*Published online 1 February 2022* *Nucleic Acids Research, 2022, Vol. 50, No. 4* **2223–2239**
*https://doi.org/10.1093/nar/gkac056*

# Commonality and diversity in tRNA substrate recognition in t$^6$A biogenesis by eukaryotic KEOPSs

**Jin-Tao Wang[1,†], Jing-Bo Zhou[1,†], Xue-Ling Mao[1], Li Zhou[2], Meirong Chen[3], Wenhua Zhang[2], En-Duo Wang[1,4,*] and Xiao-Long Zhou[1,*]**

[1]State Key Laboratory of Molecular Biology, CAS Center for Excellence in Molecular Cell Science, Shanghai Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences, University of Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, China, [2]School of Life Sciences, Lanzhou University, 222 South Tianshui Road, Lanzhou 730000, Gansu, [3]School of Pharmacy, China Pharmaceutical University, 639 Longmian Avenue, Nanjing 211198, Jiangsu and [4]School of Life Science and Technology, ShanghaiTech University, 393 Middle Hua Xia Road, Shanghai 201210, China

Received December 16, 2021; Revised January 16, 2022; Editorial Decision January 17, 2022; Accepted January 25, 2022

## ABSTRACT

***N*$^6$-Threonylcarbamoyladenosine (t$^6$A) is a universal and pivotal tRNA modification. KEOPS in eukaryotes participates in its biogenesis, whose mutations are connected with Galloway-Mowat syndrome. However, the tRNA substrate selection mechanism by KEOPS and t$^6$A modification function in mammalian cells remain unclear. Here, we confirmed that all ANN-decoding human cytoplasmic tRNAs harbor a t$^6$A moiety. Using t$^6$A modification systems from various eukaryotes, we proposed the possible coevolution of position 33 of initiator tRNA$^{Met}$ and modification enzymes. The role of the universal CCA end in t$^6$A biogenesis varied among species. However, all KEOPSs critically depended on C32 and two base pairs in the D-stem. Knockdown of the catalytic subunit *OSGEP* in HEK293T cells had no effect on the steady-state abundance of cytoplasmic tRNAs but selectively inhibited tRNA$^{Ile}$ aminoacylation. Combined with *in vitro* aminoacylation assays, we revealed that t$^6$A functions as a tRNA$^{Ile}$ isoacceptor-specific positive determinant for human cytoplasmic isoleucyl-tRNA synthetase (IARS1). t$^6$A deficiency had divergent effects on decoding efficiency at ANN codons and promoted +1 frameshifting. Altogether, our results shed light on the tRNA recognition mechanism, revealing both commonality and diversity in substrate recognition by eukaryotic KEOPSs, and elucidated the critical role of t$^6$A in tRNA$^{Ile}$ aminoacylation and codon decoding in human cells.**

## INTRODUCTION

tRNAs are the most heavily modified RNA species in the cell, considering localizations of modified positions and modification diversity. tRNA modification plays critical functions in genome decoding by either guaranteeing the unique structure and stability of tRNA architectures or regulating translation fidelity and efficiency during ribosomal decoding (1,2).

Among all positions of a given tRNA molecule, those in the anticodon loop, in particular positions 34 and 37, are the most extensively modified due to their critical roles in codon-anticodon base pairing strength and accuracy (3,4). *N*$^6$-Threonylcarbamoyladenosine (t$^6$A) is a modified nucleotide found exclusively at position 37 adjacent to the anticodon in ANN-decoding tRNAs across all three domains of life (5,6). Enzymes for t$^6$A biogenesis have been studied in bacteria, archaea, and eukaryotes and in organelles. t$^6$A modification occurs in two consecutive steps: YRDC/Sua5 family proteins catalyze the formation of the L-threonylcarbamoyladenylate intermediate (TC-AMP) from threonine, bicarbonate and ATP (7), and subsequently, TsaD/Kae1/Qri7 family proteins transfer the L-threonylcarbamoyl moiety of the TC-AMP intermediate to A37 (8–13). Despite the conservation of catalytic subunits, the components of t$^6$A modification machinery, specifically that performing the second step, differ significantly among various species/organelles. In bacteria, TsaC (in the YRDC/Sua5 family), TsaD, TsaB and TsaE jointly mediate t$^6$A biogenesis (10), while in archaea and eukaryotic cytoplasm, YRDC/Sua5 and a four- or five-subunit KEOPS (Kinase, Endopeptidase, and Other Proteins of Small size) complex cooperatively generate t$^6$A (11,12,14). Eukaryotic mitochondria seemingly use the minimalistic enzymes Sua5 and Qri7 in yeast or YRDC and OSGEPL1

*To whom correspondence should be addressed. Tel: +86 21 5492 1247 Fax: +86 21 5492 1011; Email: xlzhou@sibcb.ac.cn
Correspondence may also be addressed to En-Duo Wang. Tel: +86 21 5492 1241; Fax: +86 21 5492 1011; Email: edwang@sibcb.ac.cn
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

in mammals (15–17) (Supplementary Table S1). In the KEOPS complex, Gon7 (GON7), Pcc1 (LAGE3), Kae1 (OSGEP), Bud32 (TP53RK) and Cgi121 (TPRKB) are arranged linearly (14,18). In addition to the catalytic subunit Kae1 (OSGEP), Bud32 (TP53RK) is an ATPase that catalyzes the hydrolysis of ATP to ADP and Pi (19); however, the exact role of this subunit in $t^6A$ biogenesis remains unclear. Furthermore, the function of both Gon7 (GON7) and Pcc1 (LAGE3) remains unknown (5). TsaC has been reported to interact with the TsaD-TsaB-TsaE complex (10); however, Sua5/YRDC seems to be independent of the KEOPS complex in the cytoplasm or of OSGEPL1 in mitochondria (15,19).

Substrate selection and binding during the second step have long been unknown. Only recently the binding mode of TC-AMP analog by bacterial TsaD-TsaB has been clarified (20). However, only limited reports have described tRNA recognition by $t^6A$ modification enzymes, in part owing to difficulties in reconstituting $t^6A$ modification activity *in vitro* in earlier studies. Based on a *Xenopus laevis* oocyte *in vivo* modification system, only U36 was absolutely required for effective $t^6A$ modification (21). Our two recent works have revealed that (C/A>U>G)32-N33-(U/G>A)34-N35-U36-A37-A38 was the nucleotide sequence requirements in the anticodon loop for human mitochondrial $tRNA^{Thr}$ to be efficiently modified by Sua5-Qri7 (22) or YRDC-OSGEPL1, and $Lys^{203}$ in OSGEPL1 seems to be a critical tRNA-binding element (15). Our initial finding of A38 being a critical element has been confirmed by others (17). Using the archaeal KEOPS complex as a model, it has been recently found that Cgi121 in an archaeal KEOPS complex (human TPRKB homolog) binds the CCA terminus of substrate tRNAs and that this interaction is essential for $t^6A$ biogenesis by the archaeal KEOPS complex (18).

$t^6A$ modification in bacteria and yeast stabilizes the anticodon loop architecture by preventing intraloop interactions, facilitates codon-anticodon pairing by mediating base-stacking interactions at the ribosomal decoding site to prevent +1 frameshifting and promotes downstream modification at other sites (6,23). In addition, reduced $tRNA^{Ile}$ aminoacylation levels by isoleucyl-tRNA synthetase (IleRS) have been observed in *Escherichia coli* cells lacking the TsaC or TsaD gene (24). However, yeast IleRS does not use the $t^6A$ moiety as a positive determinant in tRNA charging (12,24). These observations likely explain why the $t^6A$ modification apparatus is essential in some bacteria but not in yeast. In mammalian cells, knockdown of *OSGEP* or *TP53RK* leads to impaired protein synthesis, endoplasmic reticulum stress, signaling in response to DNA damage, and apoptosis (25). Owing to the pivotal role of $t^6A$ modification in mRNA translation and protein homeostasis, it is not surprising that deletions or mutations in $t^6A$ modification enzymes in yeast or in humans lead to cellular dysfunctions and disorders in humans (25,26). Yeast cells in which the Sua5 gene is deleted exhibit delayed growth and sensitivity to various stresses, including heat, ethanol and salt (27). Deletion of the Kae1 gene in yeast also causes severe growth retardation (16). Several independent studies have shown that genetic mutations in each gene of the human $t^6A$ biogenesis pathway (YRDC, OSGEP, TP53RK, TPRKB, GON7 and LAGE3) cause severe effects associated with Galloway–Mowat syndrome (GAMOS), characterized by the combination of early onset nephrotic syndrome and microcephaly with brain anomalies (25,28). Effects of *OSGEP* mutations on other phenotypes, such as neurodegeneration, have also been observed (29). Despite obvious advances in the mechanism and significance of $t^6A$ in bacteria, yeast and archaea, however, in mammalian cells, the basic understanding of the molecular mechanism of $t^6A$ biogenesis, including tRNA selection and recognition by the eukaryotic cytoplasmic KEOPS complex, the potential contribution of $t^6A$ modification to tRNA abundance and the aminoacylation level of ANN-decoding tRNAs, codon-anticodon decoding, and +1 frameshifting restriction, has thus far been very limited.

In the present work, using several KEOPS complexes and bacterial, yeast and human tRNAs, we clarified how eukaryotic KEOPS complexes select tRNA substrates, highlighting the critical role of the anticodon loop and two base pairs in the D-stem in determining the $t^6A$ modification level. Furthermore, we revealed the important or determinative role of $t^6A$ modification in tRNA aminoacylation and in preventing +1 frameshifting. Our results help to understand the basic knowledge of the $t^6A$ modification mechanism and function in human cells.

## MATERIALS AND METHODS

### Materials

Anti-FLAG (F7425) and horseradish peroxidase (HRP)-conjugated secondary antibodies were purchased from Merck (Darmstadt, Germany). Anti-GAPDH (60004-1-Ig) and anti-OSGEP (15033-1-AP) antibodies were purchased from Proteintech (Wuhan, China). Anti-TP53RK (A14952) antibody was purchased from ABclonal Technology Co., Ltd. (Wuhan, China). [$^{14}$C]Thr was obtained from American Radiolabeled Chemicals, Inc. (St. Louis, MO, USA), and [$^{3}$H]Ile was obtained from PerkinElmer, Inc. (Hopkinton, MA, USA). KOD-plus mutagenesis kits were obtained from TOYOBO (Osaka, Japan). Lipofectamine 2000 transfection reagent, puromycin and SuperSignal West were obtained from Thermo Scientific (Waltham, MA, USA). Anti-digoxigenin-AP (11093274910), 10% blocking reagent (11096176001) and CDP-Star were purchased from Roche (Basel, Switzerland). 50× Denhardt solution (B548209-0050), 20× saline sodium citrate (SSC) (B548109) and fish sperm DNA (B548210) were purchased from Sangon Biotech Co., Ltd. (Shanghai, China). Primer synthesis, biotin- or digoxin-DNA probe synthesis and DNA sequencing were performed by Tsingke Biotechnology Co., Ltd. (Beijing, China) and Biosune (Shanghai, China).

### Plasmid construction, mutagenesis and gene expression

Genes encoding human KEOPS (hKEOPS) subunits OSGEP (UniProt No. Q9NPF4), TP53RK (UniProt No. Q96S44), TPRKB (UniProt No. Q9Y3C4), LAGE3 (UniProt No. Q14657) and GON7 (UniProt No. Q9BXV9) were amplified from cDNA obtained by reverse transcription of total RNA from human HEK293T cells. Ribosomal recognition sites (RBSs) were inserted between two adjacent coding genes into a pJ241 vector (30) with a $His_6$-

tag at its C-terminus via a Seamless Cloning Kit (Beyotime, Shanghai, China). The resulting polycistronic plasmid pJ241-hKEOPS expressed five recombinant proteins, OSGEP, TP53RK, TPRKB, LAGE3 and GON7-$His_6$, under the control of the T7 promoter. *Escherichia coli* codon-optimized DNA encoding *Caenorhabditis elegans* PCC1 (F59A2.5), KAE1 (Y71H2AM.1) plus a $His_6$-tag at the C-terminus, BUD32 (F52C12.6), CGI121 (W03F8.4) (GON7 homolog in *C. elegans* not yet identified) and a ribosomal binding site sequence between two adjacent genes was chemically synthesized and inserted into a pET24a vector between NdeI and XhoI sites. The resulting polycistronic plasmid pET24a-*Ce*KEOPS (*C. elegans* KEOPS) expressed four recombinant proteins, PCC1, KAE1-$His_6$, BUD32 and CGI121, under the control of the T7 promoter (8). The plasmid encoding *Saccharomyces cerevisiae* (*Sc*) KEOPS (pJ241-*Sc*KEOPS) expressed five recombinant proteins, KAE1, BUD32, CGI121, PCC1 and GON7-$His_6$, under the control of the T7 promoter (30). Human YRDC (hYRDC) and yeast Sua5 expression plasmids were constructed as reported previously (15,22). The construction of expression plasmids for *E. coli* TsaC (UniProt No. P45748), TsaD (UniProt No. P05852) and TsaB (UniProt No. P76256) was as described in a previous report (31). The gene encoding *E. coli* TsaE (UniProt No. P0AF67) was subcloned from pET28a-yjeE into a pET21a vector via NdeI and XhoI, yielding a pET21a-TsaE expression plasmid in which a $His_6$ tag is introduced at the C-terminus (31). The gene encoding truncated human cytoplasmic isoleucyl-tRNA synthetase (hIleRS) ($Met^1$-$Ser^{1073}$) (UniProt No. P41252) was obtained by amplifying the cDNA obtained by reverse transcription of the total RNA of human HEK293T cells and inserted into a pJ241 vector with a $His_6$-tag at its C-terminus by a Seamless Cloning Kit (Beyotime, Shanghai, China). The OSGEP gene was inserted between the EcoRI and HindIII sites of pCMV-3Tag-3A and the HindIII and EcoRI sites of pEGFP-N2. Similarly, TP53RK was inserted between the HindIII and XhoI sites of pCMV-3Tag-3A and the HindIII and BamHI sites of pEGFP-N2 and TPRKB was inserted between the XhoI and HindIII sites of pEGFP-N2. GON7 was inserted between the HindIII and BamHI sites of pEGFP-N2, and LAGE3 was inserted between the HindIII and EcoRI sites of pEGFP-N2.

All the primers used for cloning were listed in Supplementary Table S2. Gene mutations were obtained through KOD-plus mutagenesis kits (TOYOBO, Osaka, Japan) according to the manufacturer's instructions.

