Targeted identification of protein interactions in eukaryotic mRNA translation

Andrew J. Link[1,3,4]*, Xinnan Niu[1], Connie M. Weaver[1], Jennifer L. Jennings[1], Dexter T. Duncan[1], K. Jill McAfee[1], Morgan Sammons[5], Vince R. Gerbasi[1], Adam R. Farley[3], Tracey C. Fleischer[1], Christopher M. Browne[3], Parimal Samir[3], Allison Galassie[4], and Braden Boone[2]

[1]Department of Pathology, Microbiology and Immunology, [2]Department of Bioinformatics, Vanderbilt University School of Medicine, Nashville, TN 37232 [3]Department of Biochemistry, [4]Department of Chemistry, [5]Department of Biological Sciences, Vanderbilt University, Nashville, TN 37232

*Corresponding author

Andrew J. Link, PhD

Department of Pathology, Microbiology and Immunology

Vanderbilt University School of Medicine

1161 21st Ave South

Nashville, TN 37232

TEL: (615) 343-6823

FAX: (615) 343-7392

Andrew.Link@vanderbilt.edu

Received: 14/05/2019; Revised: 13/12/2019; Accepted: 04/02/2020

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/pmic.201900177.

Summary

To identify protein-protein interactions and phosphorylated amino acid sites in eukaryotic mRNA translation, we performed replicate TAP-MudPIT and control experiments targeting *S. cerevisiae* genes previously implicated in eukaryotic mRNA translation by their genetic and/or functional roles in translation initiation, elongation, termination, or interactions with ribosomal complexes. Replicate tandem affinity purifications of each targeted yeast TAP-tagged mRNA translation protein coupled with multidimensional liquid chromatography and tandem mass spectrometry analysis were used to identify and quantify copurifying proteins. To improve sensitivity and minimize spurious, nonspecific interactions, we employed a novel cross-validation approach to identify the most statistically-significant protein-protein interactions. Using our experimental and computational strategies, we validated the previously described protein composition of the canonical eukaryotic mRNA translation initiation, elongation, and termination complexes. In addition, we identified statistically-significant unpublished protein interactions and phosphorylation sites for *S. cerevisiae's* mRNA translation proteins and complexes.

Introduction

mRNA translation is the process in which ribosomes and amino acid-charged tRNAs decode mRNAs and synthesize polypeptides. It is an essential process for all organisms. In eukaryotes, the coordinated translation of mRNAs, known as translational control, is a major regulatory mechanism involved in many essential biological processes, including development [1], stress response [2], signaling [3], plasticity [4], immune response [5], and cell growth [6]. Defects in protein synthesis and translational control are major factors in human diseases, including fragile-X syndrome [7] and cancers [3c, 8]. Many bacterial and viral pathogens target the human protein translation process [9]. Drug therapies focusing on the translation machinery are being developed and are currently being used to treat various human diseases [10, 62, 63, 64, 65]. As such, a comprehensive understanding of eukaryotic mRNA translation and its control mechanisms is essential to understanding both normal and disease-altered cellular processes.

Eukaryotic mRNA translation is a highly regulated process controlled by a complex network of proteins and posttranslational modifications [3a, 11]. Composed of a small 40S and a large 60S subunit, the eukaryotic 80S ribosome is the site of protein synthesis in eukaryotic cells. The generation of 80S ribosomes is an energy intensive and highly complex process requiring the coordinate activity of a large number of trans-acting protein assembly factors [12]. The 80S ribosomes interact with and coordinate the interactions between mRNAs, amino acid-charged tRNAs, and protein translation factors to synthesize new polypeptides in the cell [13]. 80S ribosomes are found in large numbers either freely in the cytoplasm or attached to the surfaces of the endoplasmic reticulum. High resolution cryo-EM and x-ray crystallographic structures of the 80S ribosome have been determined [14]. Data show that the 40S

subunit orchestrates the recruitment, pairing, and positioning of an mRNA's codon with the cognate anticodon of an aminoacyl tRNA [11d, 15]. The 60S subunit catalyzes peptide bond formation between aminoacyl tRNAs and the growing polypeptide attached to the peptidyl tRNA [13]. A growing number of essential and nonessential proteins have been identified that interact with the ribosome, mRNAs, and tRNAs to mediate and regulate protein synthesis [11a, 11d, 16].

Translation of eukaryotic mRNAs is typically divided into three phases: initiation, elongation, and termination/recycling (**Fig. 1**). The eukaryotic translation initiation phase is a multistage process in which the 40S and 60S ribosomal subunits along with the initiator Met-tRNA$_i$ are assembled by a complex network of eukaryotic initiation factors (eIFs) into an 80S ribosome at the mRNA's AUG initiation codon (**Fig. 2A**) [17]. Several reviews describe our current understanding of the eukaryotic translation initiation process and the functional roles of the initiation factors [11a, 11e, 15b, 16-18]. For most eukaryotic mRNAs, Kozak's scanning model describes our current understanding of eukaryotic translation initiation [11a, 15b, 19]. In the Kozak model, a network of interacting eIF protein complexes (**Fig. 2A**) recruit the 40S ribosomal subunit along with the methionyl initiator tRNA (Met-tRNA$_i$) to the mRNA's 5' cap structure, where it scans the 5'UTR of the mRNA for an AUG start codon in a favorable sequence context. When the Met-tRNA$_i$'s anticodon aligns with the mRNA's AUG start codon, the eIFs dissociate, and the 60S ribosomal subunit joins to form the 80S ribosome with the Met-tRNA$_i$ in the P site [11d, 11e, 15b, 20]. Alternative models of eukaryotic translation initiation have also been proposed and studied, including cap-independent translation initiation involving internal ribosomal entry sites (IRES) [21].

Following initiation, the elongation phase of mRNA translation involves the 80S ribosome traveling down the mRNA, reading codons, and recruiting cognate

aminoacyl tRNAs to the ribosome's A-site for the stepwise catalytic addition of charged amino acids to the growing polypeptide, which is attached to the peptidyl tRNA in the 80S ribosome's P-site [11a]. The eukaryotic elongation phase requires a network of eukaryotic elongation factors (eEFs) interacting with aminoacyl-tRNAs and the mRNA-80S complex (**Fig. 2B**) [11a, 22].

When the 80S ribosome reaches the mRNA's stop codon, the ribosome complex and a network of eukaryotic protein release factors (eRFs) (**Fig. 2C**) terminate polypeptide synthesis and promote the release of the nascent polypeptide and subsequent dissociation of the 80S ribosome into the 40S and 60S ribosomal subunits [11a, 22-23]. The small and large ribosomal subunits are recycled for a new round of translation initiation, elongation, and termination [11a, 22-23].

Protein-protein interactions are the physical contact of high specificity between two or more protein molecules [24]. By definition, a protein complex is a group of two or more associated polypeptide chains linked by non-covalent interactions. The interactions in the complex are physically mediated by a combination of forces including electrostatic interactions, hydrogen bonds, van der Waals attraction, and hydrophobic effects [24, 25]. Protein complexes are the foundation of many biological processes in the cell and perform a vast array of essential biological functions including mRNA translation [24]. The close proximity of the protein components in the complex can improve the rate and selectivity of binding interactions between the protein complex and its substrates, leading to higher cellular efficiency [24, 25]. Specific members of the complex may have different functions [24], and they may either activate or inhibit one or more of the complex's other protein components [24, 25].

The comprehensive mapping of the protein interactions in protein complexes provides potentially new insights into novel, unexpected protein interactions, biochemical functions, and regulation in a cellular process. A number of genetic and biochemical approaches have been developed to identify and characterize protein-protein interactions [25]. The discovery of novel or unexpected protein interactions enables putative functional or biochemical roles to be assigned to previously uncharacterized proteins in a biological process. The screening and development of pharmaceutical agents that target protein-protein interactions are being actively pursued for treating a variety of human diseases and abnormalities, including various cancers [24b, 26, 61, 62]. Eukaryotic mRNA translation has emerged as a therapeutic target for a growing number of human diseases [61, 65].

Several large-scale studies have applied epitope-tagging and mass spectrometry to identify protein-protein interactions on a global scale [27]. Alternative approaches for identifying protein-protein interactions, including the yeast 2-hybrid method, have also been used to target binary protein interactions for both specific proteins or a targeted organism's proteins on a global scale [60, 28b]. Multiple public databases catalogue the vast number of identified protein-protein interactions [28a, 29]. The Biological General Repository for Interaction Datasets (BioGRID) is a public database that archives and disseminates genetic and protein interaction data from model organisms and humans [30].

Tandem affinity purification coupled with mass spectrometry (TAP-MS) was developed as a generic protocol to purify a targeted protein expressed at its natural level under native conditions and to identify the interacting proteins copurifying with the targeted protein [31]. The epitope-tagged TAP protein is often referred to as the "bait" and the co-purifying proteins as the "prey" proteins. Despite the success of

TAP-MS experiments to identify protein interactions, nonspecific interactions between bait and prey proteins challenge investigators to distinguish in an unbiased manner between nonspecific bait-prey interactions (false interactions/false positives) and authentic *in vivo* interactions [32]. Aggregate databases of protein contaminants commonly observed in protein affinity purification-mass spectrometry experiments have been assembled [33]. Various statistical and computational tools to process and analyze mass spectrometry data such as *QPROT* [34], *CompPASS* [35], and *MSFragger* [36] have been developed to process and analysis mass spectrometry-derived protein interaction data to reduce the false discovery rate while maintaining the sensitivity to identify true interactions between the bait and prey proteins, especially transient or weak interactions [32].

We originally identified the *S. cerevisiae* protein Asc1p as a novel component of the yeast 40S ribosomal subunit [37]. Given the high percentage of uncharacterized genes in the sequenced genomes of eukaryotic organisms, we hypothesized that unexpected and unpublished proteins are associated with eukaryotic mRNA translation complexes. To test this hypothesis, we performed a large-scale, systematic, replicate tandem affinity purification (TAP) and mass spectrometry analysis (**Fig. S1**) on *S. cerevisiae* translation initiation, elongation, termination, and ribosome-associated proteins (**Table S1**). TAP-tagged yeast strains for each targeted protein were either obtained from previous projects designed to study the *S. cerevisiae* proteome or generated in this project [38]. Using our library of TAP-tagged yeast strains targeting previously identified mRNA translation protein factors (**Fig. 2A, B, C** and **Table S1**), we performed replicate protein purifications under native conditions. To identify and quantity the purified proteins from each TAP-tagged strain, we employed two-dimensional microcapillary liquid chromatography coupled

with micro-electrospray ionization and automated tandem mass spectrometry (MudPIT) on the trypsin-digested purified protein complexes [37a]. To identify the peptide sequences of the tryptic protein fragments, each acquired tandem mass spectrum was computationally compared to the entire *S. cerevisiae* proteome using the *Sequest* algorithm [39]. The identified peptides were then reassembled into a list of proteins and abundance factors calculated using the mass spectrometry data [37b, 40a, 40b, 40c-n]. Multiple computational approaches were then used to identify the most statistically significant prey proteins interacting with each bait protein. Using a novel strategy, we merged independent statistical approaches to identify the most statistically significant protein-protein interactions focusing on unexpected, unpublished interactions. In this study, the identification of previously identified protein interactions (**Fig. 2A, B, C**) validated the published data and supported our experimental and bioinformatics approaches to identify new, unpublished protein interactions.

Protein phosphorylation is a reversible post-translational modification of proteins in which an amino acid residue typically a serine, threonine, or tyrosine residue is covalently modified by a protein kinase and the addition of phosphate group [41]. Phosphorylation alters the structural conformation of a protein, causing it to become either activated, deactivated, or modifying its function[1]. Protein phosphorylation has been shown to be an important regulatory mechanism for reversibly modulating mRNA translation activity [42]. In this study, the copious amounts of acquired mass spectrometry data obtained identifying the targeted *S. cerevisiae* mRNA translation bait and interacting prey proteins were also analyzed to identify unexpected and unpublished phosphorylated amino acid residues.

The identification of both unexpected and unpublished proteins interacting with the canonical *S. cerevisiae* translation initiation, elongation, or termination proteins and complexes as well as the identification of potentially reversible phosphorylation sites expands our fundamental knowledge of this essential biological process. The unexpected and previously unpublished protein interactions and protein phosphorylation sites discovered in this study are expected to drive future functional, mechanistic, and structural studies to dissect their roles in the essential eukaryotic biological process of mRNA translation.

