Published online 28 November 2016

Nucleic Acids Research, 2017, Vol. 45, No. 2 ***805–817***
doi: 10.1093/nar/gkw1181

Proteomic analysis of the human KEOPS complex identifies C14ORF142 as a core subunit homologous to yeast Gon7

Leo C.K. Wan[1,2], Pierre Maisonneuve[1], Rachel K. Szilard[1], Jean-Philippe Lambert[1], Timothy F. Ng[1,2], Noah Manczyk[1,3], Hao Huang[1,4], Rob Laister[5], Amy A. Caudy[2,6], Anne-Claude Gingras[1,2], Daniel Durocher[1,2,*] and Frank Sicheri[1,2,3,*]

[1]Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON M5G 1X5, Canada, [2]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 3E1, Canada, [3]Department of Biochemistry, University of Toronto, Toronto, ON M5S 3E1,Canada, [4]School of Chemical Biology and Biotechnology, Shenzhen Graduate School of Peking University, Shenzen, 518055, China, [5]Department of Medical Oncology and Hematology, Princess Margaret Cancer Centre, Toronto, ON M5G 2M9, Canada and [6]Terrence Donnelly Centre for Cellular & Biomolecular Research, University of Toronto, ON, M5S 3E1, Canada

Received August 19, 2016; Revised November 01, 2016; Editorial Decision November 14, 2016; Accepted November 15, 2016

ABSTRACT

The KEOPS/EKC complex is a tRNA modification complex involved in the biosynthesis of N^6-threonylcarbamoyladenosine (t^{6}A), a universally conserved tRNA modification found on ANN-codon recognizing tRNAs. In archaea and eukaryotes, KEOPS is composed of OSGEP/Kae1, PRPK/Bud32, TPRKB/Cgi121 and LAGE3/Pcc1. In fungi, KEOPS contains an additional subunit, Gon7, whose orthologs outside of fungi, if existent, remain unidentified. In addition to displaying defective t^{6}A biosynthesis, *Saccharomyces cerevisiae* strains harboring KEOPS mutations are compromised for telomere homeostasis, growth and transcriptional co-activation. To identify a Gon7 ortholog in multicellular eukaryotes as well as to uncover KEOPS-interacting proteins that may link t^{6}A biosynthesis to the diverse set of KEOPS mutant phenotypes, we conducted a proteomic analysis of human KEOPS. This work identified 152 protein interactors, one of which, C14ORF142, interacted strongly with all four KEOPS subunits, suggesting that it may be a core component of human KEOPS. Further characterization of C14ORF142 revealed that it shared a number of biophysical and biochemical features with fungal Gon7, suggesting that C14ORF142 is the human ortholog of Gon7. In addition, our proteomic analysis identified specific interactors for different KEOPS subcomplexes, hinting that individual KEOPS subunits may have additional functions outside of t^{6}A biosynthesis.

INTRODUCTION

Living organisms rely on a core set of universally conserved genes to sustain life and cellular function, with a majority of essential genes partaking in transcriptional and translational processes. The post-transcriptional modification of tRNAs is one such process, whereby specific tRNA nucleotides, in particular those at positions 34 and 37, are modified to modulate tRNA structure and function (1). An essential modification of tRNAs at nucleotide position 37, known as N^6-threonylcarbamoyladenosine (t^{6}A) or its derivative cyclic-t^{6}A, is found on all ANN-codon-recognizing tRNAs in the three domains of life (2–5). The biosynthesis of t^{6}A is catalyzed by two universally conserved protein families: Sua5/YrdC (alternatively known as Tcs2/Tcs1) and Kae1/Qri7/YgjD (alternatively known as Tcs3/Tcs4/TsaD) (6–8). Using threonine, bicarbonate and adenosine triphosphate (ATP) as substrates, Sua5/YrdC first catalyzes the formation of a threonylcarbamoyladenylate intermediate (9–11), which is subsequently used by Kae1/Qri7/YgjD to catalyze the transfer of a threonylcarbamoyl moiety onto substrate tRNAs. While Sua5/YrdC family members function independently as monomers, Kae1/Qri7/YgjD members functions as part of related but distinct protein complexes in the different domains of life (12–14). In the mitochondria of eukaryotes, Qri7 operates as an isolated homodimer (10,15). In bacteria, YgjD operates in a ternary complex with the inactive YgjD structural ortholog YeaZ (alternatively known as TsaB) and

*To whom correspondence should be addressed. Tel: +416 586 8471; Fax: +416 586 8869; Email: sicheri@lunenfeld.ca
Correspondence may also be addressed to D. Durocher. Tel: +416 586 4800 (Ext. 2544); Fax: +416 586 8869; Email: durocher@lunenfeld.ca

the ATPase YjeE (alternatively known as TsaE) (16–19). In archaea and eukaryotes, Kae1 operates as part of the KEOPS/EKC complex (referred to here simply as KEOPS) with the ATPase Bud32 (alternatively known as Tcs5), the ATPase regulator Cgi121 (alternatively known as Tcs7) and the dimerization module Pcc1 (alternatively known as Tcs6). In budding yeast, KEOPS contains a fifth subunit, Gon7 (alternatively known as Tcs8). Whether this subunit is unique to yeast or present but yet-to-be discovered in other eukaryotes and archaea remains to be determined.

KEOPS was originally discovered in two independent genetic screens conducted in the budding yeast *Saccharomyces cerevisiae*. In the first study, KEOPS subunit Cgi121 was discovered as a genetic suppressor of the *cdc13-1* allele that causes a telomere-capping defect (12). In support for Cgi121 functioning as a novel telomere regulator, deletion of *CGI121* in a *cdc13-1* strain reversed the accumulation of single-stranded DNA at telomeres, which is a hallmark of telomere dysfunction imparted by the *cdc13-1* allele. Tandem affinity purification and mass spectrometry analysis of Cgi121 revealed that it formed a protein complex with Kae1, Bud32 and Gon7 proteins. Subsequent deletion of each KEOPS subunit resulted in extreme slow growth and shortened telomeres phenotypes in yeast. In the second study, KEOPS subunit Pcc1 was discovered as a genetic suppressor of a U1snRNP splicing defect allele that causes a cold-sensitivity phenotype (13). Subsequent analyses demonstrated that defects in U1snRNP function abolished proper splicing of the Pcc1 transcript and that depletion of Pcc1 was causative for the cold-sensitive phenotype. Similar to the previous study, tandem affinity purification and mass spectrometry analysis of Pcc1 revealed that it formed a protein complex with Kae1, Bud32, Cgi121 and Gon7. The authors additionally showed that Pcc1 localized to the chromatin of transcriptionally active genes, suggesting that Pcc1 played a role in transcriptional activation. In agreement with this hypothesis, yeast strains harboring a temperature sensitive *pcc1-4* allele demonstrated defects in recruiting the transcriptional co-activators Mediator and SAGA to transcriptionally active chromatin. How the slow growth, shortened telomere and transcriptional defect phenotypes of KEOPS in *S. cerevisiae* relate to the underlying biochemical function of KEOPS in t^6A biosynthesis remains a mystery.

Although sequence analysis suggests that Gon7 is absent from archaea and multicellular eukaryotes, it is nonetheless essential for life and necessary for *in vivo* t^6A biosynthesis in yeast (10). Structurally, Gon7 in isolation is an intrinsically disordered protein and only partially folds into a discrete conformation when it heterodimerizes with Pcc1 (20). One possible explanation for the lack of readily identifiable Gon7 orthologs is that the absence of an autonomous fold in Gon7 has allowed the primary protein sequence to diverge to such a degree that distant orthologs are no longer detectable using first-pass bioinformatic approaches. In this study, we conducted an affinity purification-mass spectrometry (AP-MS) analysis on the four recognizable KEOPS subunits in HEK293 cells to explore whether a human Gon7 ortholog can be found, and to identify additional novel interacting proteins that might provide a link between the role of KEOPS in t^6A biosynthesis and the diverse set of phenotypes associated with its depletion. In total, we identified 152 high-confidence KEOPS interactors, which in concordance with its recognized role in modifying tRNAs, were enriched in proteins implicated in protein translation. We identified a single small protein, C14ORF142, which prominently interacted with all four KEOPS subunits, suggesting that C14ORF142 may be an integral core component of the human KEOPS complex. We conducted follow on characterizations of C14ORF142 and showed that C14ORF142, like Gon7, is intrinsically disordered in isolation (20), binds directly to LAGE3 (human ortholog of Pcc1) and to the LAGE3-OSGEP subcomplex (human ortholog of Pcc1-Kae1) within KEOPS (20), and forms a 1:1:1:1:1 complex with OSGEP, PRPK, TPRKB and LAGE3 proteins (human orthologs of Kae1, Bud32, Cgi121 and Pcc1, respectively) (20). Lastly, we showed that C14ORF142 potentiates the t^6A catalytic activity of human KEOPS *in vitro*. Together, these findings identify C14ORF142 as the human ortholog of fungal Gon7.

MATERIALS AND METHODS

Generation of Flp-In T-REx HEK293 stable cell lines

The cDNA for human OSGEP, PRPK, TPRKB, LAGE3 and C14ORF142 were cloned into pcDNA5-FRT-FLAG (21) and co-transfected into Flp-In T-REx-293 cell lines with pOG44, a commercially available Flp-Recombinase expression vector (ThermoFisher Scientific) in a 60 mm dish. On the following day, all cells were passaged onto a 150 mm dish, and stably integrated cells were selected with hygromycin and blasticidin. Colony formation occurred over a span of 14 days, at which point cells were further passaged for experimentation. Tetracycline treatment for 24 h was used to induce expression of each FLAG-tagged KEOPS subunit.

