Title
In vivo Proximity labeling of translation initiation complexes and repressive RNA granules by APEX-Seq

Author
Padron, Alejandro J.

Publication Date
2019

Peer reviewed|Thesis/dissertation

In vivo Proximity labeling of translation initiation complexes and repressive RNA granules by APEX-Seq

by

Alejandro Padrón

A dissertation submitted in partial satisfaction of the
requirement for the degree of
Doctor of Philosophy
in
Molecular and Cell Biology
and the Designated Emphasis
in
Computational and Genomic Biology
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Nicholas T. Ingolia, Chair
Professor Abby F. Dernburg
Professor Hernan G. Garcia
Professor Mohammad R. K. Mofrad

Spring 2019

Abstract

In vivo Proximity labeling of translation initiation complexes and repressive RNA granules by APEX-Seq

by

Alejandro Padrón

Doctor of Philosophy in Molecular and Cell Biology

and the Designated Emphasis in

Computational and Genomic Biology

University of California, Berkeley

Professor Nicholas Ingolia, Chair

RNA localization is a fundamentally important regulatory mechanism for the control of gene expression. I developed an *in vivo* subcellular RNA proximity labeling technique called APEX-Seq to monitor RNA localization and organization in the cell. In my approach, a proximity labeling enzyme is fused to a query protein and rapidly (<1 minute) biotinylates nearby RNAs *in vivo* upon addition of a labeling substrate, allowing their subsequent purification and analysis. We show first that APEX-Seq can distinguish the localization of transcripts to different regions of the cell. Further, we can detect the enrichment of specific transcripts in proximity to translation initiation proteins, showing that our system can capture more subtle, RNA-protein localization patterns. Finally, we present a high-resolution time course of protein and RNA condensation into stress induced RNA granules. A powerful aspect of APEX-Seq is the ability to match the spatial transcriptome with quantitative spatial proteomics, allowing for a more complete picture of the spatial landscape of the cell. Additionally, I rationally engineered a split APEX enzyme for conditional spatial proteomics and transcriptomics. I found evidence of unique translation initiation complexes through the use of split APEX on the initiation factors eIF1A/eIF4H, and eIF4A/eIF4B. Overall, these tools open the door for comprehensive and high throughout spatial transcriptomics, and proteomics with subcellular resolution.

Table of Contents

Dedication

To myself.

For working when you did not want to work, and for thinking when you did not want to think. For focusing on yourself. For making the best out of each situation. For loving yourself, and for your devotion to your craft. For wanting to be the best version of yourself. For always showing up, and for never giving up.

Chapter 1: RNA localization and organization in the cell

RNA localization is a fundamental process used by cells to build better information processing devices. Localized information processing is so crucial to a cell, that the instructions for localization of RNA is hardcoded in the genome. Several mechanisms exists in place to send RNAs to their required sites. Localization through random diffusion relies on local entrapment of RNAs or RNA anchoring at target sites through stochastic encounters, which may arise through cytoplasmic streaming to target sites (Suter 2018; Weil 2014). Although random diffusion and streaming appear to be energetically parsimonious methods of localizing RNA, they are likely wasteful, given that a relatively small percentage of RNAs will become localized at the target sites of interest whereas nonlocalized RNAs are targeted for degradation (Wolke et al. 2002). More efficient, targeted, mechanisms exist for RNA localization; namely, through the use of molecular motors and the actin cytoskeleton. Indeed, all three families of molecular motors (dynein, kinesin, and myosins) are involved in RNA localization (Gagnon and Mowry 2011). For instance, in highly polarized cells like neurons, dendritic mRNAs have been shown to move in both retro and anterograde fashion, however, movement is more biased toward the distal dendritic end, which allows them to be delivered to their target sites, allowing RNAs to participate in local translation when required (Das, Singer, and Yoon 2019). A classic example of axonal RNA localization is the beta-actin mRNA, whose cDNA was detected in squid axoplasm (Bassell and Singer 2001). The *trans*-acting RNA binding protein, Zipcode binding protein (ZBP1), associates with beta-actin in the perinuclear space of the soma and localizes beta-actin mRNA into translationally repressed transport granules for proper axonal transport (Suter 2018; Nicastro et al. 2017). The ZBP1-beta actin mRNA ribonucleoprotein complex (mRNP) is moved along microtubules through interactions between the RRM domains of ZBP1 and the KIF11 molecular motor (Song et al. 2015). Once the beta-actin RNA reaches its target site, post translational modifications on ZBP1 causes beta actin RNA release and subsequent translation (Hüttelmaier et al. 2005; Wu et al. 2015; Nicastro et al. 2017). Molecular motors like kinesin and dynein also serve as local anchors to tether RNAs in place after reaching either the plus or minus end of microtubules. In simpler cells like fibroblasts, beta-actin localization is implicated in cell motility (Kislauskis, Zhu, and Singer 1997). Localized protein synthesis of actin in neurons may function in filopodial protrusion and process outgrowth (Bassell and Singer 2001).

Mechanisms of localizing RNA. Patterns of RNA localization are particularly important at establishing the spatially restricted patterns of gene expression seen during development. A classic developmental model for RNA localization is the *Drosophila* oocyte where the fly syncytium leverages structural scaffolds for macromolecular transportation and localization. A classic example of this are microtubules, which are essential for RNA localization and transportation (Pokrywka and Stephenson 1991). In the developing fly, *bicoid* RNA is localized toward the anterior end of the embryo (Berleth et al. 1988). The localization of bicoid RNA results in an asymmetric localization pattern resulting in the synthesis of bicoid protein wherein most of the protein is toward the anterior end and protein concentration exponentially decreases toward the posterior end of the embryo. These high anterior bicoid concentrations are crucial for proper anterior cell fate decisions (Driever and Nüsslein-Volhard 1988). During mid-oogenesis, *bicoid* RNAs are synthesized in the nurse cells of the fly. The *bicoid* RNA is then transported into the anterior of the fly oocyte. Through a series of genetic and biochemical experiments, the *bicoid* RNA has been shown to interact with the double-stranded RNA binding protein Staufen through regulatory elements in the 3'-UTR (MacDonald, Leask, and Kerr 1995; Dominique Ferrandon et al. 1994). Staufen itself is implicated in the control of localization of certain RNAs through the formation of transportation granules (D. Ferrandon et al. 1997). The *bicoid* RNA itself is moved toward the anterior end of the fly through interactions between RNA binding proteins as well as microtubule-dependent motor proteins (Theurkauf and Hazelrigg 1998). The types of motor proteins and RNA binding proteins involved depends on where the RNAs are localized, and which RNAs are being targeted. The *bicoid* RNA is also regulated through interactions with the RNA binding protein Swallow. The Swallow protein has been shown to directly associate with the dynein motor protein, which moves in a minus-end directed manner on microtubules. Moreover, Swallow colocalizes with *bicoid* as well as with dynein in the anterior region during early fly development (Schnorrer, Bohmann, and Nüsslein-Volhard 2000; Lipshitz and Smibert 2000).

Localization in bacteria*.* RNA localization has classically been studied and considered in developing embryos or in highly polarized cells like neurons. Much less attention has been given to RNA localization patterns in relatively smaller cells (~1 μm), such as bacteria. This lack of attention to bacterial RNA localization stems from the spatial and temporal overlap of transcription and translation in bacterial cells. More strikingly, bacterial cells lack classic membrane bound organelles, and traditionally have been assumed to lack complex subcellular

RNA localization, and even protein localization mechanisms. Despite these outstanding differences, previous work has reported co-transcriptional localization of certain mRNAs, namely, *groESL, and ptsC* mRNAs in *C. crescentus*, and *lacZ* in *E. coli (Montero Llopis et al. 2010; Fei and Sharma 2018)*. Indeed, additional work has shown asymmetric patterns of RNA diffusion through the use of the MS2 reporter system (Fei and Sharma 2018; Henkin 2004). Although most transcript copies visualized through the MS2 reporter system appear to move randomly near their sites of transcription, a minor population of the reporter RNA was seen diffused throughout the cell or moving in helical patterns resembling chains (Henkin 2004; Fei and Sharma 2018). Additionally, other studies have reported on subcellular patterns of RNA localization in bacteria. A clear example comes from fluorescence in situ hybridization (FISH) studies showing distinct patterns of localization at both poles for the *nifH* RNAs, which encode for a dinitrogenase reductase protein in *A. vinelandii (Pilhofer et al. 2009)*. Moreover, in *E. coli* the *lacY* RNA, encoding for a membrane bound lactose permease, was localized near cytoplasmic membranes, whereas *cat* RNAs were shown to exhibit a helix-like pattern of localization (Nevo-Dinur et al. 2011). Translation can occur at distinct sites in the cell, like near membranes during cotranslational secretion of inner membrane proteins (Fei and Sharma 2018). In bacteria as in eukaryotes, the localization of where RNAs can control the sites of translation. Gaining control over where enzymatic or certain protein interactions occurs are complex, and evolved control mechanisms that are present even in simple bacteria. Gene expression itself is impacted by the spatial distribution of RNAs. Indeed, mRNAs are localized prior to the production of the proteins they encode for. The localization of RNAs, rather than protein, likely serves as a translational robustness mechanism, since it allows many proteins to be produced *in situ*, and on-demand, while minimizing the per molecule energy expenditure in localization, and reducing loss or degradation due to stochasticity. Moreover, combining RNA localization with translational control allows a cell to control protein synthesis independently in different parts of the cell (Alberta et al. 2008). RNA localization is at the heart of cell biology and development processes, such as symmetry breaking during development, cell migration, and cellular polarity (Medioni, Mowry, and Besse 2012).

Phase transitions and RNA localization*.* Cells are extremely dense and concentrated enclosures filled with large amounts of small molecules, RNAs, and proteins. The crowding phenomenon is in part a reflection of the diverse genetic programs cells have evolved to perform in order to combat complex, and changing evolutionary pressures. Molecular crowding

also influences enzymatic reactions by increasing the effective concentration of their substrates, and in some cases, forcing higher degrees of specificity for substrates. In *E. coli* for instance, macromolecular concentrations in the cytoplasm are approximately 300 mg/mL, of which ~ 220 mg/mL is made up of protein (Cayley et al. 1991). Despite the benefits to crowdedness, there are major drawbacks all cells have evolved to deal with. One major drawback to having a crowded cytoplasm is the potential for protein misfolding and aggregate formation, which are often detrimental to cells; in particular, to neuronal cells where protein misfolding manifests phenotypically as neurodegenerative diseases like Amyotrophic lateral sclerosis (ALS) (Parakh and Atkin 2016). Intriguingly, cells appear to have developed ways of combating protein solubility through the use organic hydrotropes. A hydrotrope is an amphiphilic molecule that improves the solubility of hydrophobic molecules. Important work from the Hyman lab has now shown that ATP appears to behave as a hydrotrope, increasing the solubility of proteins and preventing their aggregation (Patel et al. 2017). Indeed, ATP itself is present at concentrations of 5-10 mM, whereas the K_m of most enzymes is below 100 μM, which suggests that the surplus of ATP in the cell is being used for some other purpose; namely, assisting in protein solubility according to convincing evidence from the Hyman lab (Patel et al. 2017). The work raising several questions, one being how are the other highly concentrated small molecules influencing the solubility properties of macromolecules? This observation stems from other small molecule with hydrotrope-like properties like glutamate, which is present in ~100 mM (Milo and Phillips 2015) in the cell – ten times more concentrated than ATP.

Having densely crowded environments has resulted in cells leveraging the physical chemical properties that arise from this crowding. An example of this are liquid-liquid phase separations (Brangwynne et al. 2009). In the worm, P granules named after the embryonic lineage that gives rise to the germline (J. T. Wang and Seydoux 2014), form during development and are necessary for proper sex tissue development. Mutations in individual P granule proteins appears to result in sterility, and impaired translation of some mRNAs. However, mutants that fail to partition P granules to the P lineage appear to be viable and fertile, which suggests the phenomenological formation of P granules is not essential to distinguish soma from germline embryo (J. T. Wang and Seydoux 2014). However, these conclusions have been drawn on imperfect experimental design due mainly to the complexities of studying phase separated condensates. Expanding on this further, it is challenging to disambiguate whether the mutation under study is influencing P granule formation alone or whether other functions of the protein

being interrogated is also being impacted. It is also unclear whether other macromolecular condensates are moonlighting for P granule function. To better ascertain the roles of P granules and other macromolecular condensates, new tools are needed to decouple the manipulation of condensate formation and protein function.

The Brangwynne lab has now developed a synthetic granule called the OptoDroplet with properties similar to those of natural macromolecular condensates (Shin et al. 2017). This new tool should help shed light on what the functional consequence of forming macromolecular condensates is for a cell, and whether different phase transitions functionally matter for cells. Recent work has now suggested that liquid-like condensates exhibit a form of long-term spatial memory for a cell; in other words, forming condensates in one side of a cell can drive asymmetric patterns of protein localization in minutes, and these patterns can persist for hours (Dine et al. 2018).

RNA granules during stress. If phase separations are misregulated they can lead to a phase transition, where the liquid liquid-like properties of the condensate transitions into a solid, less dynamic structure. A classic example of a phase transition are found in neurodegenerative diseases such as ALS (Parakh and Atkin 2016), where protein misfolding leads to aggregate formation. Interestingly, examples of known phase transitions are thought to be deleterious to cells. In contrast to this observation, prion aggregates are heritable and are thought to be a common mechanism for phenotypic inheritance. Interestingly, a physiological function of the Sup35 prion domain was recently teased out. The prion domain of Sup35 proteins appears to monitor cellular pH levels, through a cluster of negatively charged amino acids, and forms Sup35 condensates that are liquid-like initially, and subsequently solidify to form protective protein gels (Franzmann et al. 2018).

Cellular stress forms liquid-liquid like condensates in mammalian cells, namely stress granules. When a cell senses stress, translation initiation is inhibited, and stress granules form through interactions between key RNA binding proteins (RBPs). The interactions that drive stress granule formation are mediated by the intrinsically disordered regions (IDRs) of certain RBPs. Moreover, the material properties of a stress granule can change, in addition to their composition, and dynamic behavior depending on whether the ALS-associated variants of SOD1 accumulate in stress granules. Cells therefore appear to have evolved surveillance

mechanisms to monitor the composition of stress granules and prevent the conversion into an aberrant form (Mateju et al. 2017).

Interestingly, there appears to be an inverse correlation between stress granule formation and translation. If translation is inhibited at the stage of translation initiation, through the phosphorylation of eIF2α or the inhibition of eIF4A, robust stress granules form within minutes (Panas, Ivanov, and Anderson 2016). However, if elongating ribosomes are directly immobilized on an RNA through cycloheximide, stress granules fail to form suggesting the presence of the ribosome on mRNAs itself prevents stress granule formation. Although much work has now shed light on how stress granules form through IDRs and protein protein interactions, less attention has been paid to how RNAs are contributing to stress granule formation. Recent work has shown that RNAs become highly structured upon ribosome poison and disassembly from mRNAs (Adivarahan et al. 2018). Perhaps by allowing for RNA structure, particularly in the CDS, which would otherwise be prevented through active translation, mRNAs themselves scaffold protein-RNA interactions for key RBPs to help nucleate stress granule formation.

Moreover, RNA granules are notoriously challenging to study given their fluid and dynamic nature. RNA granules are also quite large, ranging from a few hundred nanometers to ~ 3 μm in the worm. Therefore conventional techniques like crosslinking, which work over the angstrom length scale are insufficient to capture their full complexity. Immunoprecipitation is also insufficient, due to the liquid-like properties of RNA granules, resulting in disassembly during cell lysis and partial recovery of RNA granule components. Therefore, new tools are needed in order to ask questions such as what makes up these dynamic structures, and what

Proximity labeling. Over the years a number of different groups have engineered several proximity labeling enzymes that catalyze the biotinylation of surrounding proteins wherever the enzyme was localized. The radius of biotinylation varies by enzyme. Biotin ligase (BirA), and its variants (Roux et al. 2012; Varnaitė and MacNeill 2016; Kim et al. 2016; Oostdyk et al. 2019), catalyze protein biotinylation through an adenylate ester intermediate which has a half life in the minute timescale. Ascorbate peroxidase (APEX) (Hung et al. 2014; Lam et al. 2015; Hung et al. 2016; Lee et al. 2016; Martell et al. 2017) on the other hand also catalyzes a biotinylation reaction through a radical based reaction. Although conceptually similar, the reaction mechanism is fundamentally different with APEX than BirA. Indeed, APEX produces a short

lived (millisecond timescale) biotin-phenoxyl radical intermediates resulting in a tighter labeling radius of ~ 20-40 nm (Varnaitė and MacNeill 2016).

Although proximity labeling has been quite informative and has served as a useful tool for understanding subcellular organization, methods that increase specificity at target sites are useful to engineer and develop for a more precise understanding of the biology in question. Several labs have now engineered split proximity labeling enzymes (Isabel Myriam Schopp et al. 2017; De Munter et al. 2017; Isabel M. Schopp and Béthune 2018; Han et al. 2019; Martell et al. 2016). These valuable tools now give the user quadratic specificity, since enzymatic activity depends on two, rather than a single, component.

Given that APEX labels neighboring proteins through a radical chain reaction, and knowing that nucleic acids are amenable to free radical based chemistry APEX could in principle be leverage not only for protein proximity labeling, but also for RNA proximity labeling. Indeed, a power aspect of APEX is the ability to capture both the spatial proteome as well as the spatial transcriptome. Combining both of these aspects allows for the a more complete picture of an organelle of interest. Not only that, but APEX now serves as a tool to monitor liquid-liquid phase separation of molecular condensates over both time and space.

Chapter 2: Proximity labeling by APEX-Seq reveals the organization of translation initiation complexes and repressive RNA granules

Contents of this chapter were originally posted as a preprint on bioRxiv on October 25, 2018 as:

Padrón A. Iwasaki S. Ingolia NT. “Proximity RNA labeling by APEX-Seq Reveals the Organization of Translation Initiation Complexes and Repressive RNA Granules” (2018) *bioRxiv* 454066

Abstract

Diverse ribonucleoprotein complexes control messenger RNA processing, translation, and decay. Transcripts in these complexes localize to specific regions of the cell and can condense into non-membrane-bound structures such as stress granules. It has proven challenging to map the RNA composition of these large and dynamic structures, however. We therefore developed an RNA proximity labeling technique, APEX-Seq, which uses the ascorbate peroxidase APEX2 to probe the spatial organization of the transcriptome. We show that APEX-Seq can resolve the localization of RNAs within the cell and determine their enrichment or depletion near key RNA-binding proteins. Matching the spatial transcriptome, as revealed by APEX-Seq, with the spatial proteome determined by APEX-mass spectrometry (APEX-MS), obtained from exactly the experimental setup, provides new insights into the organization of translation initiation complexes on active mRNAs, as well as unanticipated complexity in stress granule composition. Our novel technique allows a powerful and general approach to explore the spatial environment of macromolecules.

Introduction

Proximity labeling has emerged as a valuable approach for understanding patterns of protein interaction and localization within cells. Proximity labeling techniques rely on enzyme-catalyzed *in vivo* reactions that mark targets near the labeling enzyme — which is typically fused to a query protein — and enable later, *ex vivo* analysis. The labeling reaction occurs within living cells and often acts over tens of nanometers. Proximity labeling is thus particularly well suited to capture transient, dynamic, and heterogeneous structures, complementing biochemical purifications that rely on direct and stable interaction.