Materials and Methods

Yeast techniques

Media and protocols used for growing yeast in this study have been previously described [43].

Yeast strains

The *S. cerevisiae* TAP strains used in this project have been previously described [38a] and were obtained from Open Biosystems, Inc. To construct a TAP-tagged Sui3 yeast strain (AL085), the yeast strain CVY1 (*MATα ura3-52 his3-Δ200 lys2-80 trp1-Δ901)* was transformed with the PCR product generated from the amplification of the TAP cassette in the plasmid pFA6a-kanMX6-CTAP4 using the primers shown below [44].

5'-

ATTAAAACCGGTTTCCAAGCTACCGTTGGTAAGAGAAGGAGAATGCGGATCCCCGGGTTAA

TTAA-3'

and

5'-

AAATCCGTATTTATTATATATATGCTAACAGGTAAAGCACCAACAGAATTCGAGCTCGTTTAA

AC-3'

Tandem affinity purification

For the purification of translation complexes, each TAP-tagged strain and an untagged yeast strain (control) were grown in 2 L of YPD medium to an $O.D._{600}$ of 2-4 and processed (**Fig. S1**) [40g]. Yeast cells were pelleted at 2,300xg for 10 min at 4°C and washed with ice-cold water. Cells were broken open using a BioSpec bead beater and ice-cold 0.4-0.6 mm glass beads in lysis buffer (1% NP-40, 6 mM Na_2HPO_4, 4 mM NaH_2PO_4, 150 mM NaCl, 2 mM EDTA, 50 mM NaF, 4 ug/ml leupeptin, 0.1 mM Na_3VO_4, and 1x Roche Complete Protease Inhibitor). The crude lysates were centrifuged at 2,300xg for 5 min at 4°C. The partially-cleared lysates were mixed with 1 ml of a 1:1 slurry of IgG-Sepharose resin (Amersham) equilibrated in lysis buffer without protease inhibitors and incubated for 1 h on a nutator at 4°C. The IgG-Sepharose

resin was pelleted by centrifuging at 200xg for 2 min at 4°C. The lysates were discarded, and the IgG-Sepharose beads were resuspended in 10 mM Tris-HCl pH 8, 150 mM NaCl, 0.1% NP-40. The resuspended beads were transferred to BioRad Poly-prep chromatography columns. The IgG-Sepharose beads were equilibrated in TEV buffer (10 mM Tris-HCl pH 8, 150 mM NaCl, 0.1% NP-40, 0.5 mM EDTA, and 1 mM DTT). After capping the bottoms of the columns, 300 units of TEV protease in 2 ml of TEV buffer were added, and the columns were incubated for 1 h on a nutator at room temperature. The elution from each column was transferred to a new, capped poly-prep column and mixed with 6 ml of calmodulin binding buffer (0.1% NP-40, 10 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM MgOAc, 1 mM imidazole, 2 mM $CaCl_2$, 10 mM β-mercaptoethanol). Each solution was mixed with 300 µl of a 1:1 slurry of calmodulin affinity beads (Stratagene) equilibrated in calmodulin binding buffer, and the beads were incubated for 1 h on a nutator at 4°C. Proteins were eluted from the beads with 1 ml of 10 mM Tris-HCl pH 8, 150 mM NaCl, 0.02% NP-40, 1 mM MgOAc, 1 mM imidizole, 20 mM EGTA, and 10 mM β-mercaptoethanol.

Precipitation of TAP elution

To concentrate the eluted complexes for mass spectrometry analysis, 85% of each TAP elution was precipitated with trichloroacetic acid, washed with acetone, and resuspended in 20 µl of 100 mM ammonium bicarbonate, 5% acetonitrile.

Trypsin digestion of protein complexes

Resolubilized TAP-isolated proteins were reduced with a 1/10 volume of 50 mM dithiothreitol at 65°C for 5 min and then alkylated with a 1/10 volume of 100 mM iodoacetamide at 30°C for 30 min in the dark. To digest the proteins, 2 µg of modified trypsin (Promega) was added, and the reactions were incubated for 18 h at 37°C.

Mass spectrometry analysis of trypsin-digested protein complexes

Each trypsin-digested TAP sample was analyzed using multidimensional protein identification technology (MudPIT) [37a, 45] (**Fig. S1**). Briefly, a fritless, 100 µm i.d. microcapillary column was packed with 9 cm of 5 µm C_{18} reverse-phase material (Synergi 4µ Hydro RP80a, Phenomenex) followed by 3 cm of 5 µm strong cation exchange material (Partisphere SCX, Whatman), and finally 2 cm of the initial C_{18} reverse-phase material. The trypsin-digested TAP sample was loaded directly onto the triphasic column equilibrated in 0.1% formic acid, 2% acetonitrile. The triphasic column was placed in-line with an LCQ-Deca-XP-Plus or LTQ-OrbitrapXL ion trap mass spectrometer (Thermo Fisher, Inc). An automated 6-cycle multidimensional chromatographic separation was performed using buffer A (0.1% formic acid, 5% acetonitrile), buffer B (0.1% formic acid, 80% acetonitrile), and buffer C (0.1% formic acid, 5% acetonitrile, 500 mM ammonium acetate) at a flow rate of 0.3 µl/min. The

first cycle was a 20 min isocratic flow of buffer B. Cycles 2-6 consisted of 3 min of buffer A, 2 min of X% buffer C, 5 min of buffer A, and a 60 min linear gradient to 60% buffer B. Cycles 2-6 used 15, 30, 50, 70, and 100% of buffer C, respectively. During the linear gradient, the eluting peptides were analyzed by one full MS scan (400-2000 m/z) followed by five MS/MS scans on the five most abundant ions in the full MS scan while operating under dynamic exclusion.

Mass spectrometry data analysis

RAW data files generated by the MudPIT experiments were converted to an ASCII peak list using the program *extractms2* to identify +1 or multiply charged precursor ions (Jimmy Eng and John R. Yates III, personal communication). For multiply charged precursor ions ($z \geq +2$), an independent search was performed on both the +2 and +3 mass of the parent ion. Initial analysis of all acquired tandem mass spectrometry data used the *SEQUEST-PVM* algorithm with the precursor peptide mass tolerance set to 3 Da and a forward and reverse *S. cerevisiae* ORF protein database (SGD.fasta.6718) with a fixed cysteine modification of 57 Da and no enzyme specificity [39, 46]. Fully-tryptic peptides identified with a 5% false discovery rate were processed into a list of identified proteins and protein abundance factor (PAF) values as previously described [37a, 40g, 40i, 40j, 47]. We define a protein's PAF value as the total number of non-redundant spectra that correlated significantly to each cognate protein, normalized by the molecular weight of the protein ($x10^4$) [40e, 40j]. We counted MS/MS spectra representing the +2 and +3 charge states of the same peptide sequence as two distinct spectra when calculating the PAF values. PAF values were used to semi-quantify the relative abundance of each identified protein from the TAP experiments to identify statistically significant protein-protein interactions [40j].

***Significance Analysis of Microarrays* (*SAM*) to identify protein-protein interactions**

To identify statistically significant protein interactions from TAP-tagged yeast strains, we employed *Significance Analysis of Microarrays* (*SAM*) and the list of mass spectrometry-identified proteins and their PAF values [48]. Each TAP-tagged translation protein was independently purified with TAP and analyzed with MudPIT in replicate (n = 2-7). As controls, we performed TAP on the untagged parental yeast strain followed by MudPIT analysis in replicate (n = 15) using the identical protocol performed on the targeted TAP-tagged yeast strains. To measure the strength of the relationship between each TAP bait and identified prey protein, *SAM* analysis was run comparing the identified proteins' PAF data from replicate TAP-tagged bait experiments to control experiments using 1000 permutations and a two-class unpaired test with the Wilcoxon test statistic [48].

Cross-validation and identification of statistically significant protein-protein interactions

To cross-validate the most statistically significant protein-protein interactions, especially the unexpected interactions, *SAM* analysis was combined with a second, independent computational approach *Significance Analysis of INTeractome* (*SAINT*) [32a] (**Fig. S3**). *SAINT* was designed to assign confidence scores to protein-protein interaction data generated from protein affinity-purifications coupled with mass spectrometry analysis [32a].

In our cross-validation approach to identify statistically significant protein-protein interactions, an independent analysis of the *Sequest* output was first performed using the *De-Noise* algorithm [49]. The *De-Noise* algorithm maximizes the number of correct peptides identified at a 5% false discovery while reducing the

number of incorrect peptide identifications [49]. Protein assembly and spectral counting of the peptides identified with *De-Noise* were performed using *ProteoIQ* v2.6 (Premier Biosoft) to derive updated PAF values for the identified proteins. Second, to significantly increase the stringency of both the *SAM* and *SAINT* analysis, we dramatically expanded the number of control samples (**Fig. S4**). In addition to comparing TAP-tagged strains to the untagged parent strain, we reasoned that a TAP-tagged strain could also be compared to results from unrelated TAP bait proteins. For example, strains with TAP-tagged elongation and termination factors can serve as negative controls for strains with TAP-tagged initiation factors. Third, the identified proteins and their updated PAF values from replicate experiments were imported and processed independently using *SAM* and *SAINT*. For the cross-validated approach, the *SAM* analysis was performed comparing each TAP-tagged translation factor to the expanded number of controls using 1000 permutations and a two-class unpaired test with the Wilcoxon test statistic. For *SAM*, the d(i) value, which is like the '*t*' value from a *t*-test, was converted to a *p*-value using the *R* programming language for statistical computing and a one-tailed test. For the *SAINT* analysis, we used previously described settings and the normalized spectra count data (PAF) from the replicate TAP bait and the expanded control data sets [50].

Because of the difference in the metrics used by *SAM* and *SAINT* to score the bait-prey pairs, the *p*-value scores derived from the *SAM and SAINT* scores were normalized using the formula shown below, where X_{ij} is the bait-prey score calculated by *SAM* or *SAINT* and ${}_{min}X_{ij}$ and ${}_{max}X_{ij}$ are the maximum and minimum of bait-prey scores from the replicate TAP-bait experiment.

$$X_{norm(ij)} = (X_{ij} - {}_{min}X_{ij}) / ({}_{max}X_{ij} - {}_{min}X_{ij})$$

The interaction pairs were then ranked by averaging the normalized scores from *SAM* and *SAINT* for the replicate experimental data sets. The average score for each bait-prey pair was converted to a corresponding *z*-score using statistical information from the data population created from the merged *SAM* and *SAINT* output. For the bait-prey pairs in a TAP experiment, we made three assumptions: 1) there are numerous nonspecific bait-prey interactions that are generated at each step of a TAP experiment; 2) the nonspecific or false bait-prey pairs follow a normal distribution; 3) the true bait-prey pairs should not fall in a normal distribution and can be identified as outliers that are statistically distinguished from the nonspecific bait-prey pairs with a high confidence level (99% or 95%). Using the *R* statistical software package, a standard *z*-score distribution was generated to identify the bait-prey outliers and bait-prey interactions at the >99% (outlier) and 95% confidence intervals.

Identification of phosphorylated residues

To identify phosphorylated amino residues in the identified proteins, we first screened the MS-MS analysis of each TAP-purified translation complex for loss of phosphoric acid from the precursor ions from. From the MudPIT mass spectrometry analysis of the purified yeast translation complexes, the five most intense fragmentation ions acquired in each MS/MS spectra were screened for a neutral *m/z* loss of either 98 (z=+1), 49 (z=+2), or 32.67 (z=+3) from the precursor ion (**Fig. S5** and **Fig. 7**). Tandem mass spectra with at least one of these motifs were re-analyzed using the *SEQUEST* algorithm assuming differential phosphorylation (+80 Da) on either serine, threonine, or tyrosine residues and a yeast protein database of translation factors derived from the *S. cerevisiae* protein database. Tandem mass

spectra significantly correlating to fully tryptic phosphopeptides were manually evaluated

Results and Discussion

Purification of Eukaryotic Translation Factor Complexes

To identify both expected and novel protein-protein interactions involving known eukaryotic mRNA translation factors, we performed a systematic, replicate affinity purification and tandem mass spectrometry analysis on TAP-tagged *S. cerevisiae* proteins implicated in eukaryotic translation initiation, elongation, or termination, or previously shown to interact with the yeast ribosome (**Fig. 2A, B, C; Fig. S1;** and **Table S1**). In addition to validating the current models of the mRNA translation initiation, elongation, and termination protein complexes, we hypothesized that unexpected protein-protein interactions would be identified that would be the starting point for new functional, mechanistic, and structural studies of eukaryotic mRNA translation. Our ultimate goal is a comprehensive functional and mechanistic understanding of the essential process of eukaryotic mRNA translation, including identification of all the protein factors and posttranslational modifications involved.