Gene expression vectors were expressed in *E. coli* BL21 (DE3) cells. The genes encoding hYRDC and Sua5 were expressed as described in a previous report (15). The overexpression of the hKEOPS complex and *Sc*KEOPS was carried out when the initial cell culture reached an absorbance at 600 nm ($A_{600}$) of 0.6–0.8, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added at a final concentration of 0.5 mM, and the cells were induced overnight at 16°C. *Ce*KEOPS gene expression was induced with 200 μM IPTG, and transformants were cultured for 3–5 h at 37°C. Expression of the hIleRS ($Met^1$-$Ser^{1073}$) gene was induced with 50 μM IPTG overnight at 18°C. Protein purification was initially performed according to a previously described method (32). After initial purification via Ni-NTA affinity chromatography, hKEOPS, hYRDC and *Ce*KEOPS were further purified by gel filtration on a Superdex S200 column with 50 mM Tris-HCl (pH 7.5) and 150 mM NaCl at the rate of 0.5 ml/min. *Sc*KEOPS and hIleRS ($Met^1$-$Ser^{1073}$) were further purified via ion exchange chromatography (Mono Q column), which was first pre-equilibrated with buffer A (containing 25 mM Tris–HCl (pH 7.8), 50 mM NaCl and 1 mM DTT). The proteins were eluted by a linear gradient from buffer A to buffer B (containing 25 mM Tris–HCl (pH 7.8), 1 M NaCl and 1 mM DTT) at the rate of 1.0 ml/min. Fractions containing *Sc*KEOPS and hIleRS protein were concentrated in 30 kDa molecular mass cut-off Amicon. The protein concentration was determined via a Protein Quantification Kit (Beyotime) according to the manufacturer's instructions.

### tRNA gene cloning and transcription

Genes encoding human cytoplasmic (hc) $tRNA^{Thr}$(AGU, CGU, UGU), $tRNA^{Ser}$(GCU), $tRNA^{Arg}$(CCU, UCU), $tRNA^{Asn}$(GUU), $tRNA^{Met}$(e), initiator $tRNA^{Met}$ ($tRNA^{Met}$(i)), $tRNA^{Lys}$(UUU), $tRNA^{Lys}$(CUU), $tRNA^{Ile}$(AAU), $tRNA^{Ile}$(UAU), $tRNA^{Ile}$(GAU), *E. coli* (*Ec*) $tRNA^{Met}$(i) and *S. cerevisiae* (*Sc*) $tRNA^{Met}$(i) were incorporated into a pTrc99b plasmid. tRNA transcripts were obtained by *in vitro* T7 RNA polymerase transcription as described previously (33,34). The primers used for template preparation were listed in Supplementary Table S2.

### Determination of *in vitro* $t^6A$ modification and aminoacylation activities

The $t^6A$ modification reaction was performed at 37°C in a 40 μl reaction mixture containing 50 mM Tris–HCl (pH 8.0), 200 mM NaCl, 15 mM $MgCl_2$, 5 mM $MnCl_2$, 50 mM $NaHCO_3$, 5 mM DTT, 4 mM ATP, 100 μM [$^{14}$C]Thr, 10 μM transcribed hctRNAs (or its variants) and 2 μM Sua5 and *Sc*KEOPS or hYRDC and hKEOPS or hYRDC and *Ce*KEOPS. The aminoacylation reaction was performed at 37°C in a 40 μl reaction mixture containing 60 mM Tris–HCl (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 5 mM DTT, 4 mM ATP, 20 μM [$^3$H]Ile, and 5 μM $t^6A$ unmodified or modified $hctRNA^{Ile}$ (AAU) or $hctRNA^{Ile}$ (UAU) or $hctRNA^{Ile}$ (GAU) transcript and 1 μM hIleRS ($Met^1$-$Ser^{1073}$). Aliquots (9 μl) of the reaction solution were added to filter pads at various time points and quenched with cold 5% TCA. The pads were washed three times for 15 min each with cold 5% TCA and then three times for 10 min each with 100% ethanol. Finally, the pads were dried under a heat lamp, and the radioactivity of the precipitates was quantified using a scintillation counter (Beckman Coulter, Atlanta, GA, USA).

### Melting temperature ($T_m$) assays of tRNAs

The specific tRNA was dissolved in a buffer containing 50 mM sodium phosphate buffer (pH 7.0), 10 μM EDTA and 100 mM NaCl. The initial absorbance of the tRNA at 260 nm was diluted to between 0.2 and 0.3. The melting temperature curve was determined at 260 nm using an Agilent Cary 100 spectrophotometer at a heating rate of 1°C per minute from 25°C to 95°C.

### Cell culture, transfection, nucleocytoplasmic separation and immunofluorescence assays

HEK293T cells were cultured in Dulbecco's modified Eagle's media supplemented with 10% fetal bovine serum in a 37°C incubator with 5% $CO_2$. Transfection was performed using Lipofectamine 2000 transfection reagent. At 24 h after transfection, the cells were harvested and washed with ice-cold phosphate-buffered saline (PBS) three times. Nucleocytoplasmic separation was performed by a Nuclear and Cytoplasmic Protein Extraction Kit (P0027, Beyotime) according to the manufacturer's protocol. For immunofluorescence assays, the cells were fixed in 4% paraformaldehyde for 30 min and then permeated in 0.2% Triton X-100 for 10 min at RT. After washing with PBS, the fixed cells were blocked in PBS containing 4% BSA and then incubated with rabbit anti-OSGEP or anti-TP53RK antibodies at a 1:200 dilution overnight at 4°C. The cells were then immunolabeled with Alexa Fluor 488-conjugated goat anti-rabbit IgG in PBS at a 1:1000 dilution for 2 h and the nuclear counterstain DAPI for 5 min at RT. Fluorescent images were taken and analyzed using a Leica TCS SP8 STED confocal microscope (Leica).

### Construction of OSGEP knockdown (KD) cell lines

shRNA sequences (Supplementary Table S3) were inserted into the lentiviral vector pLKO.1 between AgeI and EcoRI sites. The lentiviral vectors were cotransfected with the packaging vector pCMVDR8.9 and enveloped vector pCMV-VSVG into HEK293T cells using Lipofectamine 2000 reagent for lentivirus production. After 48 h, viruses were collected for 48 h of infection of HEK293T cells. Stably infected cells were selected via puromycin (2.5 μg/ml) for 48 h. Finally, the surviving cells were identified by western blotting using anti-OSGEP antibody.

### Dual-luciferase reporter assays

6× ANN codons were inserted into a pmirGLO plasmid after the firefly luciferase (F-luc) gene ATG start codon. For the frameshifting assay, a specific nucleotide (see the RESULTS section) was inserted at a specific codon of the F-luc gene. Subsequently, the plasmids were transfected into WT and OSGEP KD cells in a 24-well plate using Lipofectamine 2000. After cultivating 24 h, the cells were harvested and assayed by a Dual-Glo Luciferase Assay System (Promega). Renilla luciferase (R-luc) was used to normalize F-Luc activity to evaluate the translation efficiency of the reporter.

### tRNA isolation and ultra-performance liquid chromatography–mass spectrometry/mass spectrometry (UPLC–MS/MS) analysis

Total RNA was extracted from WT and KD cells using TRIzol reagent. The fourteen endogenous ANN-decoding tRNAs and two other tRNAs ($tRNA^{Leu}$(CAA) and $tRNA^{Ser}$(CGA)) were isolated from total RNAs using tRNA fishing by their own solid-phage complementary biotinylated DNA probes (Supplementary Table S4) using streptavidin agarose resin (20361, Thermo Scientific). The biotinylated DNA probes were designed to complement the 5′ or 3′ sequences of the tRNAs. In brief, the specific biotinylated DNA probes were incubated with total RNA at 65°C for approximately 1.5 h in annealing buffer (1.2 M NaCl, 30 mM HEPES–KOH (pH 7.5), 15 mM EDTA, 0.5 mM DTT). Streptavidin agarose beads were then added to the mixture, which was subsequently incubated for 30 min at 65°C. After binding, the agarose beads were washed three times with washing buffer (0.1 M NaCl, 2.5 mM HEPES–KOH (pH 7.5), 1.25 mM EDTA, 0.5 mM DTT). tRNAs on the agarose beads were extracted using TRIzol reagent and precipitated by ethanol. Purified tRNAs were identified via 8 M urea polyacrylamide gel electrophoresis.

Before subjected to UPLC–MS/MS, 500 ng of tRNA transcripts or specific endogenous tRNAs were digested with 1 μl of nuclease P1, 0.2 μl of benzonase, 0.5 μl of phosphodiesterase I, and 0.5 μl of bacterial alkaline phosphatase in a 20 μl solution including 4 mM $NH_4OAc$ (pH 5.2) at 37°C overnight. After complete hydrolysis, 1 μl of the solution of products was subjected to UPLC–MS/MS.

### Western blotting

Cells were lysed using RIPA (Radio Immunoprecipitation Assay) buffer (50 mM Tris–HCl (pH 7.5), 150 mM NaCl, 0.5% sodium deoxycholate, 1% Triton X-100, 0.1% SDS) and proteins separated on a 10% separating gel via SDS-PAGE. Western blotting was performed as described in a previous report (35).

### Northern blotting

For tRNA abundance determination, 3 μg of total RNA was loaded on a 10% polyacrylamide–8 M urea gel in Tris-borate EDTA (TBE) buffer at RT under 150 V for 1.5 h. For the aminoacylation assays, 5 μg of total RNA dissolved in 0.1 M NaAc (pH 5.2) was electrophoresed through an acidic (0.1 M NaAc (pH 5.2)) 10% polyacrylamide 8 M urea gel at 4°C under 18 W for 16 h. The RNAs were then transferred onto a positively charged nylon membrane (Merck) at 4°C under 250 mA for 40 min. After UV crosslinking (8000 × 100 J/$cm^2$), the membrane was pre-blocked with prehybridization solution (4 × SSC (containing 0.6 M NaCl and 0.06 M Na-Citrate), 20 m$MNa_2HPO_4$, 7% SDS, 1.5× Denhardt solution, 0.4 mg/l fish sperm DNA) at 55°C for 1 h. The membrane was subsequently hybridized with digoxin (DIG)-labeled probes (Supplementary Table S4) for specific tRNAs and 5S rRNA at 55°C overnight. The membrane was then washed with 2 × SSC buffer (containing 0.1% SDS) followed by washing buffer (0.1 M maleic acid and 0.15 M NaCl (pH 7.5)) twice for 5 min and then blocked with 1× blocking reagent (0.1 M maleic acid, 0.15 M NaCl, 1% blocking reagent (pH 7.5)) for 30 min at RT. Then, the membrane was incubated with anti-digoxigenin-AP buffer (1× blocking reagent, 1:10 000 anti-digoxigenin-AP) for 1 h and washed twice every 10 min. Finally, the membrane was treated with CDP-Star and imaged via an Amersham Imager 680 system (GE, CA, USA).

## RESULTS

### All ANN-decoding tRNAs in human cells contain $t^6A$ modification

The human genome encodes >400 high-confidence tRNA genes (2); however, the modification atlas of full sets of human cytoplasmic tRNAs has not been fully established. In the human cytoplasm, ANN codons encode seven amino acids: Arg, Asn, Ile, Lys, Met, Ser and Thr. Accordingly, fourteen ANN-decoding tRNA isoacceptors, i.e., $tRNA^{Thr}$(AGU, CGU, UGU), $tRNA^{Ile}$(AAU, UAU, GAU), $tRNA^{Ser}$(GCU), $tRNA^{Lys}$(CUU, UUU), $tRNA^{Arg}$(CCU, UCU), initiator $tRNA^{Met}$ ($tRNA^{Met}$(i)), elongator $tRNA^{Met}$ ($tRNA^{Met}$(e)) and $tRNA^{Asn}$(GUU), were found to be potential modification substrates for the hKEOPS. Thus far, some ANN-decoding tRNAs such as $tRNA^{Met}$(i) (36), $tRNA^{Met}$(e) (37) and $tRNA^{Lys}$(UUU) (38) have been experimentally demonstrated to contain $t^6A$ moieties. To understand whether all the above tRNA isoacceptors harbor $t^6A$ modifications, we purified all these tRNA species from HEK293T cells via tRNA fishing in conjunction with a solid-phase complementary biotinylated DNA probe (Supplementary Figure S1A and B; Supplementary Table S4). UPLC-MS/MS confirmed that all fourteen ANN-decoding tRNAs harbored $t^6A$ modifications (Supplementary Figure S1C). Because the non-ANN-decoding tRNA samples, including $tRNA^{Leu}$(CAA) and $tRNA^{Ser}$(CGA), showed little evidence of $t^6A$ (Supplementary Figure S1C), suggesting that samples were not cross-contaminated with other tRNA species. Thus, the above results showed that the hKEOPS was able to introduce $t^6A$ modifications to all ANN-decoding tRNAs.

### C33 acts as an anti-determinant in $t^6A$ biogenesis by the yeast modifying machinery but not the human or nematode

To understand the molecular mechanism of $t^6A$ biogenesis by the hKEOPS complex, we cloned the open reading frames (ORFs) of all five human genes (*OSGEP*, *TP53RK*, *TPRKB*, *LAGE3*, *GON7*) into the bacterial expression vector pJ241 (pJ241-hKEOPS), with each ORF having an independent ribosome-binding site (RBS) and start and stop codons, in a polycistronic gene organization, as was done for expression vector construction with respect to *Sc*KEOPS (30). Because all human ANN-decoding tRNAs harbored $t^6A$ modifications *in vivo*, we also constructed all the necessary tRNA clones and obtained all fourteen tRNA transcripts via T7 *in vitro* transcription.

We overexpressed pJ241-hKEOPS in bacteria; however, in our system, the integrity of the hKEOPS complex was always disrupted (disassociation of GON7) during the subsequent purification step (size exclusion or ion exchange chromatography) following His-tag affinity chromatography (Supplementary Figure S2A). In contrast, the *Sc*KEOPS complex was well expressed and purified with an intact composition with all subunits, as shown in a previous report (Supplementary Figure S2B) (30). Indeed, the $t^6A$ synthesis activity of the purified stoichiometrically inhomogeneous hKEOPS was much lower than that of *Sc*KEOPS under the same conditions (Supplementary Figure S2C). Based on the homogeneity, purity and activity, we preferentially used the *Sc*KEOPS complex as a eukaryotic KEOPS model to modify human cytoplasmic tRNAs to understand tRNA recognition and selection mechanism. We understand some limitations of modifying human tRNAs using a yeast complex; therefore, for modifying some key tRNA mutants, hKEOPS was also used for comparison. Furthermore, KEOPS from another multicellular organism, *C. elegans* (*Ce*KEOPS), was also cloned and purified (Supplementary Figure S2D). Fortunately, such a combination in determining the activity (modification of the same set of human tRNAs by different KEOPS complexes) provided some unexpected insights, which revealed both similarity and some undetected striking differences between KEOPS complexes from yeasts, humans and nematodes (see below).

