **380**, 454-456.

Polson, A.G., Ley III, H.L., Bass, B.L. and Casey, J.L. (1998) Hepatitis Delta Virus RNA editing is highly specific for the amber/W site and is suppressed by hepatitis delta antigen. *Mol. Cell. Biol.*, **18**, 1919-1926.

Poulsen, L.K., Larsen, N.W., Molin, S. and Andersson, P. (1992) Analysis of an *Escherichia coli* mutant strain resistant to the cell-killing function encoded by the *gef* gene family. *Mol. Microbiol.*, **6**, 895-905.

Prentki, P. and Krisch, H.M. (1984) In vitro insertional mutagenesis with a selectable DNA fragment. *Gene*, **29**, 303-13.

Qian, Q. and Björk, G.R. (1997) Structural alterations far from the anticodon of the $tRNA^{Pro}$ of *Salmonella typhimurium* induce +1 frameshifting at the peptidyl-site. *J Mol Biol*, **273**, 978-992.

Reece, K.S. and Phillips, G.J. (1995) New plasmids carrying antibiotic-resistance cassettes. *Gene*, **165**, 141-2.

Revy, P., Muto, T., Levy, Y., Geissmann, F., Plebani, A., Sanal, O., Catalan, N., Forveille, M., Dufourcq-Labelouse, R., Gennery, A., Tezcan, I., Ersoy, F., Kayserili, H., Ugazio, A.G., Brousse, N., Muramatsu, M., Notarangelo, L.D., Kinoshita, K., Honjo, T., Fischer, A. and Durandy, A. (2000) Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell*, **102**, 565-75.

Richardson, N., Navaratnam, N. and Scott, J. (1998) Secondary structure for the apolipoprotein B mRNA editing site. Au- binding proteins interact with a stem loop. *J Biol Chem*, **273**, 31707-17.

Rom, E. and Kahana, C. (1994) Polyamines regulate the expression of ornithine decarboxylase antizyme in vitro by inducing ribosomal frame-shifting. *Proc Natl Acad Sci U S A*, **91**, 3959-63.

Rould, M.A., Perona, J.J., Soll, D. and Steitz, T.A. (1989) Structure of E. coli glutaminyl-tRNA synthetase complexed with tRNA(Gln) and ATP at 2.8 A resolution. *Science*, **246**, 1135-42.

Rueter, S.M., Burns, C.M., Coode, S.A., Mookherjee, P. and Emeson, R.B. (1995) Glutamate receptor RNA editing *in vitro* by enzymatic conversion of adenosine to inosine. *Science*, **267**, 1491-1494.

Rueter, S.M., Dawson, T.R. and Emeson, R.B. (1999) Regulation of alternative splicing by RNA editing. *Nature*, **399**, 75-80.

Rueter, S.M. and Emeson, R.B. (1998) Adenosine-to-inosine conversion in mRNA. In Grosjean, H. and Benne, R. (eds.), *Modification and editing of RNA*. American Society Microbiology Press, Washington D.C., pp. 343-361.

Rusche, L.N., Cruz-Reyes, J., Piller, K.J. and Sollner-Webb, B. (1997) Purification of a functional enzymatic editing complex from Trypanosoma brucei mitochondria. *EMBO J*, **16**, 4069-81.

Salavati, R., Panigrahi, A.K., Morach, B.A., Palazzo, S.S., Igo, R.P. and Stuart, K. (2002) Endoribonuclease activities of Trypanosoma brucei mitochondria. *Mol Biochem Parasitol*, **120**, 23-31.

Sambrook, S., Fritsch, E.F. and Maniatis, T. (1989) *Molecular cloning- a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Saneyoshi, M. and Nishimura, S. (1971) Selective inactivation of amino acid acceptor and ribosome-binding activities of Escherichia coli tRNA by modification with cyanogen bromide. *Biochim Biophys Acta*, **246**, 123-31.

Schade, M., Turner, C.J., Lowenhaupt, K., Rich, A. and Herbert, A. (1999) Structure-function analysis of the Z-DNA-binding domain Zalpha of dsRNA adenosine deaminase type I reveals similarity to the (alpha + beta) family of helix-turn-helix proteins. *EMBO J*, **18**, 470-9.

