capCLIP: a new tool to probe protein synthesis in human cells through capture and identification of the eIF4E-mRNA interactome.

Kirk B. Jensen,[1,2,*] B. Kate Dredge,[3] John Toubia,[3,4] Xin Jin,[1,5] Valentina Iadevaia,[6] Gregory J. Goodall[3,7] and Christopher G. Proud[1,2]

[1]Lifelong Health, South Australian Health & Medical Research Institute, North Terrace, Adelaide 5000, Australia
[2]School of Biological Sciences, Faculty of Sciences, University of Adelaide, Adelaide 5005, Australia
[3]Centre for Cancer Biology, SA Pathology and University of South Australia, Adelaide, SA 5000, Australia
[4]ACRF Cancer Genomics Facility, Centre for Cancer Biology, SA Pathology and University of South Australia, Frome Road, Adelaide SA 5000, Australia
[5]School of Medicine and Pharmacy, Ocean University of China, 5 Yushan Road, Qingdao, China;
[6]School of Biosciences and Medicine, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom.
[7]Adelaide Medical School, Faculty of Health and Medical Sciences, University of Adelaide, Adelaide 5005, Australia

* Corresponding author: telephone +61 8 8128 4908; kirk.jensen@sahmri.com

Abstract

Translation initiation of eukaryotic mRNAs commences with recognition of the m^7G cap by the cap-binding protein eIF4E and the subsequent recruitment of additional translation initiation factors to the mRNA's 5′ end. eIF4E's essential role in translation suggests that the cellular eIF4E-mRNA interactome (or 'eIF4E cap-ome') may serve as a faithful proxy of cellular translational activity. Here we describe capCLIP, a novel method to systematically capture and quantify the eIF4E cap-ome. To validate capCLIP, we identified the eIF4E cap-omes in human cells ± the partial mTORC1 inhibitor rapamycin. TOP (terminal oligopyrimidine) mRNA representation is systematically reduced in rapamycin-treated cells, consistent with the well-known effect of mTORC1 inhibition on translation of these mRNAs. capCLIP tag data also permits the identification of a refined, 7-nucleotide consensus motif for TOP mRNAs (5′-CUYUYYC-3′). In addition, we apply capCLIP to probe the consequences of phosphorylation of eIF4E, a modification whose function had remained unclear. Phosphorylation of eIF4E drives an overall reduction in eIF4E-mRNA association, and strikingly, the mRNAs most sensitive to phosphorylation possess short 5′ UTRs. capCLIP provides a sensitive and comprehensive measure of cellular translational activity. We foresee its widespread use as a high-throughput means for assessing translation in contexts not amenable to existing methodologies.

Introduction

The control of protein synthesis, or mRNA translation, is a key regulatory mechanism employed by eukaryotic cells to control the composition of the cellular proteome (1). The translational control mechanisms used by mammalian cells are very diverse and offer the ability to regulate specific mRNAs or subsets of mRNAs. Such mechanisms act primarily at the initiation phase of protein synthesis, where ribosomes are recruited to mRNAs (1). eIF4E (eukaryotic initiation factor 4E) is of central importance both for the process of translation initiation and for its regulation (2). eIF4E binds to the 7-methyl-GTP (m^7GTP) moiety of the so-called 5′-cap structure, which is present on all cellular cytoplasmic mRNAs. This interaction is crucial for the efficient translation of almost all mRNAs (3). Cap-bound eIF4E associates with the scaffold protein eIF4G, which in turn interacts with other translation initiation factors, including the oligomeric factor eIF3. This eIF4E-eIF4G-eIF3 complex then recruits 40S ribosomal subunits to the 5′ end of the mRNA, permitting scanning of 40S-associated initiation complexes to the AUG, 60S subunit joining, and the subsequent translation of the message (4). While haplo-insufficient eIF4E mice are viable, $eIF4E^{+/-}$ mouse embryonic fibroblasts (MEFs) are resistant to oncogenic transformation (5). This data indicates that cellular translational activity is especially sensitive to the amount of eIF4E protein, and importantly also suggests that there exists a direct correlation between eIF4E protein levels and translational activity within the cell.

As the cap-eIF4E interaction is essential for cap-dependent translation of mRNA, it seems reasonable that identification of the mRNAs bound to eIF4E (the eIF4E 'cap-ome') under specific cellular conditions should function as a 'readout' of cellular translational activity. To test this idea, we developed capCLIP, which provides a methodologically unique means to comprehensively measure mRNA translational activity cell-wide. capCLIP complements existing high-throughput methodologies in this area, such as ribosome profiling. In addition, capCLIP also provides unprecedented and previously inaccessible insights into individual eIF4E-mRNA interactions. Thus, capCLIP is also ideal for probing cellular regulatory pathways which influence, either directly or indirectly, the binding of eIF4E to mRNA. As capCLIP measures eIF4E binding to individual cellular mRNAs, it can provide a range of detailed insights into how cellular perturbations which affect eIF4E-mRNA binding shape translation. In this report, we describe the capCLIP method in full, and describe two distinct experimental applications of the methodology as demonstrations of capCLIP's potential to provide new insights into of the cellular mechanisms of translational control.

capCLIP is a novel cross-linking immunoprecipitation (CLIP)-based methodology (6,7) that takes advantage of the observation that eIF4E can be photo-crosslinked to the m^7G cap of a mRNA (8). Thus, UV irradiation of living cells (or in principle, tissue) can be used to covalently link eIF4E to the cap of bound mRNAs. The *in-vivo* covalent linkage of eIF4E with an individual capped mRNA species permits the stringent co-purification of the linked RNA fragment though the immunoprecipitation and subsequent Bis-Tris PAGE purification of the eIF4E protein, and allows the selective identification of the eIF4E-mRNA binding events which occur within the living cell.

We first use the capCLIP methodology to identify the eIF4E cap-omes in Hela cells with or without treatment with rapamycin, a partial inhibitor of mTORC1 (mechanistic target of rapamycin, complex 1). mTORC1 is a multi-subunit protein kinase that is activated by amino acids, hormones, growth factors and other trophic stimuli, and is a major regulator of translation initiation and elongation and of ribosome biogenesis (9). A major consequence of mTORC1 inhibition by rapamycin or by drugs which act as direct inhibitors of mTOR's kinase activity (known as 'mTOR-KIs') is the marked impairment

of the synthesis of ribosomal proteins (10), which are encoded by mRNAs that possess a 5′-terminal tract of oligopyrimidines (5′-TOP) (11). While the precise features of a functional TOP sequence remain to be established, there is general experimental agreement that TOP mRNA transcripts commence with an invariant C nucleotide, followed by an uninterrupted sequence of 4-15 pyrimidines (11). Furthermore, there is considerable understanding of the translational consequences of mTORC1 signalling, notably through the application of ribosome footprinting/profiling, a set of techniques that can determine the distribution and density of ribosomes across the open reading-frames (ORFs) of individual mRNAs (12). As such techniques are now widely used for estimating the translational activity of individual mRNAs, determining how the eIF4E cap-ome responds to rapamycin will allow us to compare the two methodologies, and to evaluate how well capCLIP data performs as a quantitative readout of cellular translational activity.

Using capCLIP ± rapamycin treatment, we identify in Hela cells an eIF4E cap-ome of approximately 3,500 unique mRNAs, of which 86 are significantly depleted from the eIF4E cap-ome of rapamycin-treated cells. As expected, the majority of these mRNAs (62, or 72%) are TOP mRNAs. The fold-reduction in binding between eIF4E and these 86 mRNAs, and the fact that the majority of mRNAs exhibiting this magnitude of loss contain TOPs, suggest that capCLIP does indeed provide a readout of cellular translational activity. Furthermore, using capCLIP and Hela CAGE (cap analysis of gene expression) data, we mapped the transcription start sites (TSS) of 52 TOP mRNAs. This in turn permitted us to conduct a completely unique analysis where to probe, in depth, the relationship between a TOP motif's 'strength' (the fold-change in eIF4E-mRNA binding caused by rapamycin) and its sequence, and in particular where the TOP motif lies with regard to the the TSS. Our analysis indicates that mRNAs exhibiting strong TOP functionality all share a highly-conserved, seven pyrimidine nucleotide consensus sequence, which is consistent with the emerging model of TOP mRNA regulation by the protein LARP1. As capCLIP provides the means to link a mRNA's 5′ end sequence to its eIF4E binding behaviour, it provides us with a completely novel means to determine the influence of 5′ end sequence on a mRNA's translational activity. We expect that the ability to link 5′ end sequence with mRNA function may prove valuable in a number of additional experimental contexts.

Secondly, we also examine the suitability of capCLIP to reveal new insights into areas of translational regulation that are *not* well understood. One application of the method, for which it is particularly suitable, is to understand the translational and biological consequences of phosphorylation of eIF4E itself. eIF4E undergoes a single phosphorylation event on residue Ser209, catalysed by the MNK kinases (MAP kinase interacting kinases). The MNKs are serine/threonine kinases which are activated downstream of the Ras/MAPK signalling pathway and are directly activated by p38 MAP kinase and/or ERK. The two mammalian MNKs, MNK1 and MNK2 (*MKNK1* and *MKNK2* are the human genes) each give rise to two splice forms, which in turn produce differing protein C-termini that affect the subcellular distribution of the corresponding MNK proteins and their association with upstream MAP kinases. Both mouse and human MNK2 are constitutively active and phosphorylate eIF4E under basal conditions, while MNK1 is strongly activated by upstream MAPK signalling and therefore plays a much greater role in inducible phosphorylation of eIF4E (reviewed in (13)). eIF4E is the only known substrate of the MNKs *in vivo* (13).

While a number of studies have shown an association between elevated P-eIF4E levels and cancer (13), and while eIF4E$^{S209A/S209A}$ knock-in mice are greatly deficient for cellular transformation in a mouse PTEN KO model of high-grade prostate intraepithelial neoplasia (PIN), the translational consequences

of eIF4E Ser209 phosphorylation remain unclear (14). While a small number of candidate mRNAs have been identified whose translation and/or polysomal association appears to increase with P-eIF4E, we understand neither the mechanistic basis for these selective translational differences nor the biological consequences of the changes in synthesis of specific proteins as they relate to upstream Ras/ERK signalling. Using the highly-selective MNK inhibitor eFT-508 (15) to block the phosphorylation of eIF4E in human cells, we employed capCLIP to identify the cellular cap-omes in cells with no detectable phospho-eIF4E and in cells maximally stimulated for eIF4E phosphorylation. Intriguingly, there is an average ~1.7 fold *decrease* in eIF4E-mRNA binding across the entire ~5,000 unique mRNAs of the eIF4E cap-ome upon eIF4E phosphorylation. Of these mRNAs, 256 exhibit a statistically significant reduction in binding to eIF4E, and we suggest approaches to validate the reduction in their translation implied by this data. We also show that those mRNAs which are most strongly decreased upon eIF4E phosphorylation possess significantly shorter 5′ UTRs than average. Our data as a whole is completely consistent with *in vitro* measurements showing that phosphorylation of eIF4E lowers its affinity for capped mRNA and cap analogues (16-18); however, a mechanistic explanation for the preferential loss of mRNAs with short 5′ UTRs will require further study.

In sum, we show that eIF4E capCLIP provides a novel and successful pathway to characterise cap-dependent translational activity in human cells, and can be used to comprehensively and quantitatively measure the impact of drug inhibition of two distinct cellular signalling systems that impact the translational machinery. We predict that the method will see widespread use solely for its ability to provide comprehensive data on the translational output of cells, and potentially in tissue. An additional and particularly promising application of the method is in probing the functional role(s) of other cellular cap-binding proteins. For example, eIF4E2, a related cap-binding protein, is widely expressed, yet its function is still poorly understood. capCLIP could provide a deep understanding of the eIF4E2 cap-ome, and can show exactly how its cap-ome changes in response to perturbations in the cellular environment.

Materials and Methods

CRISPR/Cas9-based gene editing

To introduce a 3X flag epitope tag at the N-terminus of the endogenous eIF4E gene in HeLa cells, an sgRNA guide sequence targeting the first exon of human eIF4E was designed, with the Cas9 endonuclease cut site lying 3 bases upstream of the AUG of the eIF4E open-reading frame (scheme is shown in Figure 1B). A 200 nucleotide asymmetric, single-stranded homology-directed repair (HDR) DNA template (19) containing two modified phosphorthioate linkages at each end (20) was synthesised to permit the introduction of a 3X-flag/1X-myc epitope tag at the extreme N-terminus of the eIF4E protein. The HDR template, along with a GeneArt CRISPR Nuclease vector (Life Technologies) expressing the eIF4E guide RNA, and a Cas9 nuclease-human CD4 pre-protein, was introduced into Hela cells using nucleofection (Lonza). 48 hours post-nucleofection, Cas9-expressing cells were isolated using anti-CD4 Dynabeads, and the purified cells were plated at 1 cell/well in 4 96-well TC dishes. Following approximately 7 days of growth, wells containing clonal cell colonies were replica-plated, to allow further growth of each colony and provide cells for genomic DNA analysis. To check for insertion of the epitope tag, PCR products spanning the editing region of eIF4E were screened for length and for digestion with BspHI, a restriction site provided by the HDR template. Colonies positive in these screens were cultured further, and genomic DNA was again isolated for additional PCR amplification of the 5′-region on the eIF4E coding sequence, followed by forward and reverse Sanger sequencing of the entire edited region of eIF4E. Clones passing sequencing verification were further analysed by western blot, using both anti-flag and anti-eIF4E antibodies (Figure 1C).

capCLIP

A brief account of the pilot capCLIP experiments and 'preparative' capCLIP experiments is provided here. A full account of capCLIP reagents and methodology, along with detailed commentary of aspects of the method have been included as **Supplemental Method 1**. capCLIP was initially piloted on untreated Hela cells endogenously expressing flag-eIF4E. Cells were exposed to 350 mJ/cm^2 of short-wavelength (~254 nm) UV irradiation, followed by cell lysis in 1X PXL [1X PBS supplemented with 0.1% SDS, 0.5% deoxycholate and 0.5% NP-40]. Both a 'high' [0.10 U/µl lysate] and 'low' [0.02 U/µl lysate] concentration of Ambion RNase I was tested to determine the optimal RNase I amount for generation of RNA tags of length ~30-80 nt. Anti-flag mAb bound to protein G Dynabeads was used to immunoprecipitate flag-eIF4E and crosslinked mRNAs, followed by three washes in 1X PXL and one wash using 5X PXL (1X PXL with 5X PBS). [γ-^{32}P]ATP was used to 5′ end-label the RNA molecules, followed by a final wash, elution and Bis-Tris PAGE size separation of the labelled eIF4E:RNA complexes (Fig. 2A). Based on the initial RNase concentration experiments, an RNase I concentration of 0.02 U/µl of cell lysate was used for all subsequent capCLIP experiments.

For the preparative capCLIP experiments ± rapamycin, Hela cells endogenously expressing flag-eIF4E were treated for 2 h with 100 nM rapamycin, or with the equivalent amount of DMSO vehicle, and then UV irradiated as above. For the eFT-508 capCLIP experiment, flag-eIF4E Hela cells that had been serum-starved overnight were pre-treated with either DMSO only or with 0.1 µM eFT-508 for 1 h. Cells were subsequently stimulated with 500 nM PMA for 30 min immediately prior to UV irradiation. Cell pellets for all samples were lysed in 1X PXL supplemented with 0.02 U of RNase I per µl of cell lysate and DNAse I. Cell lysates were subject to anti-flag immunoprecipitation and washed, as above. A 5′-^{32}P-labelled RNA adapter was ligated to the 3′ end of the eIF4E-crosslinked RNA 'on-bead,' followed by a final wash, elution and Bis-Tris PAGE size separation of the ^{32}P-labelled eIF4E-RNA complexes. The material resolved by Bis-Tris PAGE was electrophoretically transferred to a nitrocellulose

membrane, and the radioactive RNA detected using phosphorimaging (Fig. 2C). The ~37-60 kDa region of the membrane, corresponding to eIF4E:RNA complexes with RNAs approximately 30-80 nt in length, was excised, protease treated, and the liberated RNA precipitated. cDNA synthesis, followed by UMI (unique molecular identifier)-containing DNA adapter ligation to the 3′ end of the cDNA were conducted, and the resulting material was purified and size-selected on denaturing PAGE (Fig. 2D). Gel slices corresponding to ssDNAs of approximately 80-130 nt were cut from the gel, and the cDNA was eluted and precipitated. This DNA was subsequently amplified by PCR, and individual capCLIP samples were barcoded prior to sequencing. Following a final size-purification of the amplified PCR product, and QC checks of library concentration and length, all sample libraries were subject to high-throughput single-end sequencing on the Nextseq 500 with a read length of 80.