We found that all fourteen human tRNAs (except $tRNA^{Met}$(i)) were effectively modified by *Sc*KEOPS, despite varying modification levels (Supplementary Figure S3). $tRNA^{Thr}$(UGU), $tRNA^{Lys}$(CUU), $tRNA^{Arg}$(UCU) and $tRNA^{Ser}$(GCU) were among the best substrates (Supplementary Table S5). The inability (or inefficiency) of $tRNA^{Met}$(i) modification by *Sc*KEOPS was unanticipated and puzzling because it readily harbored a $t^6A$ modification based on our UPLC–MS/MS data (albeit with the lowest detection scale among fourteen ANN-decoding tRNAs) (Supplementary Figure S1C) and previous data obtained by others (36).

We then compared the $tRNA^{Met}$(i) sequence with thirteen other tRNA sequences, especially the anticodon sequence, which is in proximity to the modification site A37. Indeed, we noticed that only $tRNA^{Met}$(i) contains C33, which is otherwise U33 in other tRNAs (Figure 1A), suggesting that C33 is a potential anti-determinant for $t^6A$ modification by *Sc*KEOPS. To explore this possibility, C33 of $tRNA^{Met}$(i) was mutated to U33. The modification assay indeed showed robust modification of $tRNA^{Met}$(i)-C33U by *Sc*KEOPS (Figure 1B). Therefore, these data suggested that C33 prevents $tRNA^{Met}$(i) from being efficiently modified by *Sc*KEOPS.

Considering that $tRNA^{Met}$(i) harbors $t^6A$ modification *in vivo*, it also implies that C33 is no longer an anti-determinant for hKEOPS. To answer this question, both $tRNA^{Met}$(i) and $tRNA^{Met}$(i)-C33U were modified by hKEOPS. In contrast to *Sc*KEOPS, hKEOPS introduced $t^6A$ modifications to both tRNAs with similar efficiency (Figure 1C). Moreover, comparable $t^6A$ modifications of $tRNA^{Met}$(i) and $tRNA^{Met}$(i)-C33U were observed using *Ce*KEOPS, with a higher efficiency (Figure 1D).

Above all, these data suggested that C33 functions as an anti-determinant in $t^6A$ biogenesis by the yeast modifying machinery but not the human or nematode.

### Possible coevolution of position 33 of $tRNA^{Met}$(i) and the $t^6A$ modification machinery

We further checked position 33 of $tRNA^{Met}$(i) from other species in addition to humans, including representative model species. Interestingly, those from either prokaryotes (e.g. *E. coli*, *Staphylococcus aureus*, *Aquifex aeolicus*, *Bacillus subtilis*) or single-cellular organisms (e.g. *S. cerevisiae*, *Schizosaccharomyces pombe*, *Candida glabrata*) contain U33, while those from multicellular organisms

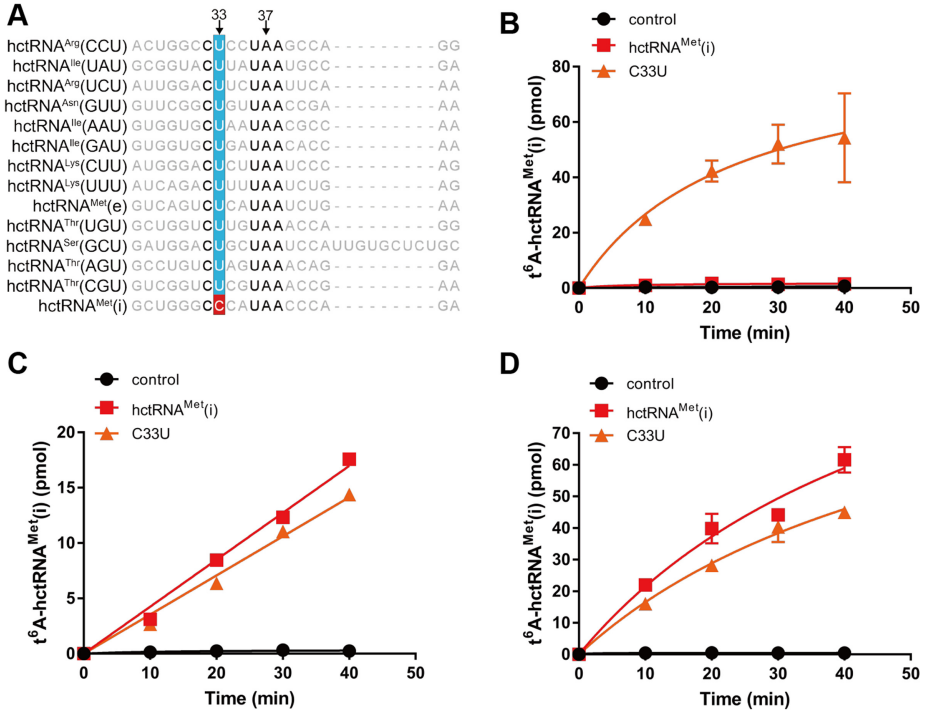


**Figure 1.** C33 is an anti-determinant for t$^6$A biogenesis in yeast but not humans or nematodes. (**A**) Sequence alignment of anticodon stem and loop regions of 14 ANN-decoding hctRNAs. hctRNA$^{Met}$(i) contains C33, while the other tRNAs contain U33. Absolute conserved nucleotides are shown in bold. Time course curves of the t$^6$A modification of hctRNA$^{Met}$(i) (red filled squares) and hctRNA$^{Met}$(i)-C33U (orange filled triangles) by Sua5 and *Sc*KEOPS (**B**); by hYRDC and hKEOPS (**C**) or by hYRDC and *Ce*KEOPS (**D**). The controls (black filled circles) represent assays in which no tRNAs were added. The data represent the averages of three independent experiments and the corresponding standard deviations. The error bars were masked by the symbols in (C).

(e.g. humans, mice, *C. elegans*, *Drosophila melanogaster* and even *Arabidopsis thaliana*) all contain C33 (Figure 2A) (39). Considering that the hKEOPS and *Ce*KEOPS complexes but not the *Sc*KEOPS complex could modify human tRNA$^{Met}$(i) with C33, these observations raise the question of whether the t$^6$A modification apparatus from species with U33-containing tRNA$^{Met}$(i) uses C33 as an anti-determinant.

To explore this possibility, we cloned four genes (encoding TsaC, TsaB, TsaD and TsaE) of *E. coli* t$^6$A modification enzymes and purified individual subunit from *E. coli* expression system (Supplementary Figure S4). Using *E. coli* tRNA$^{Met}$(i) (*Ec*tRNA$^{Met}$(i)), t$^6$A modification activity was reconstituted using four *E. coli* proteins (designated *Ec*TsaBCDE). We found that after mutation of U33 to C33 in *Ec*tRNA$^{Met}$(i), *Ec*TsaBCDE was indeed unable to modify *Ec*tRNA$^{Met}$(i)-U33C (Figure 2B). Furthermore, we also transcribed *S. cerevisiae* tRNA$^{Met}$(i) (*Sc*tRNA$^{Met}$(i)) and the *Sc*tRNA$^{Met}$(i)-U33C mutant. Consistently, the *Sc*KEOPS complex modified *Sc*tRNA$^{Met}$(i) with a t$^6$A moiety but not *Sc*tRNA$^{Met}$(i)-U33C (Figure 2C).

In combination with modification capacities of U33- and C33-containing homogeneous and/or heterogeneous tRNA$^{Met}$(i) by four t$^6$A modification complexes (*Ec*TsaBCDE, *Sc*KEOPS, hKEOPS and *Ce*KEOPS), we proposed that coevolution of position 33 of tRNA$^{Met}$(i) and t$^6$A modification apparatuses likely occurred; that is, bacteria and yeast (with U33-containing tRNA$^{Met}$(i)) t$^6$A modification enzymes employed C33 as an anti-determinant, while those from higher eukaryotes (with C33-containing tRNA$^{Met}$(i)), including humans and nematodes, displayed a relaxed specificity for position 33.

## C32 is absolutely required for t$^6$A modification by various KEOPS complexes

Subsequently, we compared all fourteen ANN-decoding tRNAs in the context of their cloverleaf structure. The consensus positions were mainly located in four regions: the anticodon loop, D-arm, TψC-arm and CCA terminus (Figure 3A).

We initially targeted the anticodon loop considering its proximity to the modification site A37. tRNA$^{Thr}$(UGU) was one of the best substrates for *Sc*KEOPS (Supplementary Figure S3; Supplementary Table S5). In addition, tRNA$^{Thr}$(UGU) harbors no other modification in the anticodon loop besides t$^6$A37 and m$^3$C32, which requires t$^6$A37 as a prerequisite (33,40); however, in the anticodon loop, tRNA$^{Thr}$(AGU) contains I34 (41), and tRNA$^{Lys}$(UUU) contains mcm$^5$s$^2$U34 (42). Therefore, tRNA$^{Thr}$(UGU) was selected for studying the recognition mechanism at the anticodon loop. Previously, U36 was shown to be a determinant for t$^6$A modification in *Xenopus laevis* oocytes (21). Its importance in modification was further confirmed using the human mitochondrial t$^6$A modification enzyme OSGEPL1 (15). Thus, U36 and modification site A37 were not included in the present assays. For other positions in the anticodon loop, each was mutated to the other three

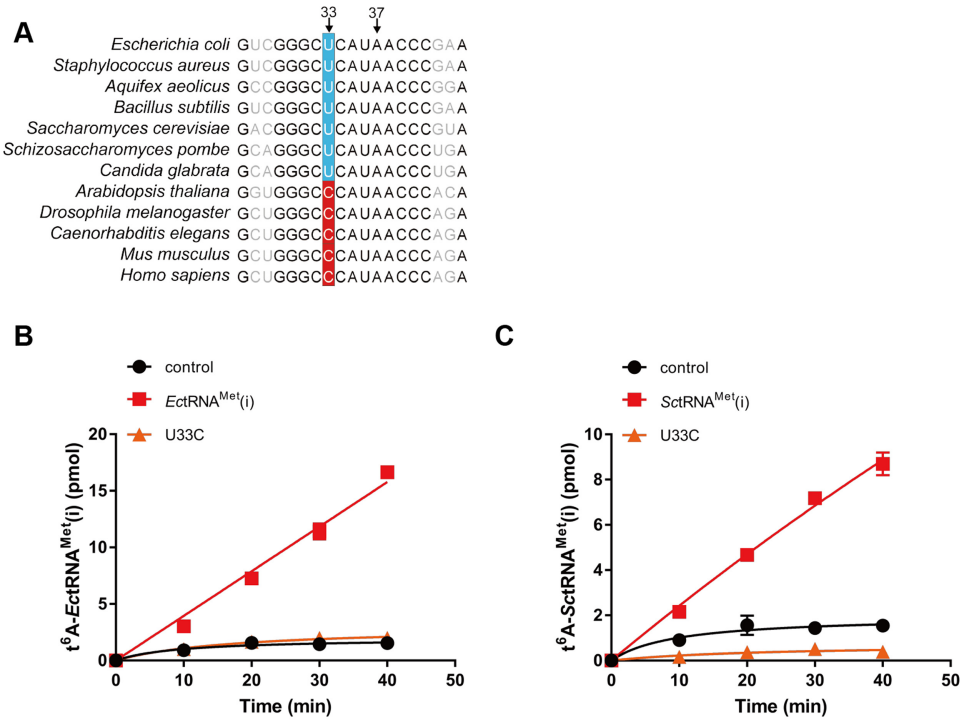


**Figure 2.** Possible coevolution of position 33 of the initiator $tRNA^{Met}$ and the $t^6A$ modification machinery. (**A**) Sequence alignment of anticodon stem and loop regions of $tRNA^{Met}$(i) from different species. Absolute conserved nucleotides are shown in bold. Time course curves of the $t^6A$ modification of *Ec*$tRNA^{Met}$(i) (red filled squares) and *Ec*$tRNA^{Met}$(i)-U33C (orange filled triangles) by *Ec*TsaBCDE (**B**) and of *Sc*$tRNA^{Met}$(i) (red filled squares) and *Sc*$tRNA^{Met}$(i)-U33C (orange filled triangles) by Sua5 and *Sc*KEOPS (**C**). The controls (black filled circles) represent assays in which no tRNAs were added. The data represent the averages of three independent experiments and the corresponding standard deviations. The error bars were masked by the symbols in (B).

nucleotides. We found that *Sc*KEOPS was unable to modify all C32 mutants of $tRNA^{Thr}$ (UGU) (Figure 3B). Consistent with the failure to modify $tRNA^{Met}$(i) with C33, $tRNA^{Thr}$(UGU)-U33A, -U33C and -U33G were not modified by *Sc*KEOPS (Figure 3C). However, mutations at positions U34 and G35 had little effect on $t^6A$ biogenesis (Figure 3D, E). Last, *Sc*KEOPS was unable to introduce $t^6A37$ in all mutants at A38 ($tRNA^{Thr}$(UGU)-A38C, -A38G and -A38U) (Figure 3F). Therefore, these results clearly showed that C32, U33 and A38 constitute essential determinants for $t^6A37$ biogenesis catalyzed by *Sc*KEOPS.

To reveal potential conservation across species, we also performed modifications of all the mutants by *Ce*KEOPS. Failure to modify all the mutants of C32 and A38 by *Ce*KEOPS was consistently observed (Figure 3G, H). Furthermore, the modification levels of all the mutants of U34 and G35 were comparable to that of wild-type $tRNA^{Thr}$(UGU) by *Ce*KEOPS (Supplementary Figure S5A, B). In sharp contrast, $tRNA^{Thr}$(UGU)-U33A, -U33C and -U33G were readily modified by *Ce*KEOPS despite variable efficiencies, suggesting that position 33 is not a critical base for *Ce*KEOPS (Figure 3I). We also determined the modification of $tRNA^{Thr}$(UGU)-U33A, -U33C and -U33G by hKEOPS. Again, all the U33 mutants were modified by hKEOPS with variable efficiency (Figure 3J), suggesting that mutations at position 33 have little effect on the recognition and catalytic activity of hKEOPS.