Schaefer, K.P., Altman, S. and Soll, D. (1973) Nucleotide modification in vitro of the precursor of transfer RNA of Escherichia coli. *Proc Natl Acad Sci U S A*, **70**, 3626-30.

Schimmel, P. and Alexander, R. (1998) Diverse RNA substrates for aminoacylation: clues to origins? *Proc Natl Acad Sci U S A*, **95**, 10351-3.

Schimmel, P. and Wang, C.C. (1999) Getting tRNA synthetases into the nucleus. *Trends Biochem Sci*, **24**, 127-8.

Schnaufer, A., Panigrahi, A.K., Panicucci, B., Igo, R.P., Jr., Wirtz, E., Salavati, R. and Stuart, K. (2001) An RNA ligase essential for RNA editing and survival of the bloodstream form of Trypanosoma brucei. *Science*, **291**, 2159-62.

Schock, D., Kuo, S.R., Steinburg, M.F., Bolognino, M., Sparks, J.D., Sparks, C.E. and Smith, H.C. (1996) An auxiliary factor containing a 240-kDa protein complex is involved in apolipoprotein B RNA editing. *Proc Natl Acad Sci USA*, **93**, 1097-1102.

Schulman, L.H. and Pelka, H. (1988) Anticodon switching changes the identity of methionine and valine transfer RNAs. *Science*, **242**, 765-8.

Schulman, L.H. and Pelka, H. (1989) The anticodon contains a major element of the identity of arginine transfer RNAs. *Science*, **246**, 1595-7.

Schwartz, R. and Curran, J.F. (1997) Analyses of frameshifting at UUU-pyrimidine sites. *Nucleic Acids Res*, **25**, 2005-11.

Seeburg, P.H. (1996) The role of RNA editing in controlling glutamate receptor channel properties. *J. Neurochem.*, **66**, 1-5.

Seeburg, P.H., Higuchi, M. and Sprengel, R. (1998) RNA editing of brain glutamate receptor channels: mechanism and physiology. *Brain Res. Rev.*, **26**, 217-219.

Seiwert, S.D., Heidmann, S. and Stuart, K. (1996) Direct visualization of uridylate deletion *in vitro* suggests a mechanism for kinetoplastid RNA editing. *Cell*, **84**, 831-841.

Semenov, E.P. and Pak, W.L. (1999) Diversification of Drosophila chloride channel gene by multiple posttranscriptional mRNA modifications. *J Neurochem*, **72**, 66-72.

Senger, B., Auxilien, S., Englisch, U., Cramer, F. and Fasiolo, F. (1997) The modified wobble base inosine in yeast tRNAIle is a positive determinant for aminoacylation by isoleucyl-tRNA synthetase. *Biochemistry*, **36**, 8269-75.

Seno, T., Agris, P.F. and Soll, D. (1974) Involvement of the anticodon region of Escherichia coli tRNAGln and tRNAGlu in the specific interaction with cognate aminoacyl-tRNA synthetase. Alteration of the 2-thiouridine derivatives located in the anticodon of the tRNAs by BrCN or sulfur deprivation. *Biochim Biophys Acta*, **349**, 328-38.

Shah, R.R., Knott, T.J., Legros, J.E., Navaratnam, N., Greeve, J.C. and Scott, J. (1991) Sequence requirements for the editing of apolipoprotein B mRNA. *J Biol Chem*, **266**, 16301-4.

Shapiro, T.A. and Englund, P.T. (1995) The structure and replication of kinetoplast DNA. *Annu Rev Microbiol*, **49**, 117-43.

Shaw, J.M., Campbell, D. and Simpson, L. (1989) Internal frameshifts within the mitochondrial genes for cytochrome oxidase subunit II and maxicircle unidentified reading frame 3 of Leishmania tarentolae are corrected by RNA editing: evidence for translation of the edited cytochrome oxidase subunit II mRNA. *Proc Natl Acad Sci U S A*, **86**, 6220-4.