Using a library of *S. cerevisiae* TAP strains, we targeted the canonical eukaryotic translation initiation, elongation, termination proteins and *S. cerevisiae* proteins previously shown to interact with the yeast ribosome (**Fig. 2A, B, C** and **Table S1**). The TAP-tagged strains were grown and harvested under identical conditions in biological replicates (see **Materials and Methods**). An equivalent number of yeast cells from each strain was used for the purification of the TAP-tagged bait protein and

its associated prey proteins under nondenaturing, relaxed stringency [31]. Each targeted yeast strain was broken open using glass beads (see **Materials and Methods**) and low speed centrifugation was used to remove unlysed cells from the crude lysates. The partially cleared extracts were used for the first affinity capture with IgG-coupled beads (see **Materials and Methods**). A TEV protease solution was next used to release the TAP-tagged complexes from the IgG column under nondenaturing conditions. The TEV-released protein flow through was mixed with calmodulin-coupled beads for a second affinity capture. The captured protein complexes were released from the calmodulin column using an EDTA solution [31]. As a negative control to identify nonspecific bait-prey interactions, we used fifteen replicates (n=15) of an untagged, control yeast strain prepared in parallel with the TAP-tagged translation gene yeast strains.

Identification of Purified Proteins

To identify and quantify the purified proteins, 85% of the final TAP-purified protein complexes released from the calmodulin column were analyzed using multidimensional microcapillary liquid chromatography coupled with tandem mass spectrometry (MudPIT) (**Fig. S1**). Each TAP-purified complex was trypsin-digested and loaded directly onto a triphasic microcapillary liquid chromatography (LC) column and fractionated using a five step MudPIT gradient to obtain both precursor and MS/MS fragmentation data on selected precursor ions [37a]. The mass spectrometry data were computationally compared to a forward and decoy *S. cerevisiae* protein database using the *SEQUEST* algorithm to identify significant tryptic peptide hits at a 5% false discovery rate [51]. The identified peptides were assembled into proteins as

described in earlier studies [37b, 40a, 40b, 40c-n]. To measure the abundance of the purified proteins, a protein abundance factor (PAF) was calculated for each identified protein as previously described [40g, 40i, 40j]. Our PAF quantification approach relies on the direct relationship between a protein's abundance in the sample and the frequency its peptides are selected for MS/MS analysis normalized by the molecular weight of the protein [40i, 40j]. In multiple, independent published studies, we have successfully used PAFs to identify unexpected interactions in yeast transcription factor complexes and ribosome complexes [37b, 40e, 40g, 40i, 40j, 40n, 52]. For each TAP-tagged bait translation factor and the control experiments, the output from the analysis of the MS/MS results was processed into a list of identified prey proteins and PAFs pairs, similar to the experimental data generated from a DNA microarray experiment [40j] [40j]. The paired lists of identified proteins and PAF values were used for statistical analysis to identify the most significant bait-prey protein-protein interactions.

Identification of Statistically Significant Protein-Protein Interactions in mRNA translation

We first attempted to identify significant protein-protein interactions in mRNA translation by simply comparing the PAF values of prey proteins identified from the TAP-tagged bait strains to the prey proteins identified from the control strain. For each identified protein, a relative abundance factor (RAF) was calculated by dividing the average PAF of a protein in the TAP purifications by the average PAF of same protein in the control purifications [40i, 40j]. We screened for proteins with a >2-fold enrichment. However, this initial approach did not measure the statistical significance of the protein interactions with the targeted bait genes [40i]. In addition, replicate TAP and mass spectrometry analysis of the same tagged gene occasionally produced variable results.

To address the problems of nonspecific background noise and data variability, we assumed that each prey protein in mRNA translation interacting with a bait protein would show variability with a normal distribution and a standard deviation for both the replicate TAP bait-prey and control data sets. Based on this presupposition, we initially used hierarchical clustering to identify significant protein interactions. An example of this analysis focusing on the known *S. cerevisiae* eIF2B initiation complex is shown in **Fig. 3**. eIF2B is a guanine nucleotide exchange factor and an essential eukaryotic translation initiation complex [53]. The eIF2B protein complex catalyzes the conversion of inactive eIF2-GDP to active eIF2-GTP, which mediates the binding of the initiator $tRNA_i^{Met}$ to the ribosome in a GTP-dependent manner [15b, 16]. The *S. cerevisiae* eIF2B translation initiation factor is composed of five protein subunits Gcd1p, Gcd2p, Gcd6p, Gcd7p, and Gcn3p [15b, 53]. Using the mass spectrometry-identified proteins and PAF values from replicate TAP purifications of each of the five known eIF2B components and a negative control data set, we employed hierarchical clustering to identify statistically significant proteins interacting with the eIF2B complex (**Fig. 3**). Visual analysis of clustered proteins identified the 5 known components of eIF2B (Gcd6p, Gcn3p, Gcd7p, Gcd1p, and Gcd2p) all clustering together separate from the controls. Surprisingly, a group of six unexpected proteins also clustered with the known eIF2B proteins (YBR159Wp, Cat2p, YAR010Cp/YBR012Wp, Faa1p, Arf2/1p, and Dpm1p).

While the graphical clustering experiment clearly showed the known eIF2B protein components grouping together as a complex (**Fig. 3**), we reasoned that confident identification and statistical confirmation of the unexpected proteins interacting with eIF2B required a more rigorous statistical analysis to validate the unexpected protein interactions. Therefore, we employed *Significance Analysis of*

Microarray (*SAM*) to statistically analyze the mass spectrometry data for each TAP-tagged translation protein to identify interacting proteins. *SAM* was originally developed to analyze DNA microarray data to identify genes with statistically significant changes in mRNA transcription during a response [48]. *SAM's* statistical algorithm uses repeated permutations of the data and non-parametric statistics to determine if the change in expression of a gene is significantly related to the stimulus [48]. The application employs a modified *t*-test to compare the two sets of data with variances that are close to zero [48b]. In this project, we employed *SAM* to identify statistically significant bait-prey interactions by comparing protein identifications and PAF pairs from the replicate purifications of the bait proteins to 15 replicate control experiments **(Revised Table S2)**. Our goals were first to validate existing models of the composition of the eukaryotic translation complexes (**Fig. 2** and **Revised Table S2**), and second to identify unexpected, statistically significant bait-prey protein interactions (**Revised Table S2**).

Our application of SAM to identify statistically significant protein-protein interactions using the TAP and control datasets proved to be successful. We found that the interacting proteins were specific to one of the three phases of eukaryotic protein translation (**Fig. 1** and **Revised Table S2**). The data showed no translation initiation, elongation, or termination protein factors were overlapping or shared during the three phases of protein translation (**Fig. 1 and Revised Table S2**). This result supports models showing initiation, elongation, and termination are distinct processes in translation that do not share protein factors, except for the ribosome. In addition, for the targeted eukaryotic mRNA translation initiation, elongation, and termination complexes (eIF, eEF, and eRF), our SAM analysis successfully identified the previously described eukaryotic canonical protein interactions (**Fig. 2A, B, C** and

Revised Table S2). Our application of SAM analysis also identified unexpected protein interactions (**Revised Table S2)**. As seen with the hierarchical clustering analysis, *SAM* analysis showed all five of the known eIF2B protein subunits copurifying (Gcd1p, Gcd2p, Gcd6p, Gcd7p, and Gcn3p) and revealed a group of proteins interacting at a statistically-significant level with each protein component of the eIF2B complex (**Fig. 4** and **Fig. S2**). One protein YBR159Wp significantly associated with all five subunits of the eIF2B complex. TAP purification from a YBR159W-TAP strain followed by MS and SAM analysis showed all five eIF2B components copurifying with TAP-tagged YBR159Wp (**Fig. 4** and **Fig. S2**). Based on this discovery, we named the YBR159W locus IFA38 and experimentally tested its functional role in *S. cerevisiae* translation [54]. The functional roles of the remaining unexpected proteins found interacting with eIF2B (Cat2p, YAR010Cp/YBR012Wp, Faa1p, Arf2/1p, and Dpm1p) will need to be experimentally dissected in future studies.

One of our statistically significant interactions was between the termination factor eRF1/Sup35p and the protein Rnq1p. Previous studies have shown a *rnq1* null yeast strain encodes a nonessential gene [67]. A *rnq1* null yeast strain leads to loss of the [PIN+] prion and decreased *de novo* generation of the [PSI+] prion phenotype [66]. To experimentally validate the Sup35p-Rnq1p interaction, we purified TAP-tagged Rnq1p and used Western blots with an anti-Sup35p antibody to show that Sup35p copurifies with Rnq1p (**Fig. 5**).

When examining the interacting proteins for each bait protein using *SAM* (**Revised Table S2**), we were excited to find additional unexpected and unpublished interactions with the canonical translation proteins. The five known core components of the eIF3 complex (Rpg1p, Nip1p, Prt1p, Tif35p, and Tif34p) all significantly interacted with each other as previously described (MCB 18:4935, 1998).

Interestingly, all five of the eIF3 core proteins also showed significant but lower scoring interactions with the components of the multifactor complex (MFC): Hcr1p, Tif5p, and Sui1p [55]. Furthermore, two of the five components of eIF3, Rpg1p and Prt1p, both showed statistically significant interactions with the eIF5B initiation factor Fun12p.

Refined Statistical Analysis of the MudPIT Results to Cross Validate Protein Interactions

To statistically support the unexpected protein-protein interactions in this project and improve the reliability and sensitivity to detect and identify bait-prey interactions, we developed a novel strategy that combines the outputs from different statistical approaches together to identify and cross-validate the most significant protein-protein interactions (**Fig. S3**). *Significance Analysis of Interactome* (*SAINT*) is a computational approach used to identify and analyze protein-protein interactions based on label-free quantitative spectral counts [32a, 56]. Unlike *SAM*'s calculation of the relative difference between negative control and experimental test data, *SAINT* uses the spectral count information from the experimental and control data sets to compute the probability of a true interaction using a Bayesian approach [32a]. It assumes the probabilities of spectral counts to be Poisson distributions representing either true or false bait-prey pairs. Using a Poisson distribution, *SAINT* infers the probability of a spectral count distribution for a false bait-prey interaction directly from negative controls. To estimate the probability of the spectral count for true bait-prey interaction, it uses joint modeling of the entire bait-prey association matrix built from multiple replicates of test samples. The prior probability of a true interaction is

the proportion of true interactions in the dataset. With the inferred parameters, *SAINT* calculates the average probability score of a true bait-prey pair for a given prey across all bait replicates [32a]. In this study, we merged the statistical outcomes from the *SAM* and *SAINT* algorithms to cross validate the identified protein interactions (**Fig. S3**).

To enhance our cross-validation strategy to identify the most statistically significant protein interactions, we employed four components. First, we used a more rigorous statistical algorithm *DeNoise* to validate the peptide identifications from the *Sequest* output results [49c, 57]. The *DeNoise* algorithm employs a novel strategy to identify the most statistically significant proteins from the *Sequest* output files [49c, 57]. Second, since our initial SAM analysis of the bait-prey protein interaction data detected no overlapping protein components between translation initiation, elongation, and termination phases, we greatly expanded the control data set to include results from both the untagged yeast strain and the unrelated TAP-tagged strains (**Fig. S4)**. Third, to computationally validate true bait-prey interactions, we employed a strategy that combines the outputs from both *SAM* and *SAINT* (**Fig. S3**). This combined analysis involved 3 steps: (a) normalizing of the scoring metric from *SAM* and *SAINT*, (b) identifying the overlapping bait-prey pairs between *SAM* and *SAINT*, and (c) ranking and statically validating the overlapping bait-prey pairs. The normalized data from the two approaches were merged together to cross-validate bait-prey pairs. The bait-prey pairs were then ranked by averaging the normalized scores from *SAM* and *SAINT*. Finally, we developed a statistical approach in which the average score for each bait-prey pair was converted to a corresponding *z*-score using statistical information from the data population created from the whole output of *SAM* and *SAINT*. In doing so, we made three assumptions: 1) there are numerous

nonspecific bait-prey interactions that are generated and reported due to systematic errors in each step of a TAP experiment; 2) the nonspecific bait-prey pairs follow a normal distribution; 3) the true bait-prey pairs should not fall in the nonspecific bait-prey pair distribution and can be identified as outliers and statistically distinguished from nonspecific bait-prey pairs with high confidence. With the converted *z*-score, a percentile value for each bait-prey pair could be calculated, and the confidence level for the right tail could be determined (e.g. outliers at 99% and 95% confidence levels) (**Revised Table S3**).