Altogether, these results showed that *Sc*KEOPS has a stricter substrate recognition mechanism at the anticodon loop region, requiring both the presence of C32 and U33, than do the *Ce*KEOPS and hKEOPS complexes (in which U33 is nonessential), highlighting the critical role of C32 in $t^6A37$ biogenesis across all eukaryotic KEOPS complexes.

## Different requirements of the CCA end in $t^6A37$ biogenesis among KEOPS complexes

All tRNAs have a CCA end for amino acid attachment. A recent report revealed that the CCA end is essential for $t^6A$ modification by the archaeal *Methanocaldococcus jannaschii* KEOPS complex for binding to the Cgi121 subunit (18). However, yeast genetic data from a Cgi121 gene knockout strain showed that $tRNA^{Ile}$(AAU) readily harbors $t^6A37$ (12). In addition, *Pyrococcus abyssi* Pcc1, Kae1 and Bud32 form a minimal functional unit that can generate $t^6A$ modification (19). These seemingly contradictory observations suggest possible divergence in the role of the CCA end in $t^6A$ modification. To precisely understand the role of the CCA end in $t^6A$ modification by eukaryotic KEOPS complexes, we transcribed a CCA end-truncated $tRNA^{Thr}$(UGU) mutant ($tRNA^{Thr}$(UGU)-ΔCCA). Modification by *Sc*KEOPS showed that $tRNA^{Thr}$(UGU)-ΔCCA could be $t^6A$ modified; however, the efficiency decreased to approximately 30% of that of wild-type $tRNA^{Thr}$(UGU) (Figure 4A), suggesting that the CCA end is important but not essential for $t^6A$ modification. However, *Ce*KEOPS was completely inactive in catalyzing $t^6A$ modification at $tRNA^{Thr}$(UGU)-ΔCCA (Figure 4B). Furthermore, we

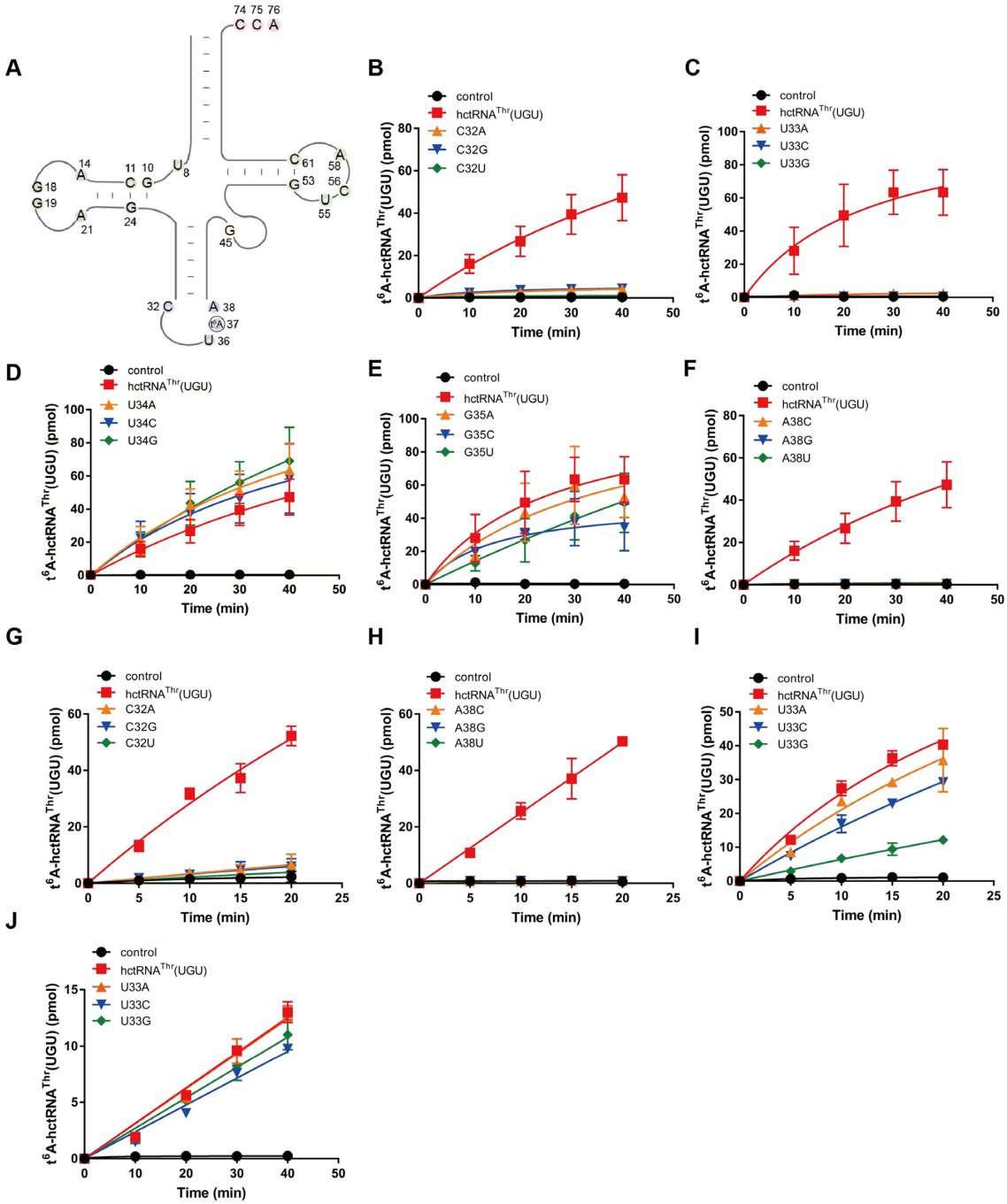


**Figure 3.** C32 is absolutely required for $t^6A$ modification by various KEOPS complexes. (**A**) Secondary structure showing conserved nucleotides of 14 ANN-decoding hctRNAs. Time course curves of the $t^6A$ modification of hctRNA$^{Thr}$(UGU) (red filled squares) together with hctRNA$^{Thr}$(UGU)-C32A (orange filled triangles), -C32G (blue filled inverted triangles) and -C32U (green filled diamonds) by Sua5 and *Sc*KEOPS (**B**); of hctRNA$^{Thr}$(UGU) (red filled squares) together with hctRNA$^{Thr}$(UGU)-U33A (orange filled triangles), -U33C (blue filled inverted triangles) and -U33G (green filled diamonds) by Sua5 and *Sc*KEOPS (**C**); of hctRNA$^{Thr}$(UGU) (red filled squares) together with hctRNA$^{Thr}$(UGU)-U34A (orange filled triangles), -U34C (blue filled inverted triangles) and -U34G (green filled diamonds) by Sua5 and *Sc*KEOPS (**D**); of hctRNA$^{Thr}$(UGU) (red filled squares) together with hctRNA$^{Thr}$(UGU)-G35A (orange filled triangles), -G35C (blue filled inverted triangles) and -G35U (green filled diamonds) by Sua5 and *Sc*KEOPS (**E**); and of hctRNA$^{Thr}$(UGU) (red filled squares) together with hctRNA$^{Thr}$(UGU)-A38C (orange filled triangles), -A38G (blue filled inverted triangles) and -A38U (green filled diamonds) by Sua5 and *Sc*KEOPS (**F**). $t^6A$ modification levels of hctRNA$^{Thr}$(UGU) (red filled squares), hctRNA$^{Thr}$(UGU)-C32A (orange filled triangles), -C32G (blue filled inverted triangles) and -C32U (green filled diamonds) by hYRDC and *Ce*KEOPS (**G**); of hctRNA$^{Thr}$(UGU) (red filled squares), hctRNA$^{Thr}$(UGU)-A38C (orange filled triangles), -A38G (blue filled inverted triangles) and -A38U (green filled diamonds) by hYRDC and *Ce*KEOPS (**H**); and of hctRNA$^{Thr}$(UGU) (red filled squares), hctRNA$^{Thr}$(UGU)-U33A (orange filled triangles), -U33C (blue filled inverted triangles) and -U33G (green filled diamonds) by hYRDC and *Ce*KEOPS (**I**). $t^6A$ modification levels of hctRNA$^{Thr}$(UGU) (red filled squares), hctRNA$^{Thr}$(UGU)-U33A (orange filled triangles), -U33G (blue filled inverted triangles) and -U33G (green filled diamonds) by hYRDC and hKEOPS (**J**). The controls (black filled circles) represent assays in which no tRNAs were added. The data represent the averages of three independent experiments and the corresponding standard deviations.

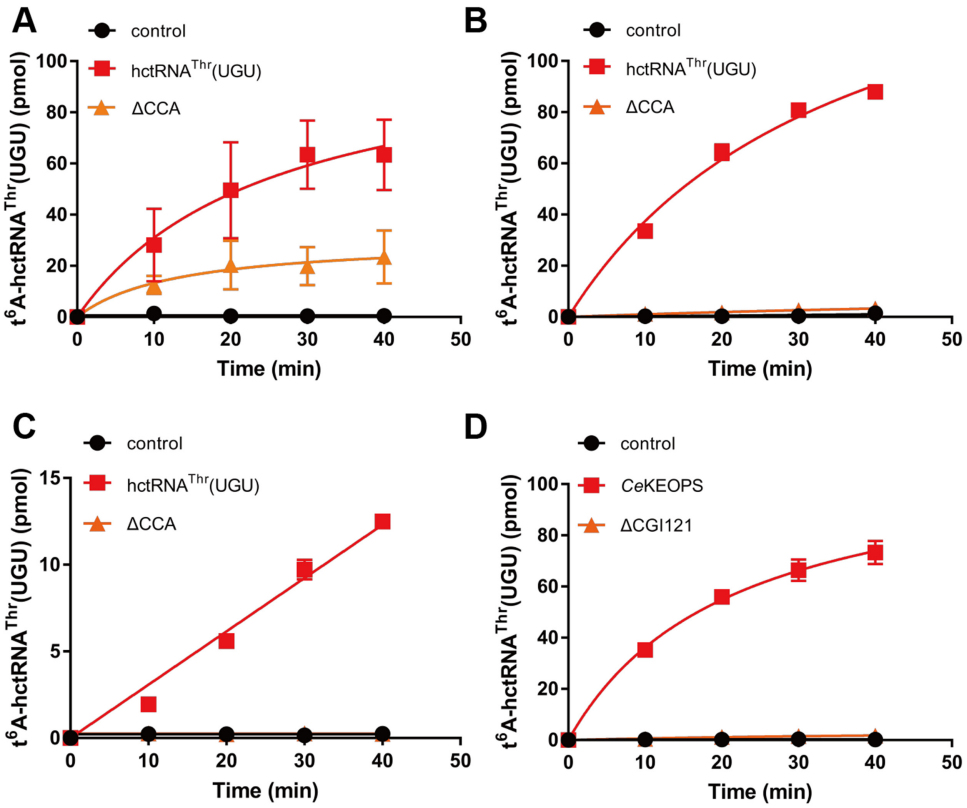


**Figure 4.** Different requirements of the CCA end in $t^6A37$ biogenesis among KEOPS complexes. Time course curves of the $t^6A$ modification of $hctRNA^{Thr}(UGU)$ (red filled squares) and $hctRNA^{Thr}(UGU)$-ΔCCA (orange filled triangles) by Sua5 and *Sc*KEOPS (**A**); of $hctRNA^{Thr}(UGU)$ (red filled squares) and $hctRNA^{Thr}(UGU)$-ΔCCA (orange filled triangles) by hYRDC and *Ce*KEOPS (**B**); of $hctRNA^{Thr}(UGU)$ (red filled squares) and $hctRNA^{Thr}(UGU)$-ΔCCA (orange filled triangles) by hYRDC and hKEOPS (**C**); and of $hctRNA^{Thr}(UGU)$ by hYRDC and *Ce*KEOPS (red filled squares) or by hYRDC and *Ce*KEOPS-ΔCGI121 (orange filled triangles) (**D**). The controls (black filled circles) represent assays in which no tRNAs were added (A–C) or in which no enzymes were added (D). The data represent averages of three independent experiments (A, B, D) or two independent experiments (C) and the corresponding standard deviations. The error bars were masked by the symbols in (B).

found that $tRNA^{Thr}(UGU)$-ΔCCA was similarly hypomodified by hKEOPS (Figure 4C), suggesting consistency between hKEOPS and *Ce*KEOPS.

Because the CCA end is bound by Cgi121 subunit in the archaeal KEOPS complex, to further investigate the role of Cgi121 in modification, a CGI121-deleted *Ce*KEOPS complex (*Ce*KEOPS-ΔCGI121) was purified (Supplementary Figure S2E). In line with the lack of $t^6A$ modification of $tRNA^{Thr}(UGU)$-ΔCCA, biogenesis of $t^6A$ by *Ce*KEOPS-ΔCGI121 with wild-type $tRNA^{Thr}(UGU)$ was not observed (Figure 4C), suggesting a crucial role of the interaction between CGI121 and the CCA end in *Ce*KEOPS.

Taken together, the above data clearly showed that the CCA end is an important but not critical motif in *Sc*KEOPS; however, the CCA determines the $t^6A$ modification capacity of both the *Ce*KEOPS and hKEOPS complexes, highlighting the evolutionary divergence of the CCA end in $t^6A$ modification among different KEOPS complexes.

### 10–25 and 11–24 base pairs are critical elements determining $t^6A$ modification levels across KEOPS complexes

In addition to the anticodon loop and CCA end, the consensus sequences among all fourteen ANN-decoding tRNAs were concentrated in the D-arm and TψC-arm (Figure 3A). Among these sites, U8 and A14 form base pairs to maintain a proper L-shaped tRNA scaffold in nearly all tRNAs. A58 functions a similar role by pairing with T54 to maintain the TψC-loop conformation (43). Therefore, the above three structural determinant sites were not included in the subsequent activity determination. The remaining elements include base pairs G18–U55, G19–C56, G10–C25 (only U25 in $tRNA^{Met}(i)$), C11–G24, G53–C61, A21 and G45 (Figure 3A).

For the G18–U55 and G19–C56 base pairs, we induced single-point mutations, G18C, G19C, U55G and C56G, to disrupt the interdomain interaction between the D-loop and TψC-loop (Figure 3A). Furthermore, G18–U55 or G16–C56 was switched to C18–G55 or C16–G56, respectively, to maintain the interaction. We found that modification of all these mutants by *Sc*KEOPS was only slightly (approximately <2-fold) decreased when compared with that of wild-type $tRNA^{Thr}(UGU)$ (Supplementary Figure S6A), suggesting that an L-shaped tRNA tertiary structure was not required for efficient $t^6A$ biogenesis. For the G53–C61 in the TψC-stem, we changed it to C53–G61 or A53–U61. We also constructed A21C or G45A mutant (Figure 3A). Similarly, *Sc*KEOPS introduced $t^6A$ modifications in these mutants with comparable efficiency (Supplementary Figure

S6B). Thus, G53–C61, A21 and G45 seem not to be critical elements in determining $t^6A$ modification levels.