The most dramatic advances in proximity labeling involve protein biotinylation, through the use of enzymes that produce diffusible reactive intermediates — either short-lived radicals(Lee et al. 2016; Lam et al. 2015; Hung et al. 2016) or longer lived adenylate esters(Kim et al. 2016; Roux et al. 2012; Rees et al. 2015; Branon et al. 2017; Choi-Rhee, Schulman, and Cronan 2004). These protein labeling tools can only indirectly address the organization of DNA in chromatin and RNA in various nuclear and cytosolic granules (Kaewsapsak et al. 2017; Myers et al. 2018). Currently, DNA can be labeled enzymatically through the action of an adenosine methyltransferase, in the DamID technique (van Steensel and Henikoff 2000). TRIBE is a similar approach for labeling RNAs with an adenosine deaminase (McMahon et al. 2016). While these base-modifying enzymes have proven valuable, for example in mapping binding sites of the heterochromatin protein CBX1 (HP1β)(Vogel et al. 2006), direct enzymatic modification operates only within a short, defined distance from the query protein and suffers from steric restrictions and other technical limitations.

Direct RNA proximity labeling promises new insights into the dynamic behavior of RNA. Translating RNAs move dynamically through the cytosol and often localize to specific regions of the cell (Morisaki et al. 2016; Wu et al. 2016; Yan et al. 2016; C. Wang et al. 2016). Inactive RNAs can be sequestered into protein-RNA granules through a process of liquid-liquid phase separation (LLPS)(Brangwynne et al. 2009; Molliex et al. 2015; Hyman and Simons 2012). Stress granule formation dynamically alters macromolecular localization, and mutations that increase stress granule formation or limit stress granule clearance are implicated in neurodegenerative diseases(Li et al. 2013; Ramaswami, Paul Taylor, and Parker 2013; Protter and Parker 2016). Recent work has argued that stress granules comprise two distinct components: a stable “core” surrounded by a concentration dependent “shell”(Jain et al. 2016), which contain distinct macromolecules. Therefore, labeling and isolation of RNAs from these sub-cellular compartments unveils the detailed mechanism and physiological relevance of this RNA localization.

One powerful and distinctive approach for proximity labeling employs an engineered ascorbate peroxidase enzyme (APEX2) to convert a cell-permeable biotin-tyramide substrate into a highly reactive free radical that labels aromatic amino acids in proteins within ~25 nanometers(Hayat

2007; Rhee et al. 2013). APEX2 labeling has already provided insight into the protein composition of stress granules (Markmiller et al. 2018). Motivated by the realization that nucleotides are also amenable to free radical-based chemistry(Chen et al. 2018), we developed an RNA proximity labeling technique, using APEX2 (APEX-Seq) as a way to probe the spatial organization of the transcriptome. We show that APEX-Seq can resolve the localization of RNAs within the cell and determine their enrichment or depletion near key RNA-binding proteins and in the stress granule. Because the same experimental setup is compatible with proximity protein labeling, we are able to probe the localization of our APEX2 fusion query relative to the ribosome and reveal unanticipated complexity within the stress granule. We then apply APEX-Seq to identify the transcripts that localize to stress granules and find that their composition varies depending on the stress applied. Our technique, which can capture RNA and protein, provides a powerful tool to explore the dynamics of RNAs and proteins in macromolecular assemblies and in LLPS.

Results

RNA biotinylation by the APEX proximity labeling enzyme

We reasoned that the radical mechanism underlying APEX proximity labeling of proteins(Lad, Mewies, and Raven 2002) would lead to similar, proximity-dependent biotinylation of RNA as well (Figure 1A). Indeed, we found that purified recombinant APEX2 enzyme (Figures S1A and S1B) biotinylated RNA *in vitro*, in a reaction that depended on both biotin-tyramide and hydrogen peroxide, reminiscent of tyramide labeling of DNA by horseradish peroxidase (HRP)(Chen et al. 2018) (Figure 1B). Importantly, the biotinylation signal that we detected was RNase sensitive (Figure 1C), and thus reflects labeling of the RNA in the reaction.

We next verified that the APEX2 labeling reaction biotinylates RNA *in vivo*. In order to test RNA biotinylation, we expressed APEX2 fused to the DEAD-box RNA helicase eIF4A1 in HEK293T cells (Figure S1C). This fusion retained its ability to bind RNA, including RocA-dependent stabilization on polypurine tracts(Iwasaki, Floor, and Ingolia 2016), which depends on the protein's normal RNA-binding interface(Iwasaki, Floor, and Ingolia 2016) (Figure S1D). Total RNA extracted from cells expressing this functional APEX2-eIF4A1 fusion was biotinylated after

biotin-tyramide pre-incubation and peroxide treatment, in contrast to RNA from naïve cells treated in the same way, which showed no detectable biotin conjugate on RNAs (Figure 1D).

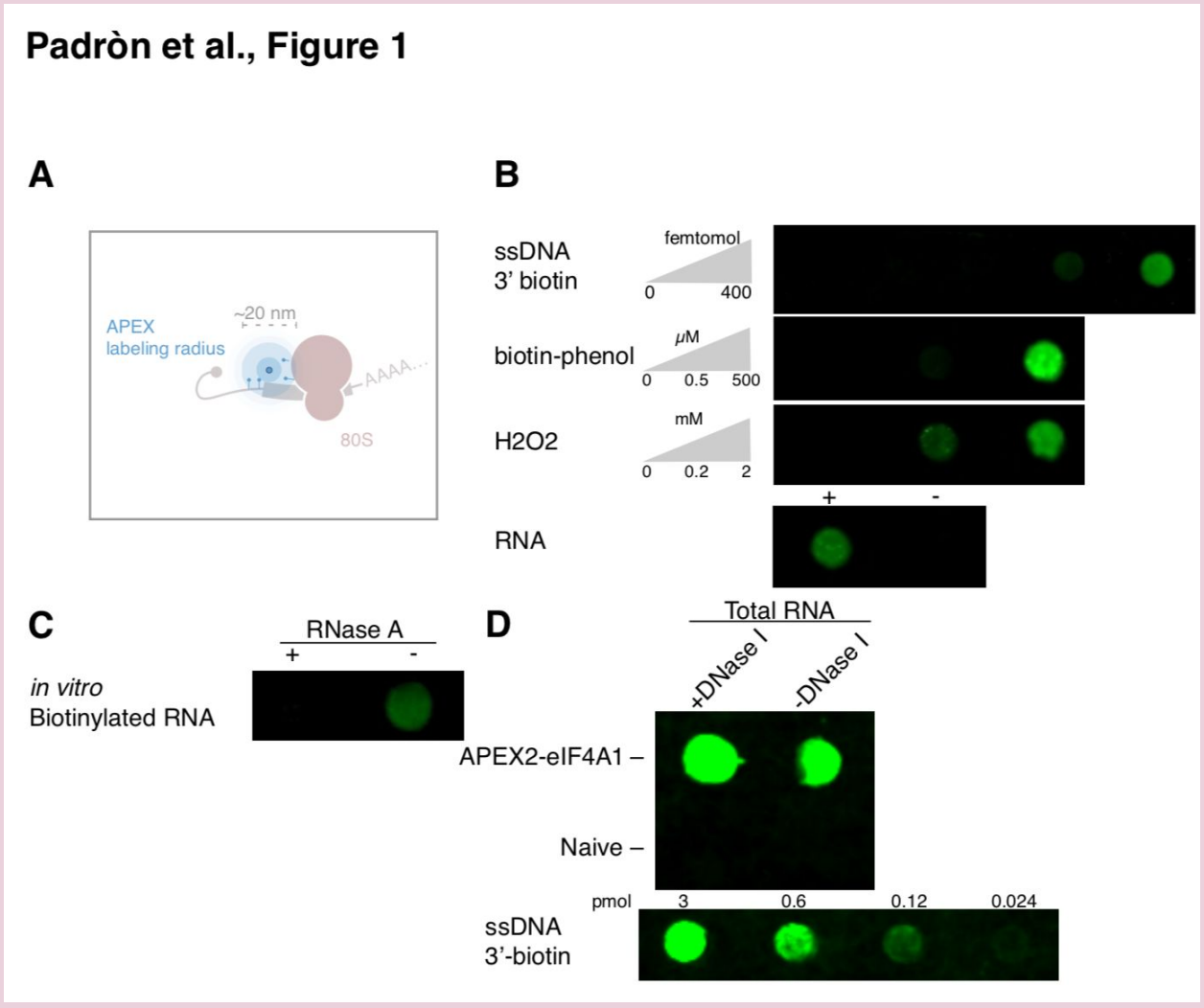
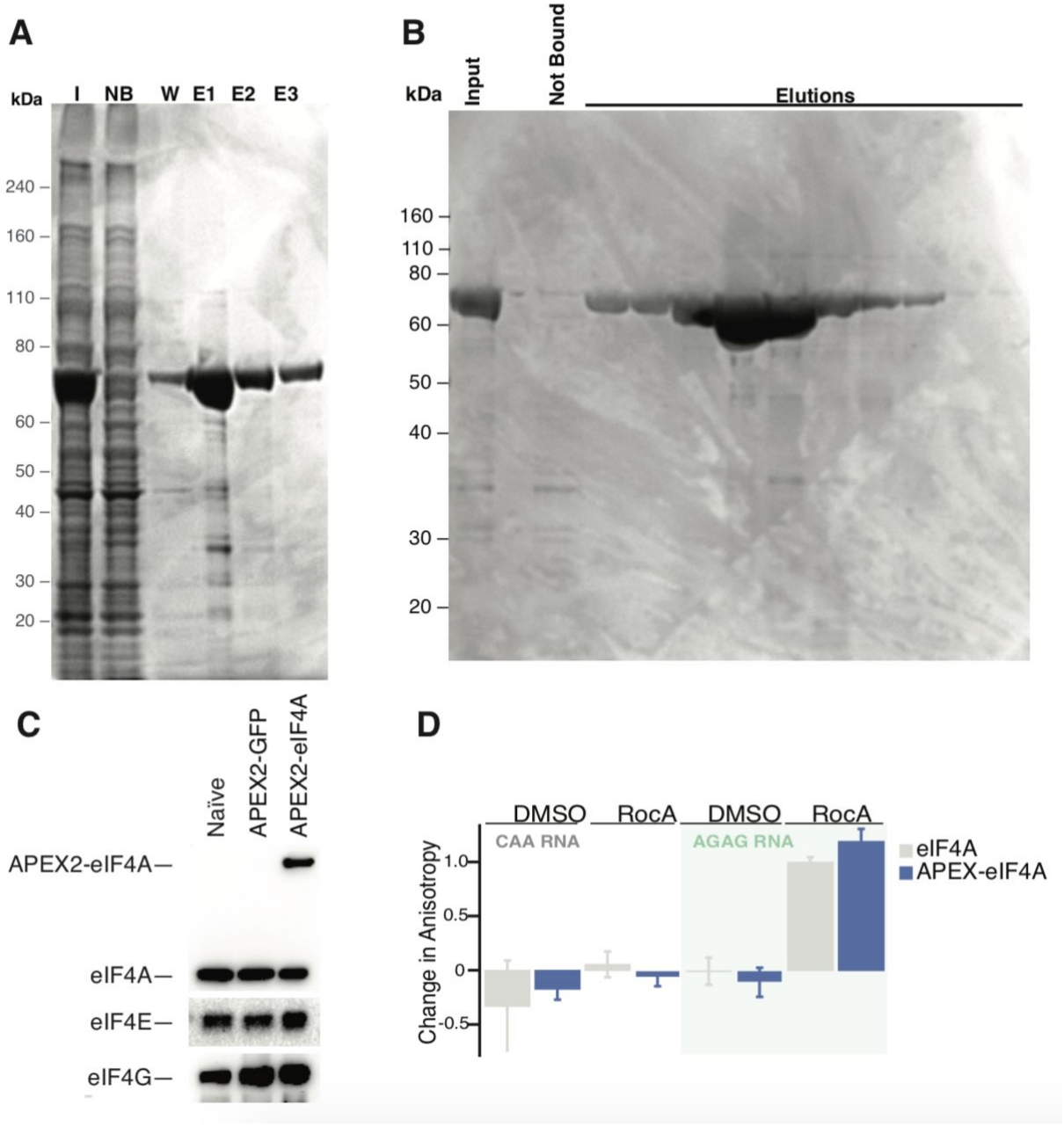


Figure 1. APEX proximity biotinylation of RNA.
(A) Diagram of APEX proximity biotinylation of both protein and RNA.
(B) *in vitro* labeling of RNA as a function of both biotin-tyramide and hydrogen peroxide. A ssDNA 3'-biotin oligo was used as a positive control.
(C) *in vitro* biotinylated RNA subjected to RNase A treatment for 30 min at 37 ºC.
(D) *in vivo* biotinylated RNA from an APEX2-eIF4A1 or Naive HEK293T cell line treated with DNase I for 30 min at 37 ºC.



Supplemental Figure 1, related to Figure 1.

(A and B) Purification of recombinant APEX2-eIF4A1 protein by (A) imidazole elution and (B) heparin chromatography.

(C) Western blot of transgenic expression of APEX2-eIF4A1 in HEK293 cells.

(D) Fluorescence polarization assay with recombinant WT eIF4A1 or APEX2-eIF4A1 using “CAA” or “AGAG” FAM-labeled oligo under DMSO or Rocaglamide A treatment.

APEX-Seq captures sub-cellular RNA localization patterns

Having shown that the APEX reaction biotinylates RNA *in vivo,* we asked whether subcellular differences in RNA localization can be detected by purifying and sequencing these biotinylated RNAs. We targeted APEX2 to three distinct locations within the cell: 1) the cytoplasm, through the use of APEX2-GFP; 2) the cytosolic face of the ER membrane, using a previously established C1(1-29)-APEX2 fusion(Lee et al. 2016) containing the 29 N-terminal residues of cytochrome P450 2C1 (rabbit CYP2C1); and 3) the nucleus, using a CBX1-APEX2 fusion that links APEX2 to heterochromatin protein 1 beta (HP1β) (Figure 2A). We carried out APEX2 labeling reactions in cells expressing each of these fusion proteins, purified the biotinylated RNA by streptavidin affinity, and analyzed both total and biotinylated RNA by deep sequencing. RNA-Seq read counts replicated extremely well in both total and streptavidin-purified samples (R^2 ~ 0.99), confirming the reproducibility of our assay (Figures 2B, 2C, and S2A-S2D). Moreover, we saw distinctive patterns of enrichment and depletion after purifying biotinylated RNA by streptavidin affinity, confirming that APEX fusions were biotinylating proximal RNA specifically (Figure S2E).

We next wanted to test whether these patterns of biotinylation reflected the sub-cellular localization of RNAs relative to the APEX fusion protein. We compared biotinylated RNA from C1-APEX2 fusions against the RNA labeled by our diffuse APEX-GFP control. We found that C1-APEX2 labeling enriched strongly for mRNAs encoding membrane-associated proteins, reflecting their localization to the surface of the ER during co-translational secretion (Figures 2D-2E and S2F-S2G). The highest enrichment scores we saw included 139x enrichment of Golgi-Associated Plant Pathogenesis-Related Protein 1 (GLIPR2) mRNA and 86x enrichment of the SLC16A7 transcript, which encodes a bidirectional transport of short-chain monocarboxylates(Lin et al. 1998). In contrast, transcripts encoding soluble cytosolic proteins such as PFDN4, a subunit of the heterohexameric chaperone prefoldin, and nuclear proteins such as the bZIP transcription factor CEBPE, were depleted in C1-APEX2 relative to APEX2-GFP (Lin et al. 1998; Vainberg et al. 1998) (Figure 2E).

Likewise, we found that the nuclear CBX1-APEX2 fusion preferentially labeled non-coding nuclear RNAs (Figure 2F and S2E). We saw a strong enrichment of the 7SK small nuclear pseudogene (RN7SKP9), as well as an enrichment of TSIX, the antisense transcript derived

from the XIST locus. By contrast, we saw depletion of the mRNA encoding for Elongation Factor 1 Alpha 2, a highly abundant coding transcript (Figure 2F). Taken together, the data from our C1-APEX2 and CBX1-APEX2 labeling show that APEX-Seq captures patterns of RNA localization across cellular compartments even when these are not separated by membranes.

