Diversity of the biosynthesis pathway for threonylcarbamoyl adenosine (t⁶A), a universal modification of tRNA

Patrick C Thiaville, Dirk Iwata-Reuyl, and Valérie de Crécy-Lagarde

Introduction

tRNAs (tRNA) are the central adaptors in the translation process responsible for decoding mRNAs. tRNAs harbor numerous post-transcriptional modifications that fine-tune their function. To date, more than 90 modifications have been identified in tRNA, and most organisms devote more genetic information to modifying tRNAs than to the tRNAs themselves. While modifications have been shown to affect different aspects of tRNA metabolism and shape interactions of tRNA molecules with the rest of the translation apparatus, most modifications to the anticodon-stem-loop (ASL) are required for accurate decoding. The diversity of tRNA modifications and how modifications affect function has been the topic of recent reviews, and for a summary of modifications to the ASL in *Escherichia coli* see the recent review by Helm and Alfonzo and Table 4 in de Crécy-Lagarde.

The first enzyme of the t⁶A pathway was discovered in 2009 and its derivatives. This complex modification of adenosine is located at position 37, next to the anticodon (t⁶A₃⁷), and is one of the few universal modifications of the ASL.

The hypermodified base t⁶A is present in nearly all ANN decoding tRNAs and has been studied in vitro and in vivo for more than 40 y. Since the first discovery of the modification by Schweizer, et al. in 1969, sporadic studies established the basic requirements for the synthesis of this universal modification, identifying the requirement for ATP, threonine and carbon dioxide intermediate, but fell short of elucidating the multi-step pathway to its formation. Subsequent studies in which native *E. coli* tRNAᵦ₃₇ (harboring an unmodified A₃₇) and yeast tRNAᵦ₃₇ transcripts were converted to t⁶A₃₇ after microinjection into *Xenopus laevis* oocytes demonstrated that the formation of t⁶A occurred in the oocyte cytoplasm and used a conserved machinery. These studies also demonstrated that A₃₇ and U₃₆ were strict determinants for t⁶A formation, and that A₃₈ enhances the efficiency of modification of A₃₇ to t⁶A₃₇. Finally, structural studies showed that t⁶A enhances anticodon-codon base-pairing by cross-strand base-stacking of the t⁶A base with the first position of the codon, and influences the structure of the ASL by preventing across the loop base-pairing between U₃₃-A₃₇, as well as stacking of bases A₃₇ and A₃₈.

Only in the last 5 y have the t⁶A biosynthesis enzymes and the pathways been elucidated, revealing both a core set of enzymes and kingdom-specific variations. This review focuses on the elucidation of the pathway, diversity of the synthesis genes, and proposes a new nomenclature for t⁶A synthesis enzymes.

Keywords: comparative genomics, modified nucleosides, translation, tRNA, t⁶A, universal proteins

The tRNA modification field has a rich literature covering biochemical analysis going back more than 40 years, but many of the corresponding genes were only identified in the last decade. In recent years, comparative genomic-driven analysis has allowed for the identification of the genes and subsequent characterization of the enzymes responsible for N₆-threonylcarbamoyl-adenosine (t⁶A). This universal modification, located in the anticodon stem-loop at position 37 adjacent to the anticodon of tRNAs, is found in nearly all tRNAs that decode ANN codons. The t⁶A biosynthesis enzymes and synthesis pathways have now been identified, revealing both a core set of enzymes and kingdom-specific variations. This review focuses on the elucidation of the pathway, diversity of the synthesis genes, and proposes a new nomenclature for t⁶A synthesis enzymes.

**Discovery of the First t⁶A Synthesis Genes**

The first enzyme of the t⁶A pathway was discovered in 2009 when it was found that a universal protein family, YrdC/Sua5 (COG0009), was involved in t⁶A modification. Based on the assumption that because t⁶A was universally conserved the t⁶A biosynthetic enzymes would also be universally conserved, this work used comparative genomic analysis to focus on universally conserved protein families of unknown function. At the time of...
In this study, 9 universally conserved protein families were of unknown function. Of these, the YrdC/Sua5 family was judged the most likely candidate for involvement in t6A biosynthesis due to: 1) its similarity to HypF, which catalyzes a carbamoylation reaction similar to a putative step in t6A biosynthesis proposed in 1974 by both Elkins and Keller and Körner and Söll; 2) mutations in the yeast yrdC ortholog SUA5 led to translation defects (initiation at non-AUG codons) and 3) E. coli YrdC was found to bind RNA and tRNA. The involvement of the YrdC/Sua5 family in t6A synthesis was experimentally validated using E. coli and Saccharomyces cerevisiae. In E. coli, yrdC is essential, but SUA5 can be deleted from S. cerevisiae, although the growth of the mutant is severely compromised. tRNAs analyzed from this mutant were devoid of t6A. The levels of t6A could be restored through complementation with SUA5Sc, yrdCEc, ywlCBs (Bacillus subtilis SUA5 homolog), and yrdCMm (Methanococcus maripaludis, an archaeal yrdC homolog). The essentiality phenotype of a E. coli yrdC deletion could be complemented by expressing orthologs from yeast, B. subtilis, and M. maripaludis in trans. Analysis of tRNAs in the complemented strains confirmed the presence of t6A. This work identified the first gene family involved in t6A synthesis, established that its function is universally conserved, and that members of the Bacteria and Eukarya / Archaea

**Figure 1.** Diversity in the synthesis of the universal tRNA modification t6A. Two types of enzymes families, TsaC (YrdC) and TsaC2 (Sua5) in Bacteria and Tcs1 (YrdC) and Tcs2 (Sua5) in Eukarya and Archaea, catalyze the formation of TC-AMP. TsaC2 and Tcs2 contain a TsaC-domain plus an additional C-terminal Sua5-domain. To transfer threonyl-carbamoyl (TC) to tRNA, Bacteria require TsaBDE, while Archaea and Eukarya use the KEOPS complex composed of Tcs3 (Kae1), Tcs5 (Bud32), Tcs6 (Pcc1) and Tcs7 (Cgi121) proteins. Tcs8 (Gon7) is found exclusively in Fungi. Mitochondria use the nuclear encoded Tcs4 (Qit7) for transfer of the TC to tRNA. Colors represent homology and correspond with Figure 2.
family could bind ATP, but because purified YrdC alone was not sufficient to produce \( t^6A \) in tRNA transcripts in vitro, it also suggested that additional enzymes were needed for \( t^6A \) synthesis, or that the role of YrdC/Sua5 family was indirect.