Affinity purification of protein complexes and peptide digestion

Affinity purification was conducted as previously described (22). Briefly, two near-confluent 150 mm dishes were used for each affinity purification experiment. Cells were collected by scraping in ice-cold phosphate buffered saline (PBS) and pelleted by centrifugation. Ice-cold lysis buffer (50 mM HEPES-NaOH pH 8.0, 100 mM KCl, 2 mM EDTA, 0.1% NP40, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM dithiothreitol (DTT) and 1x Sigma protease inhibitor cocktail P8340) was added to the cell pellet, followed by a single freeze-thaw cycle to assist cell lysis. The resultant lysate was cleared by centrifugation (20000 $\times g$, 20 min, 4°C) and the supernatant was collected. Anti-FLAG M2 magnetic beads (Sigma-Aldrich) equilibrated with ice-cold lysis buffer were added to the supernatant and the mixture was incubated at 4°C for 2 h with gentle agitation. A total of 15 μl of bead bed volume was used for each sample. After incubation, beads were pelleted by centrifugation (100 $\times g$ 1 min, 4°C) and the supernatant was removed. Beads were washed with three cycles of lysis buffer and two cycles of 20 mM Tris-HCl pH 8.0, 2 mM $CaCl_2$ by magnetization. After the last cycle of washing, beads were magnetized again and buffer was aspirated away

as completely as possible. To digest the immunopreciptated proteins, the dried magnetic beads were suspended in 5 μl of 20 mM Tris-HCl pH 8.0 containing 500 ng of sequencing grade trypsin (Sigma-Aldrich) and incubated at 37°C for 4 h. The supernatant was collected and the beads were incubated with another 5 μl of trypsin solution and incubated at 37°C overnight. Both tryptic digest supernatants were combined, and formic acid was added to a final concentration of 2%. This sample was directly used for mass spectrometry analysis.

Mass spectrometry and data analysis

AP-MS samples were analyzed by mass spectrometry in at least two biological replicates. Five microliters of each sample were directly loaded onto a 75 μm × 10 cm emitter packed with 5 μm Zorbax C_{18} (Agilent, Santa Clara, CA, USA). Spectra were acquired on an Linear Trap Quadropole (LTQ) (Thermo Fisher) mass spectrometer placed in line with an Agilent 1100 pump with split flow as described in (23). Mass spectrometry data generated were stored, searched and analyzed using the ProHits laboratory information management system platform (24). Within ProHits, the resulting RAW files were converted to an MGF format using ProteoWizard (v3.0.4468) and then searched using Mascot (v2.3.02) and Comet (v2012.02 rev.0). The spectra were searched with the RefSeq database (version 53, May 28, 2014) acquired from NCBI against a total of 34 374 human and adenovirus sequences supplemented with 'common contaminants' from the Max Planck Institute (http://141.61.102.106:8080/share.cgi?ssid=0f2gfuB) and the Global Proteome Machine (GPM; http://www.thegpm.org/crap/index.html). The database parameters were set to search for tryptic cleavages, allowing up to 2 missed cleavage sites per peptide with a mass tolerance of 3 Da for precursors with charges of 2+ to 4+ and a tolerance of ±0.6 Da for fragment ions. Variable modifications were selected for deamidated asparagine and glutamine and oxidized methionine. The results from each search engine were analyzed through the Trans-Proteomic Pipeline (TPP v4.6 OCCUPY rev 3) (25) via the iProphet pipeline (26). SAINTexpress (v3.3) (27) was used to calculate the probability value of each potential protein–protein interaction using default parameters. For SAINTexpress scoring, 10 negative controls were selected for modeling; to minimize the scoring of spurious contaminants as interactors, the 5 highest spectral count values (across the 10 selected controls) were selected for each prey protein (this strategy was described in detail in (28)). Probabilities calculated for each biological replicate were averaged to generate the final probability; a Bayesian false discovery rate of 1% or lower was required for proteins to be classified as significant interaction partners.

The raw mass spectrometry data, detailed peptide and protein identification and SAINTexpress result tables are deposited in the MassIVE repository housed at the Center for Computational Mass Spectrometry at UCSD (http://proteomics.ucsd.edu/ProteoSAFe/datasets.jsp). The data set has been assigned the MassIVE ID MSV000079909 and is available for FTP download at ftp://MSV000079909@massive.ucsd.edu. The data set was assigned the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) identifier PXD004581.

Antibody production

Rabbit antibodies to human PRPK were prepared from an *Escherichia coli* maltose binding protein (MBP)-PRPK fusion protein expressed in *E. coli* BL21(DE3) CodonPlus-RIL cells and purified on amylose resin (New England Biolabs). The crude antiserum was depleted of antibodies to MBP by passage over an immobilized MBP-LacZα column (encoded by plasmid pMAL-c2, New England Biolabs), and PRPK-specific antibodies in the flow-through were subsequently enriched by affinity purification on an immobilized MBP-PRPK column (29). Rabbit antibodies to human OSGEP were prepared from a glutathione S-transferase (GST)-OSGEP fusion expressed in *E. coli* BL21(DE3) CodonPlus-RIL cells and purified from inclusion bodies using B-PER bacterial protein extraction reagent (Pierce). Inclusion bodies were subjected to preparative SDS-PAGE (28) and the protein eluted from a gel slice was used as immunogen. OSGEP-specific antibodies were affinity purified on an immobilized MBP-OSGEP column.

Immunoprecipitation

Confluent HEK293 cells in a 100 mm dish were used for each immunoprecipitation experiment. Cells were rinsed with ice-cold PBS and lysed in 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA and 1% Triton X-100 by gentle agitation at 4°C for 15 min. The resultant lysate was cleared by centrifugation (20 000 ×*g*, 20 min, 4°C) and the supernatant was collected. To each supernatant, the appropriate antibody conjugated to DynaBeads Protein G resin (ThermoFisher Scientific) was added and the sample was incubated at 4°C for 3 h with gentle agitation. Afterward, the resin was washed three times with ice-cold 50 mM Tris-HCl pH 7.4, 150 mM NaCl and all liquid was removed from the resin after the final wash step. The resin was resuspended in 1x Laemmli buffer and used for SDS-PAGE and immunoblot analysis.

Immunofluorescence microscopy

hTERT-RPE1 cells used for immunofluorescence microscopy were obtained from American Type Culture Collection (ATCC) and routinely tested for mycoplasma infection (MycoAlert, Lonzo). Cells were cultured at 37°C/5% CO_2 in Dulbecco's modified Eagle's medium (Gibco) supplemented with 10% fetal bovine serum, non-essential amino acids and GlutaMAX (Gibco). For each experiment, cells were seeded on glass coverslips in 24-well dishes one day before fixation with 4% (w/v) formaldehyde in PBS for 15 min. Coverslips were rinsed once prior to permeabilization with 0.3% (v/v) Triton X-100 for 30 min. Fixed cells were incubated in blocking buffer containing PBS, 1% bovine serum albumin and 0.1% Triton X-100 for 1 h at room temperature. Staining was performed with the rabbit primary antibodies anti-OSGEP (1:100), anti-C14ORF142 (Aviva Systems Biology, 1:100) or anti-PRPK (1:250) and

mouse anti-β-tubulin (DM1a, Calbiochem, 1:100) antibody for 2 h at room temperature. Coverslips were then incubated with the appropriate AlexaFluor conjugated secondary antibodies (Molecular Probes, Invitrogen) and 0.25 μg/ml DAPI to counterstain the nucleus for 1 h. Stained cells were visualized on a Zeiss LSM780 confocal laser scanning microscope.

Yeast growth rate analysis

The *gon7Δ* strain was obtained as part of the EUROSCARF deletion strain collection. C14ORF142 and Gon7 were cloned into a modified pRS425-GAL vector (generously provided by Mike Tyers). The resulting vectors pRS425-GAL, pRS425-GAL-C14ORF142 and pRS425-GAL-Gon7 were transformed into parental W303 strain and *gon7Δ* using conventional techniques. Yeast growth was monitored for 40 h by measuring OD_{600} every 15 min. Briefly, an overnight culture was diluted to OD_{600} 0.05 with fresh SD-Leu medium supplemented with 2% galactose. Growth analyses were performed in clear, flat bottom 96-well microplates (Costar 3370) using a Synergy Neo Multi-Mode Reader for readout (BioTek).