Bioinformatic analysis

Raw reads were adaptor trimmed and filtered for short sequences during base-calling with bcl2fastq2 (--find-adapters-with-sliding-window --adapter-stringency 0.9 --mask-short-adapter-reads 35 --minimum-trimmed-read-length 35). The resulting FASTQ files, averaging 45 million reads per sample, were analysed and quality checked using FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc) and Picard (broadinstitute.github.io/picard). The filtered reads were mapped against the human reference genome (hg19) using the STAR alignment algorithm (21) (version 2.5.3a with default parameters and --chimSegmentMin 20, --quantMode GeneCounts) returning an average unique alignment rate of 30%. UMItools (22) was used to deduplicate reads in each sample using unique molecular identifiers (UMI's) (default settings with --method adjacency --edit-distance-threshold 1). Enriched regions of the genome were identified separately for each strand with the MACS2 peak-caller (23) (version 2.1.1.20160309) using default parameters and reporting only peaks with an FDR cut-off (-q) < 0.05. The resulting peak files from each strand were merged. Differential binding analysis was performed using R (version 3.4.3) and the DiffBind package (24). Alignments were visualised and interrogated using the Integrative Genomics Viewer v2.3.80 (25).

CAGE data

The CAGE datasets used for analysis of mRNA TSSs are ENCSR000CJJ (Hela S3 cells) and ENCSR000CKD (CD20+ B cells) from the ENCODE portal (www.encodeproject.org) (26).

capCLIP sequence datasets

Raw sequencing files for the rapamycin capCLIP and eFT-508 capCLIP samples, along with processed peak, .bed, and .tsv files for each sample, have been uploaded to the Gene Expression Omnibus (GEO) repository at the NCBI under the accession number GSE138473. Public access will be provided upon final publication of the manuscript.

Antibodies For western analysis of Hela cell lysates and immunoprecipitations, the following antibodies were used: anti-β-actin, mouse mAb, Sigma Aldrich A2228; anti-flag, mouse mAb (clone M2) Sigma Aldrich F3165; anti-eIF4E, rabbit polyclonal, Cell Signaling Technology 9742; anti-P-eIF4E, rabbit polyclonal, ThermoFisher 44-528G; anti-eIF4G, rabbit polyclonal, Cell Signaling Technology 2498; anti-4E-BP1, rabbit polyclonal, Cell Signaling Technology 9452; anti-rpS6, goat polyclonal, Santa Cruz Biotechnology sc-13007; anti-P-rpS6 (Ser240/244), rabbit polyclonal, Cell Signaling Technology 2215; anti-P-ERK (Thr202/Try204), rabbit polyclonal, Cell Signaling Technology 4370; anti-FTH1, rabbit polyclonal, Cell Signaling Technology 3998; anti-ras, rabbit polyclonal, Cell Signaling Technology 3965.

Analysis of 5′ UTR length and nucleotide composition for eFT-508 capCLIP data

In total 90 mRNAs were selected: the 45 mRNAs with the greatest *reduction* in eIF4E binding upon eIF4E phosphorylation, and correspondingly, the 45 mRNAs with the greatest relative *increase* in eIF4E binding upon eIF4E phosphorylation. All selected mRNAs were subject to a minimum mRNA abundance cut-off of 2.5 to omit low-confidence mRNAs (low tag coverage) from the analysis. Ensembl Biomart and the NCBI RefSeq collection were used to source the 5′ UTR sequence of all 90 mRNAs selected. For mRNAs where the database queries returned multiple mRNA isoforms (often due to the presence of multiple TSSs and/or alternative splicing in the 5′ UTR regions of these mRNAs), eFT-508 capCLIP tag data and Hela CAGE data were viewed in IGV to determine which sequence best fit the observed capCLIP and CAGE data. When capCLIP data did not help to identify the precise pattern of 5′ UTR alternative exon usage in Hela cells, all potential 5′ UTR isoforms were included and the average 5′ UTR length used for subsequent calculations.

Results and Discussion

Capture and identification of the eIF4E-mRNA interactome

The capCLIP method is represented schematically in Figure 1A and briefly outlined in the Methods. We have also provided a Supplemental Methods document which provides a list of materials and reagents, a full capCLIP protocol, and includes detailed commentary on critical steps of the method. A description of bioinformatic analyses performed on the capCLIP sequencing data is provided in the Methods. In brief, capCLIP uses short-wavelength (~254 nm) UV irradiation of living cells to create a 'zero-order' photocrosslink between the m^7G cap of a mRNA (8) and an adjacent residue of a bound eIF4E protein. This covalent link between eIF4E and RNA is critical to the success of the method, as it stably 'captures' this *in vivo* interaction to permit the subsequent physical separation of eIF4E and photocrosslinked mRNAs away from mRNAs that are merely bound to the protein. Indeed, the ability of CLIP to identify *bona fide, in vivo* protein:RNA binding events is achieved by ensuring that all non-covalent protein:RNA binding complexes are removed during the purification of the crosslinked protein:RNA material. While this rigorous purification process ignores authentic high-affinity *binding* events that were present in the living cell, it also ensures that any non-specific protein:RNA binding events that may occur only during cell lysis or in later steps of the work-up do not contaminate the experiment (27). Note that, immediately following UV irradiation and cell lysis, a limiting amount of RNase is employed to achieve cleavage of long RNA polymers into fragments of ~50 bases. The goal is to reduce the presence of long RNA (which behaves poorly during purification) while preventing the generation of too-short RNA (which must remain of sufficient length (~30 nt) to be able to unambiguously align the DNA sequencing information originating from each RNA fragment to the genome (and to the mRNAs which are transcribed from this genomic region) (see Fig. 1A).

As the capCLIP methodology requires the efficient and specific immunoprecipitation of eIF4E (along with crosslinked and/or bound mRNA), and as there are no commercially-available immunoprecipitating anti-eIF4E antibodies, we employed CRISPR/Cas9 genome editing to introduce a 3X-flag epitope tag at the extreme N-terminus of the chromosomal copies of the eIF4E gene in Hela cells (see Figure 1B and Methods). Individual clonal lines were screened for correct insertion of the epitope tag, and then each clone selected for further analysis was fully sequenced surrounding the CRISPR edit site. Western analysis of WT Hela cells and 3 clonal lines is shown in Figure 1C. Clone F2G3, which exhibits homozygous expression of the 3X flag-eIF4E, was used for all subsequent experiments in this study; further analysis demonstrates the similar expression levels of eIF4E, and the similar degree of phosphorylation of eIF4E on Ser209 (its only site of phosphorylation), in WT and flag-eIF4E cells (Figure 1D).

Using capCLIP to probe how the effects of mTORC1 inhibition on translation initiation.

As mentioned in the Introduction, a major consequence of mTORC1 inhibition by either the mTOR-KIs or rapamycin is the impairment of the translation of TOP mRNAs. It is generally agreed that all TOP transcripts commence with an invariant C nucleotide, followed by an uninterrupted sequence of 4-15 pyrimidines (11). While regulation of TOP mRNA translation by mTORC1 is known to occur at the level of translation initiation, there is some debate about the mechanism(s) involved. Several translation-related proteins are direct or indirect substrates of mTORC1 including the small phosphoproteins termed eIF4E-binding proteins (4E-BPs, which comprise 4E-BP1, -BP2 and -BP3; reviewed (1)) and the protein LARP1 (28,29). Active mTORC1 phosphorylates the 4E-BPs at multiple residues; hyper-phosphorylated 4E-BPs cannot bind to eIF4E, allowing eIF4E to bind other partners, including eIF4G, and thus to mediate translation initiation (30). The mTOR-KIs have been shown to inhibit

phosphorylation of 4E-BPs and consequently disrupt eIF4E-eIF4G association (31). Conversely, rapamycin, a macrolide which indirectly impairs the kinase function of mTORC1 by associating with the immunophilin FKBP12 and occluding its active site (32-34), has little effect on the phosphorylation of 4E-BP1 or, therefore, on the association of eIF4E with eIF4G (although this varies between cell types). It has been suggested that the inhibition of TOP mRNA translation by the mTOR-KIs is due to the binding of 4E-BPs to eIF4E and the loss of eIF4E-eIF4G association, which then leads somehow to the *specific* dissociation of TOP mRNAs from eIF4E (35). However, in many cell types, rapamycin does not greatly affect the 4E-BPs or diminish eIF4E-eIF4G binding (10). Accumulating evidence suggests that the mTORC1 substrate LARP1, once it becomes hypo-phosphorylated as a result of mTORC1 inhibition, competes with eIF4E for binding specifically to TOP mRNAs (36,37). Very recently, ribosome profiling of WT and LARP1-deficient human cells using the mTOR-KI Torin1 revealed that LARP1 activity is essential for translational repression of TOP mRNAs under conditions of mTORC1 inhibition (38).

In preparation for capCLIP, we evaluated the consequences of mTORC1 inhibition on the binding of flag-eIF4E to its protein binding partners in our engineered Hela cells. We treated cells with 100 nM AZD8055 (an mTOR-KI), 100 nM rapamycin or DMSO alone for 2 hours followed by immunoprecipitation of flag-eIF4E (Figure 1E) from the resulting cell lysates. Importantly, we find that eIF4E-eIF4G binding is not affected by rapamycin treatment, although it is strongly decreased by AZD8055. This is consistent with the observation that, while rapamycin treatment only slightly alters the phosphorylation of 4E-BP1 and 4E-BP2 and their binding to eIF4E, AZD8055, in contrast, strongly inhibits 4E-BP1 phosphorylation, so that more of this protein binds to eIF4E (where it competes with eIF4G for binding to eIF4E). These data are in agreement with our earlier findings for Hela cells obtained by pull-down of endogenous eIF4E on cap resin (10). We conclude from our data that our gene-editing strategy preserves the normal levels and regulation of eIF4E, and thus that it is unlikely that we have disturbed the control of eIF4E function through disruption its network of interacting proteins.

We also tested the efficiency of flag-eIF4E immunoprecipitation under the high-detergent and high-salt conditions normally used in CLIP-based experiments (Figure 2A). We find that flag-eIF4E is still quantitatively captured by anti-flag antibody under these conditions. We also show that short-wavelength UV irradiation is required for co-capture of RNA fragments in the flag-eIF4E immunoprecipitation, and that the amount and length of RNA covalently linked to flag-eIF4E is sensitive to increasing amounts of RNase I treatment. In Figure 2B, both the 'high' RNase (0.10 U/µl of lysate) and the 'low' RNase (0.02 U/µl of lysate) treatments show an appropriate 'smear' of signal from the [γ-^{32}P]ATP-labelled RNA on the membrane, with the greatest intensity ~37 kDa, ~8kDa higher than the 29 kDa flag-eIF4E protein. The low RNase lane has ~3-fold greater signal intensity that the high RNase lane; accordingly, the low RNase concentration was used for the full-scale capCLIP experiments.

For 'preparative' capCLIP, flag-eIF4E Hela cells were treated for 2 h with 100 nM rapamycin (or DMSO only), immediately UV crosslinked on ice, and then processed through the remainder of the capCLIP methodology as detailed in the Methods and Supplemental Methods sections. In brief , flag-eIF4E was immunoprecipitated, washed, and a [γ-^{32}P]-labelled RNA adapter was ligated to the 3′ end of all RNA co-immunoprecipitating with flag-eIF4E. Two biological replicates per treatment group (2 rapamycin and 2 DMSO replicates) were performed. Covalently-linked and radioactively-labelled flag-

eIF4E:RNA complexes were separated from non-crosslinked RNA using Bis-Tris PAGE. Subsequently, the material in the gel was transferred to nitrocellulose, and RNA was detected by phosphorimaging of the nitrocellulose membrane (Figure 2C). As identical amounts of lysate were IPed and loaded for Bis-Tris PAGE analysis, the signal intensity in Figure 2C is indeed an accurate measure of the amount of mRNA crosslinking to eIF4E. There is no obvious difference in the amount of crosslinked and IPed RNA upon rapamycin treatment (compare the left and right halves of Figure 2C). The material corresponding to flag-eIF4E covalently linked to RNA 'tags' of length ~30-80 nt (which is indicated by the bracket at the right of the membrane) was isolated and the eIF4E:RNA material eluted. Protease was then added to digest eIF4E, the liberated RNA purified, and cDNA synthesis was initiated using a reverse complement DNA primer to the RNA adapter at the 3′ end of the capCLIP tag RNA. The cDNA synthesis reaction, supplemented with [α-^{32}P] dATP, was followed by ligation of a UMI-containing DNA adapter to the 3′ end of the resulting cDNA. The radioactive cDNA product was then size-separated on denaturing PAGE (Figure 2D). The signal intensity of the radio-labelled cDNA is *not* a reflection of the original experimental conditions and likely reflects the variable recovery of the eIF4E-liberated RNA tags from a necessary precipitation step; unfortunately, we saw very poor recovery of the RNA in DMSO replicate 1 (the 1* lane of Figure 2D). For all replicates, cDNA of ~90-140 nt total length (indicated by the bracket at right of membrane) was excised, purified and used for subsequent PCR-based index primer addition and amplification. Finally, individually-indexed PCR libraries were combined for high-throughput sequencing on an Illumina Nextseq 500.

<u>The eIF4E cap-ome provides a comprehensive picture of cellular translation</u>

Filtered, deduplicated capCLIP tag reads were mapped to the hg19 reference genome. Peak-calling software was then used to create peak files for each replicate. (Full read processing and bioinformatic analysis of the capCLIP data are described in the Methods.) Peak data for the rapamycin capCLIP samples are shown in Table 1A. Due to the poor recovery of the DMSO replicate 1 RNA, for some analyses, two pseudo-replicate, DMSO-only control datasets were generated from a single control dataset through a random split of the aligned sequence data; these two control datasets are thus known as control-1A and control-1B. The two rapamycin-treated replicates are named rapa-1 and rapa-2. Table 1A indicates that the majority of peak-calls in both the control and rapamycin-treated datasets align within processed (spliced) mRNA sequences. As previous work (8) has shown that eIF4E photocrosslinks to mRNA through the m^{7}GTP moiety itself, we expected that the original, individual RNA capCLIP tags would include the extreme 5′ end of each mRNA up to and including the cap. Indeed, the majority of peaks (between 34-47% of all peak calls) in each dataset align to 5′ UTR segments, with $\log_2$ enrichment values of 7.1 to 7.5 for tags mapping to 5′ UTR segments. (The ~2-fold greater number of 5′ UTR peaks in the rapamycin samples (compared to the DMSO replicates) is a consequence of the split of the single good DMSO dataset into two, and not a result of rapamycin treatment.)