A second protein family involved with \( t^6A \) synthesis, YgjD/Kae1/Qri7 (COG0533), was discovered in 2011. Like YrdC/Sua5, the YgjD/Kae1/Qri7 family of proteins is universally conserved, and also exhibited similarity to HypF, which harbors a fusion of YrdC-like and YgjD-like domains. Kae1 had first been described as a member of the KEOPS complex (Kinase, putative Endopeptidase and Other Proteins of Small size) also known as EKC (Endopeptidase-like Kinase Chromatin-associated complex), and had been proposed to be involved in a variety of phenomena unrelated to RNA modification. A phylogeny of this family revealed that yeast harbored 2 members of the family. The first, Kae1 had homologs in other eukaryotes and archaea, and the second, Qri7 was targeted to mitochondria and was part of the bacterial YgjD clade.

The hypothesis that the YgjD/Kae1/Qri7 family was involved in \( t^6A \) synthesis was confirmed by extracting tRNAs from *S. cerevisiae kae1Δ* and showing they were devoid of \( t^6A \) and that \( t^6A \) levels could be restored by complementation with either *ygjDEc* or a version of *QRI7Sc* designed to remain in the cytoplasm. These results also indicated that members of the YgjD/Kae1/Qri7 family were isofunctional for \( t^6A \) synthesis, at least in yeast. To test if YgjD/Kae1/Qri7 were isofunctional in *E. coli*, a *P_{tet}::ygjD* strain was constructed (*ygjD* is only expressed when anhydrotetracycline, aTc, is added). Only the expression of the *ygjD* gene from *E. coli* allowed complementation of the essentiality phenotype in the absence of aTc. In contrast to the YrdC/Sua5 complementation results, expression of the *KAE1Sc* and *QRI7Sc* genes from yeast, the *PRPKm* from *Methanococcus maripaludis* (PRPK is a fusion of Kae1-Bud32 in Archaea) or the *B. subtilis ygjDBs* did not complement the essentiality phenotype of the absence of *ygjD*.

While the protein families TsaC/Sua5 and Kae1/Qri7/TsaD were found to be strictly required for the biosynthesis of \( t^6A \), and a homolog of at least one member of each family is found in all domains of life (Fig. 2), YrdC and YgjD failed to produce \( t^6A \) in vitro with transcript or \( t^6A \)-deficient tRNA purified from yeast *sua5Δ*, suggesting that the biosynthetic machinery for \( t^6A \) biosynthesis required more than these 2 proteins. Over the last 2 years, a flurry of papers have reported the identification of these missing proteins, and elucidated the complete bacterial, eukaryotic/archaeal, and mitochondrial biosynthetic pathways to \( t^6A \).

### Synthesis of \( t^6A \) Varies With Domains of Life

The observation that the YrdC/Sua5 family members were functionally interchangeable between domains while YgjD/Kae1/Qri7 were not lead to a model in which \( t^6A \) biosynthesis occurred in 2 steps with kingdom, species, or organelles specific partners for the second step.
Bacteria

The identity of the remaining enzymes in bacterial t^6A synthesis was predicted from 3 pieces of evidence. First, YgiD was shown to form an association network with YeaZ (a paralog of YgiD) and YjeE, based on physical interaction between the proteins and physical clustering of the genes. Second, like YrdC and YgiD, YeaZ and YjeE were essential in *E. coli* and third, complementation of the *E. coli* ygiD essentiality phenotype required expression of both *B. subtilis* ygiD and yeaZ genes, suggesting that a YgiD/YeaZ interaction was necessary for t^6A synthesis. Notably, it’s been shown that only YeaZ-YgiD pairs from closely related organisms form complexes in vitro. The final evidence that YeaZ and YjeE were the missing proteins in t^6A bacterial synthesis was provided by in vitro reconstitution experiments, which demonstrated that recombinant YrdC, YgiD, YeaZ, and YjeE proteins from *E. coli* were collectively both necessary and sufficient to generate t^6A in reactions with threonine, bicarbonate, ATP, and either *E. coli* tRNA^Thi^ or tRNA^ACG^ transcripts, or unfraccionated tRNA from yeast suA5Δ. Notably, t^6A formation was not observed when a tRNA transcript corresponding to tRNA^23^ from *Methanothermaobacter thermautotrophicus*, which does not naturally contain t^6A, or a 17-mer corresponding to an unmodified ASL of *E. coli* tRNA^ACG^ was used as a substrates. While the former was consistent with the natural lack of t^6A in this tRNA, the latter was surprising since this ASL had previously been shown to bind specifically to *E. coli* YrdC. The t^6A synthesis pathway was subsequently reconstituted using the *B. subtilis* enzymes YwIC (an ortholog of yeast SuA5), and YdhBCE (orthologs of *E. coli* YjeE, YeaZ, and YgiD, respectively), demonstrating the universality of these enzymes in bacteria. With the newly established enzymatic role for YeaZ, YrdC, YgiD, and YjeE (and their orthologs) in the biosynthesis of threonylcarbamoyl-6-adenosine (t^6A), these enzymes were renamed TsaB, TsaC, TsaD, and TsaE, respectively (Fig. 1).