Plasmid production for insect cell protein expression

All proteins and complexes were produced with the ACEMBL MultiBac^{Turbo} baculovirus/insect cell system (ATG:biosynthetics, Merzhausen, Germany). The open reading frame for human C14ORF142 as well as codon-optimized open reading frames (codon optimization carried out by ATG:biosynthesis GmbH, Germany and gene synthesis carried out by Gen9 Inc., Cambridge, MA, USA) for human OSGEP, PRPK, TPRKB and LAGE3 (fused at its N terminus to a 6xHIS affinity tag and a TEV cleavage site, denoted as H_6T) were amplified by PCR and subcloned into various MultiBac vectors. The plasmids encoding individual KEOPS components are as follows: pACEBac1-TPRKB (pDD4195), pIDK-OSGEP (pDD4198), pIDC-H_6T-LAGE3 (pDD4196), pIDS-PRPK (pDD4199) and pIDC-C14ORF142 (pDD4197). Complete codon-optimized gene sequences including restriction sites are detailed in Supplementary Table S3. Plasmid inserts were confirmed by sequencing. The PI-SceI–BstXI fragment of pIDC-H_6T-LAGE3 was subcloned into the BstXI site of pIDK-OSGEP to generate plasmid pIDK-OSGEP/H_6T-LAGE3 (pDD4200). Sequential Cre recombination of the acceptor vector pACEBac1–TPRKB with the donor vectors pIDS–PRPK and pIDK–OSGEP/H_6T-LAGE3 generated plasmid pACEBac–TPOL (TPRKB, PRPK, OSGEP, H_6T-LAGE3) (pDD4209) encoding the four proteins TPRKB, PRPK, OSGEP and H_6T-LAGE3. Further Cre recombination of pACEBac-TPOL with pIDC-C14ORF142 generated plasmid pACEBac-TPOLC (TPRKB, PRPK, OSGEP, H_6T-LAGE3, C14ORF142) (pDD4211), encoding all five proteins. Recombined plasmids were screened for single integration events by digestion with restriction endonucleases unique to each original individual KEOPS component plasmid, as well as diagnostic PCR to confirm the presence of desired genes. OSGEP and LAGE3 were also subcloned into pFASTBAC-Hta (ThermoFisher Scientific) to generate pFASTBAC–OSGEP and pFASTBAC–LAGE3, respectively, which were confirmed by plasmid sequencing.

Plasmids pACEBac-TPOL, pACEBac-TPOLC, pFASTBAC-OSGEP and pFASTBAC-LAGE3 were transformed by electroporation into *E. coli* DH10MultiBac^{Turbo} cells (ATG:biosynthetics) to generate recombinant bacmids according to the manufacturer's recommendations. Bacmid DNA was purified by phenol–chloroform extraction and screened by diagnostic PCR to confirm the presence of desired genes. Bacmid DNA was transfected into insect Sf9 cells using Effectine transfection reagent (Qiagen) according to the manufacturer's protocol to generate baculoviruses.

Plasmid production for bacterial protein expression

S. cerevisiae Gon7, as well as human C14ORF142 and TPRKB were subcloned into a modified pGEX-2T vector (harboring a tobacco etch virus (TEV) protease recognition sequence in place of the thrombin protease recognition site and the multiple cloning site of pProEX-Hta) to generate pGEX-2T-Gon7, pGEX-2T-C14ORF142 and pGEX-2T-TPRKB respectively, allowing for expression as N-terminal GST-fusion proteins in bacteria. PRPK and LAGE3 were subcloned into pETM30-2 to generate pETM30-2-PRPK and pETM30-2-LAGE3, allowing for expression as N-terminal 6xHIS-GST-fusion proteins in bacteria.

Recombinant protein expression and purification

pGEX-2T-Gon7, pGEX-2T-C14ORF142 and pETM30-2-LAGE3 were transformed into BL21-CodonPlus DE3-RIL bacteria (Agilent Technologies) for protein expression. Bacterial expression was performed in LB media or M9 minimal media supplemented with ^{15}N-NH_4Cl (Cambridge Isotopes) for biochemical experimentation or NMR spectroscopy purposes, respectively. Bacterial pellets were resuspended in 50 mM HEPES–NaOH pH 7.5, 500 mM NaCl, 5 mM EDTA and 1 mM DTT and lysed by homogenization. The proteins from clarified lysates were purified in a 2-step chromatography procedure including glutathione affinity chromatography (GE Healthcare) in 50 mM HEPES–NaOH pH 7.5, 500 mM NaCl, 5 mM EDTA and 1 mM DTT, followed by TEV protease treatment for GST-epitope removal and elution from glutathione resin. A final size exclusion chromatography step on a Superdex75 24 ml column (GE healthcare) was performed in 20 mM HEPES–NaOH pH 7.5, 100 mM NaCl and 1 mM DTT.

pETM30-2-LAGE3 and pGEX-2T-C14ORF142 plasmids were co-transformed into BL21-CodonPlus DE3-RIL bacteria (Agilent Technologies) and grown in LB media with concurrent kanamycin and ampicillin selection. Bacterial pellets were resuspended in 50 mM HEPES–NaOH pH 7.5, 500 mM NaCl, 5 mM EDTA and 1 mM DTT and lyzed by homogenization. LAGE3-C14ORF142 complexes from clarified lysate were purified in a 2-step chromatography procedure including glutathione affinity chromatography in 50 mM HEPES–NaOH pH 7.5, 500 mM NaCl, 5 mM EDTA, followed by TEV protease treatment for GST-epitope removal and elution from glutathione resin. A final size exclusion chromatography step on a Superdex 200

24 ml column (GE Healthcare) was performed in 20 mM HEPES–NaOH pH 7.5, 100 mM NaCl and 1 mM DTT.

pETM30-2-PRPK and pGEX-2T-TPRKB plasmids were co-transformed into BL21-CodonPlus DE3-RIL bacteria (Agilent Technologies) and grown in LB media with concurrent kanamycin and ampicillin selection. Bacterial pellets were resuspended in 50 mM HEPES–NaOH pH 7.5, 500 mM NaCl and 50 mM imidazole and lyzed by homogenization. PRPK–TPRKB complexes from clarified lysate were purified in a 3-step chromatography procedure including nickel affinity chromatography (GE Healthcare) in 50 mM HEPES–NaOH pH 7.5, 500 mM NaCl, 50 mM imidazole (with an elution gradient up to 500 mM imidazole), glutathione affinity chromatography in 50 mM HEPES–NaOH pH 7.5, 500 mM NaCl, 5 mM EDTA, followed by TEV protease treatment for GST-epitope removal and elution from glutathione resin. A final size exclusion chromatography step on a Superdex 200 24 ml column (GE Healthcare) was performed in 20 mM HEPES–NaOH pH 7.5, 100 mM NaCl and 1 mM DTT.

pFASTBAC-OSGEP and pFASTBAC-LAGE3 baculoviruses were used to co-infect adherent Sf9 insect cells in 150 mm dishes. Insect cells from 30 150 mm dishes were harvested by scraping in ice-cold PBS and pelleted by centrifugation. Cell pellets were resuspended in lysis buffer (50 mM HEPES–NaOH pH 7.5, 500 mM NaCl and 50 mM imidazole) and lysed by homogenization. OSGEP–LAGE3 complexes were purified from clarified lysate in a 2-step chromatography procedure including nickel affinity chromatography in 50 mM HEPES–NaOH pH 7.5, 500 mM NaCl, 50 mM imidazole (with an elution gradient up to 500 mM imidazole) and a size exclusion chromatography step on a Superdex 200 24 ml column (GE Healthcare) in 20 mM HEPES-NaOH pH 7.5, 100 mM NaCl and 1 mM DTT.

pACEBac-TPOL and pACEBac-TPOLC baculoviruses were used to infect suspension Sf9 insect cell cultures. Pelleted insect cells were resuspended in lysis buffer (50 mM HEPES–NaOH pH 7.5, 500 mM NaCl and 50 mM imidazole) and lysed by sonication. The 4- and 5-subunit KEOPS complexes were purified from clarified lysate in a 3-step chromatography procedure including nickel affinity chromatography in 50 mM HEPES–NaOH pH 7.5, 500 mM NaCl, 50 mM imidazole (with an elution gradient up to 500 mM imidazole), followed by TEV protease treatment and dialysis with lysis buffer. This was followed by a subtractive nickel affinity chromatography step and a size exclusion chromatography step on a Superdex 200 24 ml column (GE Healthcare) in 20 mM HEPES-NaOH pH 7.5, 100 mM NaCl and 1 mM DTT.

In vitro tRNA Transcription and Purification

Overlap extension PCR was used to generate human $tRNA^{LYS(UUU)}$ hepatitis delta virus ribozyme DNA fragments. The human $tRNA^{Lys(UUU)}$–hepatitis delta virus ribozyme fragment was ligated into pUC19 and linearized with BamHI prior to performing run-off *in vitro* transcription reactions using T7 RNA Polymerase. Transcription reaction conditions were as follows: 100 mM Tris-Cl pH 8.0, 4 mM ATP, 4 mM GTP, 4 mM CTP, 4 mM UTP, 10 mM DTT, 1 mM spermidine, 0.1% Triton X-100, 25 mM $MgCl_2$, 50 μg/ml linearized DNA template, 0.2 mg/ml T7 RNA polymerase, 10 U/ml thermostable inorganic phosphate (NEB) and 200 U/ml RiboLock™ RNase Inhibitor (Thermo Scientific). Transcription reactions were incubated for 4 h at 37°C. RNA was isolated using a conventional phenol–chloroform extraction protocol, followed by isopropanol precipitation and 80% ethanol washes. Air-dried RNA pellets were solubilized in 8M urea and refolded by rapid dilution into 9 volumes of 25 mM Bis-Tris-Cl pH 6.5, 0.2 mM EDTA. The refolded tRNA was purified on a Source 15Q column (GE Healthcare) using the following chromatography conditions: (i) 4 column volumes at 20% B and (ii) linear gradient over 25 column volumes from 20% to 30% B. Solutions A and B are diethylpyrocarbonate (DEPC)-treated water and 2 M NaCl prepared in DEPC-treated water, respectively. Fractions containing tRNA were pooled and treated with one volume of isopropanol to precipitate RNA. The precipitated RNA was pelleted using centrifugation, washed with 80% ethanol and then air-dried. The RNA pellet was resuspended in 10 mM Tris-Cl pH 8. Prior to use, the tRNA was refolded by (i) boiling at 95°C for 2 min, (ii) flash cooling on ice, (iii) warming to 50°C, (iv) addition of $MgCl_2$ to a final concentration of 2 mM and (v) slow cooling to room temperature.