PCA (principal component analysis) of the four datasets (Figure 3A) indicates that rapamycin treatment leads to a significant and consistent change in the composition of the dataset. To measure the impact of rapamycin treatment upon eIF4E-mRNA association, and to see if these alterations reflect a mRNA's translational activity, as we previously showed in Hela cells (10), we used DiffBind (24) to identify those mRNAs whose association with eIF4E is significantly altered by rapamycin. In total, 5′ UTR peak-calls identify the eIF4E cap-ome as a collection of 3,372 individual mRNAs. To visualise how the binding of each of these mRNAs to eIF4E responds to rapamycin we plot in Figure 3B each mRNA as a function of their mRNA abundance ($\log_2$ of peak area; x-axis), and the $\log_2$ fold change of their binding to eIF4E in rapamycin vs. DMSO treated cells (y-axis). As a population, the 3,372 eIF4E cap-ome

mRNAs exhibit essentially unchanged binding to eIF4E with rapamycin treatment (dashed yellow line; $\log_2$ fold change = 0.02). Individually too, the vast majority of mRNAs (each depicted as a grey spot) show no significant alterations in their binding to eIF4E upon rapamycin treatment. However, 86 mRNAs (in red) do exhibit significantly different binding to eIF4E with rapamycin ($\log_2$ fold-change $\lessgtr$ than 0.50; p value < 0.01; FDR ≤ 0.05), and as expected, rapamycin *decreases* the binding of *all* 86 mRNAs to eIF4E. Additional detail on these 86 mRNAs are available in Supplementary Table 1. Approximately two-thirds of these 86 mRNAs have previously been identified by ribosome profiling screens (35,39) as ones whose translation is impaired by mTORC1 inhibition.

Figure 3C displays capCLIP tag coverage from the single control and two rapamycin sequence datasets for all tags which align to a processed (spliced) mRNA. Individual mRNAs and their aligned tags are normalised with respect to mRNA length, and the coverage of all tags across this normalised mRNA 'length-space' displayed on each plot. Each plot clearly demonstrates the extreme 5′ end enrichment of tags from all three capCLIP datasets, demonstrating that eIF4E capCLIP is indeed capturing RNA tags through crosslinking of eIF4E to the mRNA cap. In Figure 3D, gene-level IGV (Integrative Genomics Viewer (25)) views are displayed for the mRNAs *FTL*, *PABPC1*, *CFL1* and *ACTG1*, and above each transcript are histograms of the control and rapamycin-treated capCLIP tags data which maps to each of the four genes. As in Figure 3C, the tag data here too clearly indicates that the vast majority of capCLIP tags which map to a particular gene align at the extreme 5′ end of the gene's transcript. We have also plotted Hela CAGE data (cap analysis of gene expression (40,41) in each gene-level IGV view in Figure 3D. While close inspection of the alignment between each gene's capCLIP tag data and the Hela CAGE data is difficult given the size of this figure, one can see that the 5′ boundary of each gene's capCLIP tag data aligns at the transcription start sites (TSSs) identified for each transcript using CAGE. While we see essentially no differences in the TSSs identified by capCLIP as compared to the Hela CAGE data, we do find that a small but significant fraction of the mRNA 5′ boundaries specified by the NCBI RefSeq reference mRNAs that appear in IGV use alternative TSSs to those identified by capCLIP and Hela CAGE, suggesting that TSS switching does occur across human cells and tissues.

Mapping TSS sites with capCLIP and CAGE data

To understand in detail the relationship between the capCLIP tags which map at the extreme 5′ end of individual mRNAs and the TSSs for these transcripts as defined by CAGE, we plotted the 5′ boundary of the capCLIP peaks which map to a mRNA 5′ end, along with the TSS site as determined from the Hela CAGE dataset (Figure 3E). We find that the TSSs predicted using capCLIP *peak* data consistently define each mRNA TSS site 7-8 nucleotides 3′ of the 'authentic' TSS site as identified by CAGE. This discrepancy is almost certainly driven by a 7-8 nucleotide-long 'decay' in capCLIP tag coverage at the transcript end, which can be seen in the close-up IGV views of six individual mRNA 5′ ends shown in Figure 4D. We speculate that this loss of signal is due to the obstruction of the reverse transcriptase as it approaches the (still present) cap structure during the synthesis of the cDNA copy of each original mRNA fragment. Given that the m^7G cap is almost certainly the site of photocrosslinking between the mRNA and eIF4E, we expect that after the protease digestion step, each cap still remains crosslinked to a small segment of eIF4E peptide of ~1-3 amino acids, and that this peptide is likely to hinder the reverse transcriptase as it attempts to synthesise a full-length cDNA copy of the transcript 5′ end.

We further speculate that it is this loss of full tag coverage at the 5′ end of the mRNA which leads the MACS2 peak-caller software to consistently define the 5′ boundary of mRNA 'cap peaks' 7-8 nucleotides downstream of the 5′ end of the transcript (the 5′ end of the transcript being the TSS

predicted by the Hela CAGE data). Regardless, the *consistency* of this 7-8 nt offset across the 5′ UTR peaks of the eIF4E cap-ome (apparent in Figure 3E) suggests that, by compensating for this characteristic mis-location of the 5′ capCLIP peak boundary, capCLIP can provide a means of predicting a mRNA's TSS that is completely independent of the TSS estimate provided by CAGE data, Thus, *together* capCLIP and CAGE allow us to attain a significantly greater level of confidence in the active TSS site (or sites) used in a particular cell or tissue than either method by itself. Most importantly, we can use this high-confidence TSS to help our search for a relationship between a mRNA's 5′ end and the magnitude of loss of its association with eIF4E under rapamycin treatment. Such an analysis was not possible before our development of capCLIP.

TOP mRNAs are specifically depleted from the eIF4E cap-ome with rapamycin treatment

Figure 4A shows a frequency plot depicting the 86 mRNAs whose eIF4E association is significantly altered by rapamycin (which we termed 'differentially regulated mRNAs' or 'DR RNAs', in blue) and the set of all mRNAs identified by capCLIP (in grey), arranged according to their $\log_2$ fold-change in eIF4E association in response to rapamycin. As previously depicted in Figure 3B, there is no overall shift in eIF4E association across the population of all mRNAs identified by capCLIP ($\mu = 0.01$). However, there is an average two-fold decrease in eIF4E association amongst the DR mRNA group ($\mu = -1.08$; $p<0.0001$). Using the high-confidence prediction of the TSS provided by capCLIP and CAGE, we analysed the 5′-end sequences of the 86 statistically significant mRNAs to identify which contained authentic TOP motifs. Our analysis reveals that 62 (72%) possess TOPs as defined by Meyuhas and Kahan (an initial C, followed by 5-14 pyrimidines) (11) while an additional 13 (15%) appear to have a multiple TSSs, including a TOP TSS and at least one non-TOP TSS.

While the vast majority of the 62 TOP mRNAs are known TOP mRNAs encoding ribosomal proteins or other constituents of the translational machinery (consistent with current understanding of the TOP mRNA family (42)), we also identify two novel TOP transcripts, *TAF1D*, and *LGALS1*. In particular, the identification of *TAF1D* as a TOP mRNA is very intriguing. As an essential component of the SL1 complex which regulates pol I-catalysed transcription of the major ribosomal RNAs (43), a process which has long been known to be driven by mTORC1 signalling (44), the translational control of *TAF1D* levels may provide a further way in which mTORC1 signalling promotes ribosome biogenesis, in parallel with facilitating synthesis of TOP-mRNA-encoded ribosome proteins, ribosomal RNAs and other mechanisms (44). *LGALS1* (galectin-1) is a secreted lectin with potential roles in apoptosis and cell growth; we are not aware of any known link between *LGALS1* with mTORC1 signalling. Finally, the remaining 11 of the 86 mRNAs (13%) lack *any* identifiable TOP motif (that is a 5′ pyrimidine-rich motif) according to our capCLIP/CAGE TSS predictions.

Overall, the <u>specificity</u> of the changes and the <u>magnitude</u> of the effects we see in the eIF4E cap-ome ± rapamycin are strikingly similar to those obtained from ribosome-profiling experiments measuring the impact of the mTOR-KIs PP242 and Torin1 on mRNA translational efficiency(35,39). The data support our original hypothesis that capCLIP should provide a sensitive and robust measurement of translational activity. Thus, capCLIP may provide a simple and powerful alternative to ribosome profiling techniques. As such sucrose gradient centrifugation techniques to measure mRNA association with translating ribosomes and polysomes can be confounded by association of mRNAs with other high-molecular weight, albeit non-translating, complexes, or with translationally-stalled ribosomes (45), capCLIP has the potential to substitute for or complement ribosome profiling methods, and it will be interesting to test these two methods side-by-side.

mRNAs sensitive to rapamycin possess a strikingly well-defined TOP motif

The distribution of the 62 TOP mRNAs according to their $\log_2$ change in association with eIF4E under rapamycin treatment is plotted in Figure 4B, and clearly shows their dramatic shift with mTORC1 inhibition. These data also reveal that there are significant differences in the magnitude of the response of individual mRNAs. Again, using capCLIP and CAGE data to map the TSSs of mRNAs with nucleotide precision, we next asked if analysis of the TOP motifs of these 62 TOP mRNAs could provide us with a deeper understanding of the relationship between TOP motif sequence and TOP 'strength;' the magnitude of the loss of a TOP mRNA from the eIF4E cap-ome in rapamycin-treated cells. For this analysis, we excluded 10 of the 62 TOP mRNAs for which we were unable to use CAGE and capCLIP data to unambiguously identify the mRNA's TSS. (The CAGE data for these mRNAs is uncharacteristically broad, and two or more (usually overlapping) TSSs may be used for these mRNAs).

As capCLIP and CAGE data predicts that the remaining 52 TOP mRNAs use a single TSS, we extracted the first 20 nt of each of the 52 TOP mRNAs and input the data into **weblogo3** (46) to examine what, if any, new sequence features are present at the 5′-ends of these rapamycin-responsive mRNAs. The resulting weblogo, which in essence defines the 'functional' TOP motif as defined by capCLIP, is shown in Figure 4C. As expected, the weblogo consensus sequence is consistent with earlier work demonstrating that functional TOP mRNAs harbour an invariant C base at position one of their transcripts. Surprisingly, the capCLIP weblogo data also suggests that 'functional' TOP motifs also exhibit several additional properties. Firstly, our data suggest that TOP motifs are (at most) seven nucleotides in length and must be positioned directly at the mRNA's 5′-end, with absolutely no preference for pyrimidine nucleotides in positions 8-20. Secondly, and unexpectedly, our data indicates that in addition to the invariant C base in position 1, there is a near-essential requirement for a U in position 2, a very strong preference for a U in position 4, and a significant preference for a C in position 7. In sum, our capCLIP data suggest that optimal TOP motifs are defined by the sequence 5′-CUYUYYC-3′, and is thus shorter than previously thought (11) (Figure 4C).

To examine whether this TOP motif consensus sequence is consistent with the behaviour of individual mRNAs in our control/rapamycin-treated capCLIP experiments, we first examined data for six selected mRNAs: two 'non-TOP' mRNAs, *CALR* & *HIST1H2AC*; three TOP mRNAs, *RPL10*, *RPS12* & *RPLP1*; and one so-called *switching*-TOP mRNA (see below), *RPL15*. These 6 mRNAs are shown in Figure 4D. The TSS and extreme 5′-end of each of the 6 mRNAs are depicted by a black bar, with aligned histogram displays of control capCLIP tag data (in blue), capCLIP tags from rapamycin-treated cells (in red), and Hela CAGE data (in grey), above the sequence. The control and rapamycin capCLIP tag histograms have been normalised to allow the histogram heights of the two treatments to be compared. The 7-nt TOP consensus sequence (if present) is underlined in magenta for the three TOP and one 'pseudo'-TOP mRNAs (see figure legend for more details).

Interestingly, the two mRNAs whose association with eIF4E is most strongly affected by rapamycin, *RPL10* and *RPS12* (Figure 4B), harbour the same 7-nucleotide TOP motif, 5′-CUCUUUC-3′. This sequence is 1 of the 8 possible TOP motif permutations described by the capCLIP TOP consensus motif 5′-CUYUYYC-3′. To ascertain if there is an association between the magnitude of the $\log_2$ fold-change of a mRNA's eIF4E association under rapamycin treatment and any of the 8 unique TOP motif permutations, we sorted the 52 TOP mRNAs into the 8 permutation groups and calculated the mean $\log_2$ fold-change for each group (Figure 4E). While most of the groups show similar changes with

rapamycin, as expected, it is interesting that 5′-CUCUUUC-3′, the TOP sequence possessed by *RPL10* and *RPS12* is the most abundant TOP sequence permutation by a significant margin with 15 of the 52 TOP mRNAs (29%) . It is also curious that there are *no* TOP mRNAs in the 5′-CUCUCCC-3′ group, suggesting that this particular sequence either does not behave as a TOP motif, or is subject to selection pressures other than TOP function. A second notable finding from our TOP motif analysis is the strong preference for a U in position 2 of our consensus motif. In fact, only one of the 52 functional TOP mRNAs lacks a U at position 2. The exception, encoding *RPLP1*, has a C at this position (Figure 4D). Interestingly, out of the 52 TOP mRNAs showing a significant drop in eIF4E association, RPLP1 is the *least* affected by rapamycin treatment (Figure 4B), suggesting that a U in this position is necessary for a strong response of a TOP mRNA to rapamycin. In sum, the correlation between an mRNA's extent of homology with our TOP consensus motif 5′-CUYUYYC-3′ and the degree to which rapamycin alters its binding to eIF4E, gives us confidence that the weblog analysis is indeed identifying salient features of the TOP motif.

The last mRNA in Figure 4D is *RPL15*, which encodes a protein of the large ribosomal subunit. While we would expect the RPL15 mRNA to behave as a TOP, it does not do so according to our data, with essentially no change in eIF4E association upon rapamycin treatment ($\mu = 0.03$; Figure 4B). Consistent with this lack of response, examination of the capCLIP tag and CAGE data for *RPL15* clearly indicates the TSS begins 5′-GGGAGTA..-3′, which is clearly a non-TOP sequence. However, a block of 13 pyrimidines does lie 11 nt 5′ of the TSS mapped by capCLIP and CAGE, and we speculate that other human cell lines or tissues may initiate transcription from within this block, and hence transcribe *RPL15* with a potential functional TOP motif. We surveyed human ENCODE CAGE data and found that in CD20+ B-lymphocytes the predominant TSS begins with the 7th pyrimidine of the 13 nt block, such that the first 7nt of the mRNA wold be 5′-CCUUUCC-3′ (Figure 4D). Thus, in CD20+ B-cells, *RPL15* is transcribed with a functional TOP, albeit one we would predict is somewhat weak, due to the non-canonical C nucleotide at position 2 of the TOP motif. We term *RPL15* a *switching*-TOP mRNA, and hypothesise that alternative TSS selection may be a used across a range of mRNAs, thereby permitting cells to add or remove certain mRNAs from mTORC1-mediated translational regulation. Such TSS switching has also been described for the *FAU* gene, a fusion protein of a ubiquitin-like peptide and *RPS30*. FAU is normally transcribed as a TOP mRNA, but in liver an alternative TSS, which makes a non-TOP isoform of the FAU mRNA, predominates (38).

Finally, a consensus is now emerging in the literature that mTORC1 inhibitors, including rapamycin, repress TOP mRNA translation through the cap- and TOP mRNA-specific RNA-binding activity of the mTORC1 substrate LARP1. mTORC1 inhibition promotes an increase in the affinity of LARP1 for TOP mRNAs, selectively inhibiting the association of TOP mRNAs with eIF4E, and depleting them from translation initiation complexes (28,29,38,47-49)). A recent ribosome profiling study of WT and LARP1-deficient human cells ± the mTOR-KI Torin-1 reveals that LARP1 is essential for translational repression of TOP mRNAs, as Torin-1 has virtually no effect on TOP mRNA translation in LARP1 KO cells (38). As capCLIP serves as a *direct* readout of the change in eIF4E association for thousands of cellular mRNAs (rather than inferring the sequestration of mRNAs from eIF4E by LARP1 under mTORC1 inhibition as ribosome profiling does), we would argue that our capCLIP results now provide an independent line of evidence for the LARP1-TOP mRNA model. Moreover, a LARP1:RNA 3D structure, incorporating the 8-nucleotide synthetic TOP motif RNA 5′-CUUUUCCG-3′ surprisingly found that the eighth nucleotide of the RNA did not contact the LARP1 protein fragment which bound to the first 7 nucleotides of the TOP motif, but instead interacted with a second LARP1 complex in an

adjacent unit cell (47). Thus, the 7 nucleotide functional TOP motif predicted by capCLIP is consistent with the finding that LARP1 can bind at most seven nucleotides at the 5′ terminus of a mRNA. This observation is now also supported by recent experiments demonstrating that LARP1 is insensitive to the nucleotides at position 8 and beyond (38).