Padròn et al., Figure 2

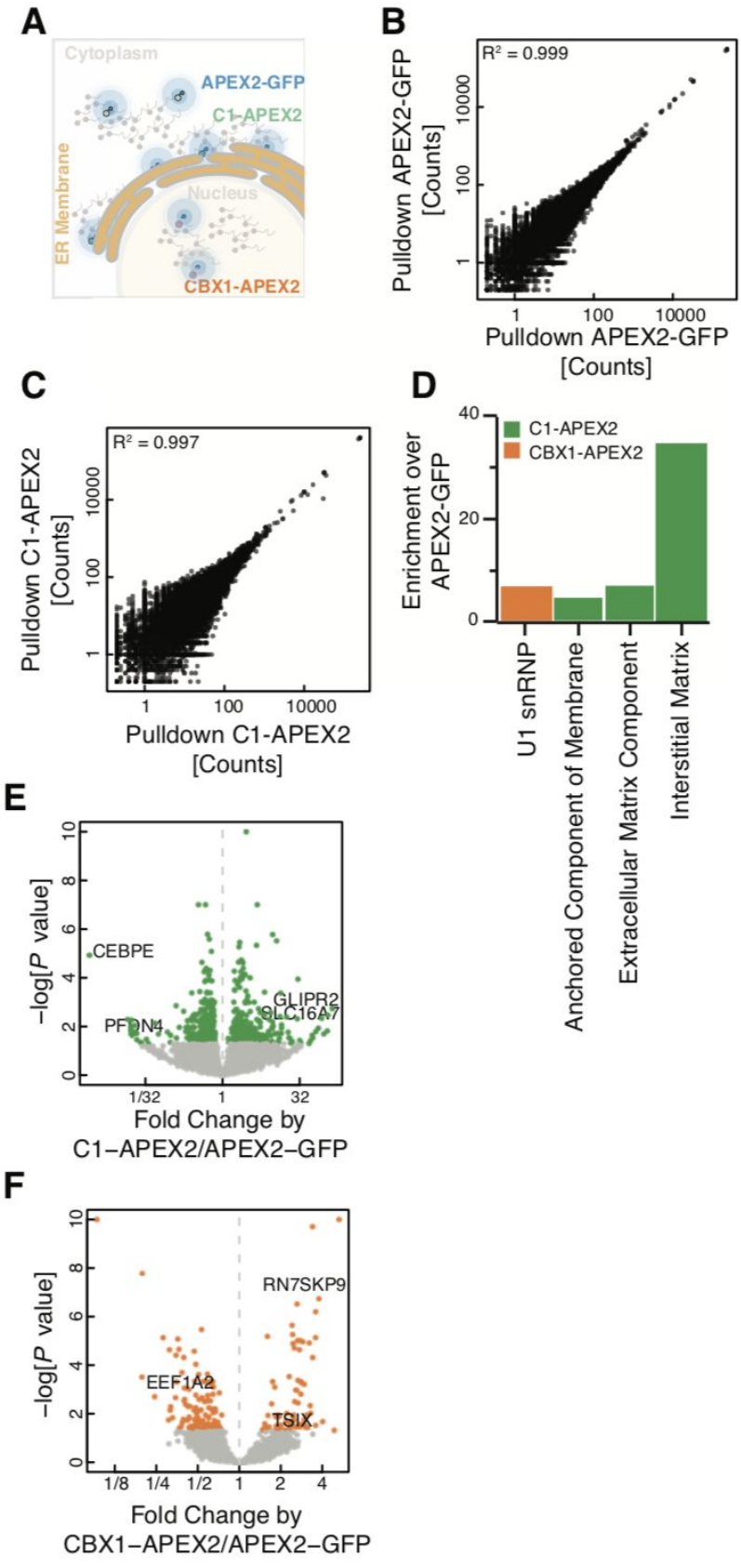
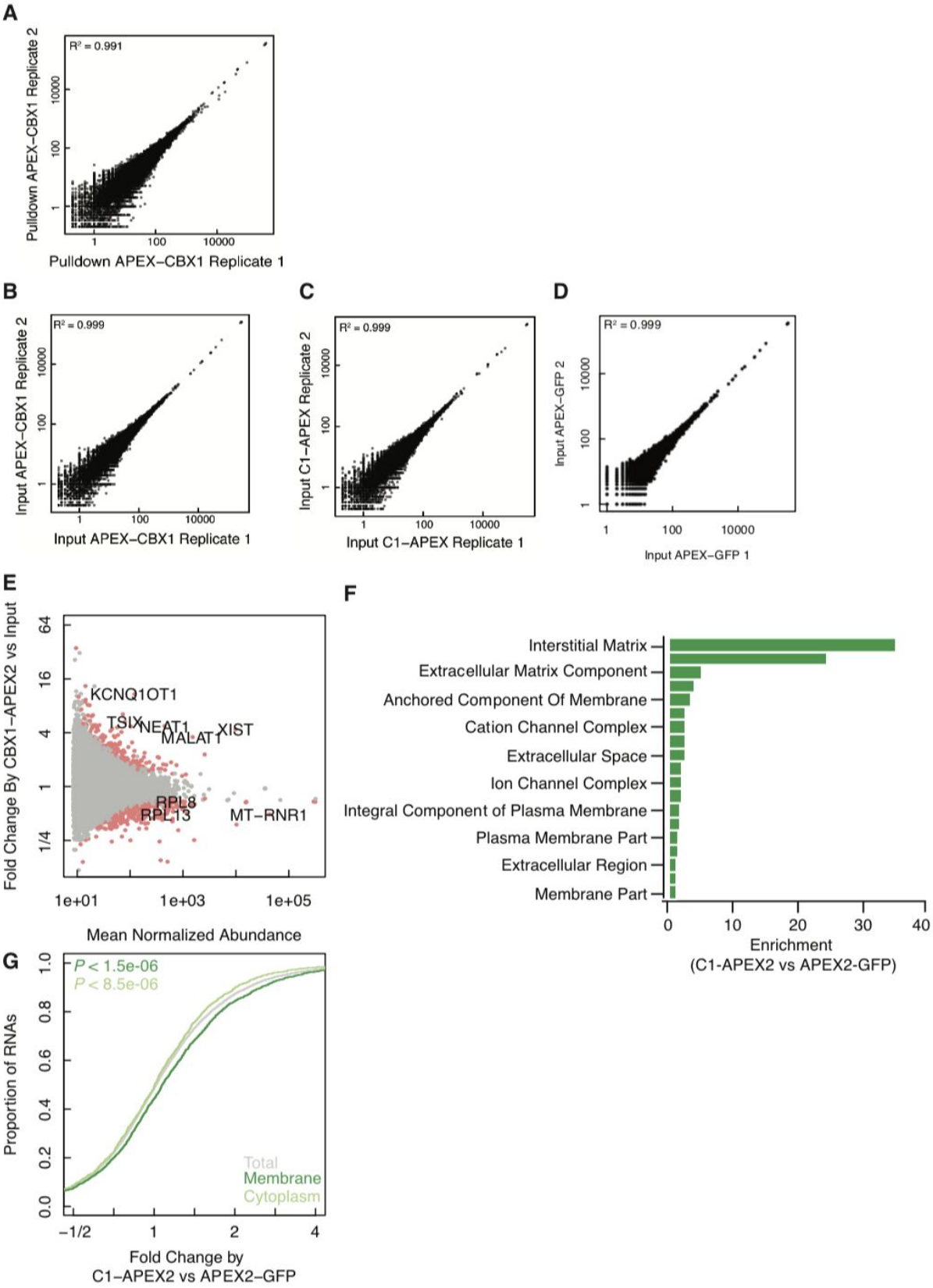


Figure 2. APEX proximity biotinylation reveals RNA sub-cellular localization.
(A) Diagram of localization using APEX-Seq.
(B and C) Correlation plot showing agreement between replicate samples for streptavidin affinity purification of biotinylated RNA from (B) an APEX2-GFP cell line and (C) streptavidin-pulldown C1-APEX2 cell line.
(D) GO term enrichment for transcripts preferentially labeled in C1-APEX2 and CBX1-APEX2 cell lines. Anchored component of the membrane n=15; extracellular matrix component n=10; interstitial matrix n=3; U1 snRNP n=9.
(E and F) volcano plot for (E) C1-APEX2 and (F) CBX1-APEX2 compared to APEX2-GFP showing enrichment and depletion for specific RNAs.

Padròn et al., Supplemental Figure S2



Supplemental Figure 2, related to Figure 2.
(A) Correlation plot showing agreement between replicate samples for streptavidin affinity purification of biotinylated RNA from CBX1-APEX2 cell line.
(B-D) As in (A), showing agreement between replicate samples for input RNA for (B) an CBX1-APEX2, (B) a C1-APEX2, and (D) an APEX2-GFP cell line.
(E) M-A plot showing a comparison between input and affinity-purified RNA for CBX1-APEX2 cell lines.
(F) GO term enrichment of C1-APEX2.
(G) C1-APEX2 cumulative distribution of membrane protein encoding RNAs.

APEX-Seq captures protein-RNA interaction patterns

We next asked whether patterns of protein-RNA interaction, occurring on an even finer length scale, could be detected by APEX-Seq enrichment. In addition to the APEX2-eIF4A1 fusion described above, we fused APEX2 to the 7-methylguanosine (m^7G)-cap binding protein eIF4E1 (Figure 3A). We compared biotin-labeled RNAs from cells expressing APEX2-eIF4E1 against total RNA and observed the depletion of non-coding RNAs, like SNORA73B, which possess 2,2,7-trimethylguanosine caps that bind eIF4E1 with far lower affinity (Niedzwiecka et al. 2002) (Figure S3A). Intriguingly, we observed the strongest enrichment for immediate early genes FOSB and EGR3 (Figure 3B). Immediate-early genes are preferentially translated when the eIF4E1-interacting scaffold eIF4G is limiting (Johannes et al. 1999), and enrichment of these transcripts may reflect their preferential association with eIF4E1. Several proto-oncogenes, including c-Myc, were also preferentially translated in conditions of limiting eIF4G(Johannes et al. 1999). Indeed, eIF4E1 is an oncogene whose overexpression promotes cellular transformation (Pelletier et al. 2015); (Graff et al. 2007). In line with these observations, we found that APEX2-eIF4E1 preferentially labels the oncogenic C-MYC, VEGF-A, and FGF13 mRNAs (Figure 3C). More broadly, protein synthesis of long, structured 5′ UTR-containing mRNAs appear to be sensitive to eIF4E1 levels (Pelletier et al. 2015); (De Benedetti and Graff 2004; Kozak 1991), and indeed we found greater enrichment of long 5′ UTR containing mRNAs upon labeling by APEX2-eIF4E1 (Figure 3D). The eIF4E1 transcript was also enriched, perhaps reflecting nascent APEX2-eIF4E1 biotinylating its own mRNA (Figure 3B).

Next, we asked whether APEX2-eIF4E1 and APEX2-eIF4A1 showed similar labeling patterns, since, along with the scaffold protein eIF4G, they form the eIF4F translation initiation complex.

Indeed, we found good correspondence between their respective labeling patterns (Figure 3E). Both proteins enriched mRNAs that encode translation machinery proteins, which are known to show a particular sensitivity to inhibition of eIF4E1 (Hsieh et al. 2012; Thoreen et al. 2012) and eIF4A1 (Iwasaki, Floor, and Ingolia 2016). While eIF4E1 and eIF4A1 labeling patterns overlapped substantially, they were not identical. For example, the mRNA encoding eIF4E1 is enriched only in APEX2-eIF4E1 labeling and not in APEX2-eIF4A1 labeling, consistent with the model that this enrichment reflects biotinylation by nascent protein. The similarity between eIF4E1 and eIF4A1 labeling patterns reflected the co-localization of these proteins, as we found a substantially weaker correlation in RNA labeling between CBX1-APEX2 and APEX2-eIF4A1 fusions (Figure S3B).

Padròn et al., Figure 3

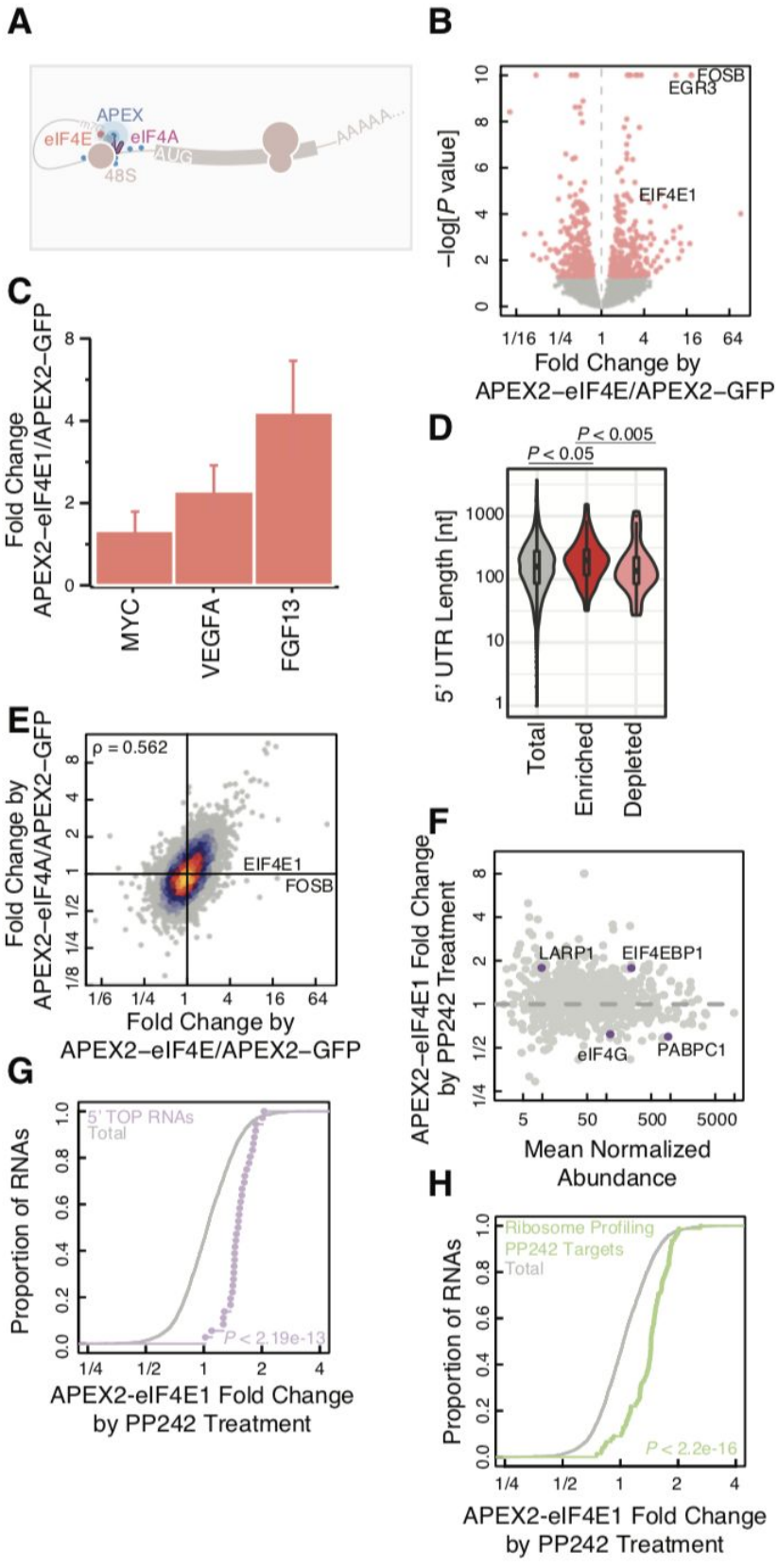
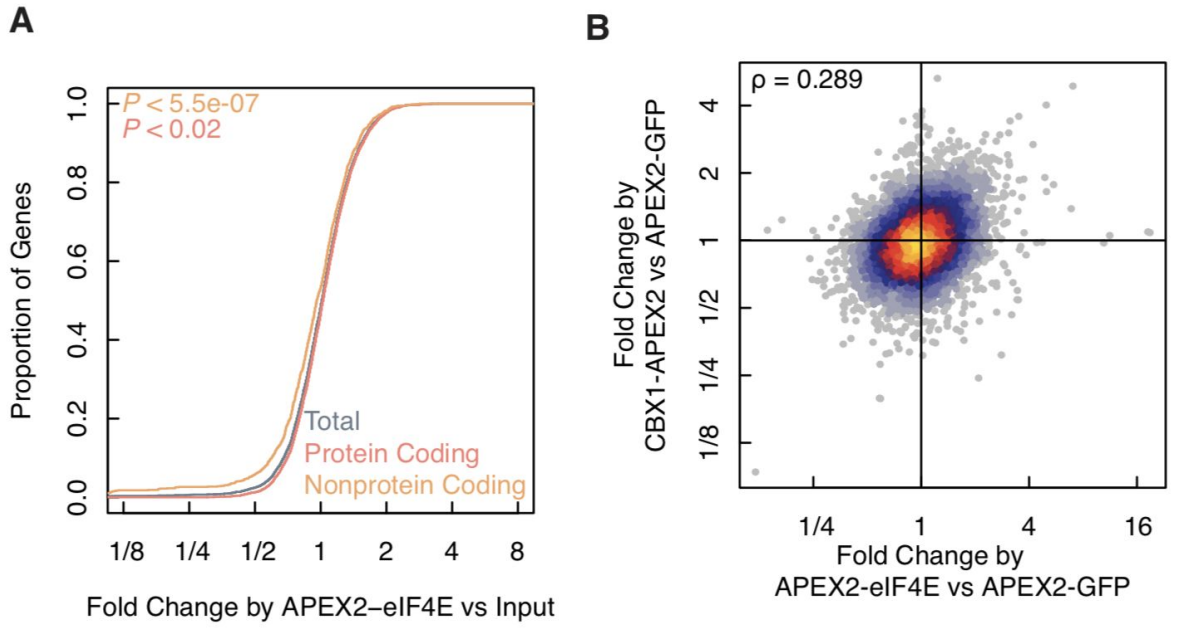


Figure 3. APEX proximity biotinylation reveals ribonucleoprotein complex composition.
(A) Diagram of the 43S preinitiation complex. Blue sticks signify biotin adducts formed from the proximity labeling reaction catalyzed by APEX2.
(B) Volcano plot for APEX2-eIF4E1 showing enrichment and depletion of specific RNAs compared to APEX2-GFP.
(C) Barplot showing enrichment for mRNAs encoding cell transforming proteins in APEX2-eIF4E1.
(D) 5' UTR length distribution for APEX2-eIF4E1 enriched or depleted RNAs. P-values calculated using the Mann-Whitney test. Enrichment and depletion is calculated as statistically significant (FDR ≤0.05) and either top or bottom 2.5% of the fold change.
(E) Correlation plot between APEX2-eIF4E1 and APEX2-eIF4A1 datasets in comparison to APEX2-GFP, highlighting uniquely enriched mRNAs with respect to APEX2-eIF4E1.
(F) APEX-MS and (G) APEX-Seq of APEX2-eIF4E1 cells upon PP242 treatment as compared to DMSO treated controls. 5' TOP RNAs were defined previously[65]. *P* value calculated using the Mann-Whitney test. (H) The proximity of the PP242 sensitive mRNAs, monitored through ribosome profiling[30], with respect to APEX2-eIF4E1. *P* value calculated using the Mann-Whitney test.

Padròn et al., Supplemental Figure S3



Supplemental Figure 3, related to Figure 3.
(A) APEX2-eIF4E1 cumulative distribution of protein-coding and noncoding RNAs.
(B) CBX1-APEX2 vs APEX2-GFP and APEX2-eIF4E1 vs APEX2-GFP correlation plot.

Proximity proteomics places eIF4A1 on the 3′ side of the 43S preinitiation complex

Since APEX2 has been used for proximity-specific protein labeling(Lam et al. 2015; Hung et al. 2016), our observation that APEX2 can label RNA now offers the dual use of APEX2 to match spatially resolved transcriptomic and proteomic data. We therefore performed quantitative tandem mass tag (TMT) mass spectrometry(Rauniyar and Yates 2014) in cells expressing APEX2-eIF4A1 (Figures S1C and S4A). Ratiometric analysis using TMT labeling showed reproducible quantitation between biological replicates (Spearman's ρ ~ 0.98) (Figures S4B and S4C). We found that eIF4A1 labeling enriches other components of the eIF4F complex — the cap-binding protein eIF4E1 and the scaffolding protein eIF4G1 — as well as the poly(A) binding protein, PABPC1, which binds eIF4G1, and several small subunit ribosomal proteins (Figures 4A and 4B). More broadly, APEX2-eIF4A1 spatial proteomics showed enrichment for translation initiation, RNP organization, and post-transcriptional regulators in gene ontology enrichment analysis (Figure 4C). This pattern of enrichment was similar to BioID analysis of eIF4A1, which relied on long duration (24 hour) labeling by an eIF4A1-birA* fusion protein(Youn et al. 2018). We saw a substantial (~25%) overlap between the proteins enriched in our rapid (<1 minute) APEX labeling experiment and those seen in long-term BioID labeling (hypergeometric $p < 0.01$) (Figure S4D).

We noticed a striking pattern of enrichment for some small subunit ribosomal proteins and depletion of others, despite their uniform presence in 43S pre-initiation complexes (Figures 4B and S4E). To better understand why different 40S proteins show enrichment or depletion in APEX2-eIF4A1 proteomics, we mapped these proteins onto the structure of the eukaryotic 43S preinitiation complex(Llácer et al. 2015). We found that the proteins enriched by APEX2-eIF4A1 were situated near the mRNA entry site of the 43S preinitiation complex, on the side towards the 3′ end of the mRNA, while the depleted proteins lie closer to the mRNA exit site (Figure 4D). The 43S preinitiation complex also binds to eIF3, a large, multi-protein complex that plays diverse roles in translation initiation (Figure 4D). Enriched eIF3 subunits were also localized

toward the mRNA entry site(des Georges et al. 2015). These data clearly place the RNA helicase eIF4A1 on the leading edge of the 43S complex, addressing an open question about the organization of this complex(Hinnebusch 2014).