We used the strategy merging *SAM* and *SAINT* to help verify the statistically significant interactions of the unexpected and unpublished proteins with the canonical translation factors initially identified with SAM **(Fig. 6)**. We identified a complex of seven proteins (Sea4p/YBL104Cp, Rtc1p/YOL138Cp, Mtc5p/YDR128Wp, Iml1p/YJR138Wp, Seh1p/YGL1001p, Rmd11p/YHL023Cp, Npr2p/YEL062Wp) that our statistical validation strategy shows interacting with Gcd11p of the conserved eIF2 translation initiation complex. Current models show eIF2 having multiple functions during initiation, including recruitment of the initiator Met-tRNA to the 40S subunit and selection of the AUG start codon [58]. Second, we identified the interaction of Rny1p with the eIF4A/Tif1p translation initiation factor. Interestingly, Rny1p is predicted to function as a nonspecific endoribonuclease (RNase) [59]. Our *SAM-SAINT* cross-validation computational strategy statistically supported many of the unexpected protein interactions initially identified using SAM **(Fig. 6**). In total, we identified 126 proteins that have novel, unexpected interactions with the translation machinery. Our cross-validation analysis statistically supported 52 of the unexpected protein interactions as significant. The statistical analysis strongly suggests that these protein interactions would be the strongest candidate genes to initially select

for functional and mechanistic experimental validation. Overall, the data and results from this study are expected to be the foundation for new studies that dissect the functional and mechanistic roles of the unexpected protein interactions in eukaryotic protein synthesis.

Identification of Phosphoproteins Associated with Translation

Reversible protein phosphorylation is major posttranslational modification regulating eukaryotic translation and signal transduction pathways [3a]. To identify phosphorylated amino acids among the purified translation proteins, we re-analyzed the acquired tandem mass spectrometry data using the SEQUEST algorithm assuming differential phosphorylation (+80 Da) on serine, threonine, or tyrosine residues using a protein database of translation factors derived from the *S. cerevisiae* ORF database **(Fig. S5** and **Fig. 7**). Tandem mass spectra significantly correlating (Cn>2) with fully tryptic phosphopeptide sequences were manually evaluated. For a *Sequest*-identified phosphopeptide to be accepted as "true," the MS/MS spectra of the peptide were required to have a phosphoric acid neutral loss ion of either 98 (z=+1), 49 (z=+2), or 32.67 (z=+3) from the precursor ion m/z value as one of the 5 most intense ions in the MS/MS spectrum. In addition, the majority of the major fragment ions in the MS/MS spectrum had to be identified as either a b- or y-fragment ion **(Fig. S5** and **Fig. 7**). Using these criteria, we identified unexpected phosphorylation sites for 27 translation initiation proteins (**Table S4**). The functional role of the identified phosphorylation sites in global or transcript-specific translation activity is unknown. Future experiments will be needed to identify the phenotype and/or functional role of the modified amino acids in translational activity.

Conclusion

Protein-protein interactions are the physical contact of high specificity between two or more protein molecules [24]. Mapping a protein's interactions provides potentially new insights into its biochemical functions and regulation [25]. The discovery of novel or unexpected protein interactions enables putative functional or biochemical roles to be assigned to previously uncharacterized proteins in a biological process. In this study, we performed an in-depth proteomic analysis on *S. cerevisiae* genes previously shown to be involved in eukaryotic mRNA translation to identify both expected and unexpected, unpublished protein interactions and phosphorylated amino acid residues. Using replicate tandem affinity purification and tandem mass spectrometry analysis of targeted genes previously shown to be involved in mRNA translation initiation, elongation, and termination combined with novel computational strategies, we identified both previously published interactions and unexpected, unpublished protein interactions and phosphorylated amino acids. The data and results from this study are expected to be the foundation for new studies that dissect the functional, regulatory, and mechanistic roles of the unexpected protein interactions and phosphorylated amino acids in eukaryotic mRNA translation.

ACKNOWLEDGEMENTS

We thank Elizabeth M. Link for critical comments in the preparation of this manuscript. This project was supported by NIH grants GM64779. K.J.M. was supported by NIH grants ES11993 and GM64779. C.M.W. and J.L.J. are supported by NIH grants GM64779 and HL68744. D.T.D. was supported by NIH grant ES11993. T.C.F. is supported by ACS postdoctoral fellowship PF-05-148-01-MBC. A.R.F is supported by NIH training grant T32 CA009385. V.R.G. is supported by NIH grant GM64779. A.J.L. was supported by NIH grants GM64779, HL68744, ES11993, and CA098131. In addition, K.J.M. and A.J.L. were supported in part with federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN266200400079C/N01-AI-40079.

Figure Legends

Fig. 1. Canonical phases of eukaryotic mRNA translation. Shown are the 40S and 60S ribosomal subunits interacting with a typical eukaryotic 5'-capped and 3'-polyA mRNA during the three phases of eukaryotic translation, resulting in the synthesis of an mRNA-encoded polypeptide. In the initiation phase of cap-dependent initiation, the 40S subunit interacts with the 5' end of the mRNA and scans the 5' UTR until it recognizes the AUG start codon. Alternatively, in cap-independent translation initiation, the 40S subunit interacts directly with the start codon, typically in the context of an internal ribosome entry site (IRES). After the AUG start codon is recognized, the 60S subunit joins the 40S subunit to form the 80S ribosome to start the elongation phase. The 80S ribosome then begins to synthesize an mRNA-encoded polypeptide, covalently linking amino acids until it reaches the stop codon. In the termination phase, the stop codon of the mRNA's protein coding sequence enters the 80S ribosome. The completed polypeptide is hydrolyzed from the peptidyl tRNA and released from the 80S ribosome. The 80S ribosome separates into the 40S and 60S subunits, and the mRNA and deacylated tRNA are released. For the three phases of eukaryotic translation, separate networks of initiation, elongation, and termination protein factors transiently interact with the mRNA, aminoacyl tRNAs, and ribosomal subunits (see Figure 2A, 2B, 2C).

Fig. 1

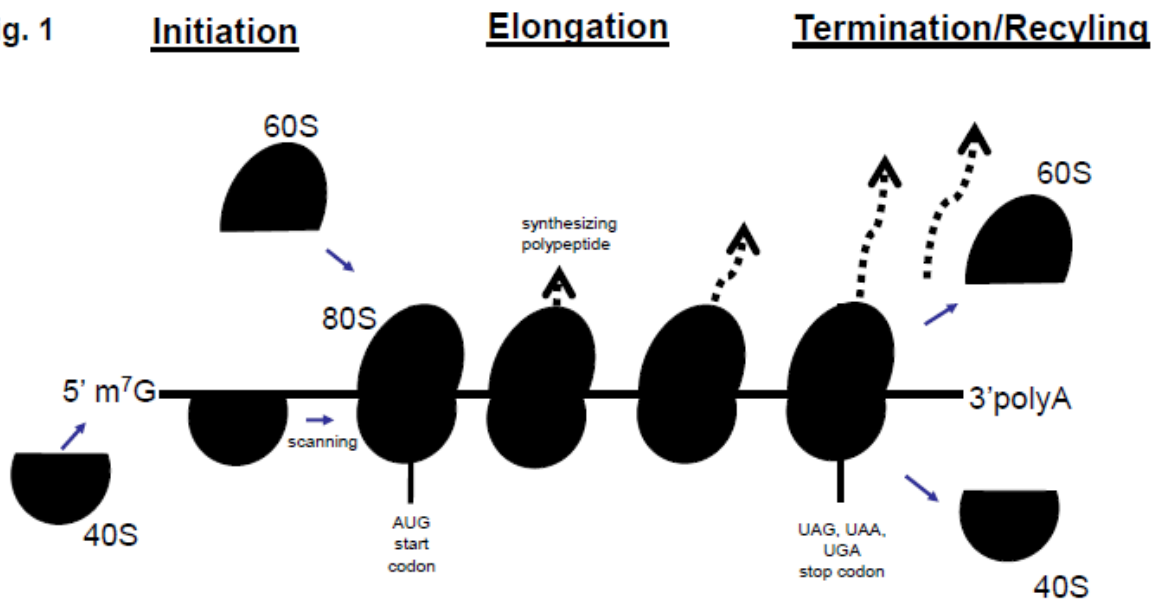


Fig. 2A, B, C. Canonical translation factors and interactions during eukaryotic mRNA translation initiation, elongation, and termination. For the underlined *S. cerevisiae* proteins in Fig. 2 and the *S. cerevisiae* genes listed in Table S1, TAP-tagged yeast strains were either created or obtained and used for TAP to affinity purify the tagged protein and interacting proteins. The mass spectrometry approach 2-D LC-MS/MS (MudPIT) was used to purify, analyze, identify, and quantify the expected and unexpected peptides and proteins interacting with each targeted TAP-tagged protein. **Fig. 2A.** The canonical yeast translation initiation factors and their previously identified protein interactions. **Fig 2B:** The canonical yeast translation elongation factors and their previously identified protein interactions. **Fig 2C:** The canonical yeast translation termination factors and their previously identified protein interactions.