For the remaining G10–C25 and C11–G24 base pairs, we constructed C10–G25, A10–U25, G11–C24 and A11–U24 mutants. Remarkably, *Sc*KEOPS was unable to introduce $t^6A$ modifications in the A10–U25, G11–C24, A11–U24 mutants; and the modification efficiency of C10–G25 was significantly decreased (by approximately one order of magnitude) (Figure 5A), suggesting that the G10-C25 and C11-G24 base pairs in the D-stem are crucial motifs for $t^6A$ modification. To confirm these results, C10–G25, A10–U25, G11–C24 and A11–U24 mutants were tested with hKEOPS. We found that hKEOPS was totally unable to introduce $t^6A$ modifications to all these mutants (Figure 5B), suggesting critical role of the G10–C25 and C11–G24 base pairs in $t^6A$ biogenesis. Consistently, *Ce*KEOPS modified A10–U25, G11–C24, A11–U24 mutants with a very reduced efficiency but that of C10–G25 was only slightly decreased (Figure 5C). To explore whether C10–G25, A10–U25, G11–C24 and A11–U24 mutations induced potential structural folding defects, we compared the $T_m$ values of wild-type $tRNA^{Thr}$(UGU) and the four mutants (Supplementary Table S6). A10–U25 mutant displayed a slight decrease in $T_m$, possible due to the weak interaction of A–U when compared with that of G10–C25 in wild-type tRNA. $T_m$ value of G11–C24 was even higher than wild-type tRNA, suggesting more compact folding. Overall, no significant decrease in $T_m$ values suggested little possibility of defective folding of these mutants.

In summary, these results clearly showed that the two base pairs (G10–C25 and C11–G24) in the D-stem, especially the C11–G24 pair, were among the key determinants for $t^6A$ modification by various eukaryotic KEOPS complexes.

### $t^6A$ is biosynthesized on mature tRNAs in cytoplasm

Notably, in human cells, two of the 14 ANN-decoding tRNAs ($tRNA^{Ile}$(UAU) and $tRNA^{Arg}$(UCU)) contain introns (39). The above data clearly showed that $t^6A$ modification enzymes require a well-organized tRNA anticodon loop for catalysis. Thus, we proposed that $t^6A$ modification likely occurs after the removal of introns. To explore this proposal, we transcribed $tRNA^{Ile}$(UAU) and $tRNA^{Arg}$(UCU) precursors (based on sequences of $tRNA^{Ile}$(UAU)-3-1 and $tRNA^{Arg}$(UCU)-3-1). Modification analyses using hKEOPS showed that $t^6A$ was only introduced to mature $tRNA^{Ile}$(UAU) and $tRNA^{Arg}$(UCU) but not their precursors (Figure 6A). We further performed similar $t^6A$ modifications using *Sc*KEOPS with the same results (Figure 6B). Taken together, these data showed that $t^6A$ is biosynthesized on mature tRNAs.

Due to mature human tRNAs are mainly localized in the cytoplasm under physiological conditions (38) (note that yeast intron-containing pre-tRNAs are also localized in the cytoplasm for splicing by the mitochondrial surface-localized tRNA splicing endonuclease) (44), the data suggested that $t^6A$ modification occurs in the cytoplasm. To further explore the localization of $t^6A$ biogenesis, we separated the cytoplasmic and nuclear fractions of HEK293T cells. Western blot analyses involving OSGEP and TP53RK antibodies showed that most endogenous OSGEP and TP53RK proteins were localized in the cytoplasm (Figure 6C). Similar cytoplasmic distribution results were obtained by western blot analyses with cytoplasmic and nuclear fractions of cells overexpressing either FLAG-tagged OSGEP (OSGEP-FLAG) (Figure 6D) or TP53RK (TP53RK-FLAG) (Figure 6E). Cytoplasmic localization of both endogenous OSGEP and TP53RK was also confirmed by immunofluorescence analysis in which anti-OSGEP or anti-TP53RK antibodies were used (Figure 6F). Moreover, each component of the hKEOPS was fused with a C-terminal EGFP and overexpressed in HEK293T cells. Fluorescence determination showed that OSGEP-EGFP, TPRKB-EGFP and LAGE3-EGFP were almost localized in the cytoplasm, while GON7-EGFP and TP53RK-EGFP were distributed in both the cytoplasm and the nucleus (Figure 6G).

Altogether, these data showed that the hKEOPS complex is mainly localized in the cytoplasm and introduces $t^6A$ modification to mature tRNA species.

### $t^6A$ is a critical positive determinant in aminoacylation of $tRNA^{Ile}$(AAU) and $tRNA^{Ile}$(GAU) but not $tRNA^{Ile}$(UAU) isoacceptors by human cytoplasmic IleRS

To understand the *in vivo* function of cytoplasmic $t^6A$ modification, we initially tried to delete the OSGEP gene in HEK293T cells by using CRISPR/Cas9-mediated gene editing. However, despite extensive efforts, we failed to obtain a null allele of OSGEP, suggesting that $t^6A$ modification is essential to cell viability. We then constructed two HEK293T cell lines stably overexpressing two independent shRNAs targeting the OSGEP gene. The protein level of OSGEP was obviously downregulated by individual shRNAs, as evidenced via western blot analysis using anti-OSGEP antibody (Figure 7A). To confirm a decrease in $t^6A$ content in tRNAs due to *OSGEP* knockdown, $tRNA^{Lys}$(UUU) was purified from WT and the two knockdown (KD) cells (shRNA1 and shRNA2) and then hydrolyzed to mononucleosides for UPLC–MS/MS analysis. Clearly, the $t^6A$ content of $tRNA^{Lys}$(UUU) in KD cells was significantly lower than that in WT cells (Figure 7B), demonstrating a reduction in $t^6A$ modification levels.

We initially explored the effects of $t^6A$ biogenesis impairment on steady-state tRNA levels for all ANN-decoding tRNAs via northern blots. The results clearly showed that the amounts of all ANN-decoding tRNAs were comparable between WT and shRNA1 cells (Supplementary Figure S7A). Furthermore, aminoacylation assays via acid northern blot analysis, in which condition the amino acid moiety remained uncleaved on tRNAs, revealed that charging levels for most tRNAs were not influenced in shRNA1 cells (Supplementary Figure S7B). However, the aminoacylation levels of two $tRNA^{Ile}$ isoacceptors ($tRNA^{Ile}$(AAU) and $tRNA^{Ile}$(GAU)) were obviously and significantly downregulated in repeated analyses (Figure 7C). Unexpectedly, alterations in the charging of $tRNA^{Ile}$ (UAU) were not ob-

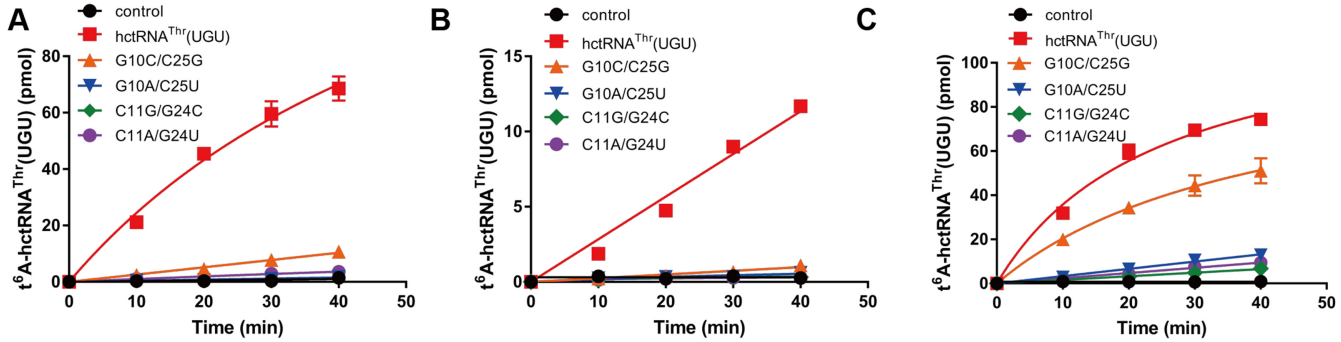


**Figure 5.** 10–25 and 11–24 base pairs are critical elements determining $t^6A$ modification levels across KEOPS complexes. Time course curves of the $t^6A$ modification of $hctRNA^{Thr}(UGU)$ (red filled squares), $hctRNA^{Thr}(UGU)$-G10C/C25G (orange filled triangles), -G10A/C25U (blue filled inverted triangles), -C11G/G24C (green filled diamonds) and -C11A/G24U (purple filled circles) by Sua5 and *Sc*KEOPS (**A**), by hYRDC and hKEOPS (**B**) or by hYRDC and *Ce*KEOPS (**C**). The controls (black filled circles) represent assays in which no tRNAs were added. The data represent the averages of three independent experiments and the corresponding standard deviations. The error bars were masked by the symbols in (B).

served (Figure 7C). Taken together, these data suggested that $t^6A$ modification has little effect on determining tRNA charging levels of most ANN-decoding tRNAs but functions as a key element in charging $tRNA^{Ile}$ (AAU) and $tRNA^{Ile}$ (GAU) isoacceptors *in vivo*.

To further explore whether $t^6A$ biogenesis determines $tRNA^{Ile}$ aminoacylation, we transcribed human cytoplasmic $tRNA^{Ile}$(AAU), $tRNA^{Ile}$(GAU) and $tRNA^{Ile}$(UAU). $t^6A$-modified transcripts were then obtained by incubating transcribed tRNAs with Sua5-*Sc*KEOPS *in vitro*. The modified $tRNA^{Ile}$(UAU) product was then subjected to hydrolysis to mononucleosides and subsequently analyzed via UPLC–MS/MS, confirming that the tRNA transcript was successfully loaded with the $t^6A$ moiety (Figure 7D). Full-length human cytoplasmic isoleucyl-tRNA synthetase (hIleRS, encoded by *IARS1*) forms a protein aggregate due to the presence of a C-terminal appended domain containing two repeats, which is not essential for $tRNA^{Ile}$ aminoacylation (45) and probably mediates hIleRS into the cytoplasmic multiple tRNA synthetase complex (46). We overexpressed a gene encoding a truncated hIleRS ($Met^1$-$Ser^{1073}$) devoid of the C-terminal domain (Supplementary Figure S8). *In vitro* aminoacylation assays clearly showed that hIleRS was completely incapable of catalyzing aminoacylation of $tRNA^{Ile}$(AAU) and $tRNA^{Ile}$(GAU) transcription; however, charging of the $tRNA^{Ile}$(UAU) transcript was robust and obvious. In contrast, $t^6A$ modification readily afforded aminoacylation of $tRNA^{Ile}$(AAU) and $tRNA^{Ile}$(GAU) by hIleRS (Figure 7E, F); however, aminoacylation of $tRNA^{Ile}$(UAU) was obviously inhibited by the presence of a $t^6A$ moiety by ∼2-fold (Figure 7G). These *in vitro* data were consistent with those obtained from the acid northern blot analyses.

Taken together, these data showed that knockdown of *OSGEP* led to an obvious decrease in $t^6A$ modification levels in human tRNAs. $t^6A$ modification defects have no role in determining ANN-decoding tRNA abundance or the tRNA charging level of most tRNAs, excluding two $tRNA^{Ile}$ isoacceptors. $t^6A$ modification is a critical determinant for $tRNA^{Ile}$ isoacceptor-specific aminoacylation by hIleRS.

### Role of $t^6A$ modification in the strength of codon-anticodon pairing and prevention of +1 frameshifting

We subsequently explored any potential contribution of $t^6A$ modification to codon decoding. We reasoned that if $t^6A$ modification is able to control A1 (of codon) and U36 base pairing in human cells, $t^6A$ deficiency would cause alteration in codon-anticodon base pairing strength and subsequent decoding efficiency. To this end, we designed a dual-luciferase reporter system in which 6× ANN codons were inserted downstream of the F-luc gene ATG start codon in a pmirGLO plasmid, which simultaneously contained a separate R-luc gene as a control (Figure 8A). By determining and comparing the fluorescence densities of F-luc and R-luc in WT and OSGEP KD cells, we quantified the decoding efficiency of each codon by a $t^6A$-harboring tRNA. We initially observed no alteration in translational efficiency of 6× TCT (Ser), which is decoded by non-$t^6A$-modified $tRNA^{Ser}$(AGA), between WT and KD cells. Then, the results showed that the translational efficiency of 6× AGG or AGA (Arg), 6× AAA or AAG (Lys), 6× AGC (Ser), 6× ATG (Met) and 6× ACT or ACA (Thr) in KD cells was significantly higher than that in the WT cells. However, the decoding efficiency of 6× ACG (Thr) and 6× ATT or ATC (Ile) codons in the KD cells was comparable to that in the WT cells. Note that the charging level of all tRNAs except $tRNA^{Ile}$(AAU) (decoding ATT codon) and $tRNA^{Ile}$(GAU) (decoding ATC codon) remained unchanged in the KD cells. The decoding rate of 6× ATA (Ile) codons in KD cells decoded by $tRNA^{Ile}$(UAU) with unaltered aminoacylation levels was lower than that in the WT cells (Figure 8B). Thus, these data showed that $t^6A$ modification contributes distinctly to decoding efficiency at different codons. For most ANN codons, $t^6A$ modifications of tRNA seemed to restrict their decoding efficiency; however, $t^6A$ modification of $tRNA^{Ile}$(UAU) probably stimulated its capacity to decode ATA codons. These codon-specific differential effects of $t^6A$ modification on tRNA decoding capacities are similar to observations in yeast cells (27).