TOP mRNA discovery: benchmarking of capCLIP to ribosome profiling

There is striking similarity between the set of rapamycin-responsive mRNAs identified by capCLIP and findings from previous work identifying mRNAs whose translation is regulated by mTORC1 signalling (10,35,39). The data strongly suggest that capCLIP not only provides a reliable and quantitative measurement of eIF4E binding to individual mRNAs, but also that the change in an mRNA's binding to eIF4E serves as a reliable proxy of the mRNA's translational activity, at least in the case of mRNAs regulated by mTORC1 signalling. To further benchmark capCLIP, we carried out several further comparisons of our data to the pioneering studies of Thoreen *et al.* (35) and Hsieh *et al.* (39), both of which used ribosome profiling to identify mRNAs translationally regulated by rapamycin or mTOR-KIs. Ribosome profiling data is most often employed to compute a mRNA's translational efficiency (T_E), a measure that relates the number of ribosomes occupying a mRNAs open reading frame (ORF) relative to the expression level of the mRNA (12). Treatment of cells with mTOR-KIs results in statistically significant reductions in T_E (ranging in general from 2-4 fold) across the ~90 known TOP mRNAs. While the Hsieh *et al.* study does report T_Es for individual TOP mRNAs upon rapamycin treatment (average T_E reduction of ~1.8-fold), it appears that only ~20% have an FDR value less than 0.05 (and these 20% are not identified). The Thoreen *et al.* study does not report any quantitative T_E data for rapamycin.

Notwithstanding the limitations in both the capCLIP and T_E data, we sought to determine the degree of correlation between the changes in eIF4E binding (capCLIP) and the changes in T_E (ribosome profiling) upon rapamycin treatment. Figure 5A plots the $\log_2$ fold change in capCLIP peak area (the change in eIF4E-cap binding) vs. the $\log_2$ fold change in T_E for the 45 TOP mRNAs in common between our differentially regulated capCLIP target list ($p<0.01$; FDR<0.05) and the Hsieh *et al.* rapamycin ribosomal profiling T_E dataset (whose FDR and p-values are not known). A Spearman r-value calculation shows that there is indeed a significant correlation ($r = 0.3216$; $p = 0.031$) between eIF4E binding and T_E. While the correlation is encouraging, effective validation that the capCLIP can serve can substitute as a quantitative indicator of translational activity for ribosome profiling will require direct comparisons of the two methods under identical conditions and with identical cell lines or tissues.

As both the degree of translational repression and the number of mRNA targets repressed are generally much larger with mTOR-KIs, we did not attempt a similar comparison of our capCLIP data with either study's mTOR-KI T_E data. However, we were intrigued by an observation that a number of the mRNAs categorised by both the Thoreen *et al.* and Hsieh *et al.* (35,39) ribosome profiling experiments as profoundly sensitive to mTORC1 inhibition were *not* similarly categorised in our rapamycin capCLIP experiments. Examination of this set of mRNAs revealed that most possessed significantly longer primary open reading-frames (ORFs) as compared to the majority of TOP mRNAs. As accurate calculations of T_E across mRNAs with ORF lengths that differ by factors of 5-fold or more may be difficult to achieve in practice (12), we evaluated if there is a relationship between mRNA ORF length and the degree of translational repression in either our capCLIP data or in the Torin-1 ribosome profiling data of Thoreen *et al*. We plotted the $\log_2$ fold change (of capCLIP peak area) or the $\log_2$ fold change of T_E (for ribosome profiling) *versus* mRNA ORF length in amino acids, for the 29 TOP mRNAs found

in common between the rapamycin capCLIP dataset and the Torin1 ribosome profiling dataset (Figure 5B). While we do not find a significant relationship between ORF length and loss of binding to eIF4E in the capCLIP data ($r = 0.2194$; $p = 0.253$), there is a significant correlation between ORF length and T_E in the Torin-1 ribosome profiling data ($r = -0.4986$; $p = 0.006$) (35). We are unaware of any rigorous studies that have experimentally tested how ORF length impacts T_E calculations made using ribosome profiling data. However, as ribosome footprint densities are likely to vary over the length of any ORF, the uncertainty in determining the 'correct' ribosome density will tend to rise with ORF length. As the capCLIP method (and peak area calculation) are not influenced by ORF length, we suggest that capCLIP, particularly when combined with RNA-seq data to account for both changes in steady-state mRNA levels, may provide a highly accurate picture of cap-dependent translation.

capCLIP reveals significant differences in the eIF4E and P-eIF4E 'cap-omes'

While the field's understanding of TOP mRNA translation is sufficiently deep to serve an ideal system for the benchmarking of capCLIP, we also sought to use the methodology to explore new biology. Specifically, we examined the biological role played by MNK phosphorylation of eIF4E. Surprisingly, even after three decades of work, the functional consequences of phosphorylation eIF4E on Ser209 remain obscure. In contrast to mTORC1 signalling, there is very little published data from high-throughput approaches concerning the translational consequences of eIF4E phosphorylation, with only a single unbiased screen of P-eIF4E function reported (using cells in which eIF4E Ser209 has been mutated to non-phosphorylatable alanine (14)).

The flag-eIF4E Hela cell line, like WT cells, exhibits moderate levels of P-eIF4E under normal growth conditions. Serum starvation for 12 h reduced P-eIF4E levels to significantly below basal levels (lane 1, Figure 6A). We then established the timing required to obtain maximal P-eIF4E levels in these serum-starved cells by stimulating them with phorbol myristate acetate (PMA), a potent activator of ERK signalling and thus of MNK(1) activity and subsequent phosphorylation of eIF4E (50). A potent increase in P-eIF4E response was seen by 15 or 30 min after stimulation, providing the earliest point of maximal eIF4E phosphorylation (lanes 2-6, Figure 6A). To provide a population of cells with undetectable P-eIF4E to serve as a control, we used the highly-specific MNK inhibitor eFT-508 (15). Serum-starved flag-eIF4E Hela cells were pre-incubated for 1 h with concentrations of eFT-508 at concentrations from 1 to 0.01 μM, or with only DMSO only as control (Figure 6B). The 0.1 μM concentration of eFT-508 was the lowest dose of the drug to substantially inhibit eIF4E phosphorylation and was subsequently used for capCLIP.

For capCLIP, serum-starved flag-eIF4E Hela cells were treated with 0.1 μM eFT-508 or DMSO only for 1 h, followed by 30 min treatment with 500 nM PMA. Full experimental details can be found in the Methods and Supplemental Methods sections. Following an essentially identical work-up to that of the rapamycin capCLIP study, ^{32}P-labelled RNA crosslinked to flag-eIF4E was separated from non-crosslinked RNA material by Bis-Tris PAGE. The phosphorimage of the Bis-Tris PAGE nitrocellulose membrane is shown in Figure 6C. A section of the membrane containing RNAs of length ~30-80 nt, as indicated by the bracket at the right of the membrane, was excised, and processed to release the covalently linked RNA from the eIF4E protein. Figure 6D shows a phosphorimage of the cDNA sizing gel, with cDNA of ~90-140 nt total length indicated by the bracket at right of membrane. This material was isolated and used for PCR amplification and sequencing. An overview of the peak-calls made from the eFT-508 capCLIP sequencing data is shown in Table 1B. Approximately 30% of all peak-calls in the two replicate PMA-only control samples and the two replicate eFT-508 treated samples lie in 5′

UTR regions (a $\log_2$ fold enrichment of ~7.3, similar to the rapamycin capCLIP samples). PCA of the four datasets (Figure 6E) indicates that the PMA-only and eFT-508 treatments are clearly differentiable on the two components of the analysis.

DiffBind (24) was again used to identify those mRNAs whose association with eIF4E was significantly altered by MNK-mediated phosphorylation of eIF4E. Of the 4,965 mRNAs identified by peak-calling in this capCLIP experiment, 256 exhibited statistically-significant differences in eIF4E binding ± eFT-508 treatment (FDR<0.05). Additional detail on these 256 capCLIP targets is provided in Supplementary Table 2. To visualise the eFT-508 capCLIP target data, Figure 7A plots each mRNA's abundance (x-axis; $\log_2$ of the capCLIP peak area) vs. the $\log_2$ fold change in eIF4E binding upon eIF4E phosphorylation (left y-axis). The right y-axis depicts the same data but with respect to the underlying PMA/eFT-508 treatment paradigm, and hence, has the opposite sign.

Figure 7A shows that surprisingly, all 256 of these statistically-significant mRNAs show a *reduction* in eIF4E binding upon stimulation of its phosphorylation. Indeed, unlike our rapamycin capCLIP experiment, where the population of eIF4E cap-ome mRNAs was essentially unchanged by the inhibitor treatment, the population of 4,965 mRNAs in this capCLIP experiment exhibit, when taken as a whole, a $\log_2$ fold change in eIF4E affinity of -0.75 upon stimulation of phosphorylation (Figure 7A, dashed yellow line). Thus, the mRNA population *as a whole* (the eIF4E cap-ome) behaves similarly to the 256 mRNAs which exhibit statistically significant reductions in eIF4E binding upon phosphorylation. However, the plot in Figure 7A also indicates that a small fraction of the cap-ome mRNAs display an apparent *increase* in binding to eIF4E upon its phosphorylation. As none of the mRNAs within this 'increasing' subpopulation exhibit statistically significant differences between the two treatment conditions, it is not possible to know the true magnitude of this increase, or indeed if there is any real increase in eIF4E binding at all. Greater capCLIP library depth will be required resolve this point. In light of the entirety of our current eFT-508 capCLIP observations, which indicate that the cap-ome as a whole exhibits a decrease in binding to eIF4E upon its phosphorylation, the most likely hypothesis is that eIF4E phosphorylation acts to diminish the intrinsic affinity of the protein for capped mRNAs. This model has support from *in vitro* biophysical studies of the effects of eIF4E phosphorylation ((16-18), discussed below). Thus, while there is a loss of binding of the mRNAs (the cap-ome) to eIF4E as a whole, why are some mRNAs particularly sensitive to eIF4E phosphorylation, and others less so?

In vivo, the consequence(s) of eIF4E phosphorylation are not well understood. As the Ser209 phosphorylation site lies on the opposite side of the protein to its binding interface with eIF4G (51), it is not clear why or whether phosphorylation of eIF4E affects any of the canonical events of translation initiation, such as eIF4F complex formation, recruitment of eIF3, recruitment of the 40S ribosomal subunit, etc. We next asked if we could identify characteristics in common (primarily length and base composition) between the 5′ UTR regions of individual mRNAs that exhibit the largest changes in eIF4E binding upon phosphorylation. In total 90 mRNAs were selected: the 45 mRNAs with the greatest *reduction* in eIF4E binding upon eIF4E phosphorylation, and correspondingly, the 45 mRNAs with the greatest relative *increase* in eIF4E binding upon eIF4E phosphorylation. Ensembl Biomart and the NCBI RefSeq collection were used to source the 5′ UTR sequences of all 90 mRNAs selected. (See Material and Method for additional details of the analysis.) As shown in the box-and-whisker plots in Figure 7B, we observed a striking difference in the median 5′ UTR length of the two groups of mRNAs. The 45 mRNAs whose binding to eIF4E *falls* with eIF4E phosphorylation have a median 5′ UTR length

of 87 nt (SD=84nt), significantly shorter than the average for human 5′ UTRs, which is estimated to be variously 218 nt (NCBI RefSeq data, (52)) and 293 nt (SD=27) (UTRdb, (53)).

In contrast, the median 5′ UTR length of the 45 mRNAs whose binding to eIF4E apparently *rises* with eIF4E phosphorylation is 242 nt (SD=230), essentially equivalent to the median 5′ UTR length for human mRNAs. Notably, and in reference to the discussion above of our uncertainty regarding their apparent increase in binding to P-eIF4E, the average $\log_2$ fold increase in eIF4E binding of this group of mRNAs is an very modest 0.61. Thus, these 45 most strongly affected mRNAs shift only ~1.5 fold in response to eIF4E phosphorylation. We also examined nucleotide composition for the two groups of 5' UTRs. While average GC content for human 5' UTRs is 58% (UTRdb, (53)), it is 68% for the 45 mRNAs which fall in binding to eIF4E upon phosphorylation; the other 45-mRNA group is unchanged from the 58% average. In sum, the mRNAs whose eIF4E binding drops the most in response to eIF4E phosphorylation exhibit two distinctive 5′ UTR features: one, much shorter than average 5′ UTR length; and two, much higher than average GC content.

C6orf89, *NRAS*, *COX5A* and *LSM4* (in blue in Figure 7A) are the 4 mRNAs with the largest reductions in eIF4E binding upon phosphorylation, ranging from decreases of approximately 4- to 6-fold. Highlighted in green are two ferritin mRNAs, which exhibit more modest changes in response to eIF4E phosphorylation but reach levels of particularly high significance. Analysis of the protein levels of NRAS, a Ras family member, and FTH1 (ferritin heavy chain) proteins, whose mRNA showed a more modest fold-decrease upon eIF4E phosphorylation, was conducted by inhibiting WT eIF4E phosphorylation in Hela cells over a 24-h time course with eFT-508 (Figure 7C). While we saw small (~1.8-fold) increases in NRAS or FTH1 steady-state protein levels at 4 and 8 h after inhibition of P-eIF4E, further validation of these and other candidate mRNAs whose translation is potentially regulated by phosphorylation of eIF4E is likely to require methods that directly measure *real-time de novo* synthesis of the corresponding protein; that is, the ability to assess the rate of production of newly-synthesised proteins. A future goal will be to apply such methods (see below) to validate the eFT-508 capCLIP dataset and thereby leverage the capCLIP data to provide a deeper mechanistic and biological understanding of the translational consequences of MNK signalling.

capCLIP indicates that phosphorylation of eIF4E generally decreases cap binding

The development of capCLIP has allowed us, for the first time, to evaluate the effect of MNK activity, and thus the phosphorylation of eIF4E, on the association of specific mRNAs with eIF4E in living cells, and to do so in an unbiased manner. Our data clearly show that stimulation of eIF4E phosphorylation causes a statistically-significant *decrease* in the association of 256 mRNAs with eIF4E implying that phosphorylation by the MNKs weakens eIF4E's binding to the cap. Our 'top' targets (*NRAS*, *COX5A*, etc.) are the mRNAs that are *most* sensitive to phosphorylation of eIF4E and whose translation is likely to be most strongly impaired by MNK activity. In addition, our finding that eIF4E phosphorylation leads to an average $\log_2$ fold change of -0.75 over the eIF4E cap-ome is consistent with several *in vitro* biophysical studies showing phosphorylation of eIF4E decreases its binding to capped RNA (16-18). This contrasts with earlier speculations, based on the 3D structure of eIF4E, that phosphorylation of eIF4E on Ser209 would actually increase its affinity for capped mRNA, due to formation of a salt bridge between phosphorylated Ser209 and Lys159 (51) or for other reasons (54). As other studies have suggested (16,18), phosphorylation of eIF4E likely decreases binding to capped mRNA due to electrostatic repulsion between the negatively-charged phosphate on Ser209 and similarly charged groups on the first nucleotides of the eIF4E-bound mRNA.