Nuclear magnetic resonance (NMR) spectroscopy

NMR-heteronuclear single quantum coherence spectroscopy (HSQC) spectra were recorded at 20°C on a Bruker Avance III 600MHz NMR spectrometer equipped with a 1.7 mm TCI probe. All NMR samples contained 400 μM (^{15}N) C14ORF142 in 20 mM HEPES-NaOH pH 7.5, 100 mM NaCl, 1 mM DTT and 5% D_2O. Data processing was conducted using NMRviewJ and NMRpipe (30) and NMR spectra were analyzed using NMRView (31).

GST pulldown assay

Bacterial lysates containing GST-Cdk2 or GST-C14ORF142 were incubated with glutathione resin for 90 min at 4°C in 50 mM HEPES–NaOH pH 7.5, 500 mM NaCl, 5 mM EDTA, 1 mM DTT and 2 mM PMSF. GST-Cdk2- and GST-C14ORF142-bound glutathione resins were then washed with 30 volumes of 50 mM HEPES–NaOH pH 7.5, 500 mM NaCl, 5 mM EDTA, 1 mM DTT and 20 volumes of 25 mM HEPES–NaOH pH 7.5, 200 mM NaCl, 2 mM DTT. The homogeneity and protein quantity of GST-Cdk2- and GST-C14ORF142-bound glutathione resin was assessed by SDS-PAGE. Twenty micrograms of OSGEP-LAGE3 complex, PRPK-TPRKB complex or both complexes were then incubated with GST-Cdk2- or GST-C14ORF142-bound glutathione resin for 90 min at 4°C in 25 mM HEPES–NaOH pH 7.5, 150 mM NaCl, 2 mM DTT, 0.1% NP-40. Resin was then washed with 50 volumes of 25 mM HEPES–NaOH pH 7.5, 150 mM NaCl, 2 mM DTT, 0.1% NP-40 and analyzed by sodium dodecyl sulfate - polyacrylamide gel electrophoresis (SDS-PAGE) and immunoblot.

SEC-MALS

Size exclusion chromatography multi-angle light scattering (SEC-MALS) analysis of the 4- and 5-subunit KEOPS complexes was conducted using miniDawn TREOS and Optilab T-rEX detectors (Wyatt Technology) coupled to a 1260 Infinity high performance liquid chromatography (HPLC) system (Agilent Technologies). A WTC-030S5 SEC column (Wyatt Technology) was used for protein separation. Chromatography experiments were conducted in 20 mM HEPES–NaOH pH 7.5, 250 mM NaCl and 1 mM DTT. All data analysis was conducted using the ASTRA software (Wyatt Technology).

In vitro t^6A biosynthesis assay

In vitro t^6A biosynthesis assays were conducted using 5 μM *S.cerevisiae* Sua5, 10 μM 4- or 5-subunit human KEOPS and 50 μM *in vitro* transcribed human $tRNA^{Lys(UUU)}$ in 25 mM Tris-Cl pH 8, 150 mM NaCl, 5 mM DTT, 5 mM $MgCl_2$, 1 mM threonine, 1 mM $NaHCO_3$, 4 mM ATP, 5 mM spermidine and 0.5 μl thermostable inorganic pyrophosphatase (NEB) at 20°C for the indicated time. The tRNA was enzymatically digested and analyzed by HPLC as described previously (32). Briefly, tRNA samples were digested using Nuclease P1 (Sigma) and dephosphorylated using calf intestinal phosphatase (NEB). The resulting mononucleosides were analyzed on a Discovery C18 (15 cm × 4.6 mm, 5 μM) reverse-phase column (Supelco Analytical) on a Dionex Ultimate 3000 HPLC Unit (Thermo Scientific). A linear gradient of 98:2 to 87.5:12.5 of 250 mM ammonium acetate pH 6.5 and 40% acetonitrile, respectively, was used to separate mononucleotides.

Small molecule mass spectrometry analysis

Instrumentation consisted of an Agilent U-HPLC system, HTS PAL autosampler and a 6230 TOF mass spectrometer with a Jet Stream ionization source (Agilent Technologies). The reverse-phase chromatography method consisted of an Agilent RRHD PFP (pentafluorphenyl) column (2.1 × 50 mm, 1.8 μm particle size, Agilent, Santa Clara, CA, USA) and an aqueous/organic gradient run as described below. Mobile phase A consisted of water with 1% acetonitrile and 0.1% formic acid. Mobile phase B consisted of acetonitrile. For the analysis of t^6A, the flow rate equaled 0.4 ml/min and the gradient extended from 0% B at 0.0 min, to 0% B at 2.67 min, to 25% B at 8.67 min, to 100% B at 9.33 min, maintained at 100% B until 11 min, then a gradient to 100% A at 12 min followed by a hold until 14 min to re-equilibrate for the new run. The column temperature was maintained at 40°C, and the autosampler at 4°C. The injection volume was 10 μl in a 5 μl loop. The time of flight mass spectrometer was used in positive ionization mode with a data storage threshold of 50. Data were collected in high resolution mode at 3 Hz. Instrument run parameters were Jet Stream gas temperature 325°C with flow at 8 l/min, sheath gas temperature 350°C, sheath gas flow 11 l/min, sample nebulizer pressure 35 psig, ESI capillary voltage 3500 V, nozzle voltage 1000, fragmentor voltage 175 V, skimmer 65 V, octopole 1 RF 750 V. The scan range was 50–1100 m/z.

RESULTS

Proteomics analysis of the human KEOPS complex

To identify KEOPS interacting proteins, we conducted AP-MS analysis of all four human KEOPS subunits in HEK293 Flp-In T-REx cells stably expressing an individual FLAG-tagged construct of a human KEOPS subunit. We identified a total of 152 high confidence KEOPS interactors with a false discovery rate cutoff lower than 1% as assessed by significance analysis of interactome (SAINT) analysis (27). Of the 152 high confidence interactors, 142 interactors were novel in that they were not reported previously in any proteomic studies involving KEOPS subunits (Supplementary Figure S1). Confirming our expectations, each of the KEOPS subunits OSGEP (Kae1 ortholog), PRPK (Bud32 ortholog), TPRKB (Cgi121 ortholog) and LAGE3 (Pcc1 ortholog) were detected as strong interactors of one another in the four respective AP-MS experiments with an average spectral count of 526 (ranging from 89 to 1099) (Figure 1A and Supplementary Figure S1). The average spectral count for all detected protein–protein interactions was 88 (ranging from 4 to 1099).

In our AP-MS analyses, we observed that PRPK and TPRKB, which bind directly within the KEOPS complex, shared a number of significant interacting proteins that were weakly or not at all detected for LAGE3 and OSGEP. These 72 proteins were highly enriched with the 'poly(A) RNA binding' GO molecular function term (*P*-value 6.33×10^{-66}; 64 terms intersect), as well as a number of molecular function, biological process and cellular compartment GO terms relating to protein translation (Supplementary Table S1). This enrichment was specific to PRPK and TPRKB and was not observed for LAGE3 and OSGEP, suggesting that PRPK-TPRKB may form sub-complexes distinct from KEOPS that function in mRNA-related processes. We note, however, that the interactions between PRPK-TPRKB and ribosomal proteins could be mediated through DNA and/or RNA, as we did not conduct AP-MS experiments with nuclease-treated samples. In addition to poly(A) binding proteins, we also detected unique interactions between PRPK-TPRKB and the TRMT6-TRMT61a complex, a tRNA modification complex involved in the methylation of adenosine 58 on initiator $tRNA_i^{Met}$. This interaction may provide additional opportunities for PRPK-TPRKB to influence tRNA function.

We also detected a number of significant LAGE3 interactors that were weakly or not at all detected for PRPK, TPRKB and OSGEP. The 57 significant LAGE3 interacting proteins were enriched with 'protein folding' GO biological process term (*P*-value 2.19×10^{-13}; 15 terms intersect) (Supplementary Table S2), suggesting that LAGE3 might have additional roles outside of KEOPS in regulating chaperones and protein folding.

C14ORF142 as a potential functional ortholog of Gon7

In our AP-MS analysis of human KEOPS, we identified a single protein, C14ORF142, which interacted strongly with all four subunits. This observation was reminiscent of the tandem affinity purification experiments of *S. cerevisiae* KEOPS, where a fifth subunit, Gon7, interacted strongly