Shaw, J.M., Feagin, J.E., Stuart, K. and Simpson, L. (1988) Editing of kinetoplastid mitochondrial mRNAs by uridine addition and deletion generates conserved amino acid sequences and AUG initiation codons. *Cell*, **53**, 401-11.

Sibler, A.P., Dirheimer, G. and Martin, R.P. (1986) Codon reading patterns in Saccharomyces cerevisiae mitochondria based on sequences of mitochondrial tRNAs. *FEBS Lett*, **194**, 131-8.

Simpson, L. (1987) The mitochondrial genome of kinetoplastid protozoa: genomic organization, transcription, replication, and evolution. *Annu Rev Microbiol*, **41**, 363-82.

Skuse, G.R., Cappione, A.J., Sowden, M., Metheny, L.J. and Smith, H.C. (1996) The neurofibromatosis type 1 messenger RNA undergoes base-modification RNA editing. *Nucl. Acids Res.*, **24**, 478-486.

Smith, H.C., Gott, J.M. and Hanson, M.R. (1997) A guide to RNA editing. *RNA*, **3**, 1105-1123.

Smith, H.C. and Sowden, M.P. (1996) Base-modification mRNA editing through deamination - the good, the bad and the unregulated. *Trends Genet.*, **12**, 418-424.

Smith, J.D. and Dunn, D.B. (1959) The occurrence of methylated guanines in ribonucleic acids from several sources. *Biochem. J.*, **72**, 294-301.

Smith, L.A., Wang, X., Peixoto, A.A., Neumann, E.K., Hall, L.M. and Hall, J.C. (1996) A Drosophila calcium channel alpha1 subunit gene maps to a genetic locus associated with behavioral and visual defects. *J Neurosci*, **16**, 7868-79.

Sodhi, M.S., Burnet, P.W., Makoff, A.J., Kerwin, R.W. and Harrison, P.J. (2001) RNA editing of the 5-HT(2C) receptor is reduced in schizophrenia. *Mol Psychiatry*, **6**, 373-9.

Sommer, B., Köhler, M., Sprengel, R. and Seeburg, P.H. (1991) RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. *Cell*, **67**, 11-19.

Sowden, M.P., Eagleton, M.J. and Smith, H.C. (1998) Apolipoprotein B RNA sequence 3' of the mooring sequence and cellular sources of auxiliary factors determine the location and extent of promiscuous editing. *Nucleic Acids Res*, **26**, 1644-52.

Spanjaard, R.A., Chen, K., Walker, J.R. and van Duin, J. (1990) Frameshift suppression at tandem AGA and AGG codons by cloned tRNA genes: assigning a codon to argU tRNA and T4 tRNA(Arg). *Nucleic Acids Res*, **18**, 5031-6.

Sprinzl, M., Horn, C., Brown, M., Ioudovitch, A. and Steinberg, S. (1998) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res.*, **26**, 148-153.

Stern, L. and Schulman, L.H. (1977) Role of anticodon bases in aminoacylation of Escherichia coli methionine transfer RNAs. *J Biol Chem*, **252**, 6403-8.

Stern, L. and Schulman, L.H. (1978) The role of the minor base N4-acetylcytidine in the function of the Escherichia coli noninitiator methionine transfer RNA. *J Biol Chem*, **253**, 6132-9.

Sylvers, L.A., Rogers, K.C., Shimizu, M., Ohtsuka, E. and Soll, D. (1993) A 2-thiouridine derivative in tRNAGlu is a positive determinant for aminoacylation by Escherichia coli glutamyl-tRNA synthetase. *Biochemistry*, **32**, 3836-41.

Szweykowska-Kulinska, Senger, B., Keith, G., Fasiolo, F. and Grosjean, H. (1994) Intron-dependent formation of pseudouridines in the anticodon of Saccharomyces cerevisiae minor tRNA(Ile). *EMBO J.*, **13**, 4636-44.