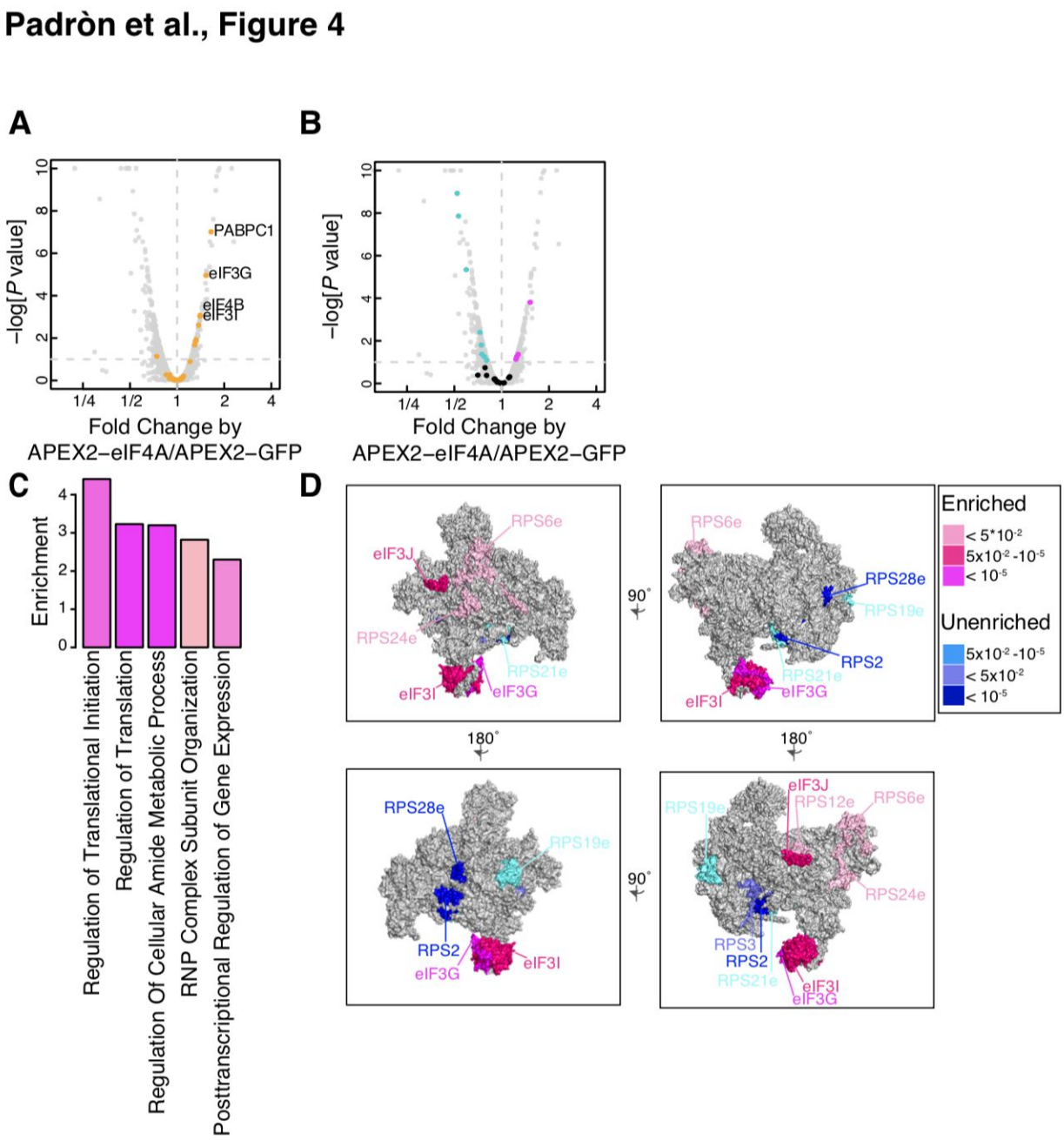
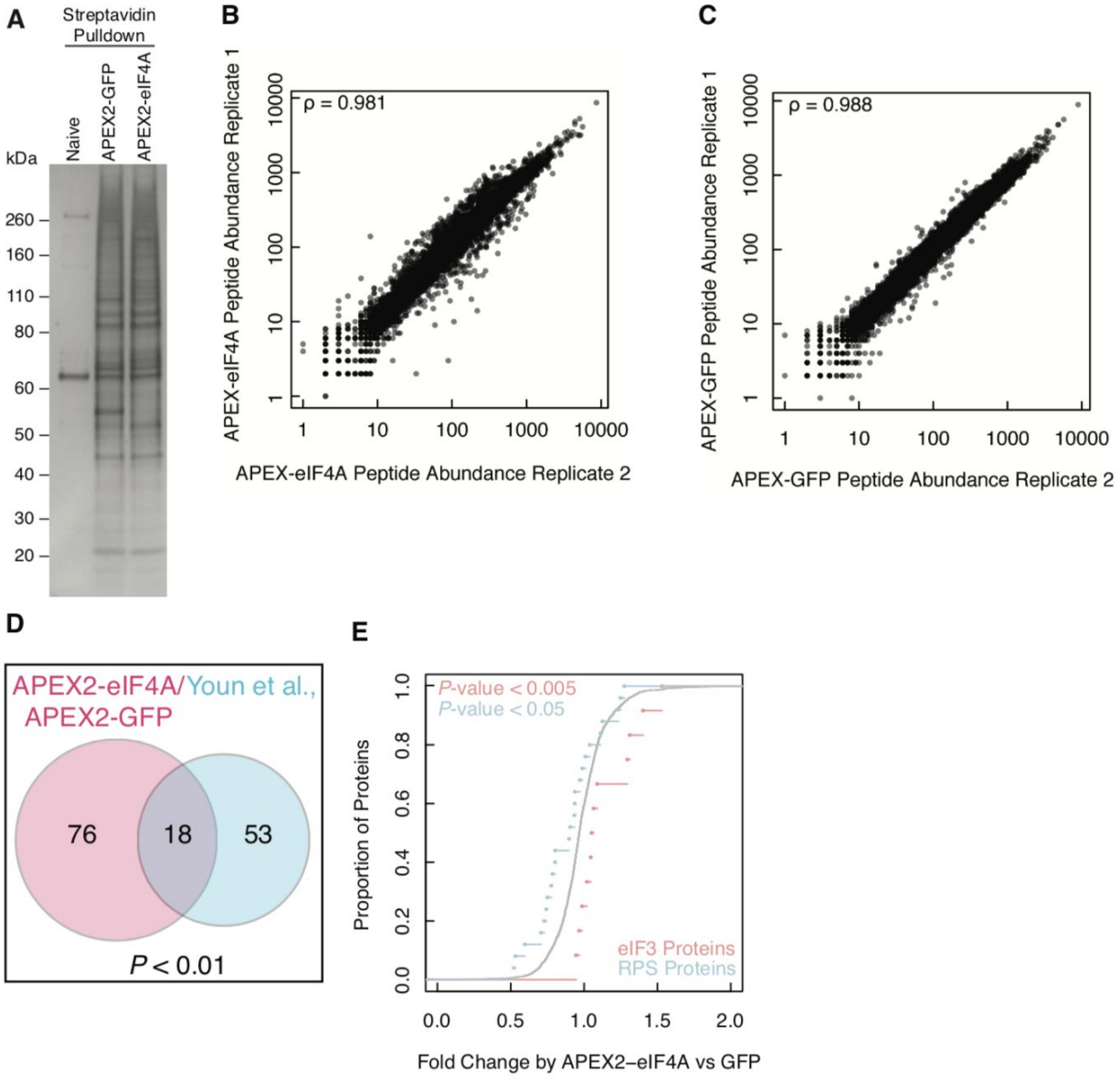


Figure 4. Proximity proteomics places eIF4A1 on the 3' side of the 43S preinitiation complex.

(A) Enrichment of translation initiation factors (orange) in APEX2-eIF4A1 proximity labeling relative to APEX2-GFP. Points in yellow signify translation initiation factors. Dashed vertical line signifies proximity to APEX-eIF4A (to the right of the line are proteins in close proximity, and to the right of the line are proteins in farther proximity). Horizontal dashed line signifies FDR corrected significance cutoff.
(B) Enrichment and depletion of small subunit ribosomal proteins in APEX2-eIF4A1 proximity labeling. Significantly enriched proteins in pink, depleted in blue, and others in black. Dashed vertical line signifies proximity to APEX-eIF4A (to the right of the line are proteins in close proximity, and to the right of the line are proteins in farther proximity). Horizontal dashed line signifies FDR corrected significance cutoff. (C) GO term enrichment for APEX2-eIF4A1 as compared to APEX2-GFP. Regulation of translational initiation n=9; Regulation of Translation n=19; Regulation of cellular amide metabolic process n=19; RNP complex subunit organization n=12; posttranscriptional regulation of gene expression n=20. (D) Mapping enriched (pink) or depleted (blue) eIF3 and RPS subunits onto the eukaryotic pre-initiation complex. Different perspectives of the 43S complex are shown. Differently colored proteins signify FDR corrected p-values.

Padròn et al., Supplemental Figure 4



Supplemental Figure 4, related to Figure 4.

(A) Silver stain for biotinylated proteins recovered after streptavidin affinity purification.

(B and C) Quantitative TMT labeling correlation plots for (A) APEX2-eIF4A1 and (B) APEX2-GFP.

(D) Venn diagram for APEX2-eIF4A1 vs APEX2-GFP compared to Youn et al., 2018. P-values were calculated using a hypergeometric test.

(E) Cumulative distribution plots highlighting enriched and depleted eIF3s and RPS subunits in a APEX2-eIF4A1 vs APEX2-GFP TMT labeling comparison. *P* values were calculated using the Mann-Whitney test.

eIF4A1 proximity labeling captures the stress granule proteome

In addition to translation initiation factors, we noticed that APEX2-eIF4A1 labeling enriched a number of stress granule proteins (Figure S5A), although we saw no evidence of stress granule formation in these unstressed cells (Figure 5A). Our observation was consistent with previous studies reporting the presence of eIF4A1 in stress granule cores(Jain et al. 2016), which are pre-existing structures present prior to SG formation(Wheeler et al. 2016), as well as the BioID proximity labeling of stress granule components after long term (24 hour) expression of eIF4A1-birA* in unstressed cells(Youn et al. 2018). Indeed, APEX2-eIF4A1 labeling after stress resulted in an ~100x enrichment of biotin signal inside granules relative to the overall cytoplasm (Figures 5B and 5C), demonstrating a strong and specific enrichment of APEX2-eIF4A1 within stress granules after heat shock. These results suggested that we could use our APEX2-eIF4A1 fusion to study the stress granule transcriptome and proteome together. Furthermore, the short (<1 minute) proximity labeling offered the possibility of tracking the dynamics of SG assembly.

We monitored the protein dynamics through a 20-minute time course of stress granule assembly following 44 ºC heat shock (Figures 5A-5D). Replicates from our quantitative TMT-labeled proteomics correlated well (Spearman's ρ ~ 0.98) (Figures 5E and 5F). The granule proteome at our final timepoint appeared broadly consistent with the proteins identified in previous proximity labeling studies, carried out after long stress induction regimes when granules were fully assembled. The quantitative changes in APEX2-eIF4A1 labeling that we saw after 20 minutes of heat shock correlated with the changes in BioID labeling (Youn et al. 2018) induced by 3 hours of arsenite treatment (ρ = 0.24) (Figure S5B). This modest correlation was enhanced by restricting our analysis to proteins with an established eIF4A1 interaction or stress granule localization (ρ = 0.54) (Figure S5B), suggesting that these two very different labeling strategies captured the same underlying stress granule proteome. We also saw a strong correlation between proteins showing stress-induced APEX2-eIF4A1 labeling in our study and previous proximity labeling studies using the granule marker G3BP1 (Markmiller et al. 2018).

While stress granules are dynamic structures *in vivo*, previous work has identified a stable "core" that can be purified biochemically(Jain et al. 2016; Wheeler et al. 2017). In order to better understand how proximity labeling patterns mapped onto stress granule assembly, we investigated these core stress granule proteins in our data. We defined a set of "stress

responsive" proteins that were statistically enriched in APEX2-eIF4A1 labeling, in comparison with APEX2-GFP, and were further enriched in a comparison between APEX2-eIF4A1 labeling of the post-shock (T_{20}) time point with respect to the pre-shock (T_0) time point. In contrast, proteins labeled "stress insensitive" were defined as those that were not significantly enriched or depleted in a comparison between APEX2-eI4FA1 labeling of the pre-shock (T_0) and latest post-shock (T_{20}) timepoints, but were enriched in a comparison between APEX2-eIF4A1 as compared to APEX2-GFP. Interestingly, we found a subset of the stress granule core components to show stress-insensitive eIF4A1 enrichment (Figure 5G), including ATXN2, a protein implicated in amyotrophic lateral sclerosis (ALS)(Elden et al. 2010).

Previous work has implicated low complexity protein domains in stress granule formation (Molliex et al. 2015). We asked if there were other protein features predictive of stress granule residency. To do this, we compared our final post-shock time point (T_{20}) with respect to our pre-shock time point (T_0). Prion propensity(Toombs et al. 2012) did not predict SG residency in our analysis (Figure 5H). On the other hand, we found the presence of RNA recognition motifs (RRMs) to be most predictive of RNA granule residency, whereas DEAD-box RNA helicases are depleted (Figures 5H and S5C-S5D). These data suggest a central role for RNA interactions in recruiting proteins to stress granules and a role for some RNA helicases in opposing granule assembly.

Padròn et al., Figure 5

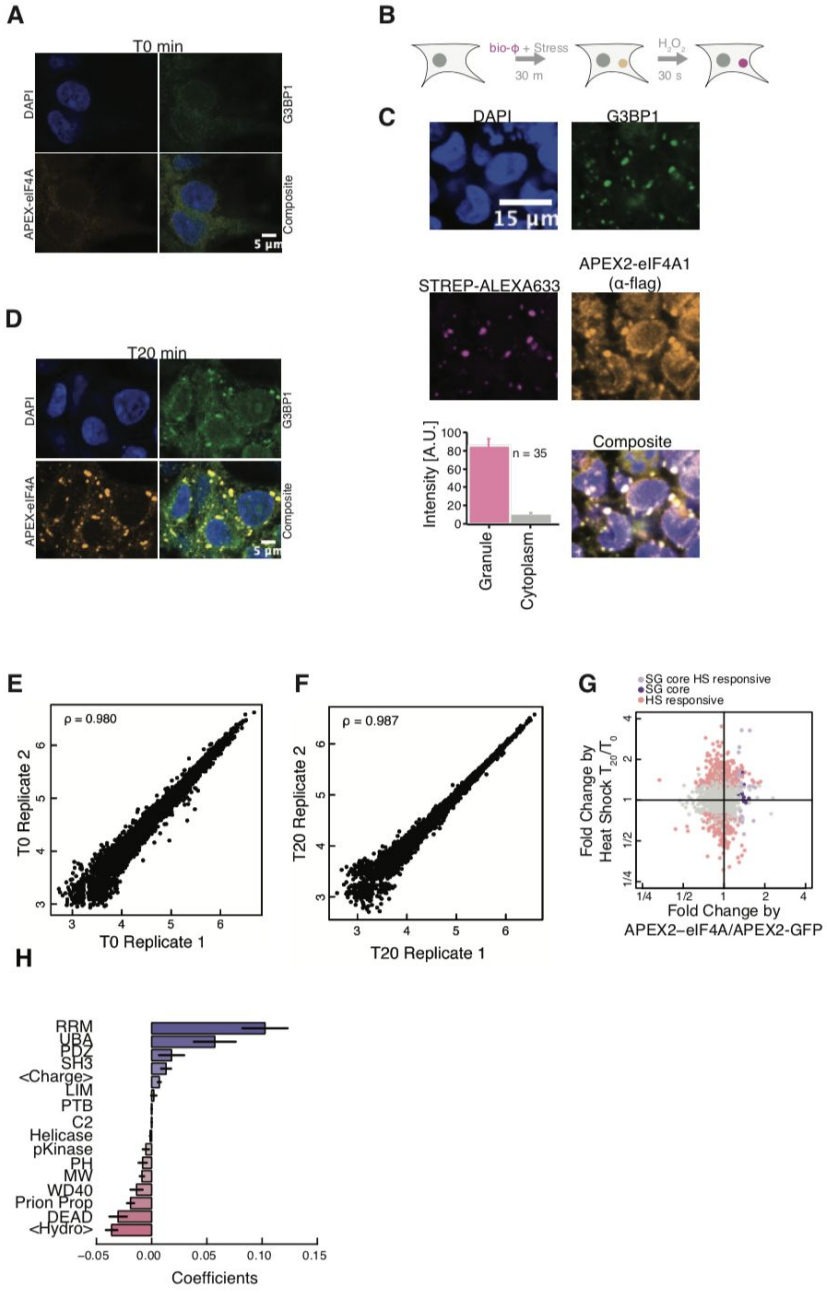


Figure 5.

(A) Immunofluorescence for flag tagged APEX2-eIF4A1 in unstressed cells.

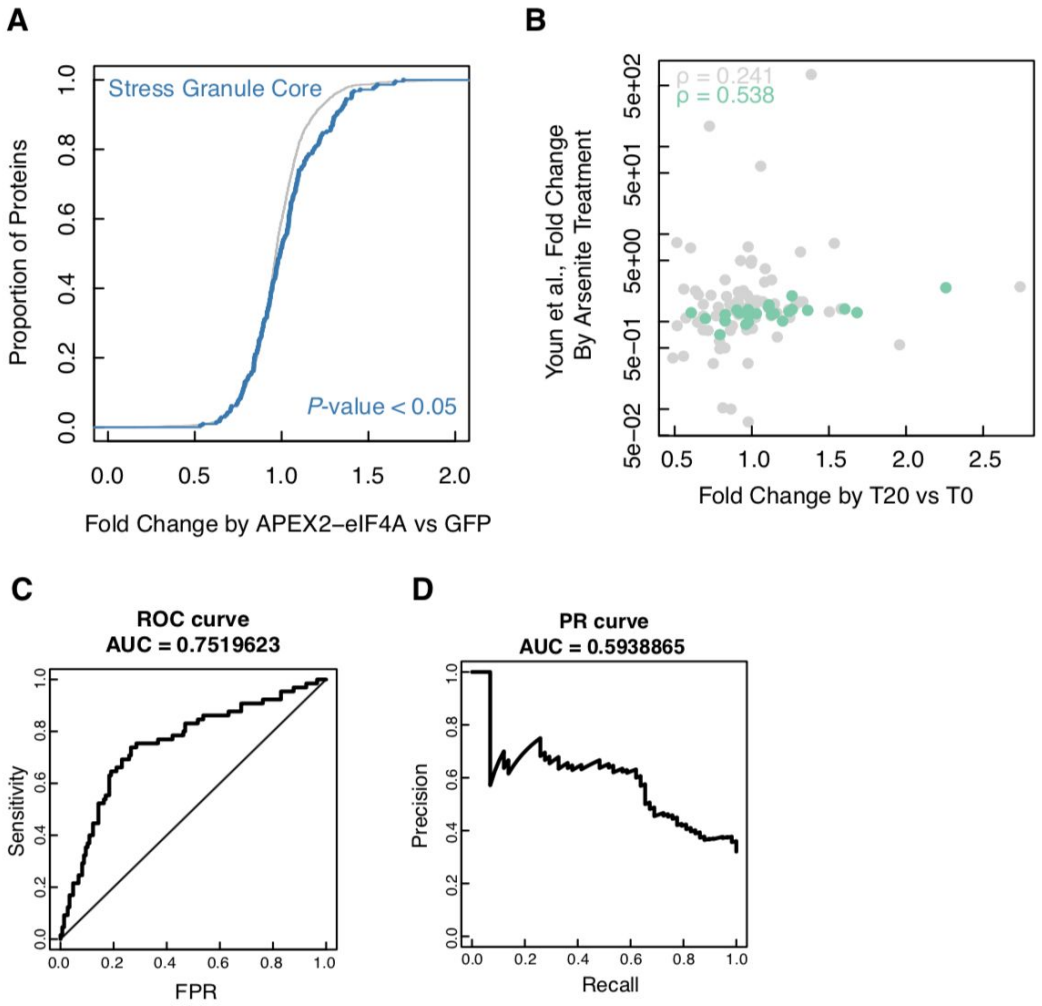
(B) Diagram of APEX2-eIF4A1 immunofluorescence assay showing the biotinylation intensity of APEX2-eIF4A1 inside and outside stress granules.