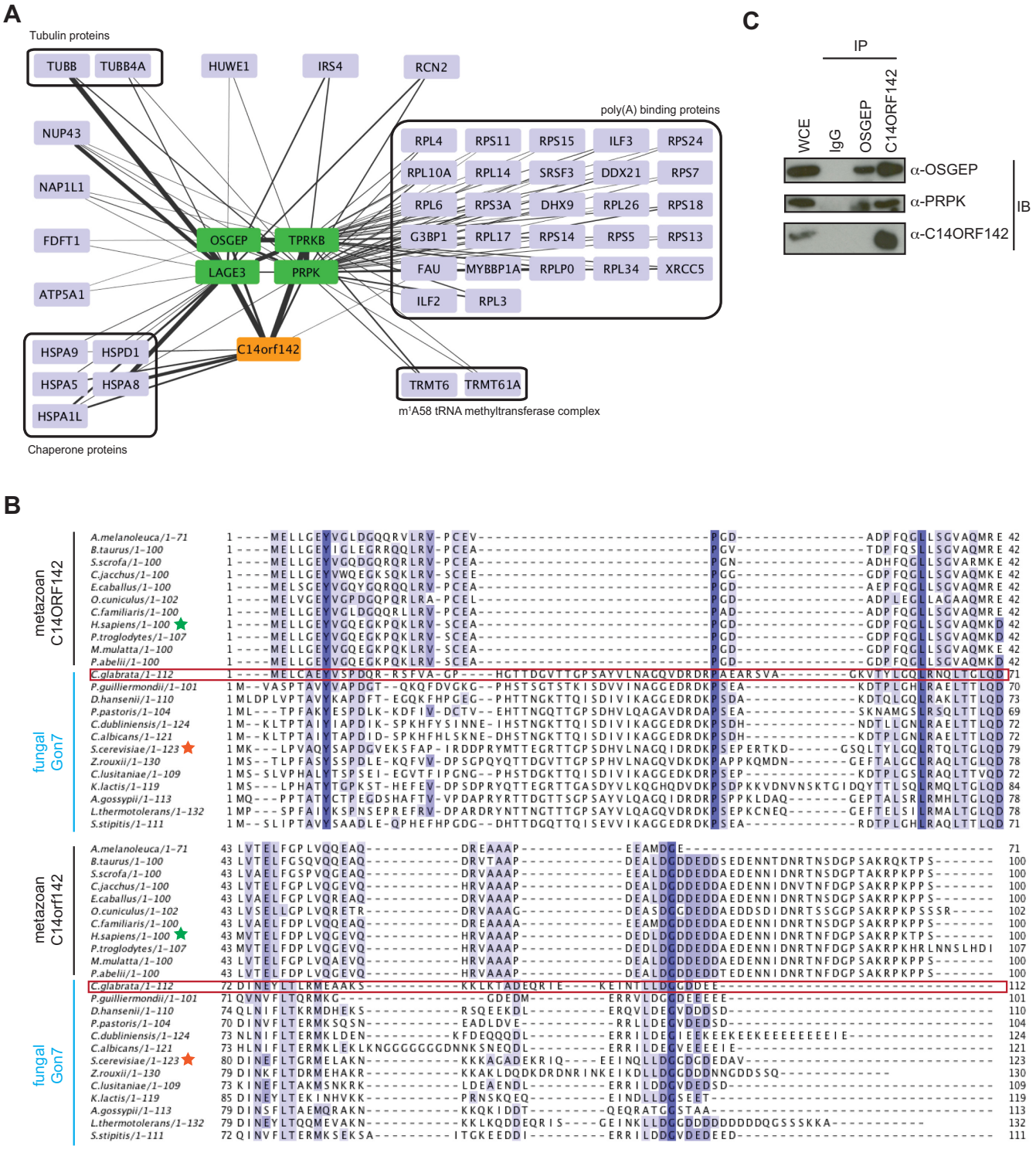

Figure 1. Affinity purification-mass spectrometry analysis of human KEOPS subunits reveals C14ORF142 as a prominent interactor. (**A**) Interaction map of OSGEP, PRPK, LAGE3 and PRPK proteins (shown in green). Each edge represents a protein–protein interaction between the two connected proteins with the line width representing the intensity of spectral counts for that particular interaction. One prominent interactor, C14ORF142 (shown in orange), strongly interacted with all four KEOPS subunits. (**B**) Co-immunoprecipitation and immunoblot analysis of endogenous OSGEP and C14ORF142 from HEK293 cells. (**C**) Multiple sequence alignment of recognizable metazoan C14ORF142 (top half) and fungal Gon7 (bottom half) orthologs. The human C14ORF142 ortholog is denoted by a green star and the *S.cerevisiae* Gon7 ortholog by a red star. The *C. glabrata* Gon7 ortholog is boxed in red.

with Bud32, Kae1, Cgi121 and Pcc1 proteins (12,13) and suggested that C14ORF142, like Gon7, might be an integral component of KEOPS. C14ORF142 is a 100-amino acid protein with no previously assigned function. However, C14ORF142 was previously identified as an interactor of three KEOPS subunits in a high throughput proteomics study, which led to speculation that C14ORF142 may be a Gon7 ortholog, although it was not experimentally confirmed (33). We first tested whether C14ORF142 could rescue the slow growth phenotype of a *gon7Δ* yeast strain. In contrast to overexpression of Gon7, expression of C14ORF142 did not functionally complement *GON7* deletion in yeast (Supplementary Figure S2). This finding was not unexpected considering the low sequence similarity between C14ORF142 and Gon7 (pairwise sequence identity of 17%) that reflects the evolutionary divergence between the two proteins (Figure 1B). Interestingly, in a multiple sequence alignment of readily recognized orthologs of Gon7 in fungal species and C14ORF142 in metazoan species (34) (Figure 1B), we observed that a *Candida glabrata* protein possessed stretches of sequence similarity that apparently link the two otherwise dissimilar protein families. Specifically, *C. glabrata* Gon7 possesses a MELCAEYV N-terminal sequence that highly resembles the MELLGEYV N-terminal sequence of human C14ORF142. Beyond the N-terminus, the pairwise identity between *C. glabrata* Gon7 and other fungal Gon7 orthologs ranges from 39.4% to 62.5%, with the *C. glabrata* and *S.cerevisiae* Gon7 orthologs having a pairwise identity of 62.5%. In contrast, the pairwise identities between *C. glabrata* Gon7 and metazoan C14ORF142 orthologs only range from 22.2% to 31.5%, with the *C. glabrata* Gon7 and human C14ORF142 orthologs having a pairwise identity of 29.2%. Together, these weak but tantalizing similarities led us to further investigate C14ORF142 as a potential ortholog of Gon7 in greater detail. Before doing so, we first validated C14ORF142 as an interactor of KEOPS by reciprocal AP-MS using HEK293 Flp-In T-REx cells stably expressing FLAG-C14ORF142 combined with immunoprecipitation and immunofluorescence analysis using antibodies recognizing endogenous C14ORF142, OSGEP and PRPK. In agreement with our initial AP-MS proteomics analyses, FLAG-C14ORF142 interacted with all four KEOPS subunits (Figure 1A), endogenous C14ORF142 co-immunoprecipitated with OSGEP and PRPK (Figure 1C), and endogenous C14ORF142 showed nuclear enrichment localization similar to OSGEP and PRPK (Supplementary Figure S3). Collectively, our data suggested that C14ORF142 is a core component of the human KEOPS complex.

C14ORF142 and Gon7 are intrinsically disordered proteins

Prior work analyzing the solution structure of yeast Gon7 using ^{1}H-^{15}N HSQC-NMR demonstrated that Gon7 is an intrinsically disordered protein (20). To examine whether C14ORF142 shares the same biophysical property as yeast Gon7, we investigated the solution structure of ^{15}N-labeled human C14ORF142 and yeast Gon7 using NMR spectroscopy (Figure 2A and Supplementary Figure S4). The ^{1}H-^{15}N-HSQC NMR spectra of both ^{15}N-labeled C14ORF142 and Gon7 revealed a similar pattern of high intensity chemical shifts between 8.0 and 8.5 ppm in the proton dimension, indicative of an intrinsically disordered protein in solution. Our observation that C14ORF142 and Gon7 are both intrinsically disordered further suggests an orthologous relationship.

C14ORF142 interacts directly with human KEOPS

If C14ORF142 is a true Gon7 ortholog, we reasoned that C14ORF142 would bind directly to KEOPS in a stoichiometric manner. To test this hypothesis, we generated a 5-subunit (OSGEP, PRPK, TPRKB, 6xHis-LAGE3 and C14ORF142) KEOPS polycistronic vector for protein expression in insect cells using the baculovirus system. We observed that C14ORF142 co-purified with OSGEP, PRPK, TPRKB and LAGE3 during nickel chelate affinity chromatography and size-exclusion chromatography as assessed by SDS-PAGE (Figure 2B). We next attempted to localize the direct binding partner of C14ORF142 within the KEOPS complex. To this end, we were able to produce a PRPK-TPRKB binary sub-complex in bacteria and a OSGEP-LAGE3 binary sub-complex in insect cells and tested each for binding to recombinant GST-C14ORF142 in pull-down reactions using glutathione resin. We observed that GST-C14ORF142 bound directly to the OSGEP-LAGE3 sub-complex but not to the PRPK-TPRKB sub-complex. However, GST-C14ORF142 did retain the PRPK-TPRKB sub-complex when incubated in the presence of the OSGEP-LAGE3 sub-complex, consistent with OSGEP-LAGE3 serving as bridging factors. As anticipated, GST-Cdk2, a fusion protein of a readily available protein kinase in our lab that was not detected as a KEOPS interactor using AP-MS, displayed only weak residual binding to OSGEP-LAGE3 and PRPK-TPRKB complexes *in vitro* (Figure 2C).

Since yeast Gon7 and Pcc1 proteins formed tight stoichiometric complexes when co-expressed in bacteria (20), we expected that C14ORF142 and LAGE3 would similarly form tight stoichiometric complexes when co-expressed in bacteria (Supplementary Figure S5). We first expressed C14ORF142 and LAGE3 as GST- and 6xHIS-GST fusion proteins, respectively, in bacteria and purified the proteins to homogeneity using glutathione affinity chromatography followed by size exclusion chromatography (SEC). During SEC purification, C14ORF142 and LAGE3 in isolation displayed elution volumes of 17 ml and 11 ml, respectively, with an overall yield of LAGE3 far lower (>1000-fold) than that observed for C14ORF142. Upon purification of C14ORF142 and LAGE3 co-expressed in bacteria, the yield of LAGE3 was dramatically improved (∼100-fold) and the two proteins co-migrated with 1:1 stoichiometry at an elution volume of 14 ml. In the same experiment, excess C14ORF142 migrated with an elution volume of 17 ml, identical to the behavior of C14ORF142 produced in isolation. Together, these observations are consistent with the hypothesis that C14ORF142 is a Gon7 ortholog that binds directly to LAGE3.