Tamura, K., Himeno, H., Asahara, H., Hasegawa, T. and Shimizu, M. (1992) In vitro study of E.coli tRNA(Arg) and tRNA(Lys) identity elements. *Nucleic Acids Res*, **20**, 2335-9.

ten Dam, E.B., Pleij, C.W. and Bosch, L. (1990) RNA pseudoknots: translational frameshifting and readthrough on viral RNAs. *Virus Genes*, **4**, 121-36.

Teng, B., Burant, C.F. and Davidson, N.O. (1993) Molecular cloning of an apolipoprotein B messenger RNA editing protein. *Science*, **260**, 1816-1819.

Thomas, G. and Favre, A. (1980) 4-Thiouridine triggers both growth delay induced by near-ultraviolet light and photoprotection. *Eur J Biochem*, **113**, 67-74.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, **22**, 4673-80.

Urbonavicius, J., Qian, Q., Durand, J.M., Hagervall, T.G. and Bjork, G.R. (2001) Improvement of reading frame maintenance is a common function for several tRNA modifications. *EMBO J*, **20**, 4863-73.

Vanhamme, L., Perez-Morga, D., Marchal, C., Speijer, D., Lambert, L., Geuskens, M., Alexandre, S., Ismaili, N., Goringer, U., Benne, R. and Pays, E. (1998) Trypanosoma brucei TBRGG1, a mitochondrial oligo(U)-binding protein that co-localizes with an in vitro RNA editing activity. *J Biol Chem*, **273**, 21825-33.

Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R. and Ficner, R. (2000) Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Mol Cell*, **6**, 1331-42.

Wagner, R.W., Yoo, C., Wrabetz, L., Kamholz, J., Buchhalter, J., Hassan, N.F., Khalili, K., Kim, S.U., Perussia, B., McMorris, F.A. and Nishikura, K. (1990) Double-stranded RNA unwinding and modifying activity is detected ubiquitously in primary tissues and cell lines. *Mol. Cell. Biol.*, **10**, 5586-5590.

Wang, Q., Khillan, J., Gadue, P. and Nishikura, K. (2000) Requirement of the RNA editing deaminase ADAR1 gene for embryonic erythropoiesis. *Science*, **290**, 1765-8.

Wilson, D.K., Rudolph, F.B. and Quiocho, F.A. (1991) Atomic structure of adenosine deaminase complexed with a transition-state analog: Understanding catalysis and immunodeficiency mutations. *Science*, **252**, 1278-1284.

Wolf, J., Gerber, A.P. and Keller, W. (2002) tadA, an essential tRNA-specific adenosine deaminase from *Escherichia coli*. *EMBO J.,* **21**, 3841-51.

Wong, S.K., Sato, S. and Lazinski, D.W. (2001) Substrate recognition by ADAR1 and ADAR2. *RNA*, **7**, 846-58.

Yamanaka, S., Balestra, M.E., Ferrell, L.D., Fan, J., Arnold, K.S., Taylor, S., Taylor, J.M. and Innerarity, T.L. (1995) Apolipoprotein B mRNA-editing protein induces hepatocellular carcinoma and dysplasia in transgenic animals. *Proc Natl Acad Sci U S A*, **92**, 8483-7.

Yamanaka, S., Poksay, K.S., Arnold, K.S. and Innerarity, T.L. (1997) A novel translational repressor mRNA is edited extensively in livers containing tumors caused by the transgene expression of the apoB mRNA-editing enzyme. *Genes Dev*, **11**, 321-33.

Yamanaka, S., Poksay, K.S., Balestra, M.E., Zeng, G.Q. and Innerarity, T.L. (1994) Cloning and mutagenesis of the rabbit ApoB mRNA editing protein. A zinc motif is essential for catalytic activity, and noncatalytic auxiliary factor(s) of the editing complex are widely distributed. *J Biol Chem*, **269**, 21725-21734.

Yang, C., Carlow, D., Wolfenden, R. and Short, S.A. (1992) Cloning and nucleotide sequence of the *Escherichia coli* cytidine deaminase (ccd) gene. *Biochemistry*, **31**, 4168-4174.