Furthermore, we designed a reporter based on the above-described dual-luciferase system to assess whether $t^6A$ modification at a given tRNA is able to prevent +1 frameshift-

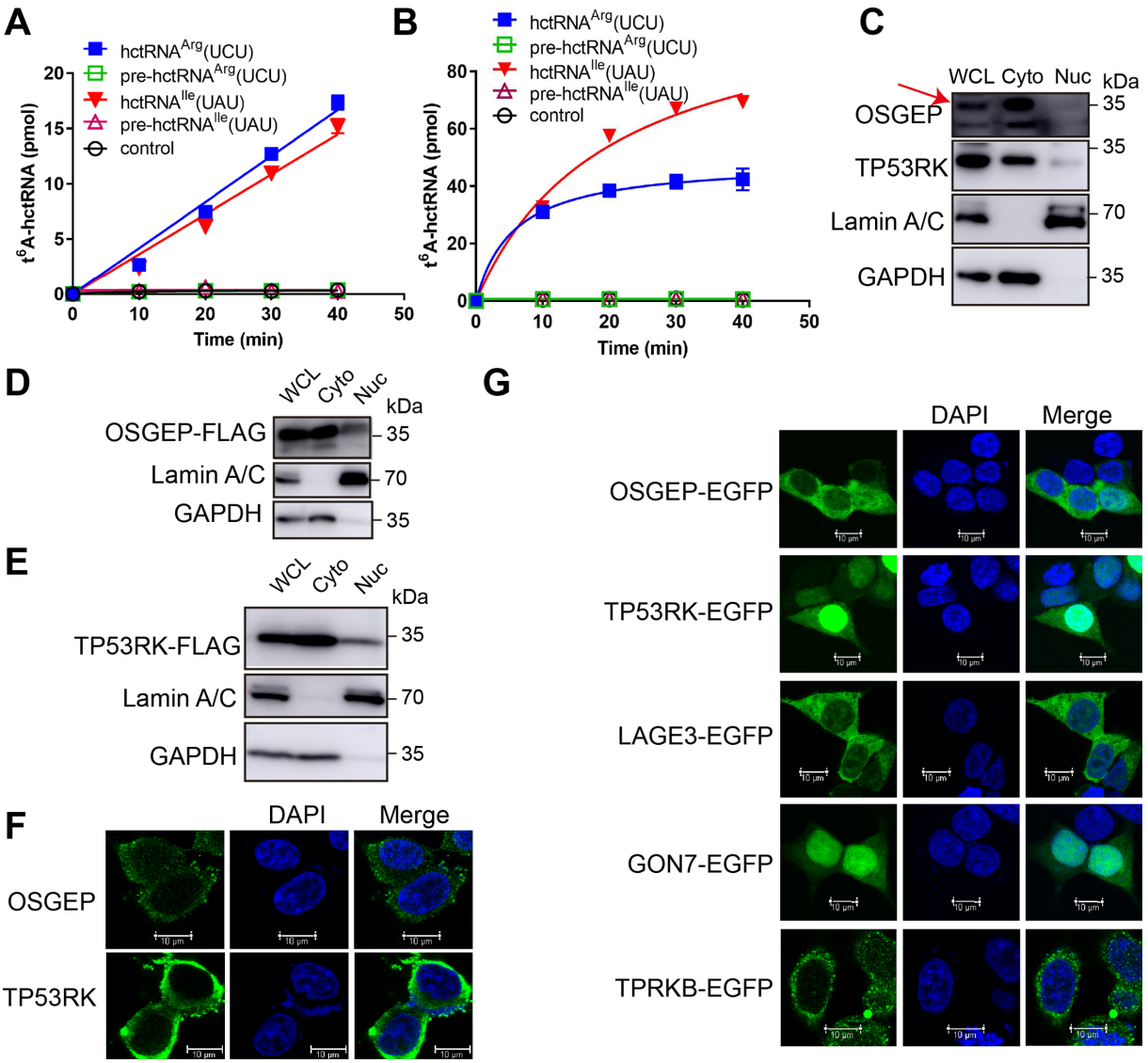


**Figure 6.** t$^6$A is introduced to mature tRNAs in the cytoplasm. Time course curves of the t$^6$A modification of four transcripts, tRNA$^{Arg}$(UCU) (blue filled squares), pre-tRNA$^{Arg}$(UCU) (green squares), hctRNA$^{Ile}$(UAU) (peach triangles), pre-hctRNA$^{Ile}$(UAU) (purplish red triangles), and controls (no tRNA addition, black circles), by YRDC and hKEOPS (**A**) or by Sua5 and *Sc*KEOPS (**B**). The data represent the average of three independent replicates and the standard deviations. Subcellular localization of endogenous OSGEP and TP53RK (**C**) or overexpressed OSGEP-FLAG (**D**) and TP53RK-FLAG (**E**) analyzed by nucleocytoplasmic separation assays. The red arrow represents OSGEP in (C). Cytoplasmic (Cyto) and nuclear (Nuc) fractions were separated from HEK293T cells. GAPDH and Lamin A/C were used as markers of the Cyto and Nuc fractions, respectively. (**F**) Immunofluorescence determination of endogenous OSGEP and TP53RK in HEK293T cells. (**G**) Fluorescence determination of the localization of the five relevant subunits (OSGEP-EGFP, TPRKB-EGFP, GON7-EGFP, TP53RK-EGFP and LAGE3-EGFP). Thu nucleus was stained with DAPI in (F) and (G).

ing. We initially selected tRNA$^{Lys}$(UUU), tRNA$^{Ile}$(AAU) and tRNA$^{Asn}$(GUU) because the codon-anticodon U-A base pairings may be more prone to +1 frameshifting due to their weak interaction strength. To measure +1 frameshifting at the ATT codon decoded by tRNA$^{Ile}$(AAU), an "A" nucleotide was inserted upstream of the ATT codon at position 7 of the Ile codon of F-luc (AATT-F-Luc) (Figure 8C). No F-luc was detected without +1 frameshifting due to premature translational termination; however, the occurrence of +1 frameshifting at this codon increased the F-luc activity in KD cells (the primary sequence of F-luc was unaltered). We detected slight expression of AATT-F-Luc in WT HEK293T cells, indicating that natural +1 frameshifting occurs at the ATT codon in human cells. Strikingly, F-luc activity was significantly elevated in KD cells compared with WT cells (by ∼2–3-fold) (Figure 8C), suggesting that a deficiency in t$^6$A among tRNA$^{Ile}$(AAU) promotes +1 frameshifting at the AUU codon. Notably, comparable increases in +1 frameshifting due to Sua5 deletion has also been observed in yeast (26). Similar designs were applied to the AAA codon (decoded by tRNA$^{Lys}$(UUU)) and AAC codon (decoded by tRNA$^{Asn}$(GUU)) (Supplementary Figure S9); however, no expression of F-luc was observed in either WT or KD cells, indicating that t$^6$A deficiency in tRNA$^{Lys}$(UUU) and tRNA$^{Asn}$(GUU) triggers no +1 frameshifting at AAA or AAC codons.

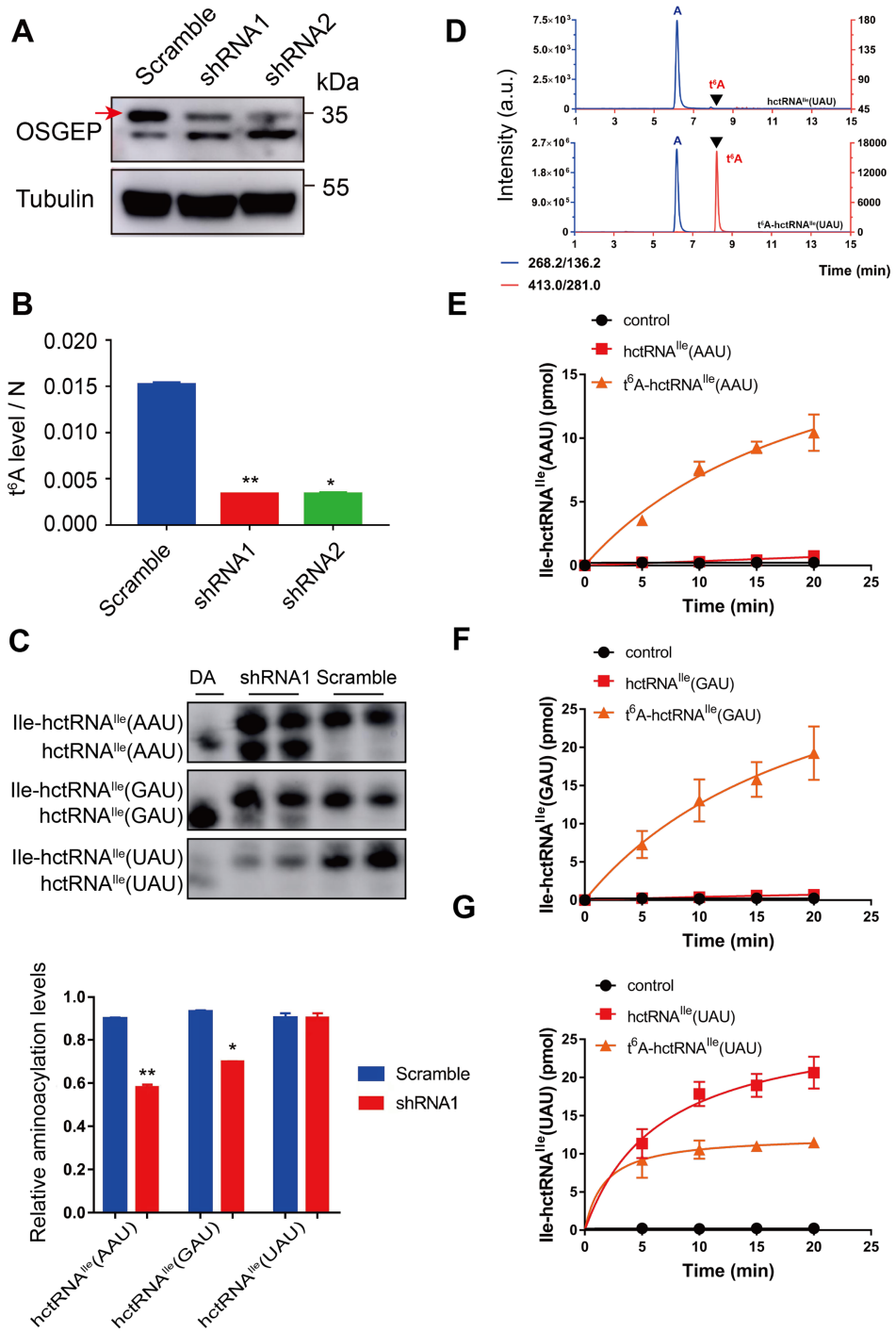


**Figure 7.** $t^6A$ is a critical positive determinant in aminoacylation of $tRNA^{Ile}(AAU)$ and $tRNA^{Ile}(GAU)$ but not $tRNA^{Ile}(UAU)$ isoacceptors by human cytoplasmic IleRS. (**A**) Western blot results showing OSGEP protein levels in HEK293T cells infected with OSGEP-specific (shRNA1, shRNA2) or control (Scramble) shRNAs. The red arrow represents OSGEP. (**B**) UPLC-MS/MS quantification of $t^6A$ modification levels in $tRNA^{Lys}(UUU)$ purified from WT and two KD (shRNA1 and shRNA2) cells. The amounts of $t^6A$ and N (= A+C+G+U) were calculated based on the area under the $t^6A$ or N peaks in the UPLC-MS/MS chromatogram. The data were two independent replicates and the corresponding standard deviations. The P values were determined using two-tailed Student's t test for paired samples. $*P < 0.05$; $**P < 0.01$. (**C**) Determination of tRNA charging levels of three $tRNA^{Ile}$ isoacceptors in wild-type (Scramble) and *OSGEP* knockdown HEK293T cell lines by acid northern blots. The relative aminoacylation levels were calculated based on two independent replicates and their corresponding standard deviations. DA, de-aminoacylated tRNA. The P values were determined using two-tailed Student's *t* test for paired samples. $*P < 0.05$; $**P < 0.01$. (**D**) UPLC-MS/MS analysis of the digested products of hypo- or modified $tRNA^{Ile}(UAU)$. Aminoacylation of $hctRNA^{Ile}(AAU)$ (red filled squares) and $t^6A$-$hctRNA^{Ile}(AAU)$ (orange filled triangles) (**E**); or $hctRNA^{Ile}(GAU)$ (red filled squares) and $t^6A$-$hctRNA^{Ile}(GAU)$ (orange filled triangles) (**F**) or $hctRNA^{Ile}(UAU)$ (red filled squares) and $t^6A$-$hctRNA^{Ile}(UAU)$ (orange filled triangles) (**G**) by hIleRS. Controls (black filled circles) represent assays in which no tRNAs were added. The data represent the averages of three independent experiments and the corresponding standard deviations.

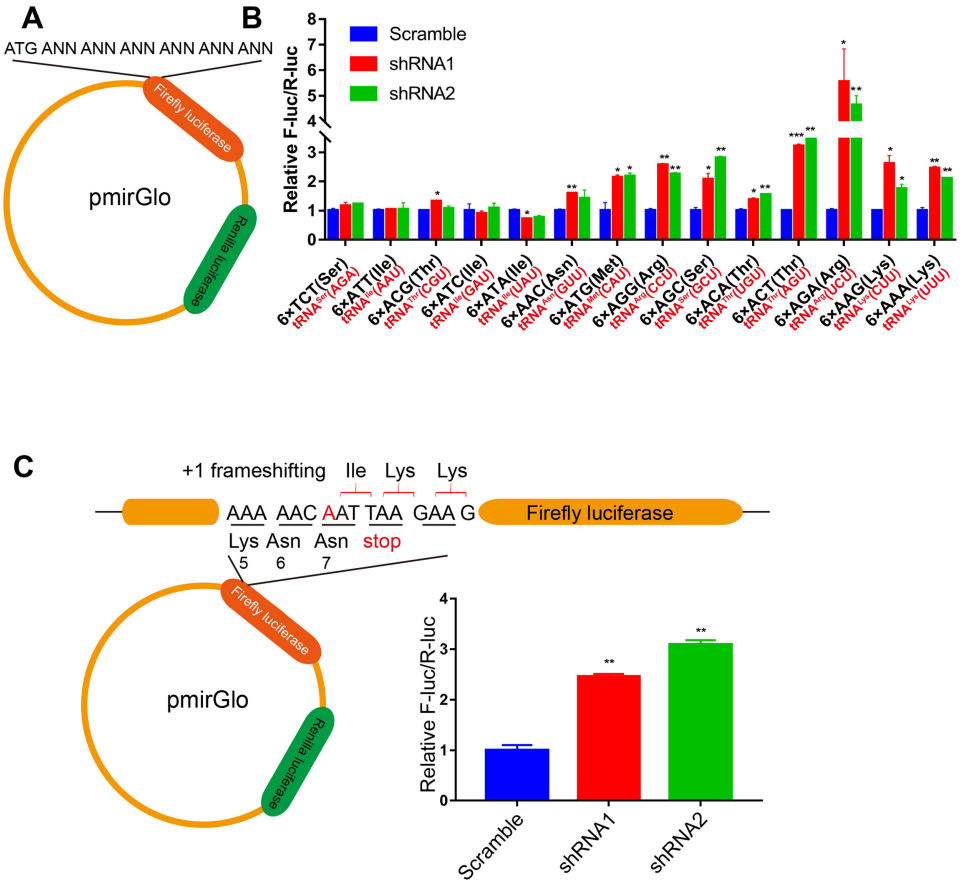


**Figure 8.** Effects of $t^6A$ deficiency on codon-anticodon pairing and +1 frameshifting. (**A**) Schematic showing the dual-luciferase system, in which 6× ANN codons were inserted downstream of the F-luc gene ATG start codon in a pmirGLO plasmid, and a separate R-luc gene was used as a control. (**B**) Effects of $t^6A$ deficiency on ANN-codon (in black in the x-axis) decoding efficiency by various tRNAs (in red in the x-axis) in WT (Scramble) and two OSGEP KD cell lines. 6× TCT codons decoded by non-$t^6A$-modified $tRNA^{Ser}$(AGA) were included as controls. The data represent the average of two independent replicates and the standard deviations. The P values were determined using two-tailed Student's t test for paired samples. $*P < 0.05$; $**P < 0.01$ and $***P < 0.001$. (**C**) Schematic showing the +1 frameshifting assay at the ATT codon decoded by $tRNA^{Ile}$(AAU). An "A" nucleotide (indicated by red) was inserted upstream of the ATT codon at position 7 of the Ile codon of F-luc (AATT-F-Luc). The data represent the average of three independent replicates and the standard deviations. The *P* values were determined using two-tailed Student's *t* test for paired samples. $*P < 0.05$; $**P < 0.01$ and $***P < 0.001$.