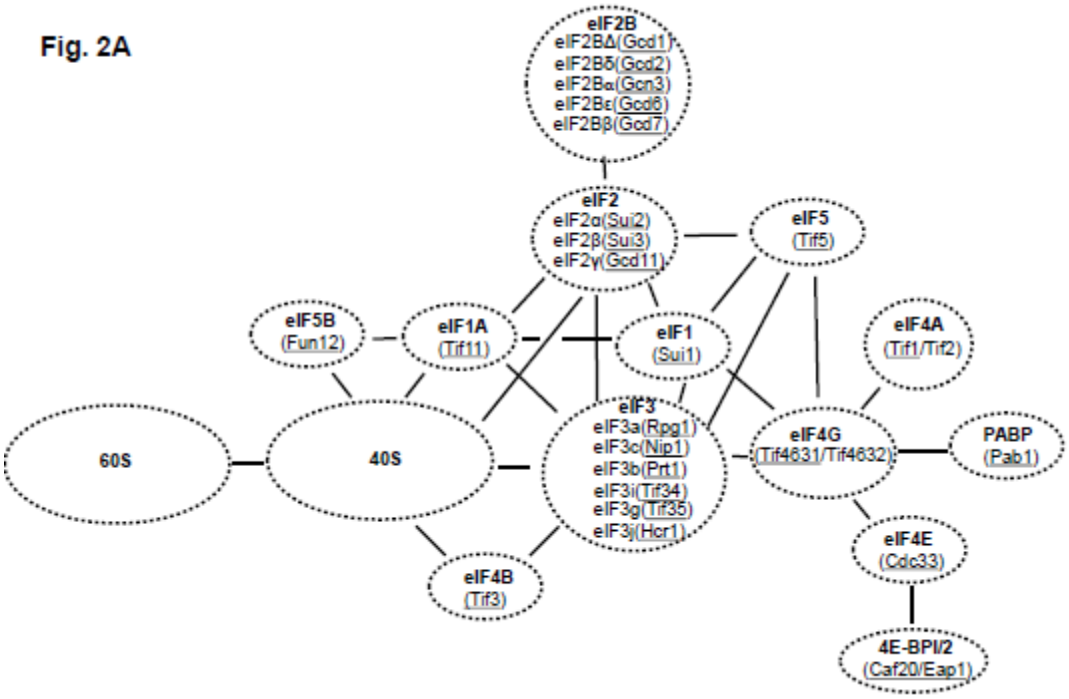


Fig 2B

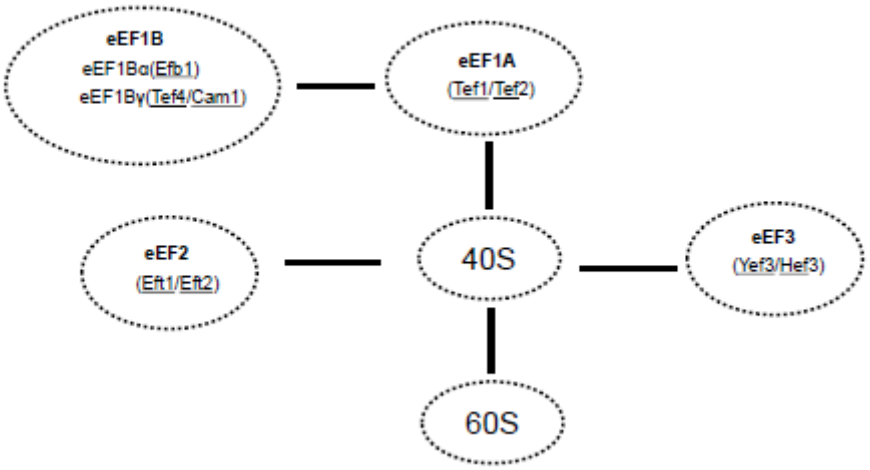


Fig 2C

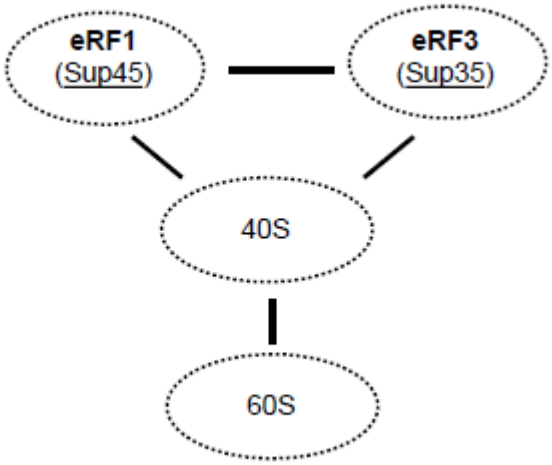


Fig. 3. Hierarchical clustering of MudPIT identified and quantified proteins purified from eIF2B TAP strains showing the unexpected *S. cerevisiae* proteins Cat2p, Tef4p, Faa1p, Dpm1p, Arf2/1p, YBR159Wp, and YAR010Cp/YBR012W-Ap copurifying with the eIF2B translation initiation complex.

Fig 3

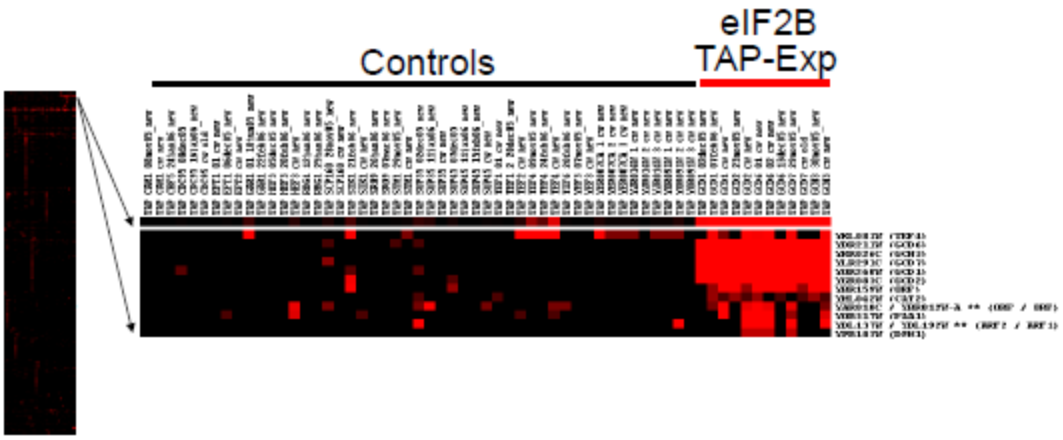


Fig. 4. *SAM* analysis of the TAP-tagged proteins in the *S. cerevisiae* eIF2B initiation factor complex. Each of the TAP-tagged components of the eIF2B complex (Gcd2p, Gcd6p, Gcn3p, Gcd1p, and Gcd7p) and untagged control yeast strains were grown and purified in replicate, and the proteins were identified and quantified using MudPIT (see Materials and Methods). *Statistical Analysis of Microarrays* (*SAM*) was then used to compare the list of identified proteins and PAF pairs for each eIF2B TAP-tagged protein to the results from 15 replicate control experiments (See Material and Methods). The top panel **(A)** shows the SAM results for the TAP-Gcd6p protein. The *SAM* graphical outputs show the expected eIF2B proteins all copurifying with Gcd6p along with YBR159Wp. To confirm these results, a TAP-tagged YBR159W yeast strain was grown, processed, and analyzed using the identical protocol. The bottom panel **(B)** shows all five eIF2B protein components copurifying with the TAP-tagged YBR159Wp protein. As supporting evidence, **Fig S2** shows the SAM results for the other canonical components of the eIF2B complex (Gcd2p, Gcn3, Gcd1 and Gcd7) and their interaction with YBR159Wp.

Revised Fig 4

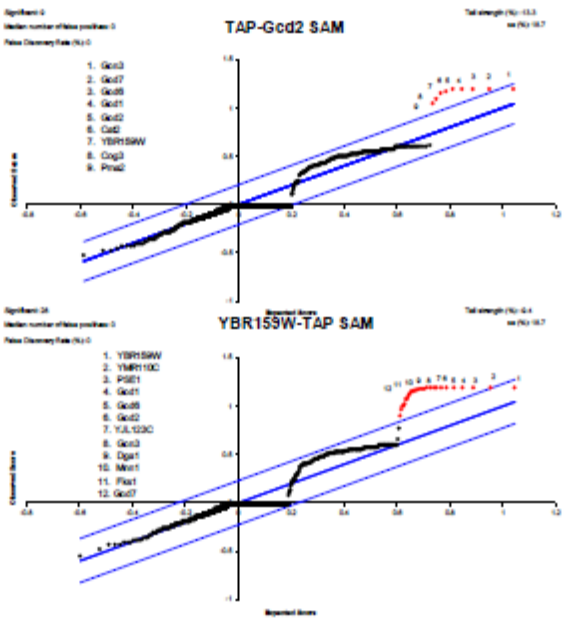


Fig. 5. TAP-Western blot validating the interaction of the yeast Rnq1-TAP complex with Sup35p. TAP was performed using a TAP-tagged Rnq1 and untagged control yeast strains up to the IgG affinity step. The IgG-captured proteins were eluted by cleavage with the TEV protease; the. calmodulin affinity purification was not done. The eluted proteins were separated by SDS-PAGE, electroblotted to PVDF, and probed with commercial rabbit polyclonal antibodies against *S. cerevisiae* Sup35p (ABNOVA). The CTRL-TAP lane contains eluted proteins from an untagged yeast strain prepared and run side-by-side with the Rnq1-TAP protein sample.

Fig 5

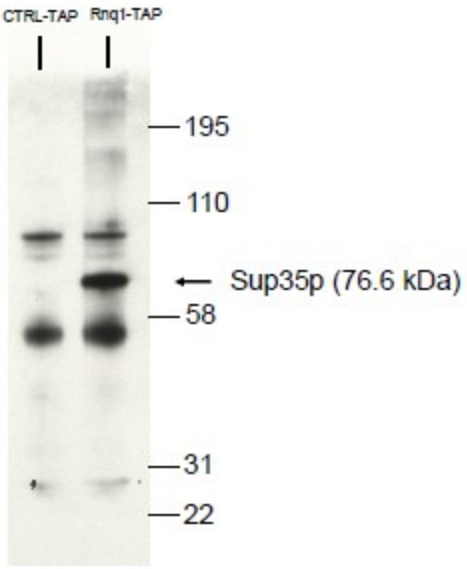


Fig. 6. Unexpected protein-protein interactions identified for canonical *S. cerevisiae* mRNA translation initiation, elongation, and termination protein complexes using the TAP-MudPIT strategy (**Fig. S1**). * indicates that the interaction was identified both with *SAM* and the *SAM-SAINT* cross-validation algorithms.