Archaea and Eukarya

The TsaE and TsaB proteins in Bacteria have no homologs in Eukarya or Archaea. The identification of the additional components of t^6A biosynthesis in these last 2 kingdoms came from the fact that Kae1 was part of the KEOPS/EKC complex. The other subunits of the KEOPS/EKC complex (Bud32, Cgi121, Pcc1, plus the fungal specific Gon7) were tested for a potential role in t^6A biosynthesis, and deletion of *pcc1* had only a small decrease (~16%) in total t^6A content. Confirmation that the KEOPS/EKC complex was responsible for t^6A formation came with in vitro reconstitution experiments. It was shown that both the KEOPS/EKC complex from *Pyrococcus abyssi* (Kae1, Bud32, Pcc1, and Cgi121), reconstituted from the individual genes expressed in *E. coli*, as well as the *S. cerevisiae* KEOPS complex (Kae1, Bud32, Pcc1, Cgi121, and Gon7), genes expressed in *E. coli* as a synthetic operon, can form t^6A in vitro, when combined with SuA5 from yeast or Archaea.

Mitochondria

Yeast mitochondrial tRNAs contain t^6A, and while none the subunits of the KEOPS/EKC complex or SuA5 have paralogs targeted to the mitochondria, the *Kae1* homolog Qr7 was found to be targeted to the mitochondria in yeast, and *Caenorhabditis elegans*, *human* and *Arabidopsis thaliana*. It was subsequently demonstrated that the nuclear encoded SuA5 can localize to both the cytoplasm and to the mitochondria in yeast through the use of alternative translation initiation at 2, in-frame AUG sites. Translation from the first AUG encoded a mitochondrial signal peptide, and SuA5 was localized to the mitochondria. SuA5 translated from the second AUG remained in the cytoplasm. Co-expression of both SuA5 and Qr7 in *E. coli* complemented the TsaD essentiality when the expression of *QRI7* alone did not, suggesting that Qr7 could substitute for the KEOPS complex or the TsaBDE proteins. In addition, expression of *QRI7* in the cytoplasm of a *bud32Δ* yeast strain restored growth defects. This was confirmed when it was demonstrated that a minimal system comprised of only SuA5 and Qr7 is sufficient to synthesize t^6A in vitro.

Thus, the enzyme families TsaC/SuA5 and TsaD/Kae1/Qr7 are shared in all organisms, and Bacteria additionally require TsaBE, while Archaea and Eukarya use the other components of the KEOPS complex. Interestingly, although TsaBDE, KEOPS, and Qr7 are functional analogs, only the TsaD/Kae1/Qr7 protein is shared among the 3 systems (Fig. 2), suggesting that this protein family along with the TsaC/SuA5 family were part of the ancestral t^6A synthesis core present in the last universal common ancestor (LUCA).

Mechanistic Analysis of the t^6A Synthesis Machineries

Experiments probing the role of ATP in bacterial t^6A formation demonstrated that both AMP and ADP were products, and that ATP consumption could be uncoupled from RNA modification, with TsaC being the source of AMP (in a threonine dependent process) and TsaD/TsaB/TsaE together producing ADP. These observations were consistent with earlier mechanistic hypotheses in which the ATP requirement in t^6A biosynthesis was rationalized on the presumed need for 2 activated acyl
intermediates during the course of t⁶A formation (i.e. acyl-phosphate and/or acyl-adenylate), the first a phosphocarboxy species (e.g. carboxyphosphate or carboxyadenylate; Fig. 3, intermediate I) activated for transfer to the nitrogen of either threonine or adenosine-37, and the second an N-carboxyphospho species activated for transfer to the remaining component (threonine or adenosine-37; Fig. 3, intermediate III). However, closer scrutiny of the TsaC reaction in the B. subtilis system revealed that the product was threonylcarbamoyl-adenylate (TC-AMP, Fig. 1), an intermediate already activated for condensation with adenosine-37 of tRNA, thus obviating ADP formation as part of the activation steps proposed to be necessary in the biosynthesis of t⁶A. Furthermore, AMP formation by TsaC was shown to arise exclusively from hydrolysis of TC-AMP,33 and that PPi was the other product of the TsaC reaction, implying that formation of TC-AMP itself proceeds through an unusual direct carboxylation of threonine by CO₂ or HCO₃⁻ (Fig. 4).

The role of ADP production in t⁶A biosynthesis remains cryptic; in the bacterial system, t⁶A can be generated without formation of ADP by supplying purified TC-AMP to a reaction containing TsaD/TsaB/TsaE in the absence of ATP,33 while in Archaea and Eukarya formation of t⁶A appears to require the reaction of ATP to ADP, although it does not appear to serve a direct role in the reaction. In Archaea, Pcc1, Kae1, and Bud32 are minimally required to produce t⁶A in vitro, with Kae1 comprising the catalytic subunit responsible for condensing TC-AMP with tRNA. Bud32 was shown to be an ATPase in the presence of Kae1, while it auto-phosphorylates when it is in complex with Cgi121. Thus, Kae1 apparently modifies the phosphotransferase activity of Bud32 and switches it from a kinase to an ATPase. It’s unclear what the specific role of this ATPase activity is, as neither Bud32 or Cgi121 participate directly in the t⁶A reaction. Cgi1321 appears to regulate activity by acting as an effector, where it’s binding significantly changes the conformation of Bud32.55