## DISCUSSION

### C32 and the D-stem are essential elements in $t^6A$ biogenesis by various KEOPSs

In this work, we studied in detail how representative KEOPS complexes from yeast, humans and nematodes accurately select tRNA substrates. We found that all three KEOPS complexes are critically dependent on common elements embedded in the anticodon loop and the D-stem. In particular, in addition to previously identified U36 and A38, C32 is a critical determinant for KEOPS complexes (21). An earlier study using the *Xenopus laevis* oocyte $t^6A$ modification system failed to determine the critical importance of C32, likely because double mutations of C32 and A38 were introduced; thus, the contribution of C32 was not definitively determined. However, C32 of human mitochondrial $tRNA^{Thr}$ is not an essential element in mitochondrial $t^6A$ formation by hYRDC and OSGEPL1 (15), suggesting that divergence in the recognition of position 32 has occurred between cytoplasmic and mitochondrial $t^6A$ modification enzymes. It is also notable that in some archaea, such as *Methanococcus jannaschii*, the $tRNA^{Thr}$(GGU) isoacceptor, which is absent in eukaryotic genomes, contains a U32 and $t^6A$-analogous $hn^6A$ modification ($N^6$-hydroxynorvalylcarbamoyl adenosine) derived from hydroxy norvaline instead of threonine, although most $t^6A$- or $hn^6A$-containing tRNAs in *M. jannaschii* contain C32 (47,48). In fact, U32 is widespread in various archaeal $tRNA^{Thr}$(GGU) isoacceptors (49), probably because position 32 is more flexible in archaea than in eukaryotes because of the lack of $m^3C32$ modification in the former (1,33,47). Moreover, $hn^6A$ is found in some thermophilic bacteria (48). It has been suggested that $hn^6A$ is likely synthesized by archaeal Sua5 and KEOPS, because archaeal Sua5 was proven to be active in hydrolyzing ATP to AMP in the presence of hydroxy norvaline (50). Indeed, a recent report also confirmed that bacterial TsaC is able to exhibit relaxed substrate specificity to produce a variety of TC-AMP analogs (51). These observations suggested that the archaeal KEOPS is likely not reliant on C32 as a key determinant for $t^6A$ or $hn^6A$ biogenesis and that dependence on C32 was likely a later event that occurred during KEOPS

evolution. The C32-binding site of the KEOPS is unclear. According to a recent KEOPS-tRNA interaction model, LAGE3 and OSGEP (human Pcc1 and Kae1, respectively) constitute the binding surface for the tRNA anticodon loop (18). Pcc1 and Kae1 have been consistently proposed to be tRNA-binding sites (19). While A37-containing one site of the loop extends into the active site of OSGEP, another site with C32 is probably bound by LAGE3. Indeed, mutation of a potential tRNA-binding residue ($Arg^{63}$ in the Pcc1 subunit of the archaeal KEOPS) abolishes the $t^6A$ modification activity of archaeal KEOPS (18). Gon7 subunit interacts with Pcc1 in an opposite side and is spatially far from tRNA anticodon loop (5); it is unlikely to be C32 binding partner. Nonetheless, the interaction between the tRNA anticodon loop and KEOPS need to be further explored in detail.

In addition, we identified here that two base pairs in the D-stem, G10–C25 and C11–G24, constitute the key elements determining efficient $t^6A$ biogenesis. Only G10-U25 is present in the human $tRNA^{Met}(i)$ isoacceptor, which is a comparably good substrate of KEOPS. Similarly, a previous report revealed the critical role of a conserved $^{10}CU^{11}$ motif in the D-arm of archaeal tRNAs in both $t^6A$ modification and stimulation of the ATPase of archaeal Bud32 (18). This evidence highlights the conserved feature of the critical role of the D-arm in both eukaryotic and archaeal KEOPS complexes. The G10–C25 and C11–G24 motifs probably interact with OSGEP and/or TP53RK according to the archaeal KEOPS–tRNA interaction model (5,18).

### Requirement of CCA end is divergent among various KEOPSs

A previous report has shown that $tRNA^{Ile}$ from yeast cells deprived of Cgi121 harbors an abundance of $t^6A$ modification, implying that the CCA end of yeast tRNA is not a determinant (12). Indeed, we found that *Sc*KEOPS is able to introduce $t^6A$ modification with CCA-truncated tRNA *in vitro*. However, both the hKEOPS and *Ce*KEOPS complexes are completely incapable of catalyzing $t^6A$ biogenesis in the absence of the CCA end. Our data are consistent with the recent findings that CCA-truncated tRNA is a poor binding substrate of hKEOPS (18). Taken together, these data suggested that hKEOPS and *Ce*KEOPS more resemble archaeal KEOPS in terms of the requirement for the CCA end than *Sc*KEOPS.

### Possible coevolution of position 33 of tRNAs and the $t^6A$ modification machinery

One of the unexpected findings is the sharp difference in the role of position 33 between the different $t^6A$ machineries. Both the hKEOPS and *Ce*KEOPS complexes have relaxed requirements for position 33. Similar results were observed with OSGEPL1 (15). However, C33 is obviously an antideterminant of both *Ec*TsaBCDE and yeast Sua5-KEOPS. Only a single mutation of C33 to U33 makes human $tRNA^{Met}(i)$ a well-qualified substrate for yeast $t^6A$ modification machinery. Interestingly, nearly all ANN-decoding tRNAs contain U33; however, $tRNA^{Met}(i)$ from multicellular organisms is an exception by later evolving a C33, possibly to finetune initiation efficiency; but the exact evolutionary advantage needs to be further studied. On the other hand, accompanying a relaxed requirement at position 33 due to evolution of U33 to C33 in $tRNA^{Met}(i)$, the CCA end seems to play a critical role in modification by KEOPS complexes from higher eukaryotes (humans and nematodes). Based on the above observations and analyses, we proposed a hierarchal evolutionary scenario that occurred from bacteria and single-cellular lower eukaryotes to higher eukaryotes. In bacteria with all U33-containing ANN-decoding tRNAs, U33 is one of the key determinants; in accordance with the fact that there is no Cgi121 counterpart in *Ec*TsaBCDE, the CCA end has little effect on $t^6A$ biogenesis; in yeast with all U33-containing ANN-decoding tRNAs, *Sc*KEOPS still depends on U33 as a determinant, while the CCA end contributes to $t^6A$ formation by binding Cgi121, as evidenced by the lower modification of the CCA end-truncated tRNA mutant; in higher eukaryotes with both U33- and C33-containing tRNAs, the requirement of position 33 is negated, but CCA end functions as a determinant.

### $t^6A$ is a critical positive determinant in aminoacylation of $tRNA^{Ile}(AAU)$ and $tRNA^{Ile}(GAU)$

We determined that all cytosolic ANN-decoding tRNAs in human cells harbor a $t^6A$ moiety. Indeed, all these tRNAs contain all necessary determining elements (G10–C25, C11-G24, C32, U36, A37, A38, the CCA end), as revealed in this study and others, except G10-U25 in $tRNA^{Met}(i)$. Our data also clearly demonstrated that $t^6A$ modification contributes little to human tRNA abundance. However, the effect on aminoacylation levels varied and was tRNA isoacceptor specific. The charging levels of most ANN-decoding tRNAs were not altered upon knockdown of *OSGEP* and upon a decrease in $t^6A$ abundance; however, $tRNA^{Ile}(AAU)$ and $tRNA^{Ile}(GAU)$, but not $tRNA^{Ile}(UAU)$ displayed obviously reduced tRNA aminoacylation levels *in vivo* due to *OSGEP* knockdown. Furthermore, the *in vitro* data clearly indicated that $t^6A$ is a critical determinant in the aminoacylation of $tRNA^{Ile}(AAU)$ and $tRNA^{Ile}(GAU)$ by hIleRS. Interestingly, the $tRNA^{Ile}(UAU)$ transcript is strongly charged in the absence of $t^6A$ modification. $t^6A$ in bacterial $tRNA^{Ile}$ has proven to be a strong positive determinant for aminoacylation by bacterial IleRS (24). However, in a bacterial IleRS-$tRNA^{Ile}$ cocrystal structure (PDB No. 1QU3) (52), A37 was shown to be distant from the anticodon-binding domain of IleRS and did not interact directly with the enzyme. Thus, the rationale of the key role of $t^6A$ modification in aminoacylation is unclear. Yeast IleRS seems to be insensitive to $t^6A$ modification when total yeast tRNAs are used as substrates (24). Consistently, yeast cells devoid of $t^6A$ modification (in which the individual genes encoding components of KEOPS were removed), are still viable despite a reduced growth rate of the cells (12). The pivotal role of $t^6A$ modification of $tRNA^{Ile}(AAU)$ and $tRNA^{Ile}(GAU)$ in aminoacylation likely explains why the OSGEP gene cannot be deleted via CRISPR/Cas9 gene editing due to uncharged $tRNA^{Ile}$ and abolished mRNA translation after deletion. However, the molecular basis for the $tRNA^{Ile}$ isoacceptor-specific requirement of $t^6A$ biogenesis in tRNA charging by hIleRS needs to be further explored.

### $t^6A$ modification finetunes translational elongation in human cells

Using a luciferase-based reporter system, we also explored the potential role of $t^6A$ modification in translation elongation. Interestingly, we found that, for most ANN codes, decreased $t^6A$ modification seems to stimulate translational elongation speed, as evidenced by the increased firefly luciferase signal. In contrast, for the Ile ATA codon, reduced amounts of firefly luciferase were observed, suggesting that the translational elongation speed at the ATA codon was downregulated due to impaired $t^6A$ modification. Overall, impaired $t^6A$ modification levels due to *OSGEP* knockdown led to various effects on translational efficiency at ANN codons. Various effects on the decoding of different codons by $t^6A$ were also observed in yeast in which the elongation rate at codons decoded by highly abundant tRNAs and I34:C3 pairs was suppressed, while that by rare tRNAs and G34:U3 pairs increased in response to $t^6A$ modification (27).

## DATA AVAILABILITY

All data presented in this study are available within the figures and in the Supplementary data.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We are grateful to Drs Xinxin Chen and Lei Wang (Institute of Biophysics, CAS) for technical assistance. We thank the core facility of molecular biology of our institute for technical support in UPLC-MS/MS analysis. We also thank Drs Herman van Tilbeurgh (Institute for Integrative Biology of the Cell, CNRS) and Jin-Qiu Zhou in our institute for providing the plasmid expressing *Sc*KEOPS.

## FUNDING

National Key Research and Development Program of China [2017YFA0504000, 2021YFA1300800, 2021YFC2700903]; Natural Science Foundation of China [91940302, 31870811, 31670801, 31822015, 81870896, 32000889, 32000847]; Strategic Priority Research Program of the Chinese Academy of Sciences [XDB19010203]; Shanghai Key Laboratory of Embryo Original Diseases [Shelab201904]; Key Laboratory of Reproductive Genetics, Ministry of Education, Zhejiang University [ZDFY2020-RG-0003]. Funding for open access charge: Natural Science Foundation of China.
*Conflict of interest statement.* None declared.