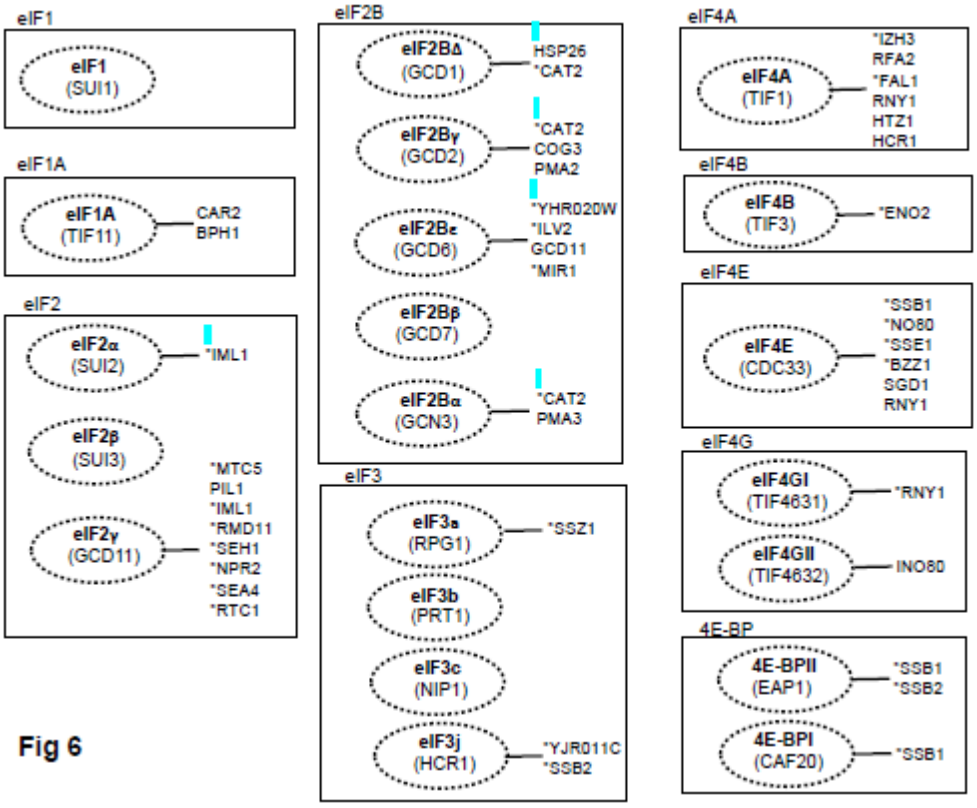


Fig 6

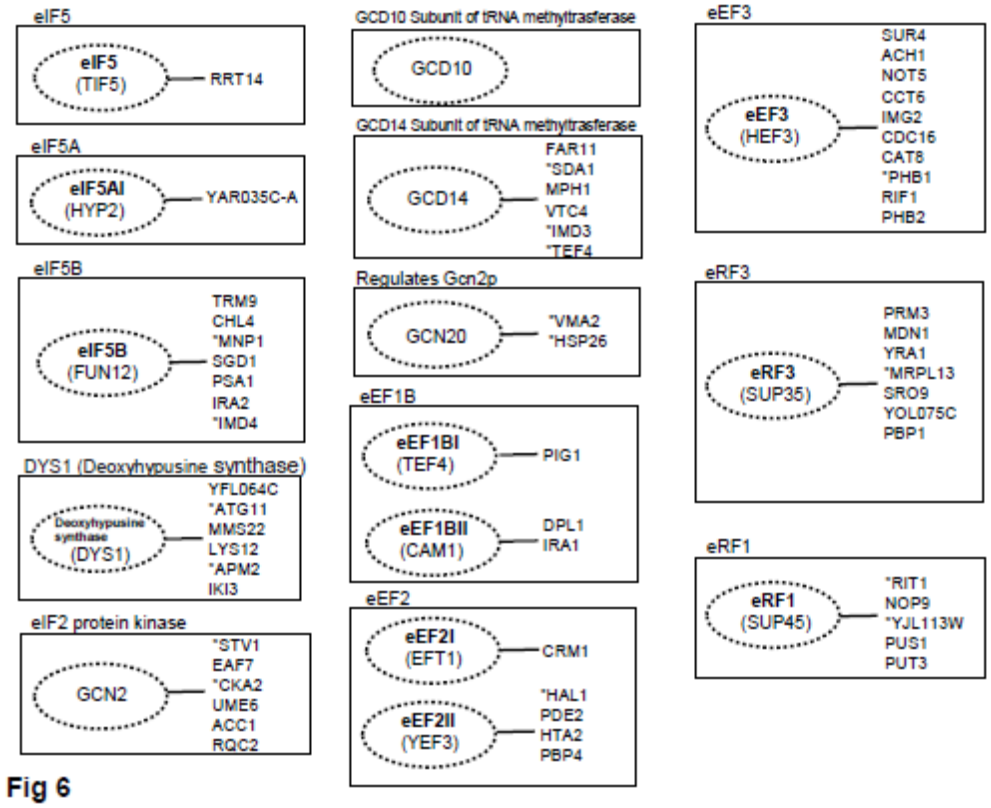


Fig 6

Fig 6

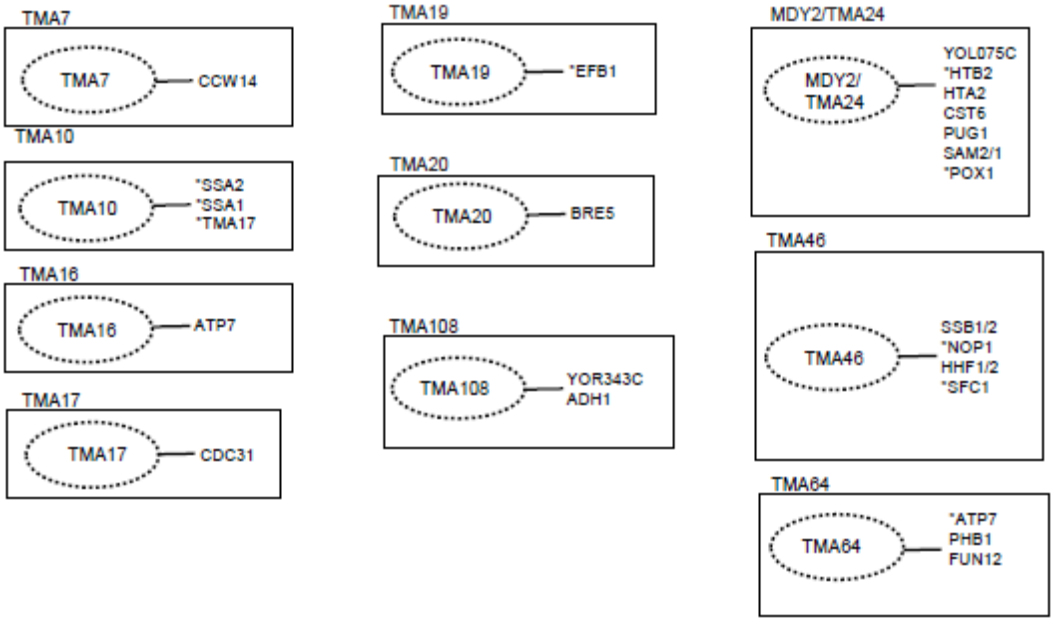
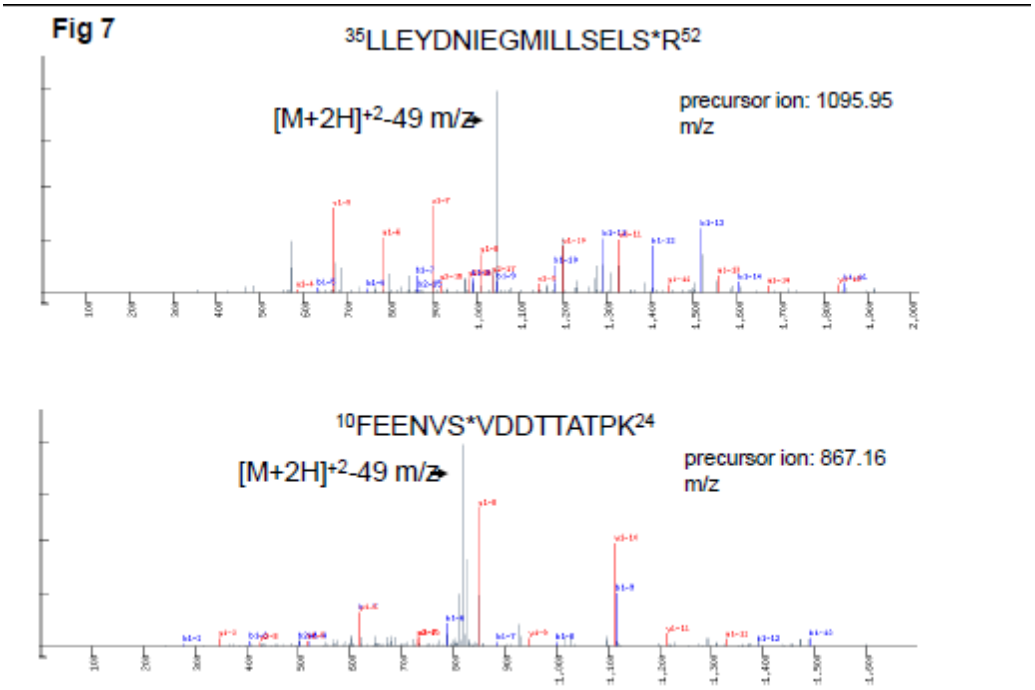


Fig. 7. Tandem mass spectrometry identification of phosphorylated peptides from the purified and trypsin-digested *S. cerevisiae* TAP mRNA translation initiation complexes. The dominant fragment ion seen in each MS/MS spectrum is the neutral loss peak (+98) caused by the loss of phosphoric acid from the selected precursor ion. (A) Tandem mass spectrum of a phosphorylated Sui2p tryptic peptide from the purified eIF2 complex. (B) Tandem mass spectrum of phosphorylated Cdc33p tryptic peptide from the purified eIF4E complex.



Bibliography

[1] a) H. Jung, C. E. Holt, *Wiley interdisciplinary reviews. RNA* 2011, 2, 153; b) W. Tadros, H. D. Lipshitz, *Dev Dyn* 2005, 232, 593; c) C. H. de Moor, H. Meijer, S. Lissenden, *Semin Cell Dev Biol* 2005, 16, 49.

[2] a) T. Pan, *Annu Rev Genet* 2013, 47, 121; b) B. Lemaitre, S. E. Girardin, *Nat Rev Microbiol* 2013, 11, 365; c) M. J. Clemens, *Semin Cell Dev Biol* 2005, 16, 13; d) A. G. Hinnebusch, *Annu Rev Microbiol* 2005, 59, 407; e) M. Holcik, N. Sonenberg, *Nat Rev Mol Cell Biol* 2005, 6, 318.

[3] a) P. P. Roux, I. Topisirovic, *Cold Spring Harbor perspectives in biology* 2012, 4; b) C. Tiedje, H. Holtmann, M. Gaestel, *Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research* 2014, 34, 220; c) M. Showkat, M. A. Beigh, K. I. Andrabi, *Molecular biology international* 2014, 2014, 686984; d) A. T. Parsa, E. C. Holland, *Trends Mol Med* 2004, 10, 607; e) M. Shmulevitz, P. Marcato, P. W. Lee, *Oncogene* 2005, 24, 7720.

[4] M. A. Sutton, E. M. Schuman, *J Neurobiol* 2005, 64, 116.

[5] a) A. Dalet, E. Gatti, P. Pierre, *FEBS Lett* 2015, 589, 1539; b) D. Li, T. Wei, C. M. Abbott, D. Harrich, *Microbiol Mol Biol Rev* 2013, 77, 253; c) I. Mohr, N. Sonenberg, *Cell Host Microbe* 2012, 12, 470; d) S. W. Lindemann, A. S. Weyrich, G. A. Zimmerman, *Trends Cardiovasc Med* 2005, 15, 9; e) L. Beretta, *Int Rev Immunol* 2004, 23, 347.

[6] a) S. Kuersten, E. B. Goodwin, *Nat Rev Genet* 2003, 4, 626; b) F. Gebauer, M. W. Hentze, *Nat Rev Mol Cell Biol* 2004, 5, 827.

[7] a) K. Garber, K. T. Smith, D. Reines, S. T. Warren, *Curr Opin Genet Dev* 2006, 16, 270; b) P. W. Vanderklish, G. M. Edelman, *Genes Brain Behav* 2005, 4, 360; c) R. B. Denman, N. Dolzhanskaya, Y. J. Sung, *Cell Mol Life Sci* 2004, 61, 1714.

[8] a) J. Pelletier, J. Graff, D. Ruggero, N. Sonenberg, *Cancer Res* 2015, 75, 250; b) R. Spilka, C. Ernst, A. K. Mehta, J. Haybaeck, *Cancer letters* 2013, 340, 9; c) A. A. Farooqi, A. Naqi, M. Z. Qureshi, A. Rana, A. Khan, A. M. Riaz, S. M. Afzal, N. Rasheed, S. Bhatti, *Cell biochemistry and function* 2012, 30, 611; d) M. A. Bjornsti, P. J. Houghton, *Cancer Cell* 2004, 5, 519; e) D. Ruggero, N. Sonenberg, *Oncogene* 2005, 24, 7426; f) G. C. Scheper, M. S. van der Knaap, C. G. Proud, *Nat Rev Genet* 2007, 8, 711; g) L. Liu-Yesucevitz, G. J. Bassell, A. D. Gitler, A. C. Hart, E. Klann, J. D. Richter, S. T. Warren, B. Wolozin, *J Neurosci* 2011, 31, 16086; h) E. T. Wang, J. M. Taliaferro, J. A. Lee, I. P. Sudhakaran, W. Rossoll, C. Gross, K. R. Moss, G. J. Bassell, *J Neurosci* 2016, 36, 11418.

[9] a) T. Honjo, Y. Nishizuka, O. Hayaishi, *J Biol Chem* 1968, 243, 3553; b) A. Gradi, Y. V. Svitkin, H. Imataka, N. Sonenberg, *Proc Natl Acad Sci U S A* 1998, 95, 11089; c) M. Lopez-Lastra, P. Ramdohr, A. Letelier, M. Vallejos, J. Vera-Otarola, F. Valiente-Echeverria, *Rev Med Virol* 2010, 20, 177.

[10] a) F. Meric, K. K. Hunt, *Mol Cancer Ther* 2002, 1, 971; b) C. F. Calkhoven, C. Muller, A. Leutz, *Trends Mol Med* 2002, 8, 577; c) A. R. Tee, J. Blenis, *Semin Cell Dev Biol* 2005, 16, 29; d) E. Diaz-Guerra, R. Vernal, W. Cantero, E. W. Mullner, J. A. Garcia-Sanz, *Infect Disord Drug Targets* 2008, 8, 252; e) R. Cencic, F. Robert, J. Pelletier, *Methods Enzymol* 2007, 431, 269; f) I. Tanaka, M. Sato, T. Kato, D. Goto, T. Kakumu, A. Miyazawa, N. Yogo, T. Hase, M. Morise, Y. Sekido, L. Girard, J. D. Minna, L. A. Byers, J. V. Heymach, K. R. Coombes, M. Kondo, Y. Hasegawa, *Cancer science* 2018, 109, 1843; g) Y. Zang, X. Zhang, L. Yan, G. Gu, D. Li, Y. Zhang, L. Fang, S. Fu, J. Ren, Z. Xu, *Journal of Cancer* 2017, 8, 3049; h) N. T. Zaveri, *Journal of medicinal chemistry* 2016, 59, 7011; i) J. C. Rojas, A. L. Boxer, *Nat Rev Neurol* 2016, 12, 74; j) S. Li, J. Fu, C. Lu, M. Y. Mapara, S. Raza, U. Hengst, S. Lentzsch, *Mol Cancer Ther* 2016, 15, 711; k) W. D. Dietrich, H. M. Bramlett, *Brain research* 2016, 1640, 94; l) A. D'Andrea, I. Gritti, P. Nicoli, M. Giorgio, M. Doni, A. Conti, V. Bianchi, L. Casoli, A. Sabo, A. Mironov, G. V. Beznoussenko, B. Amati, *Oncotarget* 2016, 7, 72415; m) S. Karaki, C. Andrieu, H. Ziouziou, P. Rocchi, *Advances in protein chemistry and structural biology* 2015, 101, 1; n) C. Demosthenous, J. J. Han, M. J. Stenson, M. J. Maurer, L. E. Wellik, B. Link, K. Hege, A. Dogan, E. Sotomayor, T. Witzig, M. Gupta, *Oncotarget* 2015, 6, 9488; o) F. Pettersson, S. V. Del Rincon, W. H. Miller, Jr., *Expert opinion on therapeutic targets* 2014, 18, 1035; p) S. Bandyopadhyay, J. T. Rogers, *Biochem Pharmacol* 2014, 88, 486; q) M. Grzmil, B. A. Hemmings, *Cancer Res* 2012, 72, 3891.