Overall, the chemistry of t⁶A formation bears similarities to the TobZ system, an O-carbamoyltransferase comprised of a TsaC-like domain fused to a Kae1-like domain that carries out the carbamoylation of tobramycin to form nebramycin.56 In the TobZ reaction, the TsaC domain catalyzes the conversion of carbamoylphosphate to carbamoyladenylate, while the Kae1 domain condenses the latter with tobramycin. Likewise, in t⁶A formation the TsaC/Sua5 homologs generate an adenylated intermediate, which then is condensed with tRNA by the

![Figure 3. Early mechanistic proposal for the formation of t⁶A. P* refers to an activated acylphosphate species, either a simple acyl monophosphate or an acyl AMP.](image1)

![Figure 4. Stepwise formation of t⁶A illustrating the intermediates in the pathway.](image2)
TsaD/Kae1/Qri7 proteins. A notable difference in the systems is that in TobZ the initial substrate is carbamoyl phosphate, which undergoes a phosphotransfer reaction to generate carbamoyl adenylate. This interchange of phosphoryl moieties is chemically not necessary for the subsequent condensation of the carbamoyl group with tobramycin, as both species are activated for the coupling reaction. In t6A biosynthesis, the initial substrates for TsaC/Sua5 are threonine and CO₂/HCO₃⁻, which react first to form N-carboxythreonine followed by reaction with ATP to form TC-AMP.

**Structural Organization of t6A Biosynthetic Proteins**

The highly reactive nature of TC-AMP is not compatible with a freely diffusible intermediate in the biosynthesis of t6A, and argues for the evolution of systems in which this intermediate is instead channeled, as in TobZ,⁵⁶ from its site of production in TsaC/Sua5 to a second active-site (presumably in TsaD/Kae1/ Qri7) where it undergoes reaction with tRNA. A number of observations are consistent with this proposal. First is the fact that the KEOPS complex from both Eukarya and Archaea is known to be a stable quaternary complex.²⁹ Second, while the bacterial proteins do not form an isolable complex analogous to KEOPS, they do interact with one another as demonstrated by the analysis of the E. coli proteins in pull-down experiments, which demonstrated binding of TsaC to both TsaB and TsaD,²⁸ and binding of TsaB to both TsaE and TsaD.²⁸,⁴⁵ Furthermore, the ability of B. subtilis YdiD (TsaD homolog) to complement this phenotype was dependent on co-expression of YdiB (TsaB homolog).³⁰ While the ability of Qri7 to complement this phenotype was dependent on co-expression of Sua5.⁴⁴ These observations are consistent with the requirement for specific physical interactions between these proteins necessary for function.

Additionally, crystallographic analysis has shown that protein-protein associations are conserved across the systems. For example, the crystal structure of Qri7 shows that dimerization is required for Qri7 function, and the dimerization surfaces for Qri7 are used by the archaeal/eukaryotic Kae1 binding to Pcc1 and the bacterial TsaD binding to TsaB.³² Interestingly, although the Pcc1 subunit of KEOPS/EKC shares no sequence similarity to Qri7 or TsaD, Pcc1 engages Kae1 in a manner surprisingly similar to dimerization of Qri7 and TsaD-TsaB.³² Thus, in all 3 systems the ability of the constituent proteins to physically interact with one another appears to be a requirement for tA biosynthesis.

**Naming Convention**

The literature is polluted with a variety of names for each t6A synthesis protein and even for the complexes. With the defined enzymatic and biological function now established it is appropriate to unify the t6A nomenclature. For all Bacteria, we recommend the following suggestions, in agreement with Ken Rudd (Curator of EcoGene, U. of Miami) and published in Deutsch, et al., of TsaB, TsaC, TsaD and TsaE, to replace YeaZ, YrdC, YgiD, and YjeE, respectively. Additionally, Sua5 in bacteria should be renamed TsaC2. TsaC2 is defined as a protein containing both a TsaC and the additional C-terminal Sua5 domain. For Eukarya and Archaea, the use of Tcs (threonyl-carbamoyl synthesis) is recommended (in yeast, TSA1 and TSA2 are in use in yeast for thiorodoxin). We recommend the following nomenclature: Tcs1 (YrdC), Tcs2 (Sua5), Tcs3 (Kae1), Tcs4 (Qri7), Tcs5 (Bud32), Tcs6 (Pcc1), Tcs7 (Cgi121), and Tcs8 (Gon7). A summary of the new and old names, as well as recommended functional descriptions can be found in Table 1. Additionally, we recommend naming the bacterial TsaBDE complex as well as the archaeal/eukaryotic KEOPS/EKC complex to Threonyl-carbamoyl Transferase Complex (TCTC), which will be in keeping with nomenclature of other members of the carbamoyl transferase family. The TCTC family can be further subdivided into bacterial (bTCTC), archaeal (aTCTC), and eukaryal (cTCTC).

**Distribution of the t6A Synthesis Genes Vary in Different Organisms**

Annotation for the first enzyme of t6A synthesis, in bacteria, is complicated by the fact that 2 forms are found (the TsaC or TsaC2) and that 50% of the genomes analyzed harbor a TsaC paralog, YciO, that does not have the same function and does not contain the conserved KRST tetrad.³¹,⁵⁷ We reannotated all members of the COG0009 family (in 9176 bacterial genomes and all contained a TsaC or a TsaC2: 6745 contain TsaC (73%), 2846 contain TsaC2 (31%), and 859 (9%) contained both. In addition, 54% contained YciO (Fig. 2, and http://tinyurl.com/t6A-bacteria). To date, no clear pattern (phylogenetic or lifestyle) has emerged in terms of presence of TsaC or TsaC2, in any given genome and the functional differences between the 2 are not understood. Most bacteria contain both TsaB and TsaE; however, TsaE can be lost in symbiotic or intracellular bacteria, such as Wolbachia or Mycoplasmas (e.g., Mycoplasma genitalium and Mycoplasma pneumoniae). To date only 2 bacteria, Mycoplasma haemofelis and Mycoplasma suis strain Illinois, are missing both TsaB and TsaE.⁵⁸ (Fig. 2 and Table 2). It seems these organisms harbor a mitochondrial like minimal t6A synthesis system (unless another unidentified protein has been recruited).