(C) As in (A), after a 20 minute heat shock treatment at 42 ºC. Biotinylated macromolecules are labeled with a Strep-ALEXA633.

(D) As in (A), after a 20 min heat shock at 44 ºC.

(E and F) Correlation plots for peptide abundance from quantitative TMT labeling proteomics at (E) T_0 and (F) T_{20} for a APEX2-eIF4A1 heat shock time course. Axes are log transformed.
(G) Comparison between T_{20} post heat shock with respect to pre heat shock (T_0) vs APEX2-eIF4A1 with respect to APEX2-GFP. SG core Heat shock responsive proteins (light pink, FDR corrected *P* value ≤ 0.05), SG core proteins that are heat shock insensitive (lavender, FDR corrected *P* value ≤ 0.05), SG core proteins that are heat shock responsive (purple, FDR corrected *P* value ≤ 0.05).
(H) Coefficients showing features that contribute positively (blue) or negatively (red) in predicting RNA granule residency in T_{20} with respect to T_0. continuous features: mean charge, <charge>, Michelitsh-Weissman (MW) score, Prion Propensity score, mean hydropathy, <hydropathy>. Categorical features: RRM n=52; PDZ n=17; UBA domain n=6; SH3 domain n=14; LIM domain n=12; PTB domain n=3; C2 domain n=16; Helicase_C domain n=3; pKinase domain n=32; PH domain n=5; WD40 domain n=28; DEAD domain n=16.

Padròn et al., Supplemental Figure 5



Supplemental Figure 5, related to Figure 5.
(A) Proportion of enriched stress granule core proteins (blue) identified with respect to APEX2-GFP. P-values calculated using the Mann-Whitney test.
(B) Arsentite treated Youn et al., 2018 dataset as compared to APEX2-eIF4A1 during heat shock (T_{20} vs T_0).
(C and D) Classifier predicting RNA granule residency. (C) Receiver operating characteristic curve. (D) Precision-recall curve.

Proximity labeling reveals the dynamic proteome of assembling stress granules

We took advantage of the high time resolution offered by rapid (<1 minute) APEX labeling in order to capture the detailed dynamics of stress granule assembly (Figures 6A and 6B). We observed substantial changes in eIF4A1 labeling across the time course, and principal component analysis of these changes revealed that the first principal component captured ~50% of variation and described an increasing trend across the time course of SG assembly (Figure 6C). The inflection point in this increase occurred after 5 minutes, which corresponds to the time when microscopically observable stress granules appear (Figure 6B). The component loadings from PC1 were dominated by increases in SG-related proteins like G3BP1 along with decreases in translation initiation and elongation factors (Figure 6D).

Hierarchical clustering segregated proteomics data from unstressed controls and early (2 - 4 minute) timepoints away from later (6 - 20 minute) samples (Figure S6A). This distinction between early and late timepoints also agreed with the time course of microscopically observable SG assembly (Figure 6B). We therefore partitioned our data into overlapping “early” (0 - 4 min), “middle” (2 - 10 min), and “late” (6 - 20 min) categories and identified proteins enriched in each of these groups. This analysis highlighted a number of important SG forming proteins in the “middle” timepoints, such as CAPRIN1 (Katoh et al. 2013; Kedersha et al. 2016) (Figures S6B-S6D). The stress granule forming protein, G3BP1, which interacts with CAPRIN1(Kedersha et al. 2016), appeared in the “late” timepoints, suggesting an order of events in interactions with respect to eIF4A1. More broadly, GO terms for RNA localization are enriched during the late time point (Figure S6E).

While the BioID data suggests that stress granule formation is driven by pre-existing interactions that remain largely unchanged after stress (Youn et al. 2018), we found a substantial and

coherent change in the pattern of eIF4A1 proximity labeling across stress granule assembly. The rapid (< 1 minute) labeling enabled by APEX2, in conjunction with our high-resolution timecourse, allowed us to resolve these dynamics more clearly. These dynamics also manifested in the correlation between proteins showing stress-induced APEX2-eIF4A1 labeling in our study and previous proximity labeling studies using the granule marker G3BP1 (Markmiller et al. 2018). This correspondence was strongest for our “middle” class of proteins ($p < 1e-10$, hypergeometric test), and was also significant for our “late” class ($p < 0.02$). These G3BP1-proximal proteins overall showed significantly higher labeling in our analysis of “middle” timepoints, relative to other proteins, as well ($p < 1e-15$, Wilcoxon test). Notably, we found G3BP1 in our “late”-enriched group and it shows heat shock responsive labeling by APEX2-eIF4A1. These results were consistent with G3BP1 and eIF4A1 co-localizing to stress granules, thus showing stronger mutual labeling and more similar overall labeling patterns after stress.

We took advantage of our high time resolution and further refined this analysis by fitting the dynamics of eIF4A1 proximity labeling with impulse models using ImpulseDE2 (Fischer, Theis, and Yosef 2018) (Figure 6E). This approach could identify proteins showing unidirectional trajectories of increased or decreased labeling, in addition to finding those showing transient changes. This analysis agreed with our hierarchical clustering results, as we found that “middle” timepoint proteins such as CAPRIN were induced early in the impulse model (cluster 2), whereas “late” proteins like G3BP1 more gradually accumulated over time (cluster 4). Overall, these data highlight the dynamic nature of the stress granule proteomic landscape as it is assembled upon heat shock.

Padròn et al., Figure 6

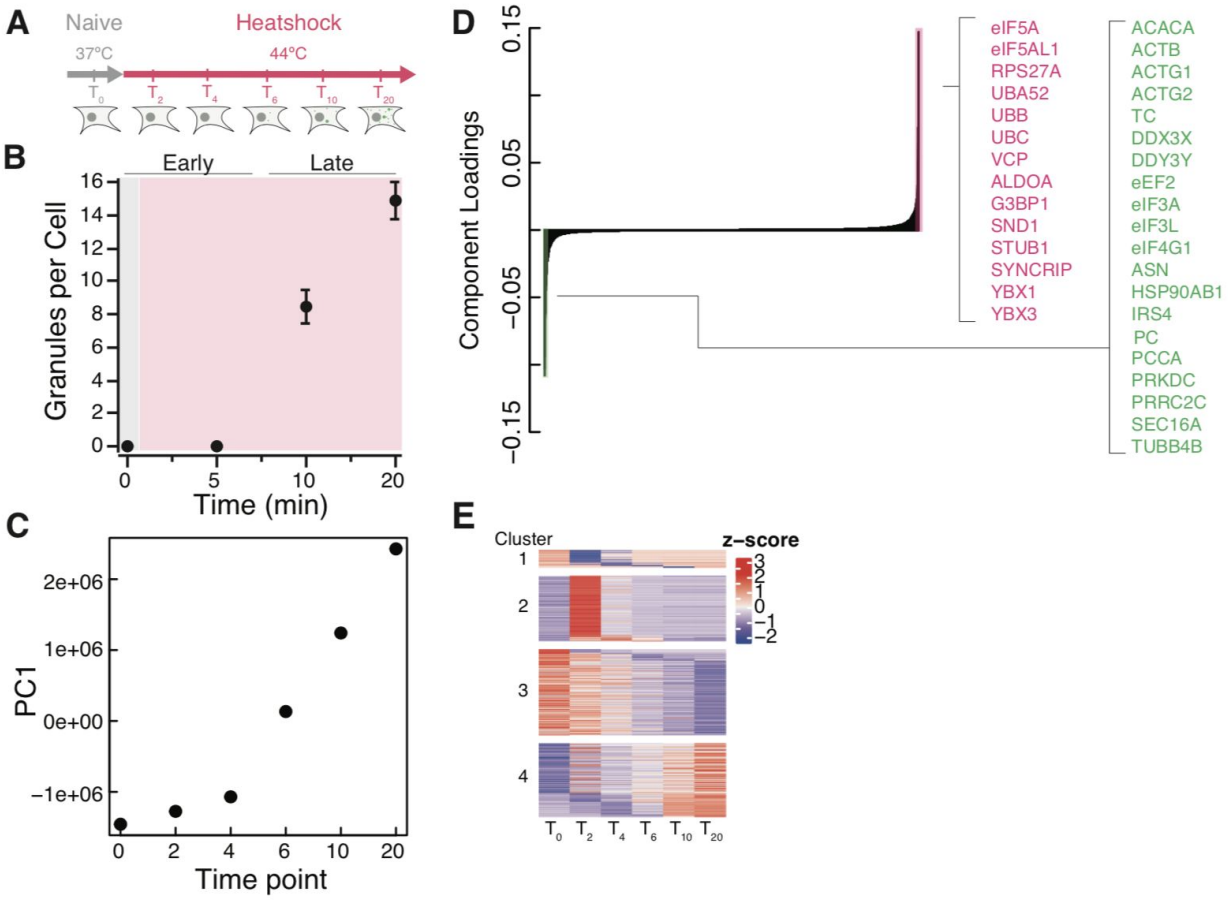


Figure 6. Timecourse of stress granule assembly.

(A) Diagram of APEX2-eIF4A1 heat shock time course.

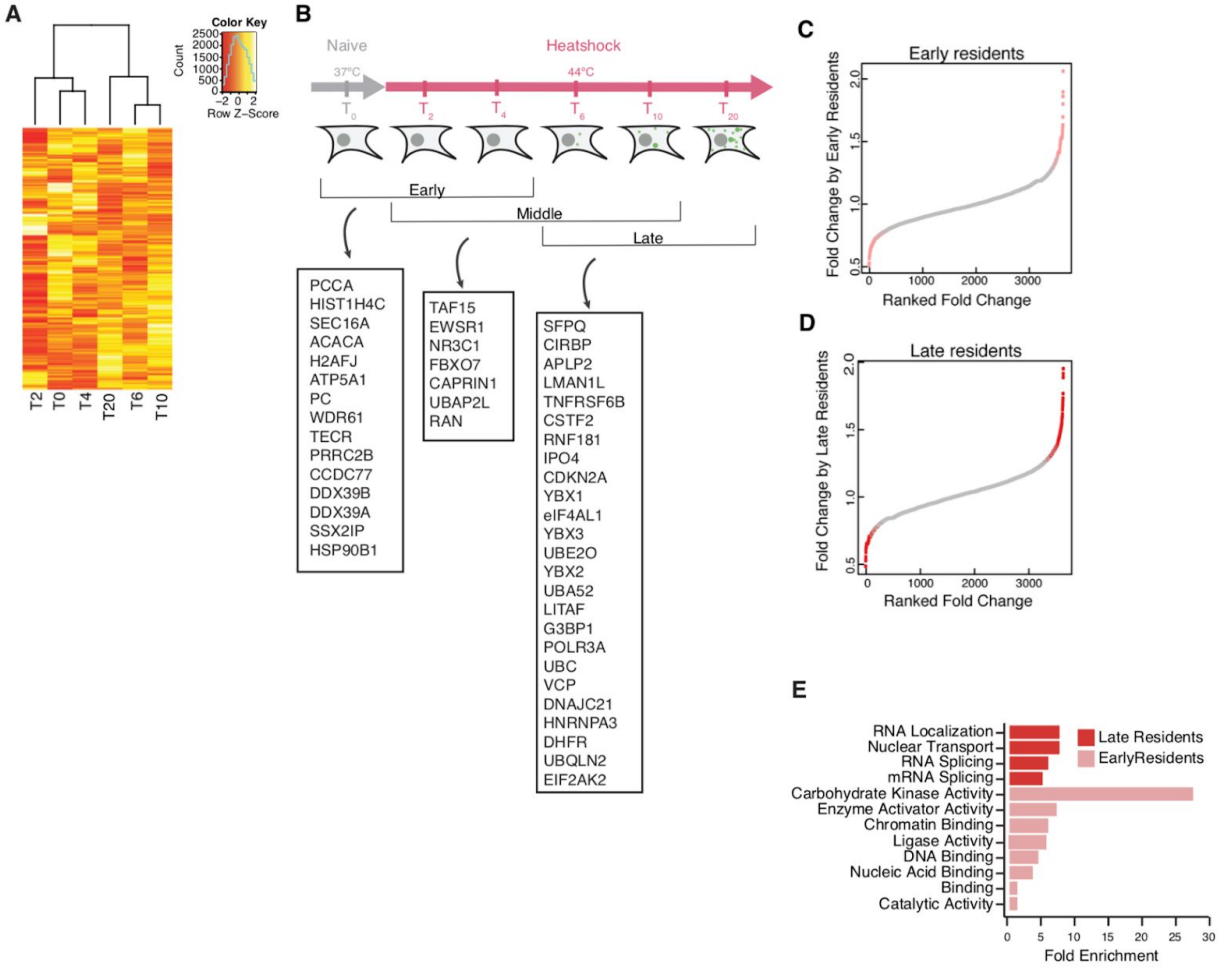
(B) Number of granules per cell in a immunofluorescence time course (n=35 per time point).

(C) First principle component as a function of heat shock stress time points.

(D) Component loadings for the first principle component of an APEX2-eIF4A1 heat shock time course.

(E) Heatmap result from ImpulseDE2 highlighting significantly changing (FDR ≤0.01) proteins that: decrease transiently with respect to heat shock (Cluster 1), increase transiently over time (Cluster 2), decrease gradually over time (Cluster 3), or increase gradually over time (Cluster 4) with respect to heat shock.

Padròn et al., Supplemental Figure 6



Supplemental Figure 6, related to Figure 6.

(A) Hierarchical clustering of an APEX2-eIF4A1 heat shock time course shows two main groups. Early residents: T_0, T_2, T_4; late residents: T_6, T_{10}, T_{20}.

(B) Diagram of heat shock time point groupings highlighting detected proteins at each grouping (early, middle, late).

(C and D) Ranked fold change plots highlighting enriched and depleted proteins (red) in either the (C) early or (D) late time points.

(E) GO term enrichment for early and late residents.

APEX-Seq reveals that different stresses yield distinct granule RNAs

The critical role for sequence-specific RNA binding proteins in SG assembly led us to analyze their RNA composition. We performed APEX-Seq on cells subjected to heat shock as well as cells treated with hippuristanol, a drug that directly targets eIF4A1 and induces RNA granule

formation through a phospho-eIF2α-independent mechanism(Panas, Ivanov, and Anderson 2016). While hippuristanol itself disrupts the eIF4A1 / RNA interaction, eIF4A1 is readily recruited into hippuristanol-induced stress granules(Mazroui et al. 2006). Both heat shock and hippuristanol cause substantial changes in the pattern of APEX2-eIF4A1 labeling (Figures S7A and S7B). Interestingly, we observed that distinct RNAs are enriched in eIF2α-dependent SGs formed after heat shock and in eIF2α-independent SGs induced by hippuristanol treatment. Heat shock enriches for longer RNAs with lower translational efficiency (Figures 7A-7B and S7C-S7E), similar to the pattern of mRNAs that co-purify with arsenite-induced stress granule cores(Khong et al. 2017). In contrast, these factors show little or no correlation with enrichment after hippuristanol treatment (Figures 7C-7D and S7F-S7H), suggesting that different transcripts may enter granules depending on the nature of the stress, especially when it modulates eIF4A1 directly.

Interestingly, while RNAs with longer ORFs were enriched in heat-induced stress granules, they took longer to enter these granules after heat shock. The length bias we saw 20 minutes after heat shock (T_{20}) was absent 4 minutes (T_4) post heat shock (Figure S7I). This change was driven by the accumulation of long transcripts in stress granules at later timepoints: ~45% of the RNAs enriched at T_4 were also enriched at T_{20}; in contrast, only ~12% of T_{20}-enriched transcripts were also enriched at T_4. The slower entry of long RNAs into stress granules may reflect the fact that elongating ribosomes take longer to finish translation on longer CDSes. Stress granule formation requires elongating ribosome to disengage from mRNAs(Buchan and Parker 2009), and run-off elongation from longer ORFs should take more time than run-off from shorter ORFs after stress-induced initiation shutoff.

Key stress granule markers such as TIA1, HNRNPA1, and G3BP1 are RNA-binding proteins, and indeed, we saw strong enrichment for RBPs in our proteomics data (Figure 5H). We thus took advantage of available enhanced RNA crosslinking and immunoprecipitation (eCLIP) data to search for correlations between SG enrichment and protein interactions across the transcriptome. In agreement with previous work, we found that RNAs enriched in either the heat shock or the hippuristanol SG transcriptome contain more TIA1 and HNRNPA1 binding sites than expected by chance(Khong et al. 2017) (Figures 7E and S7J). In contrast, G3BP1 showed enrichment only in hippuristanol SG transcripts (Figure S7K).

Padròn et al., Figure 7

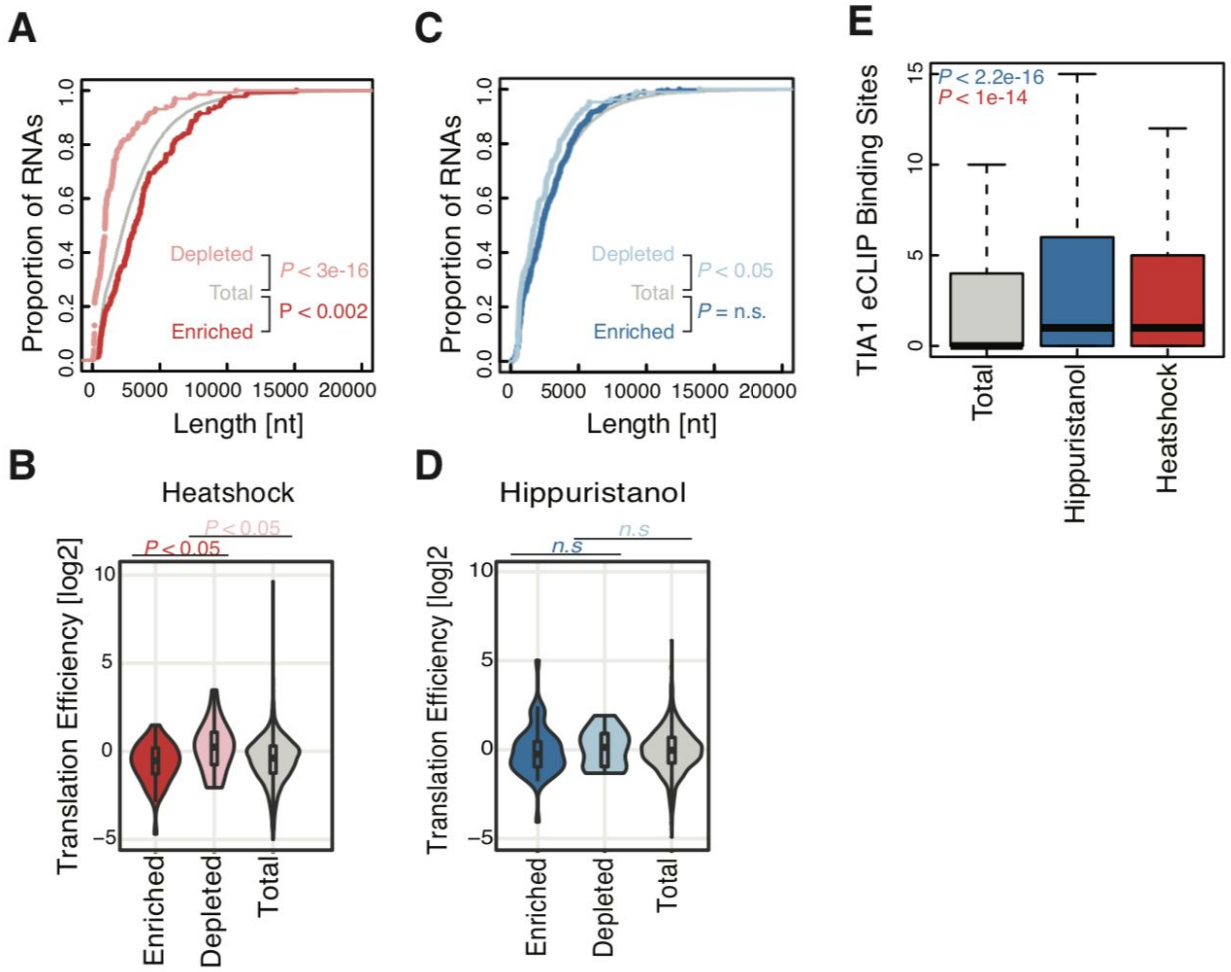


Figure 7. RNA composition of stress granules differs according to the stressor.