[11] a) L. D. Kapp, J. R. Lorsch, *Annu Rev Biochem* 2004, 73, 657; b) M. B. Mathews, N. Sonenberg, J. W. Hershey, Eds., in *Translational Control in Biology and Medicine*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY 2007; c) N. Sonenberg, A. G. Hinnebusch, *Cell* 2009, 136, 731; d) R. J. Jackson, C. U. Hellen, T. V. Pestova, *Nat Rev Mol Cell Biol* 2010, 11, 113; e) C. E. Aitken, J. R. Lorsch, *Nat Struct Mol Biol* 2012, 19, 568; f) A. Prabhakar, J. Choi, J. Wang, A. Petrov, J. D. Puglisi, *Protein Sci* 2017, 26, 1352.

[12] a) N. Takahashi, M. Yanagida, S. Fujiyama, T. Hayano, T. Isobe, *Mass Spectrom Rev* 2003, 22, 287; b) E. Thomson, S. Ferreira-Cerca, E. Hurt, *J Cell Sci* 2013, 126, 4815; c) D. Kressler, G. Bange, Y. Ogawa, G. Stjepanovic, B. Bradatsch, D. Pratte, S. Amlacher, D. Strauss, Y. Yoneda, J. Katahira, I. Sinning, E. Hurt, *Science* 2012, 338, 666.

[13] *The Ribosome*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY 2001.

[14] a) A. Verschoor, J. R. Warner, S. Srivastava, R. A. Grassucci, J. Frank, *Nucleic Acids Res* 1998, 26, 655; b) C. M. Spahn, R. Beckmann, N. Eswar, P. A. Penczek, A. Sali, G. Blobel, J. Frank, *Cell* 2001, 107, 373; c) A. Ben-Shem, N. Garreau de Loubresse, S. Melnikov, L. Jenner, G. Yusupova, M. Yusupov, *Science* 2011, 334, 1524; d) A. Ben-Shem, L. Jenner, G. Yusupova, M. Yusupov, *Science* 2010, 330, 1203.

[15] a) A. G. Hinnebusch, J. R. Lorsch, *Cold Spring Harbor perspectives in biology* 2012, 4; b) A. G. Hinnebusch, *Annu Rev Biochem* 2014, 83, 779.

[16] A. G. Hinnebusch, *Trends Biochem Sci* 2017, 42, 589.

[17] T. V. Pestova, C. U. Hellen, *Cold Spring Harb Symp Quant Biol* 2001, 66, 389.

[18] C. Schmidt, V. Beilsten-Edmands, C. V. Robinson, *J Mol Biol* 2015, DOI: 10.1016/j.jmb.2015.10.011.

[19] M. Kozak, *Gene* 2002, 299, 1.

[20] a) L. Valasek, K. H. Nielsen, A. G. Hinnebusch, *Embo J* 2002, 21, 5886; b) N. Sonenberg, T. E. Dever, *Curr Opin Struct Biol* 2003, 13, 56.

[21] a) A. A. Komar, M. Hatzoglou, *J Biol Chem* 2005, 280, 23425; b) S. D. Baird, M. Turcotte, R. G. Korneluk, M. Holcik, *Rna* 2006, 12, 1755.

[22] T. E. Dever, R. Green, *Cold Spring Harbor perspectives in biology* 2012, 4, a013706.

[23] a) R. J. Jackson, C. U. Hellen, T. V. Pestova, *Advances in protein chemistry and structural biology* 2012, 86, 45; b) C. U. Hellen, *Biochim Biophys Acta* 2009, 1789, 558.

[24] a) J. De Las Rivas, C. Fontanillo, *PLoS Comput Biol* 2010, 6, e1000807; b) S. Yadav, J. Liu, T. M. Scherer, Y. Gokarn, B. Demeule, S. Kanai, J. D. Andya, S. J. Shire, *Biophysical reviews* 2013, 5, 121.

[25] a) E. Golemis, Ed. in *Protein-protein interactions: a molecular cloning manual*, Cold Spring Harbor Laboratory Press, Woodbury, NY 2002; b) A. L. Barabasi, Z. N. Oltvai, *Nat Rev Genet* 2004, 5, 101; c) C. Herzberg, L. A. Weidinger, B. Dorrbecker, S. Hubner, J. Stulke, F. M. Commichau, *Proteomics* 2007, 7, 4032; d) D. Bergsma, S. Chen, J. Buchweitz, R. Gerszten, B. B. Haab, *Mol Cell Proteomics* 2010, 9, 446; e) K. Maclagan, R. Tommasi, E. Laurine, C. Prodromou, P. C. Driscoll, L. H. Pearl, S. Reich, R. Savva, *Future medicinal chemistry* 2011, 3, 271; f) N. Nishida, I. Shimada, *Methods Mol Biol* 2012, 757, 129; g) E. Socher, B. Imperiali, *Chembiochem* 2013, 14, 53; h) A. H. Smits, M. Vermeulen, *Trends Biotechnol* 2016, 34, 825; i) M. Frenkel-Morgenstern, A. Gorohovski, S. Tagore, V. Sekar, M. Vazquez, A. Valencia, *Nucleic Acids Res* 2017, 45, 7094; j) M. Kotlyar, A. E. M. Rossos, I. Jurisica, *Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]* 2017, 60, 8 2 1; k) M. M. Gromiha, K. Yugandhar, S. Jemimah, *Curr Opin Struct Biol* 2017, 44, 31; l) D. Vella, I. Zoppis, G. Mauri, P. Mauri, D. Di Silvestre, *EURASIP journal on bioinformatics & systems biology* 2017, 2017, 6; m) X. Peng, J. Wang, W. Peng, F. X. Wu, Y. Pan, *Brief Bioinform* 2017, 18, 798; n) K. Titeca, I. Lemmens, J. Tavernier, S. Eyckerman, *Mass Spectrom Rev* 2018, DOI: 10.1002/mas.21574.

[26] L. Jin, W. Wang, G. Fang, *Annu Rev Pharmacol Toxicol* 2014, 54, 435.

[27] a) Y. Ho, A. Gruhler, A. Heilbut, G. D. Bader, L. Moore, S. L. Adams, A. Millar, P. Taylor, K. Bennett, K. Boutilier, L. Yang, C. Wolting, I. Donaldson, S. Schandorff, J. Shewnarane, M. Vo, J. Taggart, M. Goudreault, B. Muskat, C. Alfarano, D. Dewar, Z. Lin, K. Michalickova, A. R. Willems, H. Sassi, P. A. Nielsen, K. J. Rasmussen, J. R. Andersen, L. E. Johansen, L. H. Hansen, H. Jespersen, A. Podtelejnikov, E. Nielsen, J. Crawford, V. Poulsen, B. D. Sorensen, J. Matthiesen, R. C. Hendrickson, F. Gleeson, T. Pawson, M. F. Moran, D. Durocher, M. Mann, C. W. Hogue, D. Figeys, M. Tyers, *Nature* 2002, 415, 180; b) N. J. Krogan, W. T. Peng, G. Cagney, M. D. Robinson, R. Haw, G. Zhong, X. Guo, X. Zhang, V. Canadien, D. P. Richards, B.

K. Beattie, A. Lalev, W. Zhang, A. P. Davierwala, S. Mnaimneh, A. Starostine, A. P. Tikuisis, J. Grigull, N. Datta, J. E. Bray, T. R. Hughes, A. Emili, J. F. Greenblatt, *Mol Cell* 2004, 13, 225; c) N. J. Krogan, G. Cagney, H. Yu, G. Zhong, X. Guo, A. Ignatchenko, J. Li, S. Pu, N. Datta, A. P. Tikuisis, T. Punna, J. M. Peregrin-Alvarez, M. Shales, X. Zhang, M. Davey, M. D. Robinson, A. Paccanaro, J. E. Bray, A. Sheung, B. Beattie, D. P. Richards, V. Canadien, A. Lalev, F. Mena, P. Wong, A. Starostine, M. M. Canete, J. Vlasblom, S. Wu, C. Orsi, S. R. Collins, S. Chandran, R. Haw, J. J. Rilstone, K. Gandi, N. J. Thompson, G. Musso, P. St Onge, S. Ghanny, M. H. Lam, G. Butland, A. M. Altaf-Ul, S. Kanaya, A. Shilatifard, E. O'Shea, J. S. Weissman, C. J. Ingles, T. R. Hughes, J. Parkinson, M. Gerstein, S. J. Wodak, A. Emili, J. F. Greenblatt, *Nature* 2006, 440, 637; d) A. C. Gavin, P. Aloy, P. Grandi, R. Krause, M. Boesche, M. Marzioch, C. Rau, L. J. Jensen, S. Bastuck, B. Dumpelfeld, A. Edelmann, M. A. Heurtier, V. Hoffman, C. Hoefert, K. Klein, M. Hudak, A. M. Michon, M. Schelder, M. Schirle, M. Remor, T. Rudi, S. Hooper, A. Bauer, T. Bouwmeester, G. Casari, G. Drewes, G. Neubauer, J. M. Rick, B. Kuster, P. Bork, R. B. Russell, G. Superti-Furga, *Nature* 2006; e) A. Breitkreutz, H. Choi, J. R. Sharom, L. Boucher, V. Neduva, B. Larsen, Z. Y. Lin, B. J. Breitkreutz, C. Stark, G. Liu, J. Ahn, D. Dewar-Darch, T. Reguly, X. Tang, R. Almeida, Z. S. Qin, T. Pawson, A. C. Gingras, A. I. Nesvizhskii, M. Tyers, *Science* 2010, 328, 1043; f) Y. V. Miteva, H. G. Budayeva, I. M. Cristea, *Anal Chem* 2013, 85, 749.

[28] a) B. Schwikowski, P. Uetz, S. Fields, *Nat Biotechnol* 2000, 18, 1257; b) T. Ito, K. Tashiro, S. Muta, R. Ozawa, T. Chiba, M. Nishizawa, K. Yamamoto, S. Kuhara, Y. Sakaki, *Proc Natl Acad Sci U S A* 2000, 97, 1143.

[29] a) I. Xenarios, L. Salwinski, X. J. Duan, P. Higney, S. M. Kim, D. Eisenberg, *Nucleic Acids Res* 2002, 30, 303; b) P. Pagel, S. Kovac, M. Oesterheld, B. Brauner, I. Dunger-Kaltenbach, G. Frishman, C. Montrone, P. Mark, V. Stumpflen, H. W. Mewes, A. Ruepp, D. Frishman, *Bioinformatics* 2005, 21, 832; c) D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. P. Tsafou, M. Kuhn, P. Bork, L. J. Jensen, C. von Mering, *Nucleic Acids Res* 2015, 43, D447.

[30] a) C. Stark, B. J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, M. Tyers, *Nucleic Acids Res* 2006, 34, D535; b) A. Chatr-Aryamontri, R. Oughtred, L. Boucher, J. Rust, C. Chang, N. K. Kolas, L. O'Donnell, S. Oster, C. Theesfeld, A. Sellam, C. Stark, B. J. Breitkreutz, K. Dolinski, M. Tyers, *Nucleic Acids Res* 2017, 45, D369.