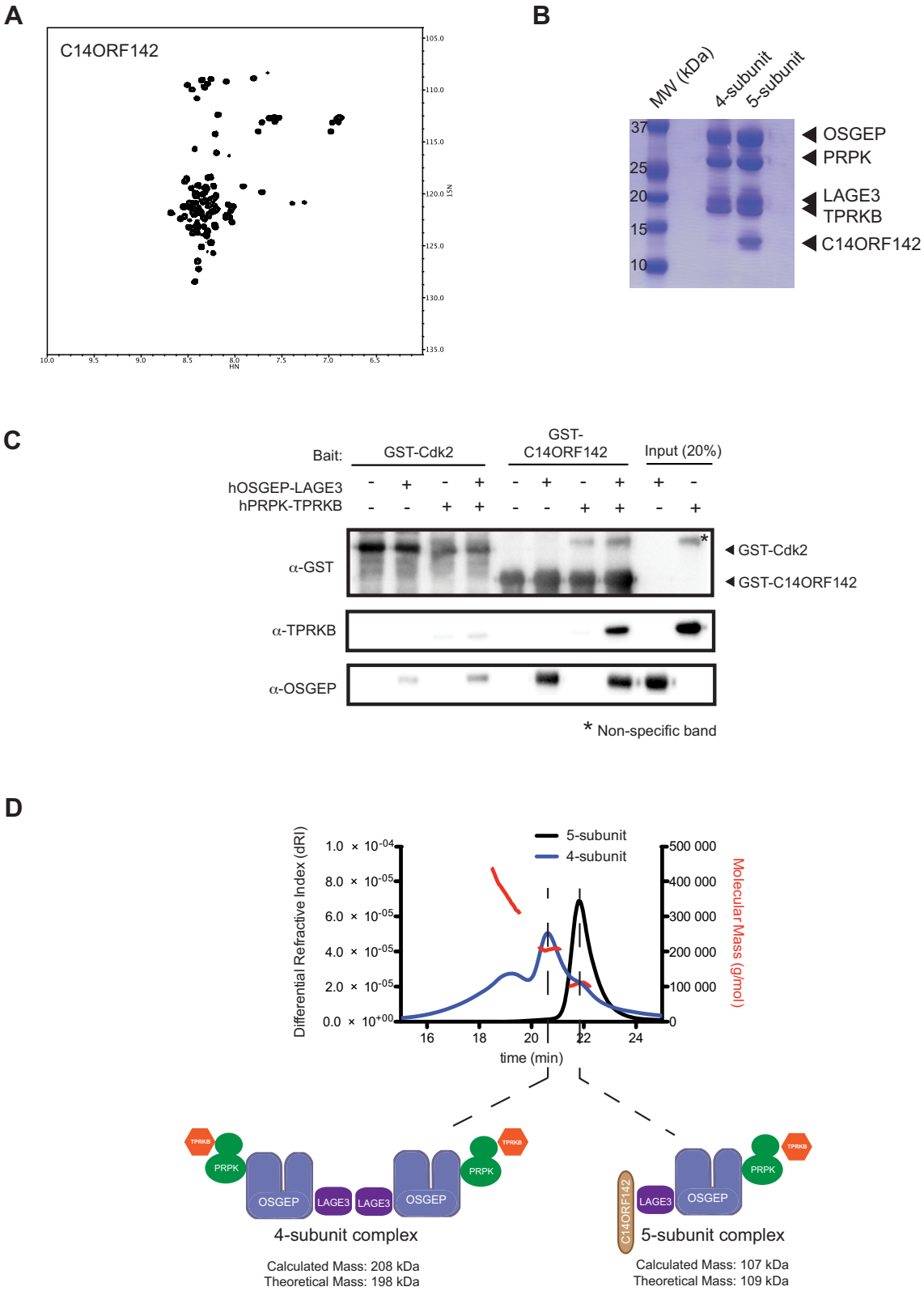

Figure 2. Biophysical and biochemical characterization of C14ORF142. (**A**) Nuclear magnetic resonance ^{1}H-^{15}N-HSQC analysis of ^{15}N-labeled C14ORF142. (**B**) SDS-PAGE analysis of a recombinant 4-subunit (OSGEP, PRPK, TPRKB and LAGE3) and 5-subunit (OSGEP, PRPK, TPRKB, LAGE3 and C14ORF142) KEOPS complex expressed in insect cells. (**C**) Glutathione S-transferase (GST) pulldown and immunoblot analysis using GST-C14ORF142 or GST-Cdk2 as bait and OSGEP-LAGE3 and PRPK-TPRKB binary complexes as prey. (**D**) SEC-MALS analysis of the 4- and 5-subunit KEOPS complexes (top). Cartoon schematic of the 4- and 5-subunit KEOPS complexes (bottom).

C14ORF142 binds KEOPS subunits with a 1:1:1:1:1 binding stoichiometry

Previous studies showed that *S.cerevisiae* Gon7 binds directly to Pcc1 on the same surface that supports Pcc1 homo-dimerization (20). In the absence of Gon7, Pcc1 homodimerization can support the super-dimerization of a 4-subunit archaeal KEOPS complex with a 2:2:2:2 binding stoichiometry (35). The disruption of superdimerization via the Gon7-Pcc1 binding interaction may account for the observation that a yeast 5-subunit KEOPS complex has a 1:1:1:1:1 stoichiometry in solution (20). We reasoned that if C14ORF142 was a Gon7 ortholog, the binding of C14ORF142 to OSGEP, PRPK, LAGE3 and TPRKB will also result in the formation of a 1:1:1:1:1 complex. To assess the precise binding stoichiometry of human KEOPS complexes in the presence and absence of C14ORF142, we performed SEC-MALS analysis on the recombinant 4-subunit and 5-subunit KEOPS complex produced from insect cells (Figure 2D). Consistent with the superdimerization disruption behavior of yeast Gon7, the 5-subunit KEOPS complex with C14ORF142 eluted as a single mono-disperse species with a calculated mass of 107 kDa, in good agreement with the theoretical mass of 109 kDa for a 1:1:1:1:1 stoichiometric complex. In contrast, the 4-subunit KEOPS complex lacking C14ORF142 eluted as two distinct peaks, the most prominent of which displayed a calculated mass of 208 kDa, in good agreement with the theoretical mass of 198 kDa for a 2:2:2:2 stoichiometric complex. Re-analysis of the pooled prominent peak by SEC revealed the identical profile as the original 4-subunit KEOPS complex sample, indicating that the 4-subunit complex resides in a equilibrium between a well defined 2:2:2:2 complex and a larger heterogeneous soluble aggregate.

C14ORF142 potentiates the t^6A biosynthesis activity of human KEOPS

We next examined whether the binding of C14ORF142 to KEOPS is important for t^6A biosynthesis activity *in vitro*. While *S. cerevisiae* Gon7 has not been tested for its effect on t^6A production *in vitro*, it does appear essential *in vivo*, with *gon7Δ* yeast strains displaying an extreme low level of intracellular t^6A comparable to that of the *kae1Δ* yeast strains (10). Using a reconstituted t^6A biosynthesis reaction with yeast Sua5, we measured the formation of t^6A products on an *in vitro* transcribed human $tRNA^{Lys(UUU)}$ in the presence and absence of the 4- and 5-subunit human KEOPS complexes produced from insect cells. As expected, KEOPS was necessary for t^6A modification of human $tRNA^{Lys(UUU)}$ while the C14ORF142-containing KEOPS complex displayed 3- to 4-fold more activity than the complex lacking C14ORF142 (Figure 3). The identity of the t^6A product was validated by mass spectrometry analysis of the t^6A peak isolated from our HPLC runs (Supplementary Figure S6).

DISCUSSION

The AP-MS analysis of the human KEOPS complex and subsequent biophysical and biochemical characterizations lead us to conclude that C14ORF142 is the human ortholog of yeast Gon7. The finding of a functional Gon7 ortholog in human raises the possibility that archaeal Gon7 orthologs may also exist, but due to an even greater evolutionary drift, these proteins cannot be detected by sequence alone. Thus, a proteomic analysis of KEOPS in archaeal species, as we performed in human cells, may be warranted. The discovery of the missing fifth regulatory subunit in human KEOPS (and by inference the missing Gon7 orthologs in other multicellular eukaryotic species) opens the door for future comprehensive studies into the biological and biochemical function of KEOPS in additional multicellular eukaryotes. We note that we did attempt to assess intracellular t^6A levels in a hTERT-RPE1 C14ORF142 CRISPR-derived knock out cell line, but the resulting slow-growth phenotype of selected clones hampered our ability to produce sufficient quantities of tRNA required for the analysis.