(A) Cumulative distribution plot for enriched and depleted RNA length for heat shock stress granule (T_{20} vs T_0). *P* values were calculated using the Mann-Whitney test.

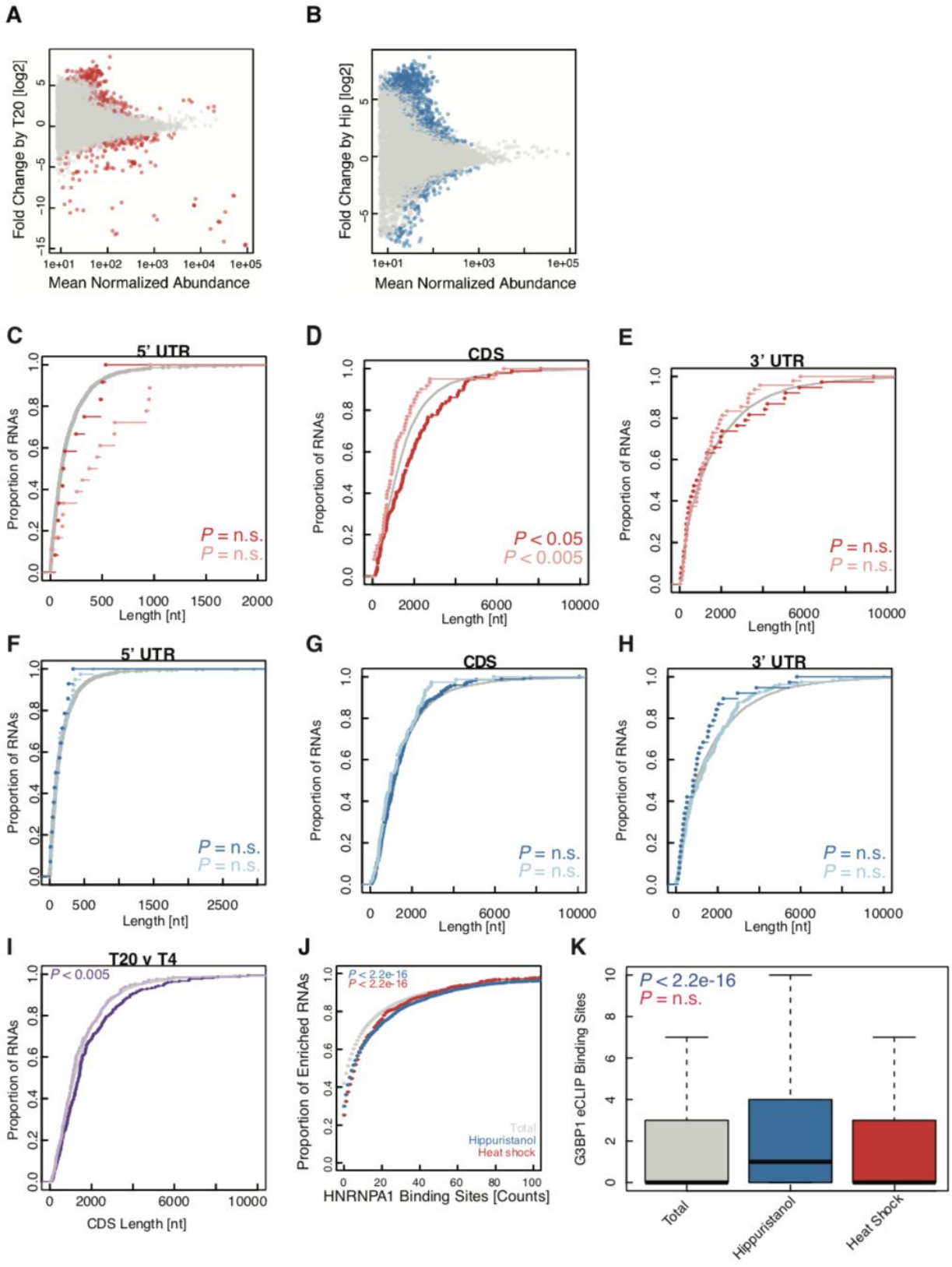
(B) Translation efficiency distributions for enriched and depleted RNAs during heat shock granule formation.

(C) As in (A) for hippuristanol stress granules.

(D) As in (B), for hippuristanol stress granule formation.

(E) TIA1 eCLIP binding site counts for either total, or enriched RNAs during hippuristanol or heat shock stress granule formation. *P*-values calculated using the Mann-Whitney test. Enrichment and depletion is calculated as statistically significant (FDR ≤0.05) and either top or bottom 2.5% of the fold change. Outliers were removed for clarity.

Padròn et al., Supplemental Figure 7



Supplemental Figure 7, related to Figure 7.

(A and B) MA plot for proximity-labeled RNAs in SGs. Ratio of stressed to unstressed cells for (A) APEX2-eIF4A1 heat shock time course (T_{20} vs T_0) and (B) hippuristanol vs DMSO. The significantly different RNAs with an FDR corrected p-value ≤0.05 are highlighted in (A) red and (B) blue.

(C-E) Cumulative distribution plots for enriched and depleted genes in (C) 5', (D) CDS, and (E) 3' UTRs for heat shock granules. *P* values were calculated using the Mann-Whitney test.

(F-H) As in (C-E) for hippuristanol-induced granules. Red and and line lines represent enriched and depleted RNAs from heat shock induced stress granules. Blue and light blue lines represent enriched and depleted RNAs from hippuristanol induced stress granules.

(I) Cumulative distribution plot comparing CDS length between APEX2-eIF4A1 T_{20} heat shock vs T_4 heat shock. Purple (T_{20} enriched), and lavender (depleted) RNAs are highlighted. *P* values calculated using the Mann-Whitney test.

(J) Cumulative distribution plot for eCLIP bindings sites of HNRNPA1 during hippuristanol (blue) or heat shock (red) stress granule formation. *P* values were calculated using the Mann-Whitney test.

(K) G3BP1 eCLIP binding site counts for either total, or enriched RNAs during hippuristanol or heat shock stress granule formation. *P* values calculated using the Mann-Whitney test. Enrichment and depletion is calculated as statistically significant (FDR ≤0.05) and either top or bottom 2.5% of the fold change.

Discussion

We have developed an approach to study sub-cellular RNA localization by proximity labeling and deep sequencing (Figure 1). The APEX-Seq approach captures patterns of localization both across (Figure 2) and within different compartments (Figure 3). Notably, the 10 - 100 nm range of APEX2 proximity labeling and the very general reactivity of radical-based APEX labeling promise powerful new capabilities for characterizing protein-RNA and protein-DNA structures in cells. It will complement chromatin immunoprecipitation (ChIP) and RNA crosslinking and immunoprecipitation (CLIP) by detecting weak or long-range interactions that are not well captured by crosslinking and circumventing the need for protein immunoprecipitation. Furthermore, the same APEX fusion protein can be used to characterize protein and RNA localization, promising new insights into dynamic and heterogeneous protein-RNA complexes within cells.

We have taken advantage of APEX-Seq in conjunction with proteomics to provide new insights into the organization of translation initiation complexes on active mRNAs and the composition of repressive RNA granules. Our data indicate that eIF4A1 assembles onto the 3′ end of the 43S preinitiation scanning complex, near the mRNA entry site (Figure 4). To date, the identification of eIF4F with respect to the preinitiation scanning complex remains unknown (Pisareva et al. 2008; Hinnebusch 2014). The similar position of other DEAD-box proteins at ribosome entry site (Oh et al. 2016; Hashem et al. 2013) suggested the similar role(s) in DEAD-box proteins in translation. We also found that eIF4A1 and eIF4E1, another component of eIF4F, show similar patterns of enrichment across the transcriptome despite their involvement in nearly all cap-dependent translation. Our results agree broadly with the transcript-specific effects of modulating eIF4F activity.

We also reported the RNA as well as the protein composition of stress granules *in vivo*. During heat shock, we find dynamic changes in the organization of the proteome that reflect the assembly of eIF4A1, along with many other proteins, into stress granules. Previous studies have argued that stable stress granule cores(Jain et al. 2016) form first, and a more dynamic outer shell then assembles onto them by liquid-liquid phase separation (LLPS) (Wheeler et al. 2016). In agreement with previous work, we found increased labeling of many stress granule core proteins early during granule formation (Wheeler et al. 2016). Interestingly, however,

certain core granule interactions appear to occur independent of stress (Figure 5), suggesting a model in which stress granule cores reflect at least in part the stabilization or enhancement of pre-existing interactions. We also found that the RNA content of stress granules can vary depending on the nature of the stress, with potential impacts on the transcriptome as well as on protein synthesis (Figure 7).

Matching both the spatial transcriptome using APEX-Seq and the spatial proteome using APEX-MS is a particularly powerful approach to better understand the organization of the cell. Critically, with the use of ratiometric comparisons, its use extends to non-membrane bound organelles that remain challenging to work with through more classic techniques, such as RNA granules. This approach addresses a need that cannot easily be met using existing techniques. Protein-RNA crosslinking has revolutionized our understanding of RNA-binding proteins and their targets, but it typically requires direct contact on the ~0.3 nm length scale of a covalent bond. Single-molecule fluorescence *in situ* hybridization is capable of measuring sub-cellular RNA localization, but it is limited by the ~100 nm resolution of optical microscopy and can target only a few, pre-chosen transcripts. The intermediate length scale of APEX-Seq complements these techniques and is well suited to address important questions about the organization of transcripts during synthesis, processing, transport, translation, and decay at sub-cellular resolution.

Materials and Methods

General procedures

HEK 293 Flp-In T-Rex cells (Invitrogen) were cultured in DMEM + GlutaMAX (ThermoFisher Scientific, 10566-016) with 10% FBS. APEX fusion plasmids were transfected along with pOG44 by X-tremeGENE 9 (Roche) and selected using 150 µg ml^{-1} of Hygromycin B and 15 µg ml^{-1} of Blasticidin to obtain stable integrants. TMT mass spec reagents (Thermo Fisher Scientific) were used for every quantitative mass spec experiment. Biotin-tyramide was purchased from Iris Biotech (catalog #: LS-3500.1000).

DNA construction

Plasmids were constructed by Gibson assembly in pCDNA5 or pET28a vectors, as indicated, using PCR amplicons generated using the primers listed below.

pCDNA5-APEX2-GFP: APEX2 was amplified from pcDNA3 APEX2-NES (Addgene # 49386).

APEX2 FWD GACTCTAGCGTTTAAACTTACACCATGGGAAAGTCTTACCCAACT

APEX2 REV TCCTCGCCCTTGCTCACCATGTCCAGGGTCAGGCGCTCCA

GFP FWD TGGAGCGCCTGACCCTGGACATGGTGAGCAAGGGCGAGGA

GFP REV CGGGCCCTCTAGACTCGAGCTCACTTGTACAGCTCGTCCA

pCDNA5-C1-APEX2: (The C1(1-29) coding sequence was ordered as a synthetic DNA from IDT)

APEX-GFP FWD AGCTTGGGTTTGCTGATGCCATGGTGAGCAAGGGCGA

APEX-GFP REV AGTACAACCACGGGGTCCATGGTGTAAGTTTAAACGCTAGAGTCCG

C1(1-29)-APEX2 FWD CTAGCGTTTAAACTTACACCATGGACCCCGTGGTTGTAC

C1(1-29)-APEX2 REV TCCTCGCCCTTGCTCACCATGGCATCAGCAAACCCAAG

pCDNA5-CBX1-APEX2: (CBX1 was amplified from human cDNA)

APEX-GFP FWD (1) CCCAAGAAAAAACGCAAGGTGGCTCGAGTCTAGAGGGCCCG

APEX-GFP REV (1)

CTTGTCGTCATCGTCTTTGTAGTCCATGGTGTAAGTTTAAACGCTAGAGTCCGG

CBX1 (ENSG00000108468) FWD

ATGGACTACAAAGACGATGACGACAAGGGGAAAAAACAAAACAAGAAGAAAGTGG

CBX1 (ENSG00000108468) REV

ACTCACAGTTGGGTAAGACTTTCCGTTCTTGTCATCTTTTTTGTCATCATCCT

APEX-GFP FWD (2)

TGACAAAAAAGATGACAAGAACGGAAAGTCTTACCCAACTGTGAGTGCT

APEX-GFP REV (2) CACCTTGCGTTTTTTCTTGGGGGCATCAGCAAACCCAAGCTC

pCDNA5-APEX2-eIF4A1: (eIF4A1 was amplified from human cDNA)

APEX-4A FWD (1)

ATGTTGCTGACCTCATCTGATGGTACCGAGCTCGGATCCACTAGTCCAGT

APEX-4A REV (1)

CTTGTCGTCATCGTCTTTGTAGTCCATGTCCAGGGTCAGGCGCTCCAGGG

APEX-4A FWD (2)
ATGGACTACAAAGACGATGACGACAAGATGTCTGCGAGCCAGGATTCCCGATCC
APEX-4A REV (2)
TGGATCCGAGCTCGGTACCATCAGATGAGGTCAGCAACATTGAGGGGCA

pCDNA5-APEX2-eIF4E1: (eIF4E1 was amplified from human cDNA)
APEX-GFP FWD AAAATAGGTTTGTTGTTTAAGCTCGAGTCTAGAGGGCCC
APEX-GFP REV TCCGGTTCGACAGTCGCCATGTCCAGGGTCAGGCGC
EIF4E (ENSG00000151247) FWD
TGGAGCGCCTGACCCTGGACATGGCGACTGTCGAACC
EIF4E (ENSG00000151247) REV
CGGGCCCTCTAGACTCGAGCTTAAACAACAAACCTATTTTTAGTGGT

pET28a-APEX2-eIF4A1: (recombinant protein expression)
APEX-eIF4A1 FWD tggacagcaaatgggtcgcgATGGGAAAGTCTTACCCAAC
APEX-eIF4A1 REV agtggtggtggtggtggtgcTCAGATGAGGTCAGCAACAT

pET28a-eIF4A1: (recombinant protein expression)
eIF4A FWD tggacagcaaatgggtcgcgATGTCTGCGAGCCAGGATTCCCGATCC
eIF4A REV AGTGGTGGTGGTGGTGGTGCTCAGATGAGGTCAGCAACAT

psiCHECK2: (*in vitro* transcription RNA reporter)
psiCHECK2 FWD taatacgactcactataggCGGAAACTGGAGCCTGAGGAGT
psiCHECK2 REV CGTCCTCCTGGCTGAAGTGGA

***in vivo* APEX labeling**

Cells were plated in 15 cm dish and cultured for 3 days with 1 μg ml^{-1} of tetracycline. Cells were treated like in[2]. Briefly, Biotin-phenol containing (500 μM final) prewarmed DMEM media was added to cells for 30 minutes prior to the start of the experiment. One mM final H_2O_2 was added to each dish for a total labeling time of 1 minute (unless stated otherwise). Cells were gently agitated for 1 minute, and quenched with 2X quenching buffer (10 mM Trolox and 20 mM sodium ascorbate in DPBS).

Protein Purification

BL21 Star (DE3) Escherichia coli cells (Invitrogen) transformed with plasmids encoding eIF4A1 (WT) or APEX2-eIF4A1 in a 1.5 L culture were cultivated at 37 °C with 50 μg ml^{-1} kanamycin and then grown at 16 °C overnight with 1 mM IPTG. The cell pellets were resuspended in His buffer (20 mM HEPES-NaOH, pH 7.5, 500 mM NaCl, 10 mM imidazole, 10 mM β-mercaptoethanol) with 0.5% NP-40, sonicated, and centrifuged at 35,000g for 20 min. The supernatant was incubated with 1.5 ml bed volume of Ni-NTA Superflow (Qiagen) for 1 h. The beads were loaded on a gravity column and washed with His buffer containing 1 M NaCl. The proteins were eluted with 50 mM Na-phosphate buffer, pH 7.5, 500 mM NaCl, 100 mM Na_2SO_4, 250 mM imidazole. Samples were run through an FPLC HiTrap Heparin HP affinity column (GE Healthcare) with no reducing agent for further purification. Fractions 8-11 were collected. Samples were mixed with 0.25 volumes of 80% glycerol, flash-frozen in liquid nitrogen, and stored at −80°C. All purification steps were performed at 4°C.

***in vitro* APEX labeling**

Recombinant APEX2-eIF4A1 (1.5 μM) was pre-incubated with hemin (4.5 μM) at room temperature for 1 hour. Excess hemin was removed by several rounds of gel filtration through a MicroSpin G-25 column (GE Healthcare). This solution was then combined with biotin-phenol, and a five molar excess of a NanoLuc reporter RNA, in addition to 1 mM ADP and 1 mM $MgCl_2$. The samples were then combined with APEX labeling buffer (10 mM Tris, pH 7.5, 150 mM NaCl, and 10% glycerol). The reaction started when peroxide was added to the mix and proceeded for a total of 1 minute *in vitro* labeling. The reactions were then stopped using either TRIzol or oligo binding buffer (Zymo). RNA was extracted using manufacturer's protocol, and loaded onto a dot-blot apparatus.

***in vitro* transcription of a reporter RNA**

A reporter RNA for *in vitro* APEX labeling experiments was synthesized using the psiCHECK-2 plasmid with an *in vitro* transcription kit (Cellscript). The RNA was later purified using an oligo clean and concentrator kit (Zymo).

Heat shock, Hippuristanol, and PP242 treatment

Cells were plated onto 10 or 15 cm dishes, and expression was induced for 3 days with 1 μg ml^{-1} of tetracycline. Media was then replaced with fresh media containing biotin-tyramide

(500 μM) for 30 minutes. Cells were then placed in a 44°C water bath for better heat transfer during heat shock experiments. APEX labeling was performed immediately after each treatment. For the hippuristanol treatment, cells were treated with biotin-tyramide containing media and 1 μM of hippuristanol for 30 min prior to the start of the APEX reaction. Labeling occurred for 30 seconds. Labeling reactions were quenched with 2X quenching buffer (see above) and cells were immediately lysed in TRIzol. For the PP242 treatment, cells were incubated with biotin-tyramide and 2.5 μM PP242-containing media for 30 minutes prior to the start of the APEX reaction. Labeling occurred for 1 minute. Reactions were quenched and samples were prepared for both APEX-MS and APEX-Seq.