While phosphorylation of eIF4E may serve as a 'switch' to alter the eIF4E's kinetic and/or binding equilibrium to capped mRNA, it is also possible there are sequence and/or structural factors in certain mRNA 5′ UTRs that serve to amplify the impact of this switch, rendering certain classes of 5′ UTRs either more or less sensitive to the influence of phosphorylation on their binding to eIF4E. Our finding that the 45 mRNAs *most* sensitive to eIF4E phosphorylation (which include the highlighted *C6orf89*, *NRAS*, *COX5A*, and *LSM4* mRNAs) possess significantly shorter, and more GC-rich, 5′ UTRs than average reveals the new insight that MNK activity and eIF4E phosphorylation provide a means to regulate cellular translational activity based on 5′ UTR composition. In the absence of P-eIF4E, the translation of mRNAs with shorter, more GC-rich 5′ UTRs may be favoured. Conversely, activation of MNK signalling and eIF4E phosphorylation would then preference translation initiation onto mRNAs with longer 5′ UTRs.

An important element of this model would be to understand the mechanistic basis by which phosphorylation of eIF4E favours initiation of translation of mRNAs with long 5′ UTRs. Intriguingly, the only unbiased screen for effects of eIF4E phosphorylation that has been reported (14) concluded that phosphorylation of eIF4E *promoted* translation of certain mRNAs. The study used a different approach and experimental system to us, substituting Ser209 in eIF4E with a non-phosphorylatable alanine in mouse embryonic fibroblasts, and employed sucrose density gradient centrifugation to identify polysome-associated mRNAs by microarray. As the Ala209 substitution may not accurately mirror the absence of a phosphate at this position (55), and as the experiment did not directly analyse translational activity or, crucially, eIF4E-mRNA interactions, it is difficult to assess this work in light of our new findings. Most other studies, which typically focus on the effect of MNK inhibition on levels of specific proteins on one or few candidates, are largely without any mechanistic rationale. Also, as they often use small molecule inhibitors of the MNKs which are now known to exert substantial off-target effects (13), they too are difficult to contextualise in relation to our new data.

Validation of our work, which is beyond the scope of this methods paper, will require analysis of whether alterations in binding of eIF4E to a given mRNA affects the rate of synthesis of the corresponding protein. This would require approaches that allow quantification of *de novo* rates of synthesis of specific polypeptides such as 'BONCAT' (56), probably coupled with pulsed stable-isotope labelling, as has been applied to study rapid changes in protein synthesis in otherwise hard-to-study cells such as primary neurons or cardiomyocytes (57,58). These techniques will likely be invaluable for validating our leading eFT-508 capCLIP candidate mRNAs. In addition, as such methods potentially serve as potentially the most direct validation of a mRNA's translational activity *in vivo*, we are eager to use them to validate capCLIP more broadly, as we hypothesise that capCLIP could serve as a method for measuring the translational activity of individual mRNA species more generally.

The way forward for capCLIP and future prospects

We envision that capCLIP can be developed into a general technique to permit the acquisition of deep, quantitative information on a cell or tissue's present translational output, the eIF4E cap-ome. As the methodology evolves, increasing capCLIP tag library depth will permit capCLIP to be employed as a general technique for the reliable and quantitative detection of changes in translational activity across even those individual mRNA species with very low expression levels. We foresee using capCLIP to identify the eIF4E cap-omes of individual tumours, especially in mouse models of tumorigenesis using patient-derived cancer cell lines which be edited to express flag-eIF4E. Especially promising is the

ability, when using flag-tagged eIF4E tumour cells, to evaluate the translational status of mRNAs in metastatic cancer cells without the need to remove surrounding tumour stroma, as the host tissue containing only WT eIF4E cannot 'contaminate' the metastatic samples. As there is little comprehensive data on metastatic proteomes/translatomes, and as such an environment is not amenable to either ribosome profiling or direct proteomic approaches, capCLIP offers a novel way to significantly enlarge our understanding of translational regulation in cancer. There are doubtless other applications for capCLIP where existing technologies are not suitable. For example, capCLIP may be especially valuable for elucidating the specific biological role(s) of other cellular cap-binding proteins, particularly the eIF4E2 (4EHP) and eIF4E3, the other two cytoplasmic cap-binding proteins in mammals. While eIF4E2 is generally thought to play a role in translational repression (see, for example (59,60)), other work suggests eIF4E2 promotes mRNA translation in certain cellular contexts (61). As capCLIP makes no assumptions about the cellular role of a cap-binding protein, it is an ideal method for elucidating eIF4E2 function. Likewise, capCLIP of LARP1, the nuclear cap-binding protein NCBP2, or of proteins with proposed cap-binding properties all have the potential to enlarge our understanding of how dynamic networks of cap-protein interactions modulate RNA processing events throughout the cell.

Acknowledgement & Funding

We gratefully acknowledge funding from the National Health & Medical Research Council [GNT1089167 and Research Fellowship GNT1118170 to GJG] and SAHMRI (to CGP).

References

1. Proud, C.G. (2019) Phosphorylation and Signal Transduction Pathways in Translational Control. *Cold Spring Harb Perspect Biol*, **11**, a033050.
2. Merrick, W.C. and Pavitt, G.D. (2018) Protein Synthesis Initiation in Eukaryotic Cells. *Cold Spring Harb Perspect Biol*, **10**, a033092.
3. Rhoads, R.E. (2009) eIF4E: new family members, new binding partners, new roles. *J Biol Chem*, **284**, 16711-16715.
4. Merrick, W.C. (2015) eIF4F: a retrospective. *J Biol Chem*, **290**, 24091-24099.
5. Truitt, M.L., Conn, C.S., Shi, Z., Pang, X., Tokuyasu, T., Coady, A.M., Seo, Y., Barna, M. and Ruggero, D. (2015) Differential Requirements for eIF4E Dose in Normal Development and Cancer. *Cell*, **162**, 59-71.
6. Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A. and Darnell, R.B. (2003) CLIP identifies Nova-regulated RNA networks in the brain. *Science*, **302**, 1212-1215.
7. Ule, J., Jensen, K., Mele, A. and Darnell, R.B. (2005) CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods*, **37**, 376-386.
8. Pelletier, J. and Sonenberg, N. (1985) Photochemical cross-linking of cap binding proteins to eucaryotic mRNAs: effect of mRNA 5' secondary structure. *Mol Cell Biol*, **5**, 3222-3230.
9. Saxton, R.A. and Sabatini, D.M. (2017) mTOR Signaling in Growth, Metabolism, and Disease. *Cell*, **168**, 960-976.
10. Huo, Y., Iadevaia, V., Yao, Z., Kelly, I., Cosulich, S., Guichard, S., Foster, L.J. and Proud, C.G. (2012) Stable isotope-labelling analysis of the impact of inhibition of the mammalian target of rapamycin on protein synthesis. *Biochem J*, **444**, 141-151.
11. Meyuhas, O. and Kahan, T. (2015) The race to decipher the top secrets of TOP mRNAs. *Biochim Biophys Acta*, **1849**, 801-811.
12. Ingolia, N.T. (2016) Ribosome Footprint Profiling of Translation throughout the Genome. *Cell*, **165**, 22-33.
13. Xie, J., Merrett, J.E., Jensen, K.B. and Proud, C.G. (2019) The MAP kinase-interacting kinases (MNKs) as targets in oncology. *Expert Opin Ther Targets*, **23**, 187-199.
14. Furic, L., Rong, L., Larsson, O., Koumakpayi, I.H., Yoshida, K., Brueschke, A., Petroulakis, E., Robichaud, N., Pollak, M., Gaboury, L.A. *et al.* (2010) eIF4E phosphorylation promotes tumorigenesis and is associated with prostate cancer progression. *Proc Natl Acad Sci U S A*, **107**, 14134-14139.
15. Reich, S.H., Sprengeler, P.A., Chiang, G.G., Appleman, J.R., Chen, J., Clarine, J., Eam, B., Ernst, J.T., Han, Q., Goel, V.K. *et al.* (2018) Structure-based Design of Pyridone-Aminal eFT508 Targeting Dysregulated Translation by Selective Mitogen-activated Protein Kinase Interacting Kinases 1 and 2 (MNK1/2) Inhibition. *J Med Chem*, **61**, 3516-3540.
16. Scheper, G.C., van Kollenburg, B., Hu, J., Luo, Y., Goss, D.J. and Proud, C.G. (2002) Phosphorylation of eukaryotic initiation factor 4E markedly reduces its affinity for capped mRNA. *J Biol Chem*, **277**, 3303-3309.
17. Slepenkov, S.V., Darzynkiewicz, E. and Rhoads, R.E. (2006) Stopped-flow kinetic analysis of eIF4E and phosphorylated eIF4E binding to cap analogs and capped oligoribonucleotides: evidence for a one-step binding mechanism. *J Biol Chem*, **281**, 14927-14938.
18. Zuberek, J., Wyslouch-Cieszynska, A., Niedzwiecka, A., Dadlez, M., Stepinski, J., Augustyniak, W., Gingras, A.C., Zhang, Z., Burley, S.K., Sonenberg, N. *et al.* (2003) Phosphorylation of eIF4E attenuates its interaction with mRNA 5' cap analogs by electrostatic repulsion: intein-mediated protein ligation strategy to obtain phosphorylated protein. *RNA*, **9**, 52-61.
19. Richardson, C.D., Ray, G.J., DeWitt, M.A., Curie, G.L. and Corn, J.E. (2016) Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol*, **34**, 339-344.
20. Renaud, J.B., Boix, C., Charpentier, M., De Cian, A., Cochennec, J., Duvernois-Berthet, E., Perrouault, L., Tesson, L., Edouard, J., Thinard, R. *et al.* (2016) Improved Genome Editing

Efficiency and Flexibility Using Modified Oligonucleotides with TALEN and CRISPR-Cas9 Nucleases. *Cell Rep*, **14**, 2263-2272.
21. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15-21.
22. Smith, T., Heger, A. and Sudbery, I. (2017) UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res*, **27**, 491-499.
23. Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P. and Burge, C.B. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature*, **456**, 470-476.
24. Ross-Innes, C.S., Stark, R., Teschendorff, A.E., Holmes, K.A., Ali, H.R., Dunning, M.J., Brown, G.D., Gojis, O., Ellis, I.O., Green, A.R. *et al.* (2012) Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature*, **481**, 389-393.
25. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G. and Mesirov, J.P. (2011) Integrative genomics viewer. *Nat Biotechnol*, **29**, 24-26.
26. Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K. *et al.* (2018) The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res*, **46**, D794-D801.
27. Darnell, J.C., Mele, A., Hung, K.Y.S. and Darnell, R.B. (2018) Mapping of In Vivo RNA-Binding Sites by Ultraviolet (UV)-Cross-Linking Immunoprecipitation (CLIP). *Cold Spring Harb Protoc*, **2018**.
28. Fonseca, B.D., Zakaria, C., Jia, J.J., Graber, T.E., Svitkin, Y., Tahmasebi, S., Healy, D., Hoang, H.D., Jensen, J.M., Diao, I.T. *et al.* (2015) La-related Protein 1 (LARP1) Represses Terminal Oligopyrimidine (TOP) mRNA Translation Downstream of mTOR Complex 1 (mTORC1). *J Biol Chem*, **290**, 15996-16020.
29. Hong, S., Freeberg, M.A., Han, T., Kamath, A., Yao, Y., Fukuda, T., Suzuki, T., Kim, J.K. and Inoki, K. (2017) LARP1 functions as a molecular switch for mTORC1-mediated translation of an essential class of mRNAs. *Elife*, **6**, e25237.
30. Gingras, A.C., Raught, B. and Sonenberg, N. (2001) Regulation of translation initiation by FRAP/mTOR. *Genes Dev*, **15**, 807-826.
31. Choo, A.Y., Yoon, S.O., Kim, S.G., Roux, P.P. and Blenis, J. (2008) Rapamycin differentially inhibits S6Ks and 4E-BP1 to mediate cell-type-specific repression of mRNA translation. *Proc Natl Acad Sci U S A*, **105**, 17414-17419.
32. Aylett, C.H., Sauer, E., Imseng, S., Boehringer, D., Hall, M.N., Ban, N. and Maier, T. (2016) Architecture of human mTOR complex 1. *Science*, **351**, 48-52.
33. Chao, L.H. and Avruch, J. (2019) Cryo-EM insight into the structure of MTOR complex 1 and its interactions with Rheb and substrates. *F1000Res*, **8**, F1000-Faculty Rev-1014.
34. Yang, H., Jiang, X., Li, B., Yang, H.J., Miller, M., Yang, A., Dhar, A. and Pavletich, N.P. (2017) Mechanisms of mTORC1 activation by RHEB and inhibition by PRAS40. *Nature*, **552**, 368-373.
35. Thoreen, C.C., Chantranupong, L., Keys, H.R., Wang, T., Gray, N.S. and Sabatini, D.M. (2012) A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature*, **485**, 109-113.
36. Xie, J., Wang, X. and Proud, C.G. (2016) mTOR inhibitors in cancer therapy. *F1000Res*, **5**, F1000-Faculty Rev-2078.
37. Kang, S.A., Pacold, M.E., Cervantes, C.L., Lim, D., Lou, H.J., Ottina, K., Gray, N.S., Turk, B.E., Yaffe, M.B. and Sabatini, D.M. (2013) mTORC1 phosphorylation sites encode their sensitivity to starvation and rapamycin. *Science*, **341**, 1236566.
38. Philippe, L., van den Elzen, A.M.G., Watson, M.J. and Thoreen, C.C. (2020) Global analysis of LARP1 translation targets reveals tunable and dynamic features of 5' TOP motifs. *Proc Natl Acad Sci U S A*, **117**, 5319-5328.

39. Hsieh, A.C., Liu, Y., Edlind, M.P., Ingolia, N.T., Janes, M.R., Sher, A., Shi, E.Y., Stumpf, C.R., Christensen, C., Bonham, M.J. *et al.* (2012) The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature*, **485**, 55-61.
40. Kodzius, R., Kojima, M., Nishiyori, H., Nakamura, M., Fukuda, S., Tagami, M., Sasaki, D., Imamura, K., Kai, C., Harbers, M. *et al.* (2006) CAGE: cap analysis of gene expression. *Nat Methods*, **3**, 211-222.
41. Takahashi, H., Lassmann, T., Murata, M. and Carninci, P. (2012) 5' end-centered expression profiling using cap-analysis gene expression and next-generation sequencing. *Nat Protoc*, **7**, 542-561.
42. Iadevaia, V., Caldarola, S., Tino, E., Amaldi, F. and Loreni, F. (2008) All translation elongation factors and the e, f, and h subunits of translation initiation factor 3 are encoded by 5'-terminal oligopyrimidine (TOP) mRNAs. *RNA*, **14**, 1730-1736.
43. Gorski, J.J., Pathak, S., Panov, K., Kasciukovic, T., Panova, T., Russell, J. and Zomerdijk, J.C. (2007) A novel TBP-associated factor of SL1 functions in RNA polymerase I transcription. *EMBO J*, **26**, 1560-1568.
44. Iadevaia, V., Liu, R. and Proud, C.G. (2014) mTORC1 signaling controls multiple steps in ribosome biogenesis. *Semin Cell Dev Biol*, **36**, 113-120.
45. Thermann, R. and Hentze, M.W. (2007) Drosophila miR2 induces pseudo-polysomes and inhibits translation initiation. *Nature*, **447**, 875-878.
46. Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res*, **14**, 1188-1190.
47. Lahr, R.M., Fonseca, B.D., Ciotti, G.E., Al-Ashtal, H.A., Jia, J.J., Niklaus, M.R., Blagden, S.P., Alain, T. and Berman, A.J. (2017) La-related protein 1 (LARP1) binds the mRNA cap, blocking eIF4F assembly on TOP mRNAs. *Elife*, **6**, e24146.
48. Philippe, L., van den Elzen, A.M.G., Watson, M.J. and Thoreen, C.C. (2020) Global analysis of LARP1 translation targets reveals tunable and dynamic features of 5' TOP motifs. *Proc Natl Acad Sci U S A*.
49. Philippe, L., Vasseur, J.J., Debart, F. and Thoreen, C.C. (2018) La-related protein 1 (LARP1) repression of TOP mRNA translation is mediated through its cap-binding domain and controlled by an adjacent regulatory region. *Nucleic Acids Res*, **46**, 1457-1469.
50. Wang, X., Flynn, A., Waskiewicz, A.J., Webb, B.L., Vries, R.G., Baines, I.A., Cooper, J.A. and Proud, C.G. (1998) The phosphorylation of eukaryotic initiation factor eIF4E in response to phorbol esters, cell stresses, and cytokines is mediated by distinct MAP kinase pathways. *J Biol Chem*, **273**, 9373-9377.
51. Marcotrigiano, J., Gingras, A.-C., Sonenberg, N. and Burley, S.K. (1997) Cocrystal Structure of the Messenger RNA 5′ Cap-Binding Protein (eIF4E) Bound to 7-methyl-GDP. *Cell*, **89**, 951-961.
52. Leppek, K., Das, R. and Barna, M. (2018) Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol*, **19**, 158-174.
53. Grillo, G., Turi, A., Licciulli, F., Mignone, F., Liuni, S., Banfi, S., Gennarino, V.A., Horner, D.S., Pavesi, G., Picardi, E. *et al.* (2010) UTRdb and UTRsite (RELEASE 2010): a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. *Nucleic Acids Res*, **38**, D75-80.
54. Tomoo, K., Shen, X., Okabe, K., Nozoe, Y., Fukuhara, S., Morino, S., Sasaki, M., Taniguchi, T., Miyagawa, H., Kitamura, K. *et al.* (2003) Structural Features of Human Initiation Factor 4E, Studied by X-ray Crystal Analyses and Molecular Dynamics Simulations. *Journal of Molecular Biology*, **328**, 365-383.
55. Beggs, J.E., Tian, S., Jones, G.G., Xie, J., Iadevaia, V., Jenei, V., Thomas, G. and Proud, C.G. (2015) The MAP kinase-interacting kinases regulate cell migration, vimentin expression and eIF4E/CYFIP1 binding. *Biochem J*, **467**, 63-76.
56. Landgraf, P., Antileo, E.R., Schuman, E.M. and Dieterich, D.C. (2015) BONCAT: metabolic labeling, click chemistry, and affinity purification of newly synthesized proteomes. *Methods Mol Biol*, **1266**, 199-215.