Yarus, M. (1982) Translational efficiency of transfer RNA's: uses of an extended anticodon. *Science*, **218**, 646-52.

Yarus, M. and Barrell, B.G. (1971) The sequence of nucleotides in tRNA Ile from E. coli B. *Biochem Biophys Res Commun*, **43**, 729-34.

Yarus, M., Cline, S.W., Wier, P., Breeden, L. and Thompson, R.C. (1986) Actions of the anticodon arm in translation on the phenotypes of RNA mutants. *J Mol Biol*, **192**, 235-55.

Yi-Brunozzi, H.Y., Easterwood, L.M., Kamilar, G.M. and Beal, P.A. (1999) Synthetic substrate analogs for the RNA-editing adenosine deaminase ADAR-2. *Nucleic Acids Res*, **27**, 2912-7.

Yi-Brunozzi, H.Y., Stephens, O.M. and Beal, P.A. (2001) Conformational Changes That Occur during an RNA-editing Adenosine Deamination Reaction. *J Biol Chem*, **276**, 37827-33.

Yokoyama, S. and Nishimura, S. (1995) Modified nucleosides and codon recognition. In Söll, D. and RajBhandary, U. (eds.), *tRNA: Structure, biosynthesis and function*. American Society for Microbiology, Washington, D.C., pp. 207-233.

# 9 Acknowlegdments

I thank Walter Keller for giving me the opportunity to do my PhD thesis in his lab and for his continuous support during all these years.

André Gerber introduced me to RNA editing and supervised me with a lot of patience especially during my diploma thesis. He contributed also to my PhD thesis with many helpful ideas and suggestions.

Mary O'Connell is a RNA editing enthusiast with many excellent ideas. It was a lot of fun listening to her stories about god and the world (and people from the RNA editing community).

I thank all the present (and of course also the past) members of the Keller lab who I had the pleasure to meet. Present members (in alphabetical order): Diana Blank, Sabine Dettwiler, Bernhard Dichtl, Monika Garas, Isabelle Kaufmann, Andrea Kyburz, Georges Martin, Bertrand Paguet, Salome Röck, Martin Sadowski, Myriam Schaub, Vreni Widmer.
Alumni: Silvia Barabino, Hildegard Etz, Wolfgang Hübner, Sabine Krause, Martin Ohnacker, Ursula Rüegsegger, Henk de Vries, Stefan Weiser, Thomas Wiederkehr.

What would life be without friends? They might not realize how important their friendship was as a contribution to this work. It was often very stimulating that many of them are not scientists.

Ä spezielle Dank geit a mini Eltere, wo mi immer unterstützt hei während minere usbiudig und immer a mi gloubt hei. Ig bi froh, dass sie mi mini eigete Idee hei la umsetze u mir nid öppis ufzwunge hei.

The last and most private thank goes to Bruno Stiefel. Thanks for being the other half of the whole.

## Der Danksagung zweiter Teil (nur für Insider)

Dies ist der inoffizielle Teil der Danksagung, der wie so häufig viel informativer sein wird als der offizielle.
Mit dem diesjährigen Frauenförderdungspreis anerkennen wir hiermit die Leistungen von Walter Keller. Seine Arbeitsgruppe ist mit einem Frauenanteil von 70% ein leuchtendes Beispiel dafür, wie konsequente Frauenförderung sich ausbezahlt. Gewisse kritische Stimmen lassen aber durchblicken, dass der hohe Frauenanteil auch negative Seiten hat und die Gruppe Gefahr läuft, nicht mehr „hormonally balanced" zu sein...
Ein gewisser Tee-Fanatismus trat vor allem in den Labors 277 und 278 auf, mit wechselnder Vorliebe für aromatisierten Schwarztee, Grüntee oder den hippen Roiboos Tee. Gewisse Verschleisserscheinungen bei den Wasserkochern führten zu wiederkehrenden Diskussionen über Superpunkte, Design by Porsche und Samowar. Elektrotechnik ist halt nicht jedermanns Sache und ist vor allem dann gefährlich wenn es unerwartet knallt. Wasserkocher sind übrigens auch hervorragend geeignet um Bettflaschen (praktisch, nicht wahr Sabine?) nachzufüllen.
Ebenso häufig wie das Wasserkochen war in diesen Labors das Telefonklingeln, was gewisse Leute dazu veranlasste, sofort alles stehen und liegen zu lassen.