Streptavidin purification of biotinylated protein

After the APEX labeling reaction, cells were quenched with 2X quenching solution (see above) once, followed by a 1X quenching solution wash. One 1X quenching solution wash was used to resuspend cells, which were then gently pelleted and lysed in 800 μl of lysis buffer containing 1X quenching reagents (1% Triton, 0.1% SDS, 20 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM MgCl2, 5 mM trolox, 10 mM sodium ascorbate, and one tablet (per 10 ml) of cOmplete Mini Protease Inhibitor Cocktail). Lysates were clarified by centrifugation for 10 min at 20,000 x g, 4°C. Streptavidin beads (Pierce) were equilibrated with lysis buffer for a total of two washes. Lysate was mixed with streptavidin beads at a volumetric ratio of 8 volumes beads per 5 volumes lysate and incubated at RT for 1 hr. Beads were washed twice with lysis buffer, once with 1M KCl solution, once with 2M urea, pH 8, and twice with lysis buffer (w/o detergent or quenching reagents). Biotinylated proteins were eluted using by boiling samples in 8M urea, pH 8 for 3 minutes at 98°C.

Streptavidin purification of biotinylated RNA and library preparation

After an *in vivo* labeling reaction, cells were lysed with TRIzol and RNA was purified by by precipitating RNA from the aqueous phase. Samples were treated with DNase I (NEB) and ~100 μg of RNA was then gently fragmented using 10^{-5} units of RNase A/T1 (Thermo Fisher Scientific) by placing 1 μl of the RNase cocktail in the lid of each tube and initiating the reaction by simultaneously centrifuging the samples. These were then incubated for 10 minutes at 37°C in a final RNA concentration of 1μg μl^{-1}. The reaction was immediately quenched using 400 μl of TRIzol. RNA was extracted by either adding 500 μl of TRIzol, followed by 200 μl of Chloroform, and by precipitating RNA from the aqueous phase, or by Direct-Zol (Zymo) purifications

following the manufacturer's instructions. C1 Streptavidin beads (10 µl per sample; Thermo Fisher Scientific) were washed three times with Buffer 1 (1 mM $MgCl_2$, 0.5% sodium deoxycholate in PBS), washed once and blocked for 30 minutes with Blocking Buffer (5X Denhardt's reagent and 150 µg/ml Poly IC (InvivoGen) in Buffer 1). Afterward, blocking Buffer was removed and replaced with extracted RNA samples in fresh Blocking Buffer. Samples were incubated at room temperature for 1 hour and then washed twice with Buffer 2 (6M Urea, pH 8, 0.1% SDS in PBS), once with Buffer 3 (2% SDS in PBS), once with Buffer 4 (750 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, in PBS), and once with Buffer 5 (150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, in PBS). RNAs were eluted from Streptavidin beads by denaturation in 300 µl TRIzol (Qiagen) and extracted using either the above mentioned TRIzol/Chloroform procedure or Direct-zol kit, following the manufacturer's instructions. RNAs were eluted in 6 µl H_2O. RNAs were then fragmented for 7 minutes (as recommended for RIN values > 2 and < 7) at 94°C, following the NEBNext Ultra II Directional RNA Library Prep Kit (NEB, catalog # E7760S) instructions.

TMT mass spec

100 µg of eluted proteins were brought up to 100 µl with 100 mM TEAB. 5 µl of 200 mM TCEP were added and samples were incubated at 55°C for 1 hour. Fresh iodoacetamide was made up in 100 mM TEAB (final concentration 375 mM). 5 µl of the 375 mM iodoacetamide were added to the sample and incubated for 30 minutes protected from light at room temperature. Six volumes (or more) (~600µL) of pre-chilled (-20°C) acetone were added. Samples were allowed to precipitate overnight at -20°C. Samples were precipitated at 8,000 x g for 10 min at 4 °C and resuspended with 100 µl of 50 mM TEAB. For the tryptic digest, 2.5 µl of trypsin (1 µg/µl) were added to 100 µg of protein and incubated overnight at 37 °C. Peptides were quantified and normalized using a quantitative colorimetric peptide assay (Thermo Fisher Scientific). 41 µl of anhydrous acetonitrile were added to each TMT labeling reagent vial and the reagent was allowed to resuspend at RT for 5 min with occasional vortexing. The reduced and alkylated protein digest were transferred to each TMT Reagent vial. Reactions were incubated for 1 hr at RT. 8 µl of 5% hydroxylamine was added to each sample and incubated for 15 minutes to quench the reaction. Samples were dried in a Speed-Vac or frozen in -80°C for storage and transportation. Samples were run on an OrbiTrap mass spectrometer (ThermoFisher Scientific) using high pH fractionation, and analyzed using Proteome Discoverer. DESeq2[72] was used to perform quantitative ratiometric comparisons between each APEX fusion and APEX2-GFP.

Translation initiation factors highlighted in Figure 4A are: EIF1AX (P47813), EIF2D (P41214), EIF2S1 (P05198), EIF2S2 (P20042), EIF3A (Q14152), EIF3CL (B5ME19), EIF3D (O15371), EIF3E (P60228), EIF3F (O00303), EIF3G (O75821), EIF3H (O15372), EIF3I (Q13347), EIF3J (O75822), EIF3K (Q9UBQ5), EIF3L (Q9Y262), EIF3M (Q7L2H7), EIF4A1 (P60842), EIF4B (P23588), EIF4E (D6RBW1), EIF4G1 (Q04637-3), EIF4H (Q15056-2), EIF5 (P55010), EIF5B (O60841), PABPC1 (P11940), PABPC4 (Q13310-3)

Western Blots

Samples were blocked with 0.6% milk in TBST (0.05% tween-20) for 1 hour at RT. Membranes were incubated with the corresponding primary antibodies in signal enhancer solution 1 (Hikari NU00101) for 1 hour at RT, and secondary antibodies in signal enhancer solution 2 (Hikari NU00102) for 1 hour at RT. Three 5 min washes were performed between antibody incubations. STREP-HRP (1:1,000, CST 3999) was used to blot for biotinylated protein. Anti-FLAG (1:1,000, CST 2368), and anti-eIF4A (1:1,000, CST 2490) antibodies were used as primary antibodies and HRP-conjugated anti-rabbit IgG (CST 7074) was used as a secondary antibody. Chemiluminescence was induced by SuperSignal West Dura Extended Duration Substrate (Thermo Scientific) and images were acquired by a FluorChem R imaging system (ProteinSimple).

Immunofluorescence

Cells were plated onto poly-lysine coated coverslips in 12 well plates and APEX fusion expression was induced by adding 1 μg ml^{-1} of tetracycline. Two days later, cells were fixed in 3.7% paraformaldehyde (PFA) in PBS for 10 minutes at room temperature, and permeabilized with 1% Triton X-100 in PBS for 10 min at room temperature. Cells were then blocked with 3% BSA in PBS for 30 min at room temperature. All primary antibodies were incubated overnight at 4 ºC and secondary antibodies at 37 ºC for 1 hour. Several PBS washes were performed in-between antibody staining. Coverslips were mounted onto slides with ProLong Gold mountant (Thermo Fisher Scientific).

Code availability

Scripts to run the analyses mentioned above are available upon request.

Accession numbers

Raw sequencing data are available for download GEO: GSE121575.

Dot Blot Assay

All dot blot assays were performed using the Dot-Blot Microfiltration Apparatus (Bio-Rad). Zeta-Probe membranes (Bio-Rad) were submerged in H_2O for 5 min. RNA samples were loaded onto the apparatus and the solution was gently pulled through the membrane by vacuum suction (setting 3). Zeta-probe membranes were then cross-linked two times at 200 μJ/cm^2 (CL-1000 Ultraviolet Crosslinker) and blocked for 2 hr using Odyssey blocking buffer + 1% SDS. 800CW Streptavidin (1:10,000 LI-COR Biosciences) was then added for 30 minutes in fresh blocking buffer with 1% SDS. Membranes were visualized in the LI-COR (Odyssey CLx). A ssDNA oligo with a 3'-biotin modification was used as a positive control (IDT).

Classifier predicting RNA granule residency

A regularized logistic regression classifier was built using biophysical features from each protein and known protein domains from PFAM. Continuous features were scaled and centered. 70% of the data was partitioned into a training set. Repeated k-folds cross validation (k = 10) was performed for alpha and lambda parameters that minimized mean squared error.

Chapter 3: Conditional proximity proteomics by a split-APEX reveals the unique proteomic landscape of translation initiation factors

Abstract

Spatial proximity labeling has shown to be a powerful tool to understand macromolecular organization in the cell. We present a conditional splitAPEX proximity labeling tool that can monitor the spatial landscape of binary interactions. The promise of splitAPEX is a refinement of the spatial landscape of macromolecules, as well as a better understanding of the shared spatial and temporal landscape of paired interactions. Since enzymatic activity of splitAPEX depends on the coincidence of two protein halves, the specificity should be proportional to the square of the localization of the components in question. To showcase the system, we have used splitAPEX on the initiation factors eIF4H (with respect to eIF4A), and eIF4B (with respect to eIF4A), as a way of describing the unique proteomic landscapes these two initiation cofactors experience. splitAPEX proximity labeling will be useful for monitoring the spatial landscape of macromolecular complexes, and conditional protein interactions.

Introduction

Advances in proximity labeling strategies have showcased the utility of these tools in better understand the spatial organization of protein complexes (Lam et al. 2015); (Roux et al. 2012); (Rhee et al. 2013). While traditional protein isolation and purification have been classically used to probe neighboring interactions, these techniques fail to capture important information due to weak protein-protein interactions that dissociate upon cell lysis and dilution. Moreover, macromolecules sometimes interact indirectly through neighboring members of a protein complex, causing these factors to be missed through traditional purification strategies. This results in the partial understanding of the neighboring components of any given protein of interest. Proximity labeling strategies using APEX2 circumvent this issue by catalyzing promiscuous enzymatic reactions that spatially probe the cellular landscape surrounding the enzymes catalytic site. Because radical half-lives are on the order of milliseconds, APEX2 can in theory produce a tighter spatial labeling radius than tools such as BirA, since the adenylate ester intermediate exists on the order of minutes (Rhee et al. 2013). However, spatial tagging proteins of interests (POI) with spatial proximity enzymes come with major limitations as POI are often not always engaged in their activity of interest, resulting in the spatial proximity results

from both the functionally engaged and functional non-engaged proximity landscapes, making spatial interpretation challenging.

The DEAD-box RNA helicase, eIF4A, is essential for translation initiation. This enzyme is part of the eIF4F complex, an important component that physically interacts with the 7-methylguanosine (m^7G) cap at the 5' ends of mRNAs required for cap-dependent translation initiation in eukaryotes. eIF4F integrates several important signaling pathways including AKT/mTOR pathway and the RAS/RAF/MEK/ERK/MNK MAPK pathway, contributing to many hallmarks of cancer (Malka-Mahieu et al. 2016). eIF4A itself is thought to unwind structure in the 5'-UTR, crucial for 40S scanning and 80S complex formation at the first AUG start codon. eIF4A alone is non-processively given the nature of DEAD-box helicases, however, coupled with eIF4G and eIF4B its enzymatic processivity increases 100-fold (Özeş et al. 2011). Interestingly, the translation initiation co-factor eIF4H shares a mutually exclusive binding site on eIF4A (Rozovsky, Butterworth, and Moore 2008), however, eIF4H differs in the degree in which it can increase the enzymatic activity of eIF4A (Özeş et al. 2011); despite being homologous to eIF4B, and associating with eIF4A(Marintchev et al. 2009). While the function of these translation initiation components have been known for almost three decades, less is known about how these initiation factors differentially assemble onto different mRNAs, and whether the biophysical properties of some mRNAs require some cofactors over others. New tools are therefore needed to address these questions.

Here, we developed split APEX (splitAPEX), a conditional spatial proximity labeling tool. splitAPEX gives the user conditional information on the spatial proximity only if the two components in question physically interact (AND-gate spatial proximity). splitAPEX gives more spatial and temporal refinement over the current proximity labeling tools, while keeping assay and mass spec approaches straightforward to use and implement. To showcase the system, we have used splitAPEX on the initiation factors eIF4H (with respect to eIF4A), and eIF4B (with respect to eIF4A), as a way of describing the unique proteomic landscapes these two initiation cofactors experience.

Results

Rational design of a splitAPEX protein

We began our rational design of split APEX by performing a multiple sequence alignment with other peroxidases in PFAM and assessing a number of metrics (Figure S1, S2): 1) how well conserved each amino acid is. The intuition being that well conserved sites are likely crucial for enzyme folding, and catalytic function. 2) How much mutual coevolutionary relationship there is speaks to how important the inter-relatedness of amino acids is. This can also highlight important functional or structural relationships that should be avoided in designing split proteins. 3) How solvent accessible is the position? The accessibility of the protein in question makes sure hydrophobic pockets are not being exposed. 4) Whether the split site was positioned inside, or outside of secondary structures. 5) And, related to 4) the b-factors of the position in question. By assessing each of these features in APEX, position 200-201 was selected, since it was in a solvent accessible region, was not a well conserved site across evolutionary time, and did not contact important catalytic residues (Fig 1B, Fig S1A). Recombinant splitAPEX proteins were made. The small fragment [201-250] was fused to MBP for protein stability, and the large fragment [1-200] was attached onto a Halo fusion protein. Upon a FAM-halo ligand conjugation, no detectable binding was measured when both halves of splitAPEX were introduced at the concentrations tested (0.1-10 μM) in a fluorescence polarization assay (data not shown), suggesting low levels of background from spontaneous association when both halves are separately expressed inside cells. Also, the large fragment [1-200] alone does not seem to show any catalytic activity Fig. S3D. Therefore, any level of background from this system likely comes from stochastic encounters of both the large and small splitAPEX fragments.

Figure 1

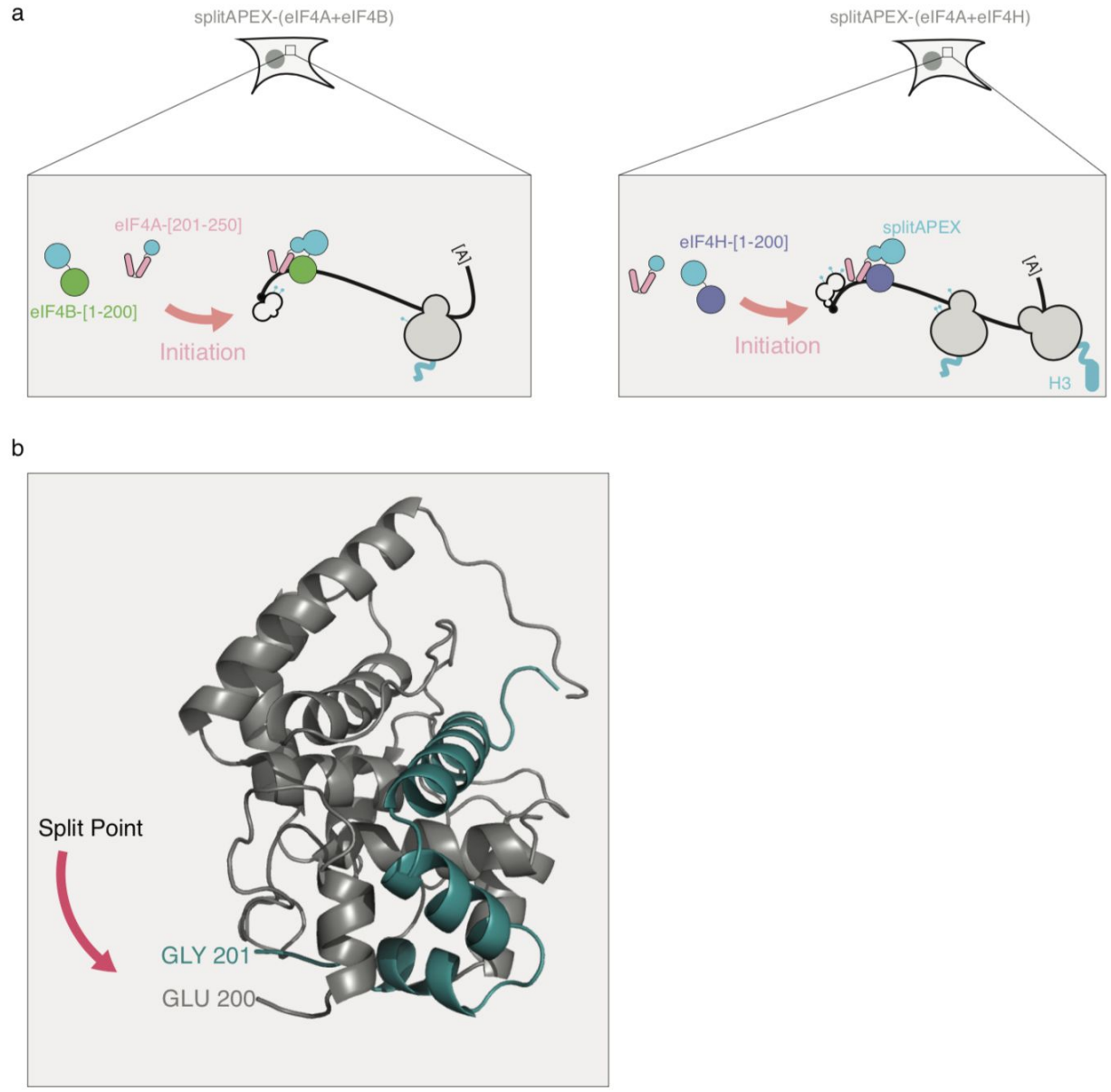
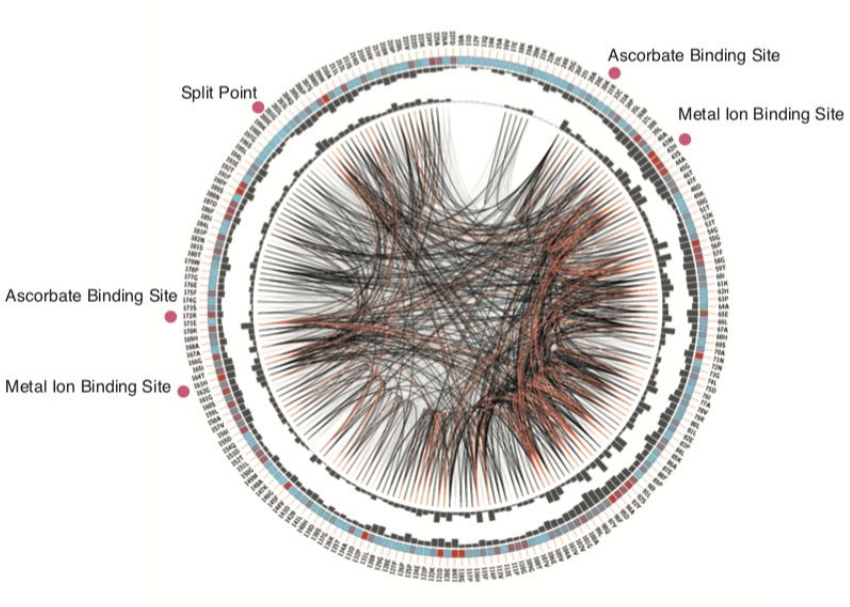


Figure 1.