Like Bacteria, all Eukarya and Archaea contain either a homolog of Tcs1 (TsaC/YrdC) or Tcs2 (TsaC2/Sua5). We have found one organism that has both, the fungi Pseudococcospora fijiensis CIRAD86, also known as Mycosphaerella fijiensis CIRAD86 (NCBI Taxonomic ID: 383855). As with bacteria, there is not a clear phylogenetic inheritance between organisms with Tcs1 or Tcs2 in Archaea or Eukarya, but a taxonomic relationship does exists in eukaryotes. Fungi exclusively contain Tcs2 (with P. fijiensis as an exception), while all Plants (including Chlamydomonas reinhardii) and all Metazoans exclusively contain Tcs1. Of the 53 Archaea analyzed, 25 contain Tcs1 and 28 contain Tcs2. The only taxonomic relationship found is in the order Halobacterales that exclusively contain Tcs1, Figure 2 and Table 3 (http://tinyurl.com/t6A-Arc-Euk).
All Archaea contain a single Tcs3 (Kae1) homolog, while Eukarya also contain a single Tcs3 homolog and also have a Tcs4 (Qri7) homolog (evolutionarily related to the bacterial TsaD), which will function in the organelles. In all genomes analyzed, both Tcs3 and Tcs4 were found in the nuclear genome and not in the organelle. Specifically, the human nuclear genome contains Tcs3 (OSGEP) for cytoplasmic t6A synthesis and Tcs4 (OSGEPL1) was shown to target to the mitochondria.54 Note to the reader, the Oberto, et al. paper incorrectly referred to OSGEPL as the Tcs4 homolog (instead of OSGEPL1), the human mitochondrial targeting protein. As an example for plants, *Arabidopsis thaliana* contains nuclear encoded Tcs3 (AT4G22720) and Tcs4 (AT2G45270). Tcs4 contains a strong chloroplast targeting signal, but has only been detected in the mitochondria.55 The human pathogen *Plasmodium falciparum* (causative agent of malaria) presents an interesting case for t6A synthesis, as the mitochondria utilize fully modified cytoplasmic tRNAs for mitochondrial translation (requirement for t6A machinery is unknown), and *P. falciparum* contains an apicoplast originating from secondary endosymbiosis of an alage.56 *P. falciparum* contains 2 nuclear encoded homologs of Tcs3 (Table 3): a Tcs3 that is similar to the yeast Tcs3, and an apicoplast-targeting Tcs3b, that is similar to Tcs3,57 but is phylogenetically distant from all known Tcs3 and from the bacterial TsaD (Thiaville and de Crécy-Lagard, unpublished). Tcs3 interacts with both Tcs5 (Bud32) and Tcs7 (Cgi121), and Tcs3b interacts with multiple proteins associated with the apicoplast ribosome (Mallari and Goldberg, personal communication). Tcs2 has not been detected in the apicoplast, and it is currently unknown how the first step in t6A synthesis occurs. (Mallari and Goldberg, personal communication).

Tcs5 (Bud32) is found in all Eukarya and Archaea sequenced to date. In the 53 Archaea analyzed, Tcs5 and Tcs3 are adjacent ORFs in 13 genomes and are fused in 25 genomes, demonstrating a strong functional linkage between the proteins of these genes. Tcs6 (Pcc1) and Tcs7 (Cgi121) are found in nearly all Archaea and Eukarya. Notable exceptions are the absence of Tcs6 in *P. falciparum* and the absence of Tcs7 in *Drosophila melanogaster*. Tcs8 (Gon7) is a fungal specific protein. Tcs8 is required for t6A formation in yeast (Thiaville and de Crécy-

### Table 1. Proposed names and functional roles for t6A synthesis genes.

| New Name | Old Names | Function |
|----------|-----------|----------|
| Bacteria |           |          |
| TsaB     | YeaZ / YdiC | tRNA adenosine(37) threonylcarbamoyltransferase complex, dimerization subunit type 1 |
| TsaC     | YrdC     | L-threonylcarbamoyladenylate synthase (EC 2.7.87.7) type 1 |
| TsaC2    | Suu5 / YwIC | L-threonylcarbamoyladenylate synthase (EC 2.7.87.7) type 2 |
| TsaD     | YgdY / YdE | tRNA adenosine(37) threonylcarbamoyltransferase complex, transferase subunit |
| TsaE     | YgdE / YdD | tRNA adenosine(37) threonylcarbamoyltransferase complex, ATPase subunit type 1 |
| Archaea / Eukaryotes | | |
| Tcs1     | YrdC     | L-threonylcarbamoyladenylate synthase (EC 2.7.87.7) type 1 |
| Tcs2     | Suu5     | L-threonylcarbamoyladenylate synthase (EC 2.7.87.7) type 2 |
| Tcs3     | Kae1 / gcp / OSGEP | tRNA adenosine(37) threonylcarbamoyltransferase complex, transferase subunit |
| Tcs4     | Qri7 / OSGEPL1 | tRNA adenosine(37) threonylcarbamoyltransferase, mitochondrial |
| Tcs5     | Bud32    | tRNA adenosine(37) threonylcarbamoyltransferase complex, ATPase subunit type 2 |
| Tcs6     | Pcc1     | tRNA adenosine(37) threonylcarbamoyltransferase, dimerization subunit type 2 |
| Tcs7     | Cgi121   | tRNA adenosine(37) threonylcarbamoyltransferase complex, regulator subunit |
| Tcs8     | Gon7     | tRNA adenosine(37) threonylcarbamoyltransferase complex, fungal specific subunit |