57. Kenney, J.W., Genheden, M., Moon, K.M., Wang, X., Foster, L.J. and Proud, C.G. (2016) Eukaryotic elongation factor 2 kinase regulates the synthesis of microtubule-related proteins in neurons. *J Neurochem*, **136**, 276-284.
58. Liu, R., Kenney, J.W., Manousopoulou, A., Johnston, H.E., Kamei, M., Woelk, C.H., Xie, J., Schwarzer, M., Garbis, S.D. and Proud, C.G. (2016) Quantitative Non-canonical Amino Acid Tagging (QuaNCAT) Proteomics Identifies Distinct Patterns of Protein Synthesis Rapidly Induced by Hypertrophic Agents in Cardiomyocytes, Revealing New Aspects of Metabolic Remodeling. *Mol Cell Proteomics*, **15**, 3170-3189.
59. Jafarnejad, S.M., Chapat, C., Matta-Camacho, E., Gelbart, I.A., Hesketh, G.G., Arguello, M., Garzia, A., Kim, S.H., Attig, J., Shapiro, M. *et al.* (2018) Translational control of ERK signaling through miRNA/4EHP-directed silencing. *Elife*, **7**, e35034.
60. Peter, D., Weber, R., Sandmeir, F., Wohlbold, L., Helms, S., Bawankar, P., Valkov, E., Igreja, C. and Izaurralde, E. (2017) GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression. *Genes Dev*, **31**, 1147-1161.
61. Uniacke, J., Holterman, C.E., Lachance, G., Franovic, A., Jacob, M.D., Fabian, M.R., Payette, J., Holcik, M., Pause, A. and Lee, S. (2012) An oxygen-regulated switch in the protein synthesis machinery. *Nature*, **486**, 126-129.

Figure 1

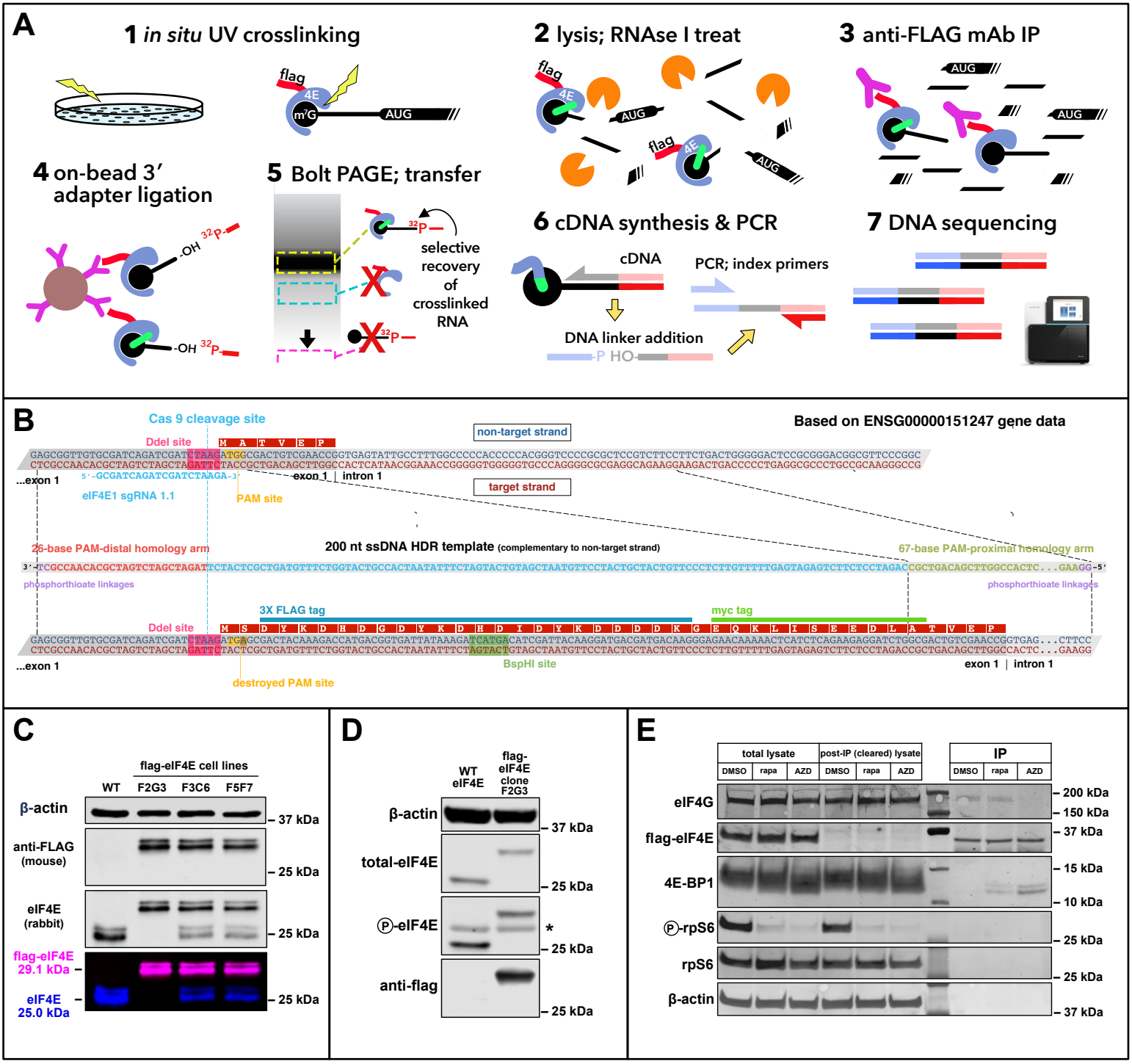

Figure 1. (A) Outline of the capCLIP method. See Results and Methods for additional details, and Supplemental Method 1 for the full capCLIP protocol. **(B)** Scheme for addition of a 3X-flag/1X-myc epitope tag (henceforth, a flag-eIF4E) to the N-terminus of human eIF4E using CRISPR/Cas9-mediated gene editing in Hela cells. **(C)** Anti-eIF4E/anti-flag westerns of WT and three CRISPR-edited flag-eIF4E Hela cell lines. At bottom is the original multi-channel LiCor image showing the 29.1 kDa flag-eIF4E and the 25 kDa WT-eIF4E bands detected by the anti-eIF4E antibody, and the 29.1 kDa flag-eIF4E band detected solely by the anti-flag antibody. **(D)** Anti-total-eIF4E/anti-P-eIF4E/anti-flag westerns of WT and 3X-flag-eIF4E Hela cells. The band marked with * observed in the anti-P-eIF4E blot is a non-specific band consistently seen when using this anti-P-eIF4E antibody. **(E)** Western blot of immunoprecipitates of flag-eIF4E from Hela cells. The IPs were analysed using anti-eIF4G, anti-eIF4E, and anti-4E-BP1 antibodies. Anti-P-rps6, and anti-rps6 antibodies were used to evaluate the ability of rapamycin and AZD to inhibit mTORC1 phosphorylation of rps6. Anti-actin is used as a loading control. Molecular weight marker lane lies between the depleted and IP samples. The relative loading volumes for total lysate, cleared lysate and IP lanes is 1:1:1 (thus the amount of protein used as *input* for each immunoprecipitation reaction is the same as the amount of protein loaded into each of the total lysate lanes).

Figure 2

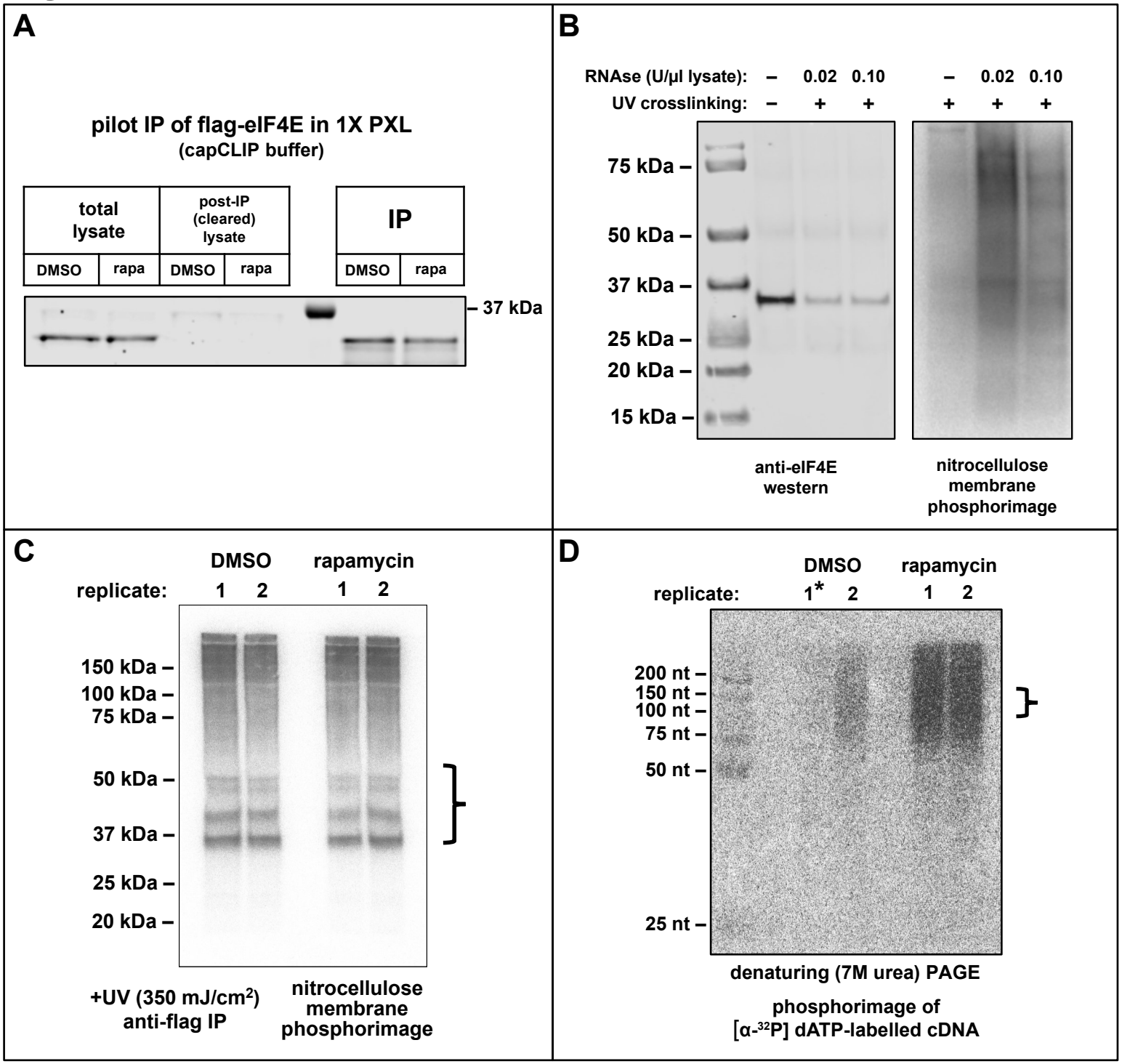

Figure 2. (A) Western analysis of flag-eIF4E immunoprecipitates from untreated (DMSO) and rapamycin-treated (rapa) Hela cells. The relative loading volumes for total lysate, cleared lysate and IP lanes is 1:1:2.5 (thus the amount of protein used as *input* for each immunoprecipitation reaction is 2.5-fold the amount of protein loaded into each of the total lysate lanes). **(B)** Left: Anti-eIF4E western of flag-eIF4E (29.1 kDa) immunoprecipitated from Hela cells in a pilot capCLIP experiment to test optimal RNase concentration. **(C & D)** 'Preparative' capCLIP experiment of flag-eIF4E Hela cells ± rapamycin treatment. **(C)** flag-eIF4E and complexed RNA was immunoprecipitated with anti-flag mAb, then ligated to a ^{32}P-labelled 3′ adapter. Phosphorimage shows ^{32}P-labelled material after PAGE separation and transfer to nitrocellulose. The bracket at right of the membrane indicates the portion of the membrane used for subsequent isolation of RNA. **(D)** Phosphorimage of a 7M urea denaturing PAGE cDNA sizing gel. Recovery of the short 30-80 nt RNA tags was inconsistent, and DMSO replicate 1 was poorly recovered. The bracket at right indicates the portion of the gel used for subsequent processing of the cDNA. The intensity of the cDNA signal is *not* an indication of the amount of crosslinked RNA present during crosslinking.