[31] G. Rigaut, A. Shevchenko, B. Rutz, M. Wilm, M. Mann, B. Seraphin, *Nat Biotechnol* 1999, 17, 1030.

[32] a) H. Choi, B. Larsen, Z. Y. Lin, A. Breitkreutz, D. Mellacheruvu, D. Fermin, Z. S. Qin, M. Tyers, A. C. Gingras, A. I. Nesvizhskii, *Nat Methods* 2011, 8, 70; b) A. I. Nesvizhskii, *Proteomics* 2012, 12, 1639.

[33] D. Mellacheruvu, Z. Wright, A. L. Couzens, J. P. Lambert, N. A. St-Denis, T. Li, Y. V. Miteva, S. Hauri, M. E. Sardiu, T. Y. Low, V. A. Halim, R. D. Bagshaw, N. C. Hubner, A. Al-Hakim, A. Bouchard, D. Faubert, D. Fermin, W. H. Dunham, M. Goudreault, Z. Y. Lin, B. G. Badillo, T. Pawson, D. Durocher, B. Coulombe, R. Aebersold, G. Superti-Furga, J. Colinge, A. J. Heck, H.

Choi, M. Gstaiger, S. Mohammed, I. M. Cristea, K. L. Bennett, M. P. Washburn, B. Raught, R. M. Ewing, A. C. Gingras, A. I. Nesvizhskii, *Nat Methods* 2013, 10, 730.

[34] H. Choi, S. Kim, D. Fermin, C. C. Tsou, A. I. Nesvizhskii, *J Proteomics* 2015, 129, 121.

[35] R. I. Sadreyev, M. Tang, B. H. Kim, N. V. Grishin, *Nucleic Acids Res* 2009, 37, W90.

[36] A. T. Kong, F. V. Leprevost, D. M. Avtonomov, D. Mellacheruvu, A. I. Nesvizhskii, *Nat Methods* 2017, 14, 513.

[37] a) A. J. Link, J. Eng, D. M. Schieltz, E. Carmack, G. J. Mize, D. R. Morris, B. M. Garvik, J. R. Yates, 3rd, *Nat Biotechnol* 1999, 17, 676; b) V. R. Gerbasi, C. M. Weaver, S. Hill, D. B. Friedman, A. J. Link, *Mol Cell Biol* 2004, 24, 8276.

[38] a) S. Ghaemmaghami, W. K. Huh, K. Bower, R. W. Howson, A. Belle, N. Dephoure, E. K. O'Shea, J. S. Weissman, *Nature* 2003, 425, 737; b) M. Altmann, P. Linder, *J Biol Chem* 2010, 285, 31907.

[39] J. K. Eng, A. L. McCormack, I. Yates, J. R., *J. Am. Soc. Mass Spectrom.* 1994, 5, 976.

[40] a) S. L. Sanders, J. Jennings, A. Canutescu, A. J. Link, P. A. Weil, *Mol Cell Biol* 2002, 22, 4723; b) S. L. Squazzo, P. J. Costa, D. L. Lindstrom, K. E. Kumer, R. Simic, J. L. Jennings, A. J. Link, K. M. Arndt, G. A. Hartzog, *Embo J* 2002, 21, 1764; c) M. S. Kobor, S. Venkatasubrahmanyam, M. D. Meneghini, J. W. Gin, J. L. Jennings, A. J. Link, H. D. Madhani, J. Rine, *PLoS Biol* 2004, 2, E131; d) A. J. Link, J. L. Jennings, M. P. Washburn, *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]* 2004, Chapter 23, Unit 23 1; e) D. W. Powell, C. M. Weaver, J. L. Jennings, K. J. McAfee, Y. He, P. A. Weil, A. J. Link, *Mol Cell Biol* 2004, 24, 7249; f) E. Sattlegger, M. J. Swanson, E. A. Ashcraft, J. L. Jennings, R. A. Fekete, A. J. Link, A. G. Hinnebusch, *J Biol Chem* 2004, 279, 29952; g) A. J. Link, T. C. Fleischer, C. M. Weaver, V. R. Gerbasi, J. L. Jennings, *Methods* 2005, 35, 274; h) G. Lominadze, D. W. Powell, G. C. Luerman, A. J. Link, R. A. Ward, K. R. McLeish, *Mol Cell Proteomics* 2005, 4, 1503; i) T. C. Fleischer, C. M. Weaver, K. J. McAfee, J. L. Jennings, A. J. Link, *Genes Dev* 2006, 20, 1294; j) K. J. McAfee, D. T. Duncan, M. Assink, A. J. Link, *Mol Cell Proteomics* 2006, 5, 1497; k) V. R. Gerbasi, A. J. Link, *Mol Cell Proteomics* 2007, 6, 1049; l) D. R. Arnett, J. L. Jennings, D. L. Tabb, A. J. Link, P. A. Weil, *Mol Cell Proteomics* 2008, 7, 2090; m) M. A. Sammons, A. K. Antons, M. Bendjennat, B. Udd, R. Krahe, A. J. Link, *PLoS One* 2010, 5, e9301; n) A. R. Farley, D. W. Powell, C. M. Weaver, J. L. Jennings, A. J. Link, *J Proteome Res* 2011, 10, 1481.

[41] a) B. M. Sefton, *Curr Protoc Cell Biol* 2001, Chapter 14, Unit 14 1; b) A. Paradela, J. P. Albar, *J Proteome Res* 2008, 7, 1809; c) A. Derouiche, C. Cousin, I. Mijakovic, *Curr Opin Biotechnol* 2012, 23, 585; d) S. J. Humphrey, D. E. James, M. Mann, *Trends Endocrinol Metab* 2015, 26, 676; e) F. Ardito, M. Giuliani, D. Perrone, G. Troiano, L. Lo Muzio, *International journal of molecular medicine* 2017, 40, 271.

[42] P. P. Roux, I. Topisirovic, *Mol Cell Biol* 2018, 38.

[43] D. C. Amberg, D. J. Burke, J. N. Strathern, *Methods in Yeast Genetics 2005 Edition*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY 2005.

[44] a) J. J. Tasto, R. H. Carnahan, W. H. McDonald, K. L. Gould, *Yeast* 2001, 18, 657; b) H. Liang, C. VanValkenburgh, X. Chen, C. Mullins, L. Van Kaer, N. Green, H. Fang, *J Biol Chem* 2003, 278, 50932.

[45] M. P. Washburn, D. Wolters, J. R. Yates, 3rd, *Nat Biotechnol* 2001, 19, 242.

[46] R. G. Sadygov, J. Eng, E. Durr, A. Saraf, H. McDonald, M. J. MacCoss, J. R. Yates, 3rd, *J Proteome Res* 2002, 1, 211.

[47] P. W. Lewis, E. L. Beall, T. C. Fleischer, D. Georlette, A. J. Link, M. R. Botchan, *Genes Dev* 2004, 18, 2929.

[48] a) M. B. Eisen, P. T. Spellman, P. O. Brown, D. Botstein, *Proc Natl Acad Sci U S A* 1998, 95, 14863; b) V. G. Tusher, R. Tibshirani, G. Chu, *Proc Natl Acad Sci U S A* 2001, 98, 5116.

[49] a) L. Jian, X. Niu, Z. Xia, P. Samir, C. Sumanasekera, Z. Mu, J. L. Jennings, K. L. Hoek, T. Allos, L. M. Howard, K. M. Edwards, P. A. Weil, A. J. Link, *J Proteome Res* 2013, 12, 1108; b) L. Jian, Z. Xia, X. Niu, X. Liang, P. Samir, A. Link, *IEEE/ACM transactions on computational biology and bioinformatics / IEEE, ACM* 2015, DOI: 10.1109/TCBB.2015.2480084; c) L. Jian, Z. Xia, X. Niu, X. Liang, P. Samir, A. J. Link, *IEEE/ACM transactions on computational biology and bioinformatics / IEEE, ACM* 2016, 13, 804.

[50] H. Choi, G. Liu, D. Mellacheruvu, M. Tyers, A. C. Gingras, A. I. Nesvizhskii, *Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]* 2012, Chapter 8, Unit8 15.

[51] J. R. Yates, 3rd, E. Carmack, L. Hays, A. J. Link, J. K. Eng, *Methods Mol Biol* 1999, 112, 553.

[52] a) A. R. Farley, A. J. Link, *Methods Enzymol* 2009, 463, 725; b) A. J. Link, C. Weaver, A. Farley, *Cold Spring Harb Protoc* 2011, 2011, pdb prot5607.

[53] a) G. D. Pavitt, K. V. Ramaiah, S. R. Kimball, A. G. Hinnebusch, *Genes Dev* 1998, 12, 514; b) G. D. Pavitt, *Biochem Soc Trans* 2005, 33, 1487.

[54] C. M. Browne, P. Samir, J. S. Fites, S. A. Villarreal, A. J. Link, *Mol Cell Biol* 2013, 33, 1041.

[55] a) K. Asano, J. Clayton, A. Shalev, A. G. Hinnebusch, *Genes Dev* 2000, 14, 2534; b) K. Asano, L. Phan, L. Valasek, L. W. Schoenfeld, A. Shalev, J. Clayton, K. Nielsen, T. F. Donahue, A. G. Hinnebusch, *Cold Spring Harb Symp Quant Biol* 2001, 66, 403.

[56] D. V. Skarra, M. Goudreault, H. Choi, M. Mullin, A. I. Nesvizhskii, A. C. Gingras, R. E. Honkanen, *Proteomics* 2011, 11, 1508.

[57] X. Liang, Z. Xia, X. Niu, A. J. Link, L. Pang, F. X. Wu, H. Zhang, *Proteome Sci* 2013, 11, S10.

[58] a) M. D. Jennings, G. D. Pavitt, *Cell Cycle* 2014, 13, 2660; b) R. Cagnetta, H. H. Wong, C. K. Frese, G. R. Mallucci, J. Krijgsveld, C. E. Holt, *Mol Cell* 2019, 73, 474; c) C. de la Parra, A.

Ernlund, A. Alard, K. Ruggles, B. Ueberheide, R. J. Schneider, *Nature communications* 2018, 9, 3068.

[59] G. C. MacIntosh, P. A. Bariola, E. Newbigin, P. J. Green, *Proc Natl Acad Sci U S A* 2001, 98, 1018.

[60] P. Uetz, L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Goodwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, S. Fields, J. M. Rothberg, *Nature* 2000, 403, 623.

[61] A. Malina, J. R. Mills, J. Pelletier, *Cold Spring Harb Perspect Biol* 2012, 4(4), a012377

[62] U. Sahin, K. Kariko, O. Türeci, *Nat Rev Drug Discov* 2014, 3(10), 759.

[63] F. DeRosa, L. Smith, Y. Shen, Y. Huang, J. Pan, H. Xie, B. Yahalom, M. W. Heartlein, *Mol Ther* 2019, 27(4), 878.

[64] J. C. Kaczmarek, A. K. Patel, K. J. Kauffman, O. S. Fenton, M. J. Webber, M. W.Heartlein, F. DeRosa, D. G. Anderson, *Angew Chem Int Ed Engl* 2016, 55, 13808.

[65] L. Liu-Yesucevitz, G. J. Bassell, A. D. Gitler, A. C. Hart, E. Klann, J. D. Richter, S. T. Warren, B. Wolozin, *J. of Neuroscience* 2011, 31(45), 16086.

[66] K.M. Keefer, K. C. Stein, H. L. True, *Scientific Reports* 2017, 7(1), 5853.

[67] G. Giaever *et al*., *Nature* 2002, 418(6896), 387.

"We require a statement of significance of the study highlighting the importance and - where appropriate - the functional significance of the research results (max. 200 words). Please insert this section between the Abstract and the Introduction of your paper."

Statement of Significance:

mRNA translation is the essential biological process in which ribosomes and amino acid-charged tRNAs decode mRNAs to synthesize polypeptides. We performed a large-scale, systematic tandem affinity purification and mass spectrometry analysis on *S. cerevisiae* translation initiation, elongation, termination, and ribosome-associated proteins to identify unexpected protein interactions and phosphorylated amino acids. The identified novel protein interactions and phosphorylation sites involving translation initiation, elongation, termination, and ribosome-associated proteins are expected to drive future functional, mechanistic, and structural studies to dissect their roles in eukaryotic mRNA translation.