### Table 2. Homologs of t6A biosynthetic genes in Bacteria.

| Organism                  | TsaC | TsaC2 (Sua5) | TsaB | TsaD | TsaE |
|---------------------------|------|--------------|------|------|------|
| *E. coli* K12             | b3282| b1807        | b3064| b4168|
| *Vibrio cholerae O1 El Tor*| VC0054| VC1079      | VC1899| VC0521| VC0343|
| *Caulobacter crescentus* NA1000 | CCNA_03501| CCNA_00057| CCNA_00069| CCNA_03648|
| *Mycoplasma genitalium* G37 | MG259* | MG208 | MG046 | N.P. |
| *Mycoplasma pulmonis*     | MYPU_6130* | MYPU_1190 | MYPU_1180 | MYPU_1200 |
| *Bacillus subtilis* subsp. subtilis str. 168 | BSU36950 | BSU05920 | BSU05940 | BSU05910 |
| *Haemophilus influenzae* Rd | HI0656 | HI0388 | HI0530 | HI0656 |
| *Acinetobacter baylyi* APD1 | ACIAD0208 | ACIAD0677 | ACIAD1332 | ACIAD2376 |
| *Salmonella Typhii* TY2 | STY395 | STY1950 | STY3387 | STY4714 |
| *Francisella novisida* U112 | FTN_0158 | FTN_1148 | FTN_1565 | FTN_0274 |
| *Pseudomonas aeruginosa* PA01 | PA0022 | PA3685 | PA0580 | PA4948 |
| *Burkholderia thailandensis* E264 | BTH_I0669 | BTH_I0066 | BTH_I2001 | BTH_I0616 |
| *Staphylococcus aureus* subsp. aureus MW2 | MW0860 | MW2040 | MW1975 | MW1973 |

*M. genitalium* MG259 is a TsaC/HemK fusion.

*M. pulmonis* TsaC (MYPU_6130) and HemK (MYPU_1060).

N.P.: Not Present.
Table 3. Homologs of t6A biosynthetic genes in Archaea and Eukarya.

| Organism                  | Tcs1 (YrdC)     | Tcs2 (Sua5)     | Tcs3 (Kae1)     | Tcs4 (Qri7)     | Tcs5 (Bud32)   | Tcs6 (Pcc1)     | Tcs7 (Gji121)   | Tcs8 (Gon7)   |
|---------------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|--------------|
| Haloflex volcanii DS2     | HVO_0253      |               | HVO_1895     |               |               |               |               |              |
| Homo sapiens              | 1p34.3        | 14q11.2       |               | 2q32.2        | 20q13.2       |               |               |              |
| Drosophila melanogaster   | CG10438       |               | CG9433        |               | CG10673       |               |               |              |
| Plasmodium falciparum     | PFL0175c      | FPD37_1030600 |               | N.P.          | N.P.          |               |               |              |
| Saccharomyces cerevisiae  | YGL169w       |               |               |               |               |               |               |              |
| Schizosaccharomyces        |               |               |               |               |               |               |               |              |
| pombe                     |               |               |               |               |               |               |               |              |
| Arabidopsis thaliana      | AT5G60590     |               |               |               |               |               |               |              |

*H. volcanii Tcs3 and Tcs5 occur as a gene fusion (HVO_1895).*

**P. falciparam PF3D7_0408900.1 (Tcs3b) targets to the apicoplast and is similar to Tcs3.**

N.P.: Not Present.

Lagard, unpublished), but the function of Tcs8 is currently unknown.

### Derivatives of t6A

Currently, there are 3 known derivatives of t6A: ct6A (cyclic t6A), m6t6A (N^6-methyl-N^6-threonylcarbamoyladenosine), and ms2t6A (2-methylthio-N^6-threonylcarbamoyladenosine),

A new twist in the t6A field was recently discovered with the identification of cyclic form of t6A (ct6A), a cyclized active ester of t6A with an oxazolone ring. Renumber starting here throughout the end of the manuscript. TcdA (previously CsdL in E. coli) catalyzes an ATP-dependent dehydration of t6A to ct6A; this reaction is performed by Tcd1 (YHR003c) and by Tcd2 (YKL027w) in yeast.

The third known derivative of t6A, ms2t6A, is found only on tRNA^{Thr}(GGU) in a subset of organisms. Particularly, ms^2t6A is found in B. subtilis, some Archaea, and in human, but not in E. coli. YqeV (MtaB) in B. subtilis and Cdkal1 in humans are responsible for the insertion of the sulfur moiety and methylation at position 2 of the adenosine containing t6A. TrmO is widely distributed throughout life and cross-kingdom functional analysis was performed to show the activity was conserved.

The third known derivative of t6A, ms^2t6A, is found only on tRNA^{Thr}(GGU) in a subset of organisms. Particularly, ms^2t6A is found in B. subtilis, some Archaea, and in human, but not in E. coli. YqeV (MtaB) in B. subtilis and Cdkal1 in humans are responsible for the insertion of the sulfur moiety and methylation at position 2 of the adenosine containing t6A. TrmO is widely distributed throughout life and cross-kingdom functional analysis was performed to show the activity was conserved.

Loss of the Cdkal1 homolog in mice is correlated to increase Type 2 diabetes. It is not clear if ct6A is the base to form ms^2t6A, or like m^6t6A, ms^2t6A is formed from t6A. (Fig. 2).