Figure 3

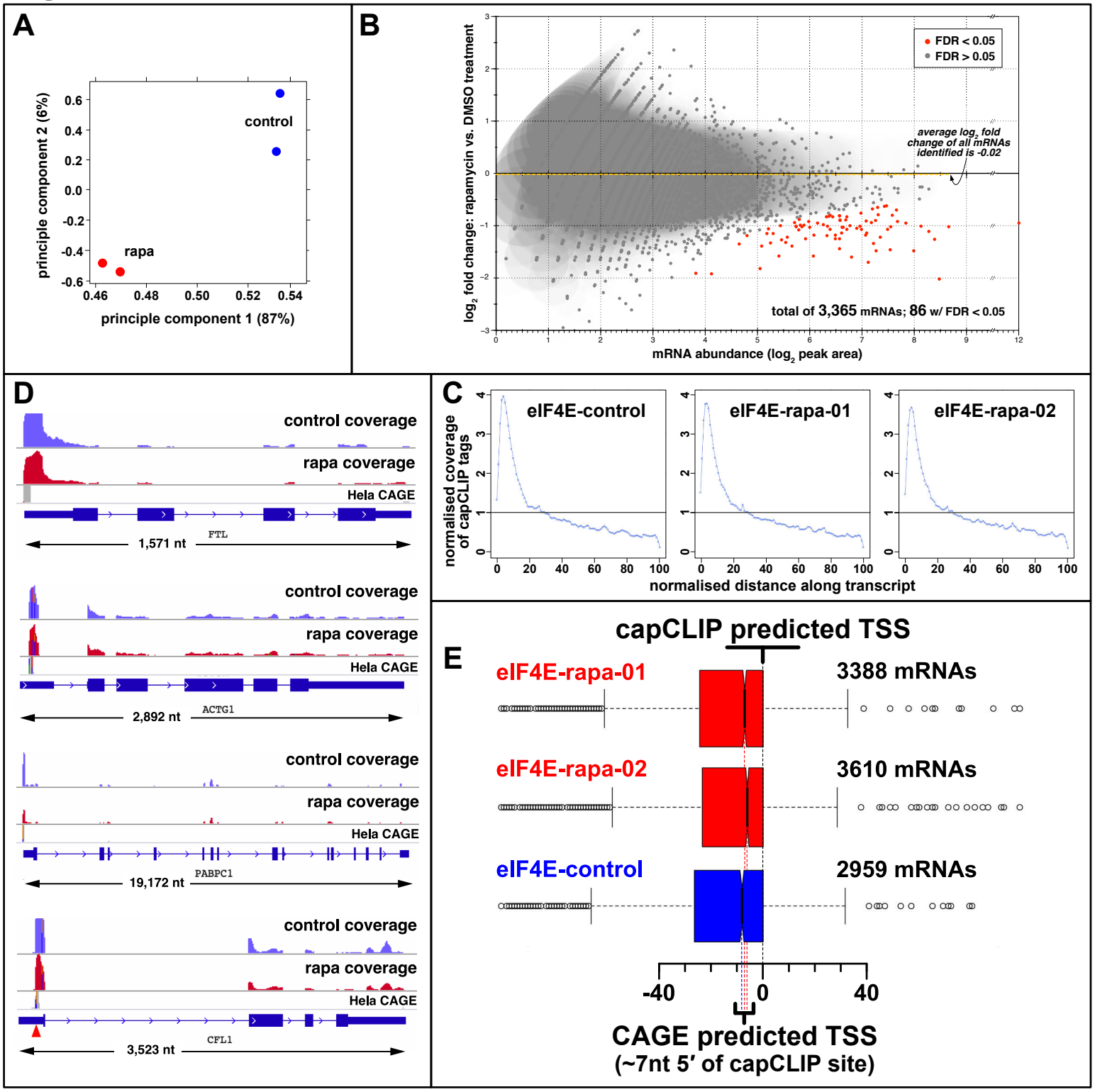

Figure 3. (A) Principal component analysis (PCA) of the two control and two rapamycin sequencing datasets. **(B)** DiffBind (24) data generated from the rapamycin capCLIP peak calculations was used to plot the $\log_2$ fold change in eIF4E affinity for individual mRNAs upon rapamycin treatment vs. the abundance of each mRNA (determined as $\log_2$ of each mRNA's peak area). Data was calculated for a total of 3,365 mRNAs, with 86 mRNAs showing statistically significant $\log_2$ fold-changes in binding to eIF4E with rapamycin treatment ($p > 0.01$; FDR>0.05). The average $\log_2$ fold-change in eIF4E binding of all 3,365 mRNAs upon rapamycin treatment, is -0.02. **(C)** The distribution of capCLIP tags as a function of the normalised distance along the transcript. Data is presented for the single, unsplit control (DMSO) replicate (see text) and two rapamycin (rapa) replicates. **(D)** Gene-level views derived from IGV showing capCLIP tag alignments from the control and combined rapamycin-treated sequencing datasets for four individual mRNAs: *FTL*, *PABPC1*, *CFL1* and *ACTG1*. IGV views of all four mRNAs have been formatted to run 5′ to 3′, left to right to aid visualisation of the capCLIP data. **(E)** Prediction of transcription start sites (TSSs) using Hela CAGE data and capCLIP peak-calling data. The box-whisker plots include, in red, the two rapamycin sequencing datasets and in blue, the unsplit control dataset. The right-hand edge of each box plot represents the position of the TSS defined by the capCLIP peak data and is defined as position 0 on the nucleotide scale below. The two red (rapa) and one blue (control) dashed lines descend from the calculated median TSSs predicted by the Hela CAGE data. The CAGE data consistently predicts each mRNA's TSS approximately 7 nt 'upstream' of the TSS predicted by capCLIP. The number of capCLIP peaks used for each box-whisker plot is indicated on the right-hand side of the figure.

Figure 4

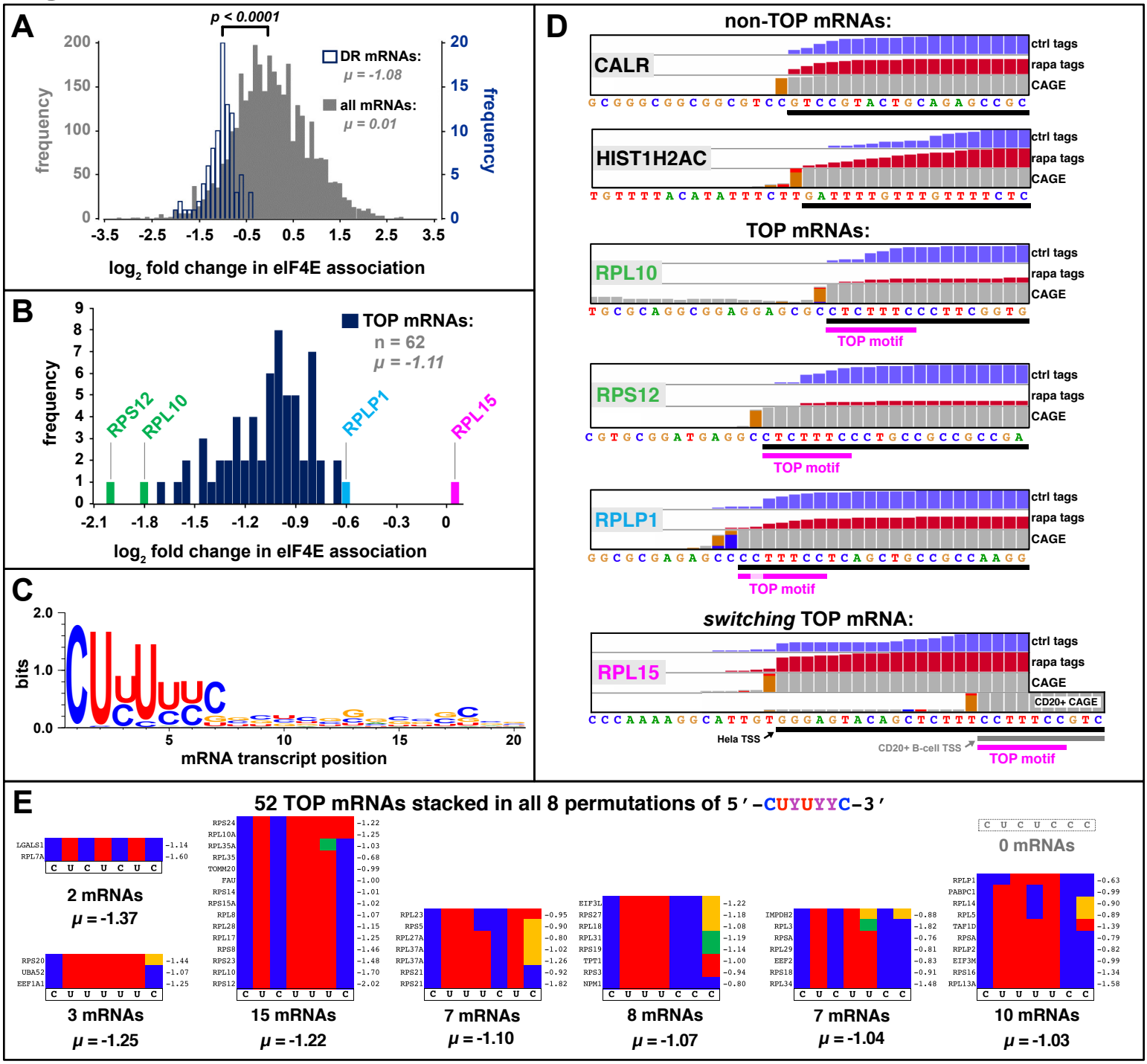

Figure 4. (A) A frequency plot of the $\log_2$ fold-change in eIF4E association upon treatment with rapamycin depicts all unchanged peak-called mRNAs (grey) and the subset of 86 mRNAs (blue) which show a statistically significant change with drug treatment (fold change <-0.5 or >0.5; FDR<0.05; p<0.01). The average fold-change of the group of 'all mRNAs' is μ= -0.01, while that of the 86 mRNAs is μ= -1.08. **(B)** The $\log_2$ fold-change in eIF4E association with rapamycin treatment is plotted for the 62 TOP mRNAs identified by capCLIP (average μ= -1.11). The TOP mRNAs exhibiting the greatest fold change with rapamycin treatment, *RPS12* and *RPL10* (in green), are perfect matches to one of the 8 permutations of the 'optimal' 7-nt TOP motif generated by weblogo/multiple sequence alignment of 52 capCLIP-identified TOP mRNAs (see text and below). *RPLP1*, the TOP mRNA which shows the weakest (but still statistically significant) change in eIF4E association with rapamycin treatment is labelled in light blue and differs from the 'optimal' TOP consensus at position 2, where it has a C instead of a conserved U. *RPL15*, a *switching*-TOP mRNA in Hela cells (see below and in text), does not change its association with eIF4E upon rapamycin treatment. **(C)** Weblogo3 analysis of the first 20 nucleotides of 52 of the capCLIP-identified TOP mRNAs where Hela CAGE data predicts the TSS with high confidence. **(D)** Views constructed from IGV data of the extreme mRNA 5′ ends of the six mRNAs *CALR*, *HIST1H2AC*, *RPL10*, *RPS12*, *RPLP1* and *RPL15*. IGV views of all gene 5′ ends have been formatted to run 5′ to 3′, left to right, to aid visualisation of the capCLIP data. The black bar immediately below the IVG genome sequence indicates the mRNA transcript for each gene, with the 5′-most nucleotide representing the TSS. At the top of each mRNA view are normalised histogram plots of control (in blue) and rapamycin treated (in red) capCLIP tag data, and the Hela CAGE data (in grey & orange) below. In general, the first grey bar to the right of an orange bar in the CAGE data line indicates the 1st nucleotide of the transcript (40). Below the black mRNA bar depicting the mRNA is the 7-nt TOP consensus sequence (in magenta) for the three TOP mRNAs and single *switching*-TOP mRNA. **(E)** Distribution of the 52 TOP mRNAs arranged into the 8 possible sequence permutations of the 7-nucleotide 'optimal' TOP motif depicted by the consensus motif in (C). Individual sequences are shown as coloured strips in 8 'stacks', with the individual nucleotide sequence of each of the eight sequence permutations listed below each stack. Sequences are coloured according to the 'classical' colouring scheme for RNA bases: red=U, blue=C, green=A and yellow=G, with individual base substitutions indicated by the corresponding alternative base colour. The identity of each TOP mRNA 'row' is indicated to the left of each stack, with the corresponding $\log_2$ fold change value (μ) of the mRNA listed to the right. No TOP mRNAs corresponding to the 8th sequence permutation (5′-CUCUCCC-3′) were identified.

Figure 5

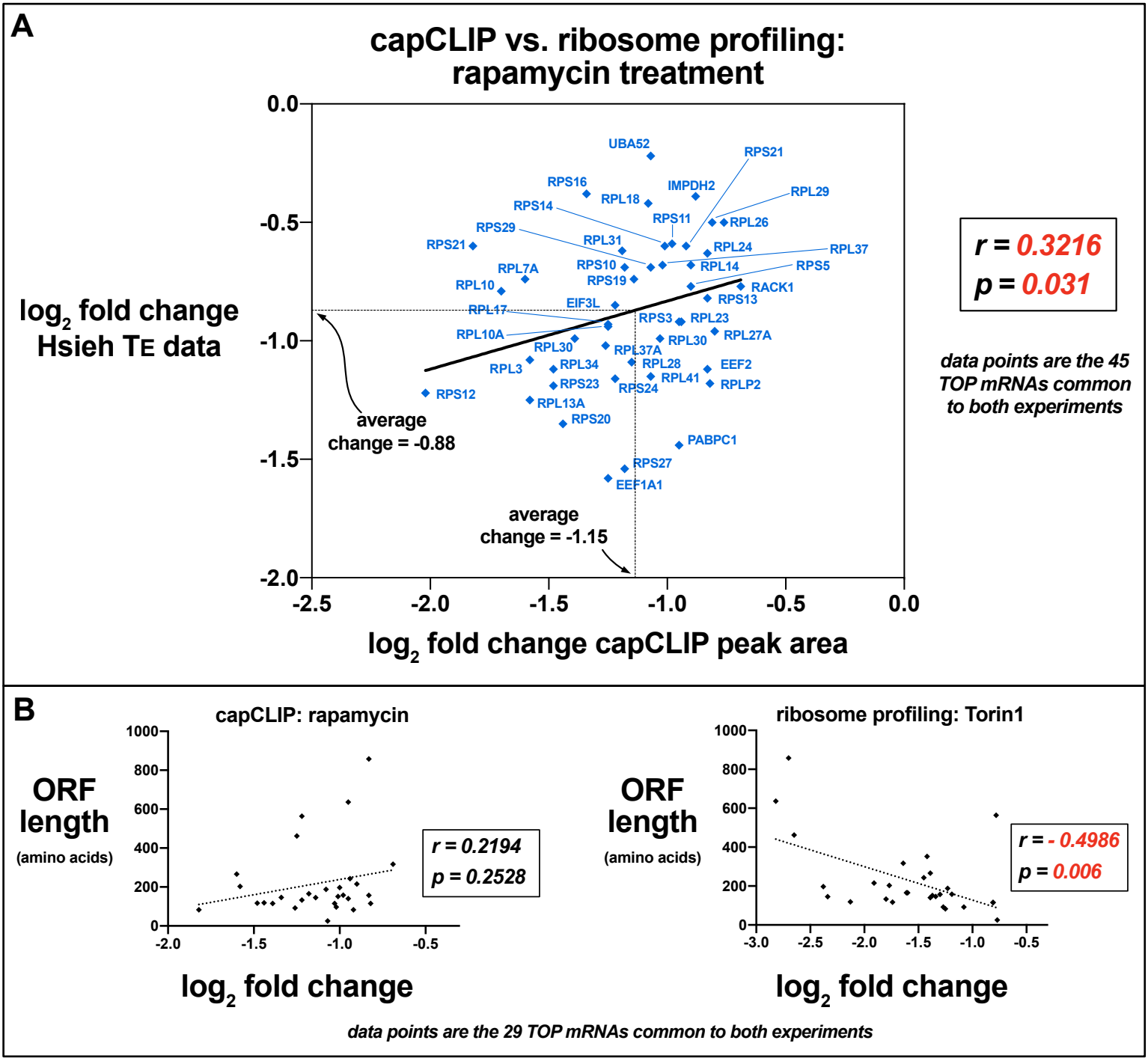

Figure 5. (A) Plot of 45 TOP mRNAs from the set of 62 TOP mRNAs identified by capCLIP for which a calculated T_E value from ribosome profiling experiments using rapamycin could be located in the data of Hsieh *et al.* (39). For each TOP mRNA the $\log_2$ fold change of T_E with rapamycin treatment is plotted vs. the $\log_2$ fold change of eIF4E binding with rapamycin treatment. Individual mRNAs are labelled. Average $\log_2$ fold change of eIF4E binding for the 45 TOP mRNAs is μ= -1.15; average $\log_2$ fold change of T_E for the 45 TOP mRNAs is μ= -0.88. Correlation of the two datasets using Spearman r is 0.3216, with p=0.0312. **(B) Left:** plot of $\log_2$ fold change of eIF4E binding ± rapamycin treatment vs. ORF length in amino acids for 29 TOP mRNAs. The 29 mRNAs are a subset of the 62 TOP mRNAs identified by capCLIP for which a calculated T_E value (ribosome profiling experiments ± Torin-1) could be located in the data of Thoreen *et al.* (35). Correlation of capCLIP eIF4E binding changes with rapamycin treatment to ORF length using Spearman r is 0.2194, with p=0.2528. **Right:** plot of $\log_2$ fold change of T_E ± rapamycin treatment vs. ORF length with the same set of 29 mRNAs Correlation of T_E change with rapamycin treatment to ORF length using Spearman r is -0.4986, with p=0.006.