## REFERENCES

1. Wolff,P., Villette,C., Zumsteg,J., Heintz,D., Antoine,L., Chane-Woon-Ming,B., Droogmans,L., Grosjean,H. and Westhof,E. (2020) Comparative patterns of modified nucleotides in individual tRNA species from a mesophilic and two thermophilic archaea. *RNA*, **26**, 1957–1975.
2. Suzuki,T. (2021) The expanding world of tRNA modifications and their disease relevance. *Nat. Rev. Mol. Cell Biol.*, **22**, 375–392.
3. Zhou,J.B., Wang,E.D. and Zhou,X.L. (2021) Modifications of the human tRNA anticodon loop and their associations with genetic diseases. *Cell. Mol. Life Sci.*, **78**, 7087–7105.
4. de Crecy-Lagard,V., Boccaletto,P., Mangleburg,C.G., Sharma,P., Lowe,T.M., Leidel,S.A. and Bujnicki,J.M. (2019) Matching tRNA modifications in humans to their known and predicted enzymes. *Nucleic Acids Res.*, **47**, 2143–2159.
5. Beenstock,J. and Sicheri,F. (2021) The structural and functional workings of KEOPS. *Nucleic Acids Res.*, **49**, 10818–10834.
6. Thiaville,P.C., Iwata-Reuyl,D. and de Crecy-Lagard,V. (2014) Diversity of the biosynthesis pathway for threonylcarbamoyladenosine (t(6)A), a universal modification of tRNA. *RNA Biol.*, **11**, 1529–1539.
7. El Yacoubi,B., Lyons,B., Cruz,Y., Reddy,R., Nordin,B., Agnelli,F., Williamson,J.R., Schimmel,P., Swairjo,M.A. and de Crecy-Lagard,V. (2009) The universal YrdC/Sua5 family is required for the formation of threonylcarbamoyladenosine in tRNA. *Nucleic Acids Res.*, **37**, 2894–2909.
8. Missoury,S., Plancqueel,S., Li de la Sierra-Gallay,I., Zhang,W., Liger,D., Durand,D., Dammak,R., Collinet,B. and van Tilbeurgh,H. (2018) The structure of the TsaB/TsaD/TsaE complex reveals an unexpected mechanism for the bacterial $t^6A$ tRNA-modification. *Nucleic Acids Res.*, **46**, 5850–5860.
9. Luthra,A., Swinehart,W., Bayooz,S., Phan,P., Stec,B., Iwata-Reuyl,D. and Swairjo,M.A. (2018) Structure and mechanism of a bacterial $t^6A$ biosynthesis system. *Nucleic Acids Res.*, **46**, 1395–1411.
10. Deutsch,C., El Yacoubi,B., de Crecy-Lagard,V. and Iwata-Reuyl,D. (2012) Biosynthesis of threonylcarbamoyl adenosine ($t^6A$), a universal tRNA nucleoside. *J. Biol. Chem.*, **287**, 13666–13673.
11. El Yacoubi,B., Hatin,I., Deutsch,C., Kahveci,T., Rousset,J.P., Iwata-Reuyl,D., Murzin,A.G. and de Crecy-Lagard,V. (2011) A role for the universal Kae1/Qri7/YgjD (COG0533) family in tRNA modification. *EMBO J.*, **30**, 882–893.
12. Srinivasan,M., Mehta,P., Yu,Y., Prugar,E., Koonin,E.V., Karzai,A.W. and Sternglanz,R. (2011) The highly conserved KEOPS/EKC complex is essential for a universal tRNA modification, $t^6A$. *EMBO J.*, **30**, 873–881.
13. Luthra,A., Paranagama,N., Swinehart,W., Bayooz,S., Phan,P., Quach,V., Schiffer,J.M., Stec,B., Iwata-Reuyl,D. and Swairjo,M.A. (2019) Conformational communication mediates the reset step in $t^6A$ biosynthesis. *Nucleic Acids Res.*, **47**, 6551–6567.
14. Mao,D.Y., Neculai,D., Downey,M., Orlicky,S., Haffani,Y.Z., Ceccarelli,D.F., Ho,J.S., Szilard,R.K., Zhang,W., Ho,C.S. *et al.* (2008) Atomic structure of the KEOPS complex: an ancient protein kinase-containing molecular machine. *Mol. Cell*, **32**, 259–275.
15. Zhou,J.B., Wang,Y., Zeng,Q.Y., Meng,S.X., Wang,E.D. and Zhou,X.L. (2020) Molecular basis for $t^6A$ modification in human mitochondria. *Nucleic Acids Res.*, **48**, 3181–3194.
16. Wan,L.C., Mao,D.Y., Neculai,D., Strecker,J., Chiovitti,D., Kurinov,I., Poda,G., Thevakumaran,N., Yuan,F., Szilard,R.K. *et al.* (2013) Reconstitution and characterization of eukaryotic $N^6$-threonylcarbamoylation of tRNA using a minimal enzyme system. *Nucleic Acids Res.*, **41**, 6332–6346.
17. Lin,H., Miyauchi,K., Harada,T., Okita,R., Takeshita,E., Komaki,H., Fujioka,K., Yagasaki,H., Goto,Y.I., Yanaka,K. *et al.* (2018) $CO_2$-sensitive tRNA modification associated with human mitochondrial disease. *Nat. Commun.*, **9**, 1875.
18. Beenstock,J., Ona,S.M., Porat,J., Orlicky,S., Wan,L.C.K., Ceccarelli,D.F., Maisonneuve,P., Szilard,R.K., Yin,Z., Setiaputra,D. *et al.* (2020) A substrate binding model for the KEOPS tRNA modifying complex. *Nat. Commun.*, **11**, 6233.
19. Perrochia,L., Guetta,D., Hecker,A., Forterre,P. and Basta,T. (2013) Functional assignment of KEOPS/EKC complex subunits in the biosynthesis of the universal $t^6A$ tRNA modification. *Nucleic Acids Res.*, **41**, 9484–9499.
20. Kopina,B.J., Missoury,S., Collinet,B., Fulton,M.G., Cirio,C., van Tilbeurgh,H. and Lauhon,C.T. (2021) Structure of a reaction intermediate mimic in $t^6A$ biosynthesis bound in the active site of the TsaBD heterodimer from *Escherichia coli*. *Nucleic Acids Res.*, **49**, 2141–2160.
21. Morin,A., Auxilien,S., Senger,B., Tewari,R. and Grosjean,H. (1998) Structural requirements for enzymatic formation of

threonylcarbamoyladenosine ($t^6A$) in tRNA: an *in vivo* study with *Xenopus laevis* oocytes. *RNA*, **4**, 24–37.
22. Wang,Y., Zeng,Q.Y., Zheng,W.Q., Ji,Q.Q., Zhou,X.L. and Wang,E.D. (2018) A natural non-Watson-Crick base pair in human mitochondrial $tRNA^{Thr}$ causes structural and functional susceptibility to local mutations. *Nucleic Acids Res.*, **46**, 4662–4676.
23. Murphy,F.V.t., Ramakrishnan,V., Malkiewicz,A. and Agris,P.F. (2004) The role of modifications in codon discrimination by tRNA(Lys)UUU. *Nat. Struct. Mol. Biol.*, **11**, 1186–1191.
24. Thiaville,P.C., El Yacoubi,B., Kohrer,C., Thiaville,J.J., Deutsch,C., Iwata-Reuyl,D., Bacusmo,J.M., Armengaud,J., Bessho,Y., Wetzel,C. *et al.* (2015) Essentiality of threonylcarbamoyladenosine (t(6)A), a universal tRNA modification, in bacteria. *Mol. Microbiol.*, **98**, 1199–1221.
25. Braun,D.A., Rao,J., Mollet,G., Schapiro,D., Daugeron,M.C., Tan,W., Gribouval,O., Boyer,O., Revy,P., Jobst-Schwan,T. *et al.* (2017) Mutations in KEOPS-complex genes cause nephrotic syndrome with primary microcephaly. *Nat. Genet.*, **49**, 1529–1538.
26. Lin,C.A., Ellis,S.R. and True,H.L. (2010) The Sua5 protein is essential for normal translational regulation in yeast. *Mol. Cell Biol.*, **30**, 354–363.
27. Thiaville,P.C., Legendre,R., Rojas-Benitez,D., Baudin-Baillieu,A., Hatin,I., Chalancon,G., Glavic,A., Namy,O. and de Crecy-Lagard,V. (2016) Global translational impacts of the loss of the tRNA modification t(6)A in yeast. *Microb. Cell*, **3**, 29–45.
28. Arrondel,C., Missoury,S., Snoek,R., Patat,J., Menara,G., Collinet,B., Liger,D., Durand,D., Gribouval,O., Boyer,O. *et al.* (2019) Defects in t(6)A tRNA modification due to GON7 and YRDC mutations lead to Galloway-Mowat syndrome. *Nat. Commun.*, **10**, 3967.
29. Edvardson,S., Prunetti,L., Arraf,A., Haas,D., Bacusmo,J.M., Hu,J.F., Ta-Shma,A., Dedon,P.C., de Crecy-Lagard,V. and Elpeleg,O. (2017) tRNA $N^6$-adenosine threonylcarbamoyltransferase defect due to KAE1/TCS3 (OSGEP) mutation manifest by neurodegeneration and renal tubulopathy. *Eur. J. Hum. Genet.*, **25**, 545–551.
30. Collinet,B., Friberg,A., Brooks,M.A., van den Elzen,T., Henriot,V., Dziembowski,A., Graille,M., Durand,D., Leulliot,N., Saint Andre,C. *et al.* (2011) Strategies for the structural analysis of multi-protein complexes: lessons from the 3D-Repertoire project. *J. Struct. Biol.*, **175**, 147–158.
31. Zhang,W., Collinet,B., Perrochia,L., Durand,D. and van Tilbeurgh,H. (2015) The ATP-mediated formation of the YgjD-YeaZ-YjeE complex is required for the biosynthesis of tRNA $t^6A$ in *Escherichia coli*. *Nucleic Acids Res.*, **43**, 1804–1817.
32. Zhou,X.L., Zhu,B. and Wang,E.D. (2008) The CP2 domain of leucyl-tRNA synthetase is crucial for amino acid activation and post-transfer editing. *J. Biol. Chem.*, **283**, 36608–36616.
33. Mao,X.L., Li,Z.H., Huang,M.H., Wang,J.T., Zhou,J.B., Li,Q.R., Xu,H., Wang,X.J. and Zhou,X.L. (2021) Mutually exclusive substrate selection strategy by human $m^3C$ RNA transferases METTL2A and METTL6. *Nucleic Acids Res.*, **49**, 8309–8323.
34. Zeng,Q.Y., Peng,G.X., Li,G., Zhou,J.B., Zheng,W.Q., Xue,M.Q., Wang,E.D. and Zhou,X.L. (2019) The G3-U70-independent tRNA recognition by human mitochondrial alanyl-tRNA synthetase. *Nucleic Acids Res.*, **47**, 3072–3085.
35. Zhou,X.L., Chen,Y., Zeng,Q.Y., Ruan,Z.R., Fang,P. and Wang,E.D. (2019) Newly acquired N-terminal extension targets threonyl-tRNA synthetase-like protein into the multiple tRNA synthetase complex. *Nucleic Acids Res.*, **47**, 8662–8674.
36. Gillum,A.M., Roe,B.A., Anandaraj,M.P. and RajBhandary,U.L. (1975) Nucleotide sequence of human placenta cytoplasmic initiator tRNA. *Cell*, **6**, 407–413.
37. Vitali,P. and Kiss,T. (2019) Cooperative 2'-O-methylation of the wobble cytidine of human elongator tRNA(Met)(CAT) by a nucleolar and a Cajal body-specific box C/D RNP. *Genes Dev.*, **33**, 741–746.
38. Watanabe,K., Miyagawa,R., Tomikawa,C., Mizuno,R., Takahashi,A., Hori,H. and Ijiri,K. (2013) Degradation of initiator $tRNA^{Met}$ by Xrn1/2 via its accumulation in the nucleus of heat-treated HeLa cells. *Nucleic Acids Res.*, **41**, 4671–4685.
39. Chan,P.P. and Lowe,T.M. (2016) GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.*, **44**, D184–D189.
40. Cui,J., Liu,Q., Sendinc,E., Shi,Y. and Gregory,R.I. (2021) Nucleotide resolution profiling of $m^3C$ RNA modification by HAC-seq. *Nucleic Acids Res.*, **49**, e27.
41. Torres,A.G., Rodríguez-Escribà,M., Marcet-Houben,M., Santos Vieira,H.G., Camacho,N., Catena,H., Murillo Recio,M., Rafels-Ybern,À., Reina,O., Torres,F.M. *et al.* (2021) Human tRNAs with inosine 34 are essential to efficiently translate eukarya-specific low-complexity proteins. *Nucleic Acids Res.*, **49**, 7011–7034.
42. Wei,F.Y., Suzuki,T., Watanabe,S., Kimura,S., Kaitsuka,T., Fujimura,A., Matsui,H., Atta,M., Michiue,H., Fontecave,M. *et al.* (2011) Deficit of tRNA(Lys) modification by Cdkal1 causes the development of type 2 diabetes in mice. *J. Clin. Invest.*, **121**, 3598–3608.
43. Suzuki,T., Nagao,A. and Suzuki,T. (2011) Human mitochondrial tRNAs: biogenesis, function, structural aspects, and diseases. *Annu. Rev. Genet.*, **45**, 299–329.
44. Yoshihisa,T., Yunoki-Esaki,K., Ohshima,C., Tanaka,N. and Endo,T. (2003) Possibility of cytoplasmic pre-tRNA splicing: the yeast tRNA splicing endonuclease mainly localizes on the mitochondria. *Mol. Biol. Cell*, **14**, 3266–3279.
45. Shiba,K., Suzuki,N., Shigesada,K., Namba,Y., Schimmel,P. and Noda,T. (1994) Human cytoplasmic isoleucyl-tRNA synthetase: selective divergence of the anticodon-binding domain and acquisition of a new structural unit. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 7435–7439.
46. Rho,S.B., Kim,M.J., Lee,J.S., Seol,W., Motegi,H., Kim,S. and Shiba,K. (1999) Genetic dissection of protein-protein interactions in multi-tRNA synthetase complex. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 4488–4493.
47. Yu,N., Jora,M., Solivio,B., Thakur,P., Acevedo-Rocha,C.G., Randau,L., de Crecy-Lagard,V., Addepalli,B. and Limbach,P.A. (2019) tRNA modification profiles and codon-decoding strategies in *Methanocaldococcus jannaschii*. *J. Bacteriol.*, **201**, e00690-18.
48. Reddy,D.M., Crain,P.F., Edmonds,C.G., Gupta,R., Hashizume,T., Stetter,K.O., Widdel,F. and McCloskey,J.A. (1992) Structure determination of two new amino acid-containing derivatives of adenosine from tRNA of thermophilic bacteria and archaea. *Nucleic Acids Res.*, **20**, 5607–5615.
49. Juhling,F., Morl,M., Hartmann,R.K., Sprinzl,M., Stadler,P.F. and Putz,J. (2009) tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res.*, **37**, D159–D162.
50. Perrochia,L., Crozat,E., Hecker,A., Zhang,W., Bareille,J., Collinet,B., van Tilbeurgh,H., Forterre,P. and Basta,T. (2013) *In vitro* biosynthesis of a universal $t^6A$ tRNA modification in Archaea and Eukarya. *Nucleic Acids Res.*, **41**, 1953–1964.
51. Swinehart,W., Deutsch,C., Sarachan,K.L., Luthra,A., Bacusmo,J.M., de Crecy-Lagard,V., Swairjo,M.A., Agris,P.F. and Iwata-Reuyl,D. (2020) Specificity in the biosynthesis of the universal tRNA nucleoside N (6)-threonylcarbamoyl adenosine (t(6)A)-TsaD is the gatekeeper. *RNA*, **26**, 1094–1103.
52. Silvian,L.F., Wang,J. and Steitz,T.A. (1999) Insights into editing from an ile-tRNA synthetase structure with $tRNA^{ile}$ and mupirocin. *Science*, **285**, 1074–1077.