### Concluding Remarks

The biosynthesis of t6A is just one example of a “rediscovery” of tRNA modifications in the genomic era, which has allowed for the discovery of globally unknown genes for enzyme reactions that were discovered more than 40 y earlier. In Bacteria, the 4 genes involved in t6A biosynthesis, due to their prokaryotic-specific essentiality and because *tahB* and *tac* are found only in bacteria, had been identified as potential antibacterial and inhibitor targets prior to the discovery of their role in t6A synthesis was even established. The unique Tcs3b found in *P. falciparum* also presents itself as an attractive anti-malarial target. For these proteins to be viable targets, it is critical to understand their distribution profile and potential range of action as well as the mechanisms underlying the essentiality phenotypes to predict resistance mechanisms. Clearly, one should use caution in

---

### Table 3. Homologs of t6A biosynthetic genes in Archaea and Eukarya.

| Organism                  | Tcs1 (YrdC)     | Tcs2 (Sua5)     | Tcs3 (Kae1)     | Tcs4 (Qri7)     | Tcs5 (Bud32)   | Tcs6 (Pcc1)     | Tcs7 (Gji121)   | Tcs8 (Gon7)   |
|---------------------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|--------------|
| Haloflex volcanii DS2     | HVO_0253      |               | HVO_1895     |               |               |               |               |              |
| Homo sapiens              | 1p34.3        | 14q11.2       |               | 2q32.2        | 20q13.2       |               |               |              |
| Drosophila melanogaster   | CG10438       |               | CG9433        |               | CG10673       |               |               |              |
| Plasmodium falciparum     | PFL0175c      | FPD37_1030600 |               | N.P.          | N.P.          |               |               |              |
| Saccharomyces cerevisiae  | YGL169w       |               |               |               |               |               |               |              |
| Schizosaccharomyces pombe |               |               |               |               |               |               |               |              |
| Arabidopsis thaliana      | AT5G60590     |               |               |               |               |               |               |              |

*H. volcanii Tcs3 and Tcs5 occur as a gene fusion (HVO_1895).*

**P. falciparam PF3D7_0408900.1 (Tcs3b) targets to the apicoplast and is similar to Tcs3.**

N.P.: Not Present.
designing drugs targeting TsAB and TsAE in *Mycoplasmas* spp. since the genes are absent. Caution would also be needed for drugs targeting TsC, due to the possibility of cross reactivity in humans, although TsAC2 and TsD may be viable options.

The discovery of the t6A pathways now allows us to address more systematically the causes of the pleiotropic phenotypes caused by the absence of t6A synthesis enzymes. Are these due to mistranslation of target proteins, to a role of t6A as a determinant for other components of the translation apparatus, or to a role of t6A or of t6A synthesis proteins in other processes than translation? Indeed, the recent discovery of a molecule similar to the t6A nucleoside in dauer signaling in nematodes 4 opens an unforeseen role for t6A derivatives in biology.

References

1. Machnicka MA, Milanowska K, Osmar Ogloa O, Purta E, Kuckowska M, Oldchowik A, Januszeiwski W, Kalinowski S, Dunin-Horkawicz S, Rother KM, et al. MODOMICS: a database of RNA modification pathways–2013 update. Nucleic Acids Res 2013; 41:D2627-71. PMID:23181484; http://dx.doi.org/10.1093/nar/gkt1007

2. Björk GR. Biosynthesis and function of modified Nucleosides. In: Söll D, Rajbhandary UL, editors. tRNA: Structure, Biosynthesis, and Function. ASM Press, Washington, DC; 1995, 165-205.

3. Yarian C, Townsend H, Czestkowski W, Sochacka E, Osman Oglou O, Alvaro EC, for sharing results ahead of publication. This work was supported by the National Institutes of Health (grant number R01 GM70641 to V. de C.-L.). P.C.T. was supported by the Chateaubriand Fellowship from the French Embassy in the United States.

4. Glavic for sharing results ahead of publication. This work was supported by the National Institutes of Health (grant number R01 GM70641 to V. de C.-L.). P.C.T. was supported by the Chateaubriand Fellowship from the French Embassy in the United States.

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Basma El Yacoubi and Jennifer Thivielle for insightful comments and editing the manuscript. We thank Jeremy Mallari, Dan Goldberg, Diego Rojas-Benitez, and Álvaro Glavic for sharing results ahead of publication.

Funding

This work was supported by the National Institutes of Health (grant number R01 GM70641 to V. de C.-L.). P.C.T. was supported by the Chateaubriand Fellowship from the French Embassy in the United States.
required for the formation of threonylcarbamoyladenosine in tRNA. Nucleic Acids Res 2009; 37:2894-909; PMID:19287007; http://dx.doi.org/10.1093/nar/gkp152.

32. Wan LCK, Mao DHL, Nielcu D, Strecer J, Chiovitti A, D, Khm, T. Poc P, Altvaterman N, VCONF, R, L, R, K, R, et al. Reconstitution and characterization of eukaryotic N6-threonylcarbamoyltransferase using a minimal enzyme system. Nucleic Acids Res 2013; 41:6352-46; PMID:23622099; http://dx.doi.org/10.1093/nar/gkt187.

33. Tominaga T, Kobayashi K, Ishii R, Ishitani R, Nureki O. Structure of saccharomyces cerevisiae mitochondrial O-galactosyltransferase with AMP. Acta Crystallogr 2014; 70:1009-14; PMID:25084372; http://dx.doi.org/10.1107/S056773941400952X.

34. Tominaga T, Kobayashi K, Ishii R, Ishitani R, Nureki O. Structure of saccharomyces cerevisiae mitochondrial O-galactosyltransferase with AMP. Acta Crystallogr 2014; 70:1009-14; PMID:25084372; http://dx.doi.org/10.1107/S056773941400952X.