Possible functions for PRPK, TPRKB, LAGE3 and C14ORF142 outside of t^6A biogenesis

The interaction network for PRPK and TPRKB is highly enriched for ribosomal proteins, suggesting that a KEOPS-independent PRPK-TPRKB sub-complex might function in a translation-related process. Interestingly, the closest structural orthologs of PRPK, namely the Rio1 and Rio2 ATPases, are both components of the ribosome and play important roles in the maturation of ribosomal RNA in the 40S ribosomal subunit (36,37). This observation hints that PRPK-TPRKB, similar to Rio1 and Rio2, may also regulate ribosomal RNA maturation. In addition to a possible role in ribosomal RNA maturation, we detected novel PRPK-TPRKB interactions that suggest that PRPK-TPRKB may regulate tRNA modifications distinct from t^6A. For example, we observed that PRPK-TPRKB strongly interacted with the m^1A58 tRNA methyltransferase complex (TRMT61a-TRMT6), which methylates adenosine 58 on initiator $tRNA_i^{Met}$ (38,39). The absence of m^1A58 results in lower steady state levels of $tRNA_i^{Met}$ and causes a slow growth phenotype in *S.cerevisiae*, similar to what is observed for KEOPS dysfunction in *S.cerevisiae* (40). Intriguingly, t^6A-deficient yeast strains show a greater tendency to initiate protein translation at upstream non-AUG codons, which further alludes to a dysfunctionality of $tRNA_i^{Met}$ molecules (41). These similarities in phenotype suggest that the PRPK-TPRKB sub-complex may regulate both t^6A and m^1A58 biosynthesis.

Similar to PRPK-TPRKB, we detected interactions that suggest that the LAGE3-C14ORF142 sub-complex may function outside of KEOPS. Specifically, we detected an enrichment of interactors implicated in protein folding for both LAGE3 and C14ORF142. This suggests that protein chaperones may be critical to the proper folding of LAGE3 and C14ORF142 or alternatively, that LAGE3 and C14ORF142 are more pervasive regulators of protein folding. Interestingly, five of the six LAGE3- and C14ORF142-interacting heat shock chaperones, HspA9 (PDB 4KB0) (42), HspA5 (PDB 3LDL) (43), HspA8 (PDB 4H5N) (44), HspA1B (PDB 4J8F) (45) and HspA1L (PDB 3GDQ) (46), belong to the ASKHA fold superfamily of proteins, which also includes OSGEP. If LAGE3 and C14ORF142 do indeed regulate protein folding, we speculate that the two pro-

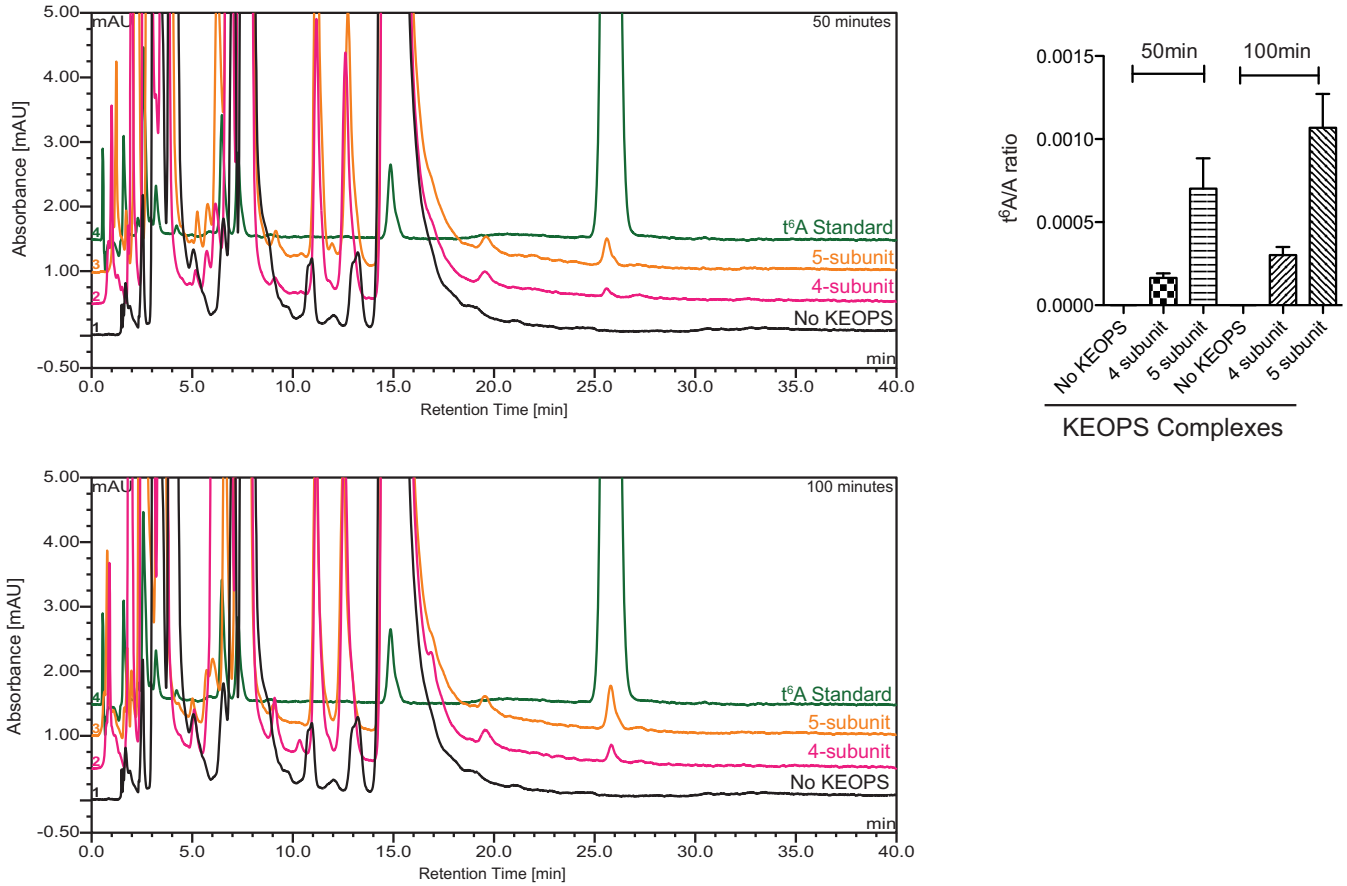

Figure 3. C14ORF142 potentiates the t^6A biosynthesis activity of KEOPS. Representative HPLC spectra of digested tRNA nucleosides from reconstituted *in vitro* t^6A reactions at 50 min (upper left) and 100 min (lower left) time points for both 4- and 5-subunit KEOPS complexes. Quantification of reconstituted *in vitro* t^6A reactions (right). Data are expressed as mean ± SEM (n = 3).

teins may engage and regulate their chaperone interactors in the same manner that they engage and regulate OSGEP.

In sum, the collection of our interaction data shed new light on the identity, structure and function of C14ORF142 as the missing fifth subunit in the human KEOPS complex. In addition, our findings provide new leads for future investigation into the diverse biological functions of KEOPS and its individual subunits.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Dr Genevieve Seabrook at the University Health Network High Field NMR facility for her assistance in collecting NMR data and her scientific and technical expertise. The 600MHz NMR spectrometer at the University Health Network High Field NMR facility is funded by the Canada Foundation for Innovation (CFI). Proteomics work was performed at the Network Biology Collaborative Centre at the Lunenfeld–Tanenbaum Research Institute, a facility supported by CFI funding and by Genome Canada and Ontario Genomics.

FUNDING

Canadian Institutes of Health Research (CIHR) Foundation [FDN 143277 to F.S., FDN 143343 to D.D., FDN 143301 to A-C.G]; Genome Canada and Ontario Genomics Funding [OGI-088 to A-C.G]. Funding for open access charge: CIHR Foundation [FDN 143277].
Conflict of interest statement. None declared.

REFERENCES

1. El Yacoubi,B., Bailly,M. and de Crecy-Lagard,V. (2012) Biosynthesis and function of posttranscriptional modifications of transfer RNAs. *Annu. Rev. Genet.*, **46**, 69–95.
2. Miyauchi,K., Kimura,S. and Suzuki,T. (2013) A cyclic form of N6-threonylcarbamoyladenosine as a widely distributed tRNA hypermodification. *Nat. Chem. Biol.*, **9**, 105–111.
3. Chheda,G.B., Hong,C.I., Piskorz,C.F. and Harmon,G.A. (1972) Biosynthesis of N-(purin-6-ylcarbamoyl)-L-threonine riboside. Incorporation of L-threonine in vivo into modified nucleoside of transfer ribonucleic acid. *Biochem. J.*, **127**, 515–519.
4. Elkins,B.N. and Keller,E.B. (1974) The enzymatic synthesis of N-(purin-6-ylcarbamoyl)threonine, an anticodon-adjacent base in transfer ribonucleic acid. *Biochemistry*, **13**, 4622–4628.
5. Kimura-Harada,F., Harada,F. and Nishimura,S. (1972) The presence of N-(9-(c-D-ribofuranosyl)purin-6-ylcarbamoyl) threonine in isoleucine, threonine and asparagine tRNAs from Escherichia coli. *FEBS Lett.*, **21**, 71–74.