Figure 6

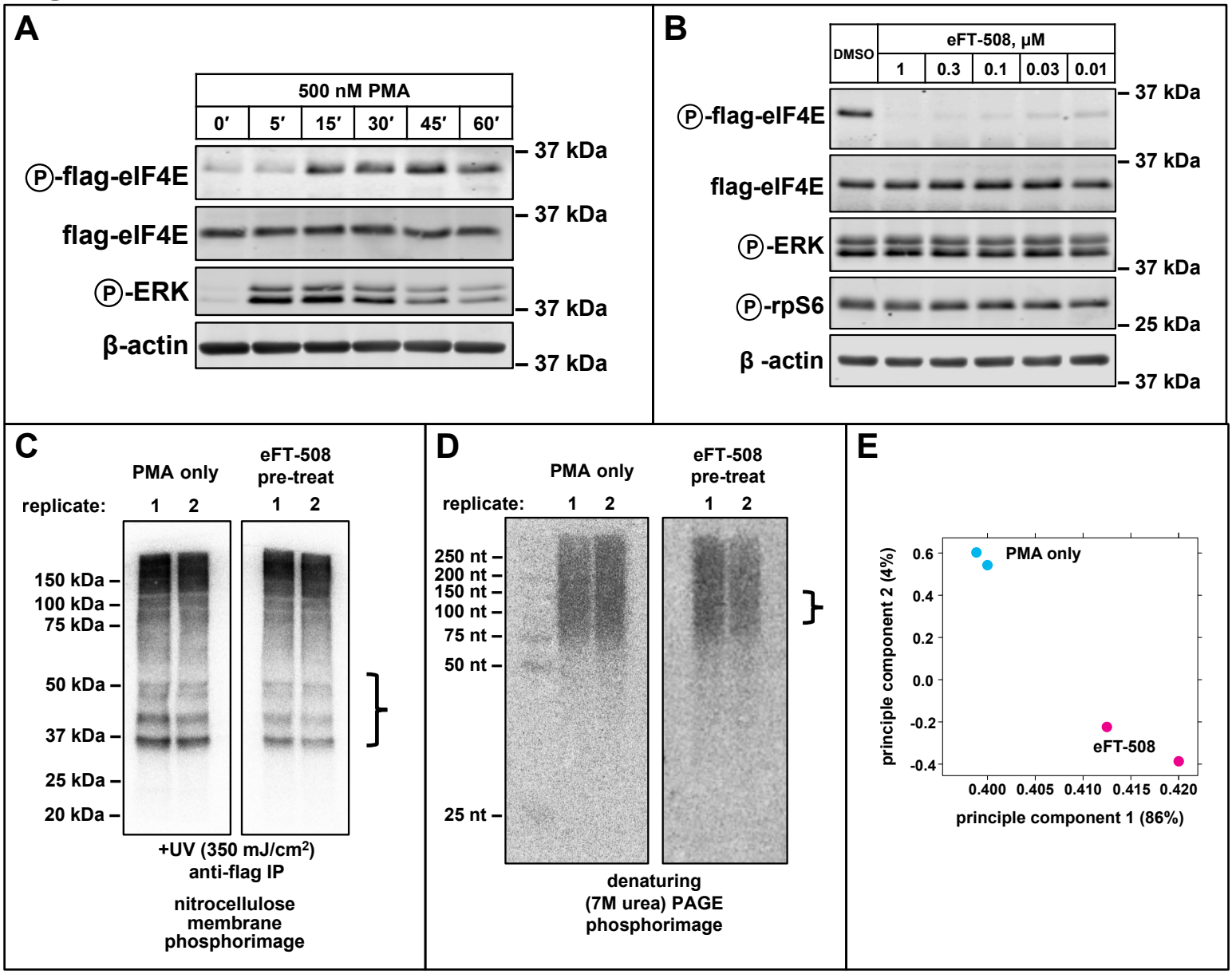

Figure 6. MNK capCLIP data **(A)** Western analysis of P-flag-eIF4E stimulation in flag-eIF4E Hela cells, Cells were serum-starved overnight and either left untreated, or treated for increasing times with 500 nM PMA. Continued activation of MAPK/ERK signalling was confirmed by probing for P-ERK. Maximal eIF4E phosphorylation was observed by 30 min. **(B)** eFT-508 dose-response experiment in flag-eIF4E Hela cells. Cells were serum-starved overnight and the pre-treated for 1 h with DMSO only (control) or with the indicated concentrations of the MNK inhibitor eFT-508. Each sample was then stimulated for an additional 30 min with 500 nM PMA. Continued activation of MAPK/ERK signalling was confirmed by probing for P-ERK, and mTORC1 activity was assayed by probing for P-rps6. Close to maximal inhibition of eIF4E phosphorylation was observed at a minimal eFT-508 concentration of 0.1 μM. **(C & D)** 'Preparative' capCLIP experiment of flag-eIF4E Hela cells ±eFT-508-mediated inhibition of eIF4E phosphorylation. **(C)** Flag-eIF4E Hela cells were serum-starved overnight, incubated for 1 h with DMSO or with 0.1 μM eFT-508, then stimulated with 500 nM PMA for an additional 30 min. Flag-eIF4E and complexed RNA was immunoprecipitated with anti-flag mAb then ligated to a ^{32}P-labelled 3′ adapter. Phosphorimage shows ^{32}P-labelled material after PAGE separation and transfer to nitrocellulose. The bracket at right of the membrane indicates the portion of the membrane used for subsequent isolation of RNA from the 2 control and 2 eFT-508 treated replicates. **(D)** Phosphorimage of a 7M urea denaturing PAGE cDNA sizing gel. The bracket at right indicates the portion of the gel used for subsequent processing of the cDNA from each replicate. **(E)** Principal component analysis (PCA) of the two control (PMA only) DNA sequencing datasets and two eFT-508 pre-treated (eFT-508) DNA sequencing datasets.

Figure 7

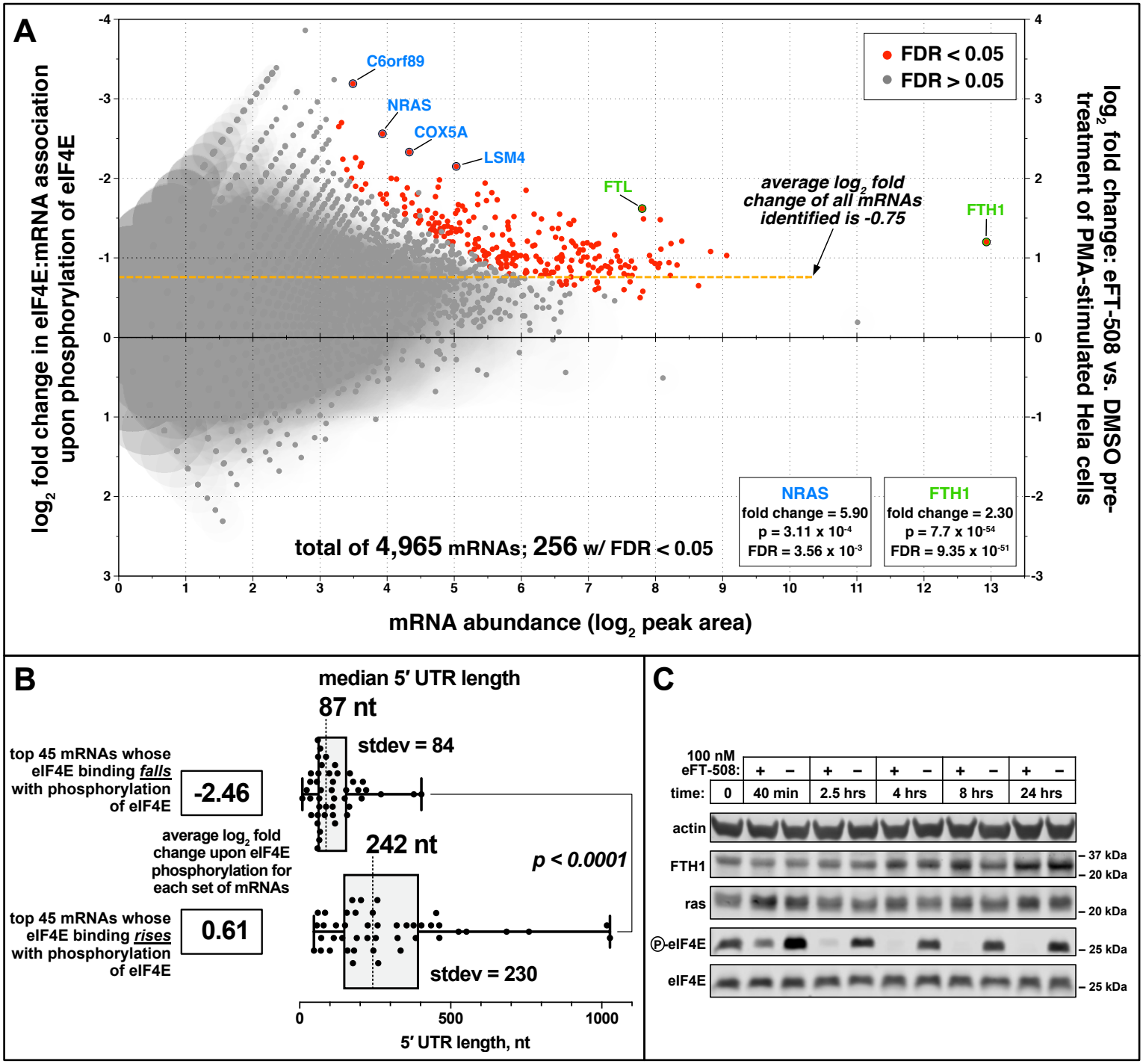

Figure 7. (A) DiffBind (24) data generated from the eFT-508 capCLIP peak calculations was used to plot each mRNA's abundance (x-axis; $\log_2$ of the capCLIP peak area) vs. the $\log_2$ fold change in eIF4E binding upon eIF4E phosphorylation (left y-axis). The right y-axis expresses the results with respect to the underlying PMA/eFT-508 treatment paradigm. Data was calculated for a total of 4,965 mRNAs, with 256 mRNAs showing a significant change in $\log_2$ fold change binding to eIF4E upon its phosphorylation ($p > 0.01$; FDR>0.05). The average shift in mRNA $\log_2$ fold change, that is, the average change in binding to eIF4E of *all* 4,965 mRNAs upon eIF4E phosphorylation, is -0.75. **(B)** Box and whisker plots showing the 5′ UTR length distribution for the 45 mRNAs with the greatest *reduction* in eIF4E binding upon eIF4E phosphorylation (top) and the 45 mRNAs with the greatest apparent *increase* in eIF4E binding upon eIF4E phosphorylation (bottom). **(C)** Western blot of protein levels in WT Hela cells growing in 10% FBS & DMEM. Cells were treated with DMSO only or with 100 nM eFT-508 for the indicated times. A T=0 'no treatment' sample was also collected. Inhibition of eIF4E phosphorylation by eFT-508 was monitored with anti-P-eIF4E antibody. Actin and eIF4E antibodies were used to control as loading controls.

Table 1

1A. rapamycin capCLIP

eIF4E-ctrl set 1A	# of peaks	log_2 enrichment	% peaks
3′ UTR	109	1.68	2.6
miRNA	11	5.01	0.3
ncRNA	60	2.70	1.4
pseudo	8	1.45	0.2
exon	585	3.62	13.9
intron	759	-1.14	18.0
intergenic	610	-2.00	14.4
5′ UTR	**1941**	**7.51**	**46.0**
snoRNA	129	11.17	3.1
snRNA	10	10.29	0.2

eIF4E-ctrl set 1B	# of peaks	log_2 enrichment	% peaks
3′ UTR	121	1.81	2.8
miRNA	8	4.52	0.2
ncRNA	54	2.53	1.3
pseudo	10	1.75	0.2
exon	573	3.56	13.3
intron	806	-1.08	18.8
intergenic	627	-1.99	14.6
5′ UTR	**1955**	**7.49**	**45.5**
snoRNA	130	11.15	3.0
snRNA	12	10.52	0.3

eIF4E-rapa set 1	# of peaks	log_2 enrichment	% peaks
3′ UTR	689	2.71	5.3
miRNA	24	4.50	0.2
ncRNA	166	2.54	1.3
pseudo	38	2.07	0.3
exon	3180	4.43	24.3
intron	2578	-1.01	19.7
intergenic	1815	-2.06	13.8
5′ UTR	**4412**	**7.06**	**33.7**
snoRNA	187	10.07	1.4
snRNA	18	9.50	0.1

eIF4E-rapa set 2	# of peaks	log_2 enrichment	% peaks
3′ UTR	713	2.82	5.7
miRNA	19	4.23	0.2
ncRNA	182	2.74	1.5
pseudo	43	2.31	0.3
exon	3045	4.43	24.3
intron	2395	-1.05	19.1
intergenic	1705	-2.09	13.6
5′ UTR	**4193**	**7.05**	**33.5**
snoRNA	194	10.19	1.6
snRNA	20	9.72	0.2

1B. eFT-508 capCLIP

eIF4E-PMA set 1A	# of peaks	log_2 enrichment	% peaks
3′ UTR	442	2.55	4.7
miRNA	19	4.64	0.2
ncRNA	150	2.87	1.6
pseudo	22	1.76	0.2
exon	2964	4.77	31.6
intron	1769	-1.07	18.8
intergenic	1071	-2.34	11.4
5′ UTR	**2764**	**7.26**	**29.4**
snoRNA	170	10.41	1.8
snRNA	19	10.06	0.2

eIF4E-PMA set 1B	# of peaks	log_2 enrichment	% peaks
3′ UTR	433	2.52	4.6
miRNA	19	4.64	0.2
ncRNA	156	2.93	1.7
pseudo	27	2.06	0.3
exon	2891	4.73	30.9
intron	1781	-1.06	19.0
intergenic	1074	-1.06	19.0
5′ UTR	**2806**	**7.28**	**29.9**
snoRNA	166	10.38	1.8
snRNA	17	9.90	0.2

eIF4E-eFT set 1	# of peaks	log_2 enrichment	% peaks
3′ UTR	503	2.23	3.8
miRNA	16	3.89	0.1
ncRNA	198	2.77	1.5
pseudo	37	2.01	0.3%
exon	3657	4.57	27.6
intron	2941	-0.84	22.2
intergenic	1538	-2.32	11.6
5′ UTR	**4202**	**7.36**	**31.7**
snoRNA	158	9.81	1.2
snRNA	14	9.12	0.1%

eIF4E-eFT set 2	# of peaks	log_2 enrichment	% peaks
3′ UTR	562	2.51	4.6
miRNA	19	4.26	0.2
ncRNA	173	2.70	1.4
pseudo	37	2.13	0.3%
exon	3193	4.50	26.1
intron	2836	-0.77	23.2
intergenic	1483	-2.25	12.1
5′ UTR	**3735**	**7.31**	**30.6**
snoRNA	159	9.94	1.3
snRNA	17	9.52	0.1

Table 1. (1A) rapamycin capCLIP sequencing statistics. Distribution of peak-calls from the two control pseudo-replicate sequencing datasets and from the two rapamycin-treated replicate sequencing datasets. (**1B) eFT-508 capCLIP sequencing statistics.** Distribution of peak-calls from the two PMA-only sequencing datasets and from the two eFT-508-treated replicate sequencing datasets.