35. Downey M, Haensch R, Mairinger L, Rollie A, Brehme G, Galica S, Guillard S, Partington M, Zubko MK, Krogan NJ, et al. A genome-wide screen identifies the evolutionarily conserved KEOPS complex as a telomere regulator. Cell 2006; 124:1155-68; PMID:16871589; http://dx.doi.org/10.1016/j.cell.2005.12.044.

36. Wan LCK, Mao DHL, Nielcu D, Strecer J, Chiovitti A, D, Khm, T. Poc P, Altvaterman N, VCONF, R, L, R, K, R, et al. Reconstitution and characterization of eukaryotic N6-threonylcarbamoyltransferase using a minimal enzyme system. Nucleic Acids Res 2013; 41:6352-46; PMID:23622099; http://dx.doi.org/10.1093/nar/gkt187.

37. Downey M, Haensch R, Mairinger L, Rollie A, Brehme G, Galica S, Guillard S, Partington M, Zubko MK, Krogan NJ, et al. A genome-wide screen identifies the evolutionarily conserved KEOPS complex as a telomere regulator. Cell 2006; 124:1155-68; PMID:16871589; http://dx.doi.org/10.1016/j.cell.2005.12.044.

38. Wan LCK, Mao DHL, Nielcu D, Strecer J, Chiovitti A, D, Khm, T. Poc P, Altvaterman N, VCONF, R, L, R, K, R, et al. Reconstitution and characterization of eukaryotic N6-threonylcarbamoyltransferase using a minimal enzyme system. Nucleic Acids Res 2013; 41:6352-46; PMID:23622099; http://dx.doi.org/10.1093/nar/gkt187.

39. Downey M, Haensch R, Mairinger L, Rollie A, Brehme G, Galica S, Guillard S, Partington M, Zubko MK, Krogan NJ, et al. A genome-wide screen identifies the evolutionarily conserved KEOPS complex as a telomere regulator. Cell 2006; 124:1155-68; PMID:16871589; http://dx.doi.org/10.1016/j.cell.2005.12.044.

40. Naor A, Thiaville PC, Altman-Price N, Cohen-Or I, Allers T, de Crecy-Lagard V, Gophna U. A genetic approach to understanding an essential eukaryotic methylthiotransferase for biosynthesis of the universal t6A tRNA modification. Nucleic Acids Res 2010; 38:6195-205; PMID:20472640; http://dx.doi.org/10.1093/nar/gkq364.

41. Arigoni F, Talabot F, Peitsch M, Edgerton MD, Melchior C, Scheel D, et al. The O-carbamoyltransferase TobZ causes the development of type 2 diabetes in mice. J Biol Chem 2014; 289(43):289-39; PMID:25204654; doi: 10.1074/jbc.M113436573.

42. Wei F, Suzuki T, Watanabe S, Kimura S, Kaitzuka T, Furumura A, Katsuki T, Fujimura A, Matsui H, Atta M, Michiue H, Fontecave M, et al. Identification of eukaryotic and prokaryotic methyltransferase for biosynthesis of 2-methylthio-N6-threonylcarbamoyladenosine in tRNA. J Biol Chem 2008; 283:24827-33; PMID:18386497; http://dx.doi.org/10.1074/jbc.M1074079.

43. Wei F, Suzuki T, Watanabe S, Kimura S, Kaitzuka T, Furumura A, Katsuki T, Fujimura A, Matsui H, Atta M, Michiue H, Fontecave M, et al. Identification of eukaryotic and prokaryotic methyltransferase for biosynthesis of 2-methylthio-N6-threonylcarbamoyladenosine in tRNA. J Biol Chem 2008; 283:24827-33; PMID:18386497; http://dx.doi.org/10.1074/jbc.M1074079.

44. Wei F, Suzuki T, Watanabe S, Kimura S, Kaitzuka T, Furumura A, Katsuki T, Fujimura A, Matsui H, Atta M, Michiue H, Fontecave M, et al. Identification of eukaryotic and prokaryotic methyltransferase for biosynthesis of 2-methylthio-N6-threonylcarbamoyladenosine in tRNA. J Biol Chem 2008; 283:24827-33; PMID:18386497; http://dx.doi.org/10.1074/jbc.M1074079.

45. Wei F, Suzuki T, Watanabe S, Kimura S, Kaitzuka T, Furumura A, Katsuki T, Fujimura A, Matsui H, Atta M, Michiue H, Fontecave M, et al. Identification of eukaryotic and prokaryotic methyltransferase for biosynthesis of 2-methylthio-N6-threonylcarbamoyladenosine in tRNA. J Biol Chem 2008; 283:24827-33; PMID:18386497; http://dx.doi.org/10.1074/jbc.M1074079.
72. Lerner CG, Hajduk PJ, Wagner R, Wagenaar FL, Woodall C, Gu YG, Searle XB, Florjancic AS, Zhang T, Clark RF, et al. From bacterial genomes to novel antibacterial agents: discovery, characterization, and antibacterial activity of compounds that bind to HI0065 (YjeE) from haemophilus influenzae. Chem Biol Drug Des 2007; 69:395-404; PMID:17581233; http://dx.doi.org/10.1111/j.1747-0285.2007.00521.x

73. Bose N, Ogawa A, von Reuss SH, Yim JJ, Ragsdale EJ, Sommer RJ, Schroeder FC. Complex small-molecule architectures regulate phenotypic plasticity in a nematode. Angew Chemie 2012; 51:12438-43; PMID:23161728; http://dx.doi.org/10.1002/anie.201206797

74. Lerunic I, Bork P. Interactive tree of life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 2007; 23:127-8; PMID:17050570; http://dx.doi.org/10.1093/bioinformatics/btl529

75. Lerunic I, Bork P. Interactive tree of life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Res 2011; 39:W475-8; PMID:21470960; http://dx.doi.org/10.1093/nar/gkr201