6. 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.
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. 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, t6A. *EMBO J.*, **30**, 873–881.
9. Lauhon,C.T. (2012) Mechanism of N6-Threonylcarbamoyladenonsine (t(6)A) Biosynthesis: Isolation and Characterization of the Intermediate Threonylcarbamoyl-AMP. *Biochemistry*, **51**, 8950–8963.
10. 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 N6-threonylcarbamoylation of tRNA using a minimal enzyme system. *Nucleic Acids Res.*, **41**, 6332–6346.
11. Parthier,C., Gorlich,S., Jaenecke,F., Breithaupt,C., Brauer,U., Fandrich,U., Clausnitzer,D., Wehmeier,U.F., Bottcher,C., Scheel,D. *et al.* (2012) The O-carbamoyltransferase TobZ catalyzes an ancient enzymatic reaction. *Angew. Chem. Int. Ed. Engl.*, **51**, 4046–4052.
12. Downey,M., Houlsworth,R., Maringele,L., Rollie,A., Brehme,M., Galicia,S., Guillard,S., Partington,M., Zubko,M.K., Krogan,N.J. *et al.* (2006) A genome-wide screen identifies the evolutionarily conserved KEOPS complex as a telomere regulator. *Cell*, **124**, 1155–1168.
13. Kisseleva-Romanova,E., Lopreiato,R., Baudin-Baillieu,A., Rousselle,J.C., Ilan,L., Hofmann,K., Namane,A., Mann,C. and Libri,D. (2006) Yeast homolog of a cancer-testis antigen defines a new transcription complex. *EMBO J.*, **25**, 3576–3585.
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. Oberto,J., Breuil,N., Hecker,A., Farina,F., Brochier-Armanet,C., Culetto,E. and Forterre,P. (2009) Qri7/OSGEPL, the mitochondrial version of the universal Kae1/YgjD protein, is essential for mitochondrial genome maintenance. *Nucleic Acids Res.*, **37**, 5343–5352.
16. Deutsch,C., El Yacoubi,B., de Crecy-Lagard,V. and Iwata-Reuyl,D. (2012) Biosynthesis of threonylcarbamoyl adenosine (t6A), a universal tRNA nucleoside. *J. Biol. Chem.*, **287**, 13666–13673.
17. Handford,J.I., Ize,B., Buchanan,G., Butland,G.P., Greenblatt,J., Emili,A. and Palmer,T. (2009) Conserved network of proteins essential for bacterial viability. *J. Bacteriol.*, **191**, 4732–4749.
18. Nichols,C.E., Lamb,H.K., Thompson,P., Omari,K.E., Lockyer,M., Charles,I., Hawkins,A.R. and Stammers,D.K. (2013) Crystal structure of the dimer of two essential Salmonella typhimurium proteins, YgjD & YeaZ and calorimetric evidence for the formation of a ternary YgjD-YeaZ-YjeE complex. *Protein Sci.*, **22**, 628–640.
19. 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 t6A in Escherichia coli. *Nucleic Acids Res.*, **43**, 1804–1817.
20. Zhang,W., Collinet,B., Graille,M., Daugeron,M.C., Lazar,N., Libri,D., Durand,D. and van Tilbeurgh,H. (2015) Crystal structures of the Gon7/Pcc1 and Bud32/Cgi121 complexes provide a model for the complete yeast KEOPS complex. *Nucleic Acids Res.*, **43**, 3358–3372.
21. Kean,M.J., Ceccarelli,D.F., Goudreault,M., Sanches,M., Tate,S., Larsen,B., Gibson,L.C., Derry,W.B., Scott,I.C., Pelletier,L. *et al.* (2011) Structure-function analysis of core STRIPAK Proteins: a signaling complex implicated in Golgi polarization. *J. Biol. Chem.*, **286**, 25065–25075.
22. Kean,M.J., Couzens,A.L. and Gingras,A.C. (2012) Mass spectrometry approaches to study mammalian kinase and phosphatase associated proteins. *Methods*, **57**, 400–408.
23. Dunham,W.H., Larsen,B., Tate,S., Badillo,B.G., Goudreault,M., Tehami,Y., Kislinger,T. and Gingras,A.C. (2011) A cost-benefit analysis of multidimensional fractionation of affinity purification-mass spectrometry samples. *Proteomics*, **11**, 2603–2612.
24. Liu,G., Zhang,J., Larsen,B., Stark,C., Breitkreutz,A., Lin,Z.Y., Breitkreutz,B.J., Ding,Y., Colwill,K., Pasculescu,A. *et al.* (2010) ProHits: integrated software for mass spectrometry-based interaction proteomics. *Nat. Biotechnol.*, **28**, 1015–1017.
25. Deutsch,E.W., Mendoza,L., Shteynberg,D., Farrah,T., Lam,H., Tasman,N., Sun,Z., Nilsson,E., Pratt,B., Prazen,B. *et al.* (2010) A guided tour of the trans-proteomic pipeline. *Proteomics*, **10**, 1150–1159.
26. Shteynberg,D., Deutsch,E.W., Lam,H., Eng,J.K., Sun,Z., Tasman,N., Mendoza,L., Moritz,R.L., Aebersold,R. and Nesvizhskii,A.I. (2011) iProphet: multi-level integrative analysis of shotgun proteomic data improves peptide and protein identification rates and error estimates. *Mol. Cell Proteomics*, **10**, M111 007690.
27. Teo,G., Liu,G., Zhang,J., Nesvizhskii,A.I., Gingras,A.C. and Choi,H. (2014) SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. *J. Proteomics*, **100**, 37–43.
28. Mellacheruvu,D., Wright,Z., Couzens,A.L., Lambert,J.P., St-Denis,N.A., Li,T., Miteva,Y.V., Hauri,S., Sardiu,M.E., Low,T.Y. *et al.* (2013) The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods*, **10**, 730–736.
29. Harlow,E. and Lane, D. (1988) *Antibodies: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, pp. 63–528.
30. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
31. Johnson,B.A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.*, **278**, 313–352.
32. Gehrke,C.W. and Kuo,K.C. (1989) Ribonucleoside analysis by reversed-phase high-performance liquid chromatography. *J. Chromatogr.*, **471**, 3–36.
33. Costessi,A., Mahrour,N., Sharma,V., Stunnenberg,R., Stoel,M.A., Tijchon,E., Conaway,J.W., Conaway,R.C. and Stunnenberg,H.G. (2012) The human EKC/KEOPS complex is recruited to Cullin2 ubiquitin ligases by the human tumour antigen PRAME. *PLoS One*, **7**, e42822.
34. Pei,J., Kim,B.H. and Grishin,N.V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.*, **36**, 2295–2300.
35. Wan,L.C., Pillon,M.C., Thevakumaran,N., Sun,Y., Chakrabartty,A., Guarne,A., Kurinov,I., Durocher,D. and Sicheri,F. (2016) Structural and functional characterization of KEOPS dimerization by Pcc1 and its role in t6A biosynthesis. *Nucleic Acids Res.*, **44**, 6971–6980.
36. Ferreira-Cerca,S., Sagar,V., Schafer,T., Diop,M., Wesseling,A.M., Lu,H., Chai,E., Hurt,E. and LaRonde-LeBlanc,N. (2012) ATPase-dependent role of the atypical kinase Rio2 on the evolving pre-40S ribosomal subunit. *Nat. Struct. Mol. Biol.*, **19**, 1316–1323.
37. Turowski,T.W., Lebaron,S., Zhang,E., Peil,L., Dudnakova,T., Petfalski,E., Granneman,S., Rappsilber,J. and Tollervey,D. (2014) Rio1 mediates ATP-dependent final maturation of 40S ribosomal subunits. *Nucleic Acids Res.*, **42**, 12189–12199.
38. Anderson,J., Phan,L. and Hinnebusch,A.G. (2000) The Gcd10p/Gcd14p complex is the essential two-subunit tRNA(1-methyladenosine) methyltransferase of Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 5173–5178.
39. Ozanick,S., Krecic,A., Andersland,J. and Anderson,J.T. (2005) The bipartite structure of the tRNA m1A58 methyltransferase from S. cerevisiae is conserved in humans. *RNA*, **11**, 1281–1290.
40. Anderson,J., Phan,L., Cuesta,R., Carlson,B.A., Pak,M., Asano,K., Bjork,G.R., Tamame,M. and Hinnebusch,A.G. (1998) The essential Gcd10p-Gcd14p nuclear complex is required for 1-methyladenosine modification and maturation of initiator methionyl-tRNA. *Genes Dev.*, **12**, 3650–3662.
41. 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 t6A in yeast. *Microb Cell*, **3**, 29–45.
42. Amick,J., Schlanger,S.E., Wachnowsky,C., Moseng,M.A., Emerson,C.C., Dare,M., Luo,W.I., Ithychanda,S.S., Nix,J.C., Cowan,J.A. *et al.* (2014) Crystal structure of the nucleotide-binding domain of mortalin, the mitochondrial Hsp70 chaperone. *Protein Sci.*, **23**, 833–842.

43. Macias,A.T., Williamson,D.S., Allen,N., Borgognoni,J., Clay,A., Daniels,Z., Dokurno,P., Drysdale,M.J., Francis,G.L., Graham,C.J. *et al.* (2011) Adenosine-derived inhibitors of 78 kDa glucose regulated protein (Grp78) ATPase: insights into isoform selectivity. *J. Med. Chem.*, **54**, 4034–4041.
44. Zhang,Z., Cellitti,J., Teriete,P., Pellecchia,M. and Stec,B. (2015) New crystal structures of HSC-70 ATP binding domain confirm the role of individual binding pockets and suggest a new method of inhibition. *Biochimie*, **108**, 186–192.
45. Li,Z., Hartl,F.U. and Bracher,A. (2013) Structure and function of Hip, an attenuator of the Hsp70 chaperone cycle. *Nat. Struct. Mol. Biol.*, **20**, 929–935.
46. Wisniewska,M., Karlberg,T., Lehtio,L., Johansson,I., Kotenyova,T., Moche,M. and Schuler,H. (2010) Crystal structures of the ATPase domains of four human Hsp70 isoforms: HSPA1L/Hsp70-hom, HSPA2/Hsp70-2, HSPA6/Hsp70B', and HSPA5/BiP/GRP78. *PLoS One*, **5**, e8625.