In diesem Abschnitt darf natürlich keineswegs Walter's pathologische Liebe zu Ramseiers Apfelsaft fehlen, sowie sein kaum übersehbarer Drang immer genügend SV-Servietten im Büro zu haben, um weinende Doktorandinnen ärztlich versorgen zu können. Vermutlich eine direkte Konsequenz des Frauenanteils (siehe oben). In der Hoffnung, dass Walter's Frau diesen Abschnitt nicht zu sehen bekommt, möchte ich nur ganz am Rand erwähnen, dass wir als Gruppe es vor ihr geschafft haben von Walter ins Restaurant Donati eingeladen worden zu sein. Vielen Dank – auch für die sonst nicht zu seltenen Einladungen ins Da Gianni (St. Johann) und andere Gourmettempel in Basel und Umgebung.
Kaum vergessen werde ich auch Walters Lieblingsstatement „Um GOTTES Willen!", welches nur er auf seine spezielle Art zu intonieren vermag. Dieses alternierend mit dem englischen Pendant – „Oh my goodness!". Unvergessen bleiben auch seine nächtlichen Alpträume, in denen in wechselnder Reihenfolge seine DoktorandInnen vorkamen und die ihn aus dem Schlaf schrecken liessen.
Walter hat es auch immer wieder geschafft, aus unerklärlichen Gründen seinen Computer abstürzen zu lassen, was ihn allerdings nicht davon abhalten konnte, nach 35 Jahren Folien-Präsentationen doch noch auf Powerpoint-Präsentationen umzusteigen.
Walter's Tendenz zur Nervosität, die sich vor Vorträgen, Vorlesungen, Meetings, vor dem Eintreffen von Referee Reports und einer Reihe weiterer Ereignisse äusserte, wurde hervorragend ausgeglichen durch die stoische Ruhe von André, der nur bei negativen Resultaten von Experimenten gewisse Emotionen durchblicken liess. Unvergessen bleiben seine Versuche, Minipreps mit dem falschen Kit zu sequenzieren oder die Transkription mit der T7 DNA Polymerase durchzuführen. Solche Patzer führten dazu, dass Prof. Gerber tatsächlich ein Experiment wiederholen musste, sehr zu seinem Erstaunen.

Ein spezieller Dank geht hiermit auch an die Vertreterinnen und Vertreter diverser Firmen, die mir immer wieder gezeigt haben, wo ich auf gar keinen Fall enden will.

# 10 Curriculum vitae

| | |
|---|---|
| **Name** | Jeannette Wolf |
| **Date of birth** | October 11, 1973 |
| **Place of birth** | Bern (BE) |
| **Nationality** | Swiss |
| **Marital status** | married |

**Education**

| | |
|---|---|
| Dec. 1993 | Matura Typus D (French, English, Italian and Spanish) at the Gymnasium Muttenz, Switzerland |
| Oct. 1994-July 1997 | Undergraduate studies in Biologie II at the Biozentrum, University of Basel |
| Aug. 1997-Oct. 1998 | Diploma thesis in the laboratory of Prof. Walter Keller: „Characterization of a bacterial adenosine deaminase that acts on tRNA and isolation of its human homologue“ |
| Jan. 1999-June 2002 | Ph.D. thesis in the laboratory of Prof. Walter Keller: „Characterization of human and bacterial tRNA-specific adenosine deaminases“ |

**Publications**

Wolf, J., Gerber, A. P. and Keller, W. (2002) tadA, an essential tRNA-specific adenosine deaminase from *Escherichia coli*. *EMBO J.*, **21**, 3841-51

Keller, W., Wolf, J. and Gerber, A. (1999) Editing of messenger RNA precursors and of tRNA by adenosine to inosine conversion. *FEBS Lett.*, **452**, 71-76