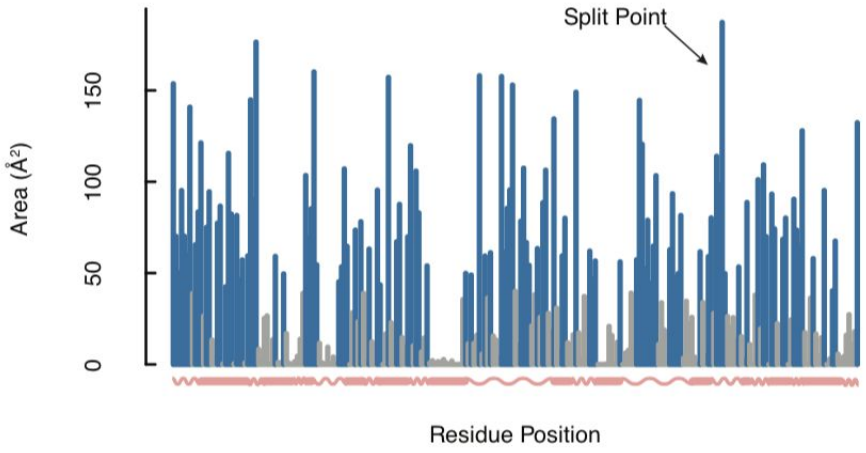
A) splitAPEX fused to the translation initiation protein eIF4A, and the accessory factors eIF4B, or B) eIF4H. C) Rational design of a split site on the APEX crystal structure. Glycine 201 and glutamine 200 were selected as the split point. The small APEX fragment (201-250) is highlighted in teal.

Figure S1

S1A



S1B

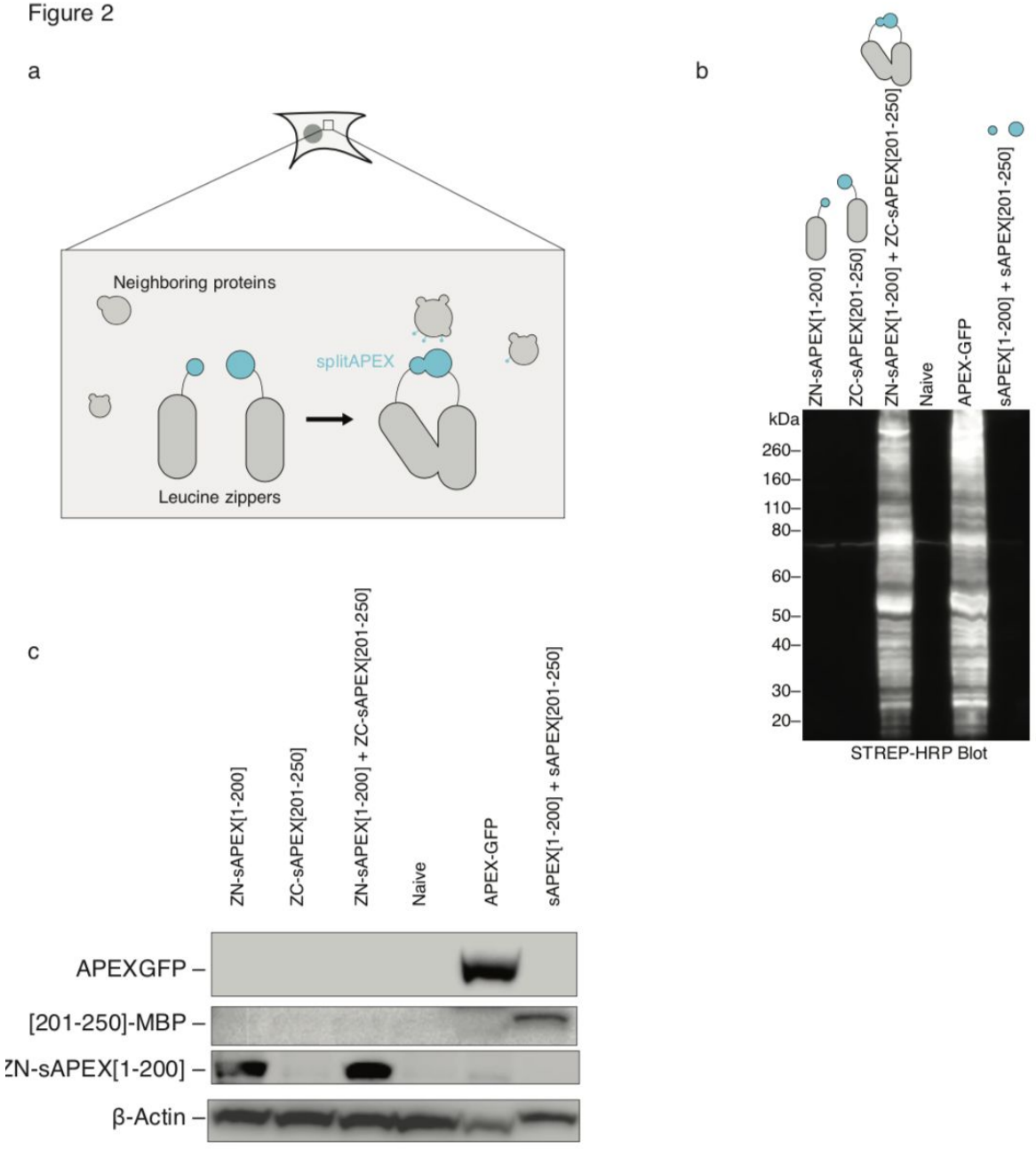


(**S1A**) Circo plot from a multiple sequence alignment of all peroxidases in PFAM. The outer numbers and letters signify amino acid position. The colored squares signify consevation score (red being more conserved).The barplots show mutual information scores between positions. This is also highlighted with the interconnected lines showing more mutual information (red) or less mutual information (gray) between positions. (**S1B**). Solvent accessibilty of APEX. In blue are regions that contain more than 40 angstrom squared of solvent accessibility. The split point is highlighted. Below is a representative diagram of the secondary structure of APEX.

in vivo protein fragment complementation assay

splitAPEX protein fragments were fused onto antiparallel leucine zippers (Ghosh, Hamilton, and Regan 2000) as a way to reconstitute APEX activity. The use of antiparallel leucine zippers can reconstitute *in vivo* enzymatic activity of splitAPEX Fig 2. The k_{off} of the antiparallel leucine zipper likely contributes in large measure to the degree to which enzymatic activity can be reconstituted; likewise, the reconstitution of enzymatic activity in any context will be in part governed by this property.

Figure 2

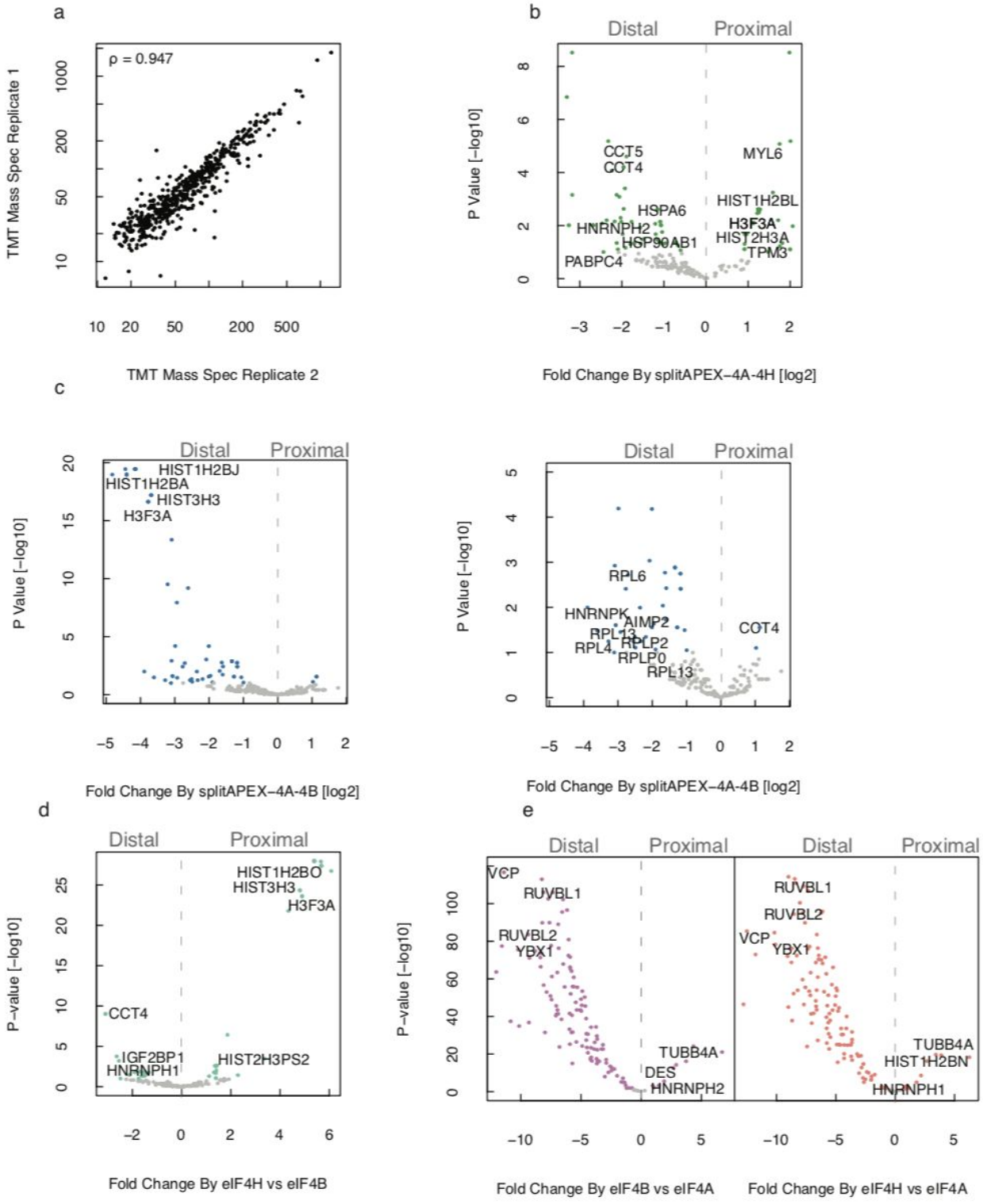


(**a**) Leucine system diagram diagram. Blue sticks indicate biotinylated neighboring proteins. (**b**) Fragment complementation assay. STREP-HRP western blot showing reconstituted activity levels of splitAPEX. (**c**) Western blot for fragment complementation fusion proteins.

Monitoring translation initiation with splitAPEX

This interesting observation raised the possibility in our minds of probing differential translation initiation programs. To gain further insight into this hypothesis, and to showcase the utility of conditional proximity labeling tool, we engineered splitAPEX tagged eIF4A-[201-250], eIF4H-[1-200], and eIF4B[1-200] proteins in order to monitor the proteomic landscape of mRNAs that either recruit eIF4H or eIF4B, alongside eIF4A and the rest of the 4F complex. Indeed, these splitAPEX fusions reconstituted activity levels upon splitAPEX-(eIF4A-eIF4H/B) complex formation (Fig S4). We used quantitative TMT mass spectrometry, and included a diffuse splitAPEX-(antiparallel leucine zipper) control in order to perform ratiometric comparisons. Overall, our quantitative TMT mass spec samples correlate well with each other Fig 3A (ρ=0.947). Cell lines expressing splitAPEX-(eIF4H+eIF4A) showed an enrichment (closer in proximity) toward histone H3 and H2B family members, likely due to co-translational proximity labeling. CCT components, which are responsible for properly folding newly translated proteins, were depleted in labeling by splitAPEX-(eIF4H+eIF4A). Conversely, PABPC4 appears to be depleted (Fig 3B). Interestingly, splitAPEX-(eIF4B+eIF4A) showed a strong depletion of histone proteins. CCT4 also appears to be in closer proximity to eIF4B-eIF4A translation complexes than eIF4H-eIF4A translation complexes, Fig. 3C. Specific RNA binding proteins appear differentially proximal to some translation initiation complexes over others. For instance, IGF2BP1 (also called CRD-BP), appears to be more proximal to translation initiation complexes containing eIF4A+eIF4B than eIF4A+eIF4H (Fig 3D). Comparing our splitAPEX translation initiation components to full length APEX-eIF4A shows a depletion of specific stress granule components, like VCP, YBX1, and RUV-like proteins (Jain et al. 2016) (Fig. 3E). These data suggest the use of specialized translation initiation complexes on elements of the transcriptome, and showcase the use of conditional proximity labeling tools over traditional ones.

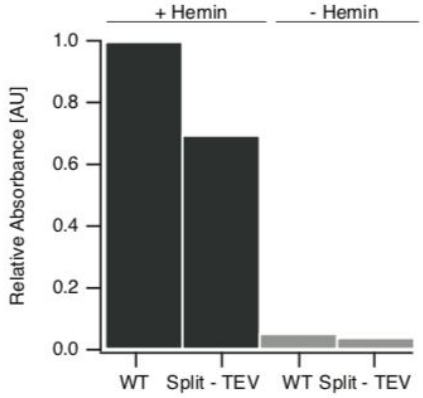
Figure 3



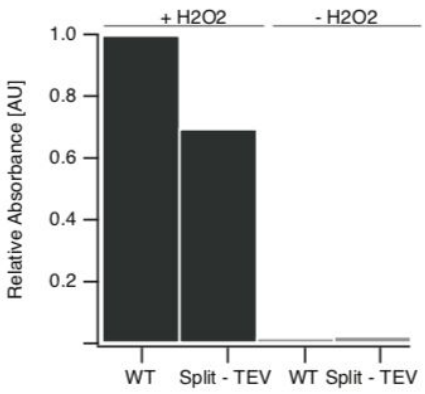
(**a**) Peptide correlation plot. (b) splitAPEX-(eIF4H+eiF4A) ratiometric comparison relative to splitAPEX-(leucine zipper). (c) splitAPEX-(eIF4B+eiF4A) ratiometric comparison relative to splitAPEX-(leucine zipper). (d) splitAPEX-(eIF4H+eIF4A) ratiometric comparison relative to splitAPEX-(eIF4B+eIF4A). (e) spitAPEX-(eIF4H/B+eIF4A) ratiometic comparison with respec to APEX-eIF4A (full length)

Figure S3. *in vitro* activity assay

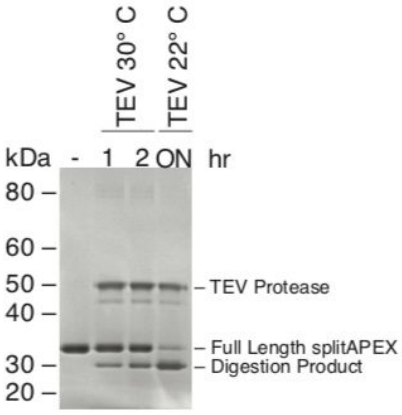
A. Dependence for Hemin between WT and splitAPEX recombinant protein



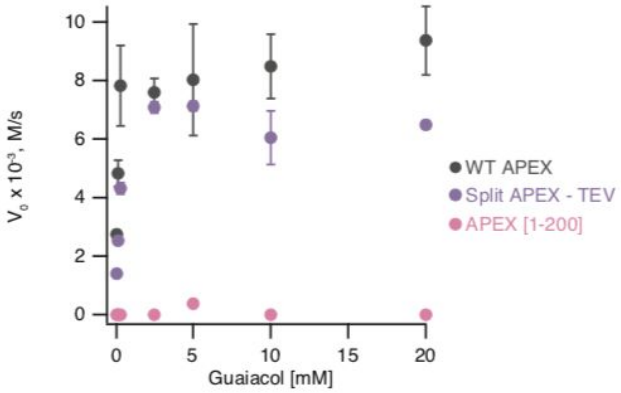
B. Dependence for H2O2 between WT and splitAPEX recombinant protein



C. Commassie stain of a TEV Protease digested splitAPEX.



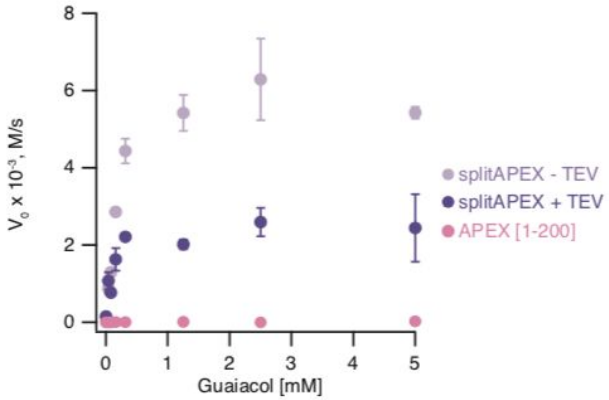
D. Guaiacol activity assay for WT and splitAPEX recombinant protein. APEX[1-200] showed no activity by itself



WT APEX
K_m = 0.113 ± 0.034
V_{max} = 8.71 ± 0.478

Split APEX -TEV
K_m = 0.230 ± 0.057
V_{max} = 6.953 ± 0.337

E. Guaiacol activity assay for WT and splitAPEX +/- TEV Protease digestion.



Split APEX - TEV
K_m = 0.188 ± 0.040
V_{max} = 6.259 ± 0.333

Split APEX + TEV
K_m = 0.286 ± 0.029
V_{max} = 2.495 ± 0.181

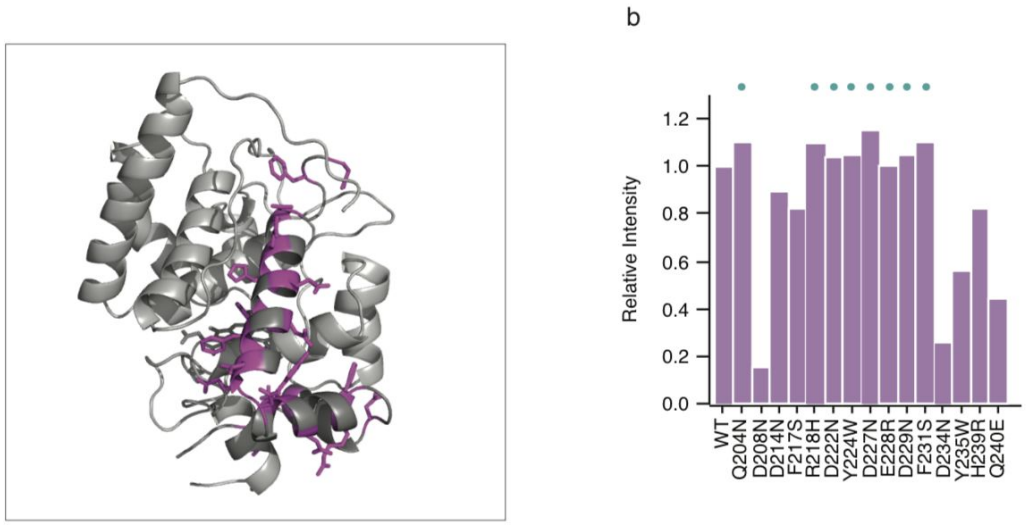
Figure S3.

A) in vitro guaiacol oxidation activity assay measuring the dependence of Hemin, and B) peroxide in the proximity labeling reaction. C) Commassie stain showing a splitAPEX with a TEV site insertion at position 200/201. Samples were subjected to TEV protease digestion, resulting in almost complete production of a fully split product. D) Guaiacol activity assay measuring the Km and Vmax of the wild type (WT), and splitAPEX recombinant protein without TEV digestion. The [1-200] APEX fragment is catalytically dead. E) Guaiacol activity assay upon TEV digestion, showing a dependency of the two fragments to guaiacol oxidation.

Targeted splitAPEX mutagenesis decreases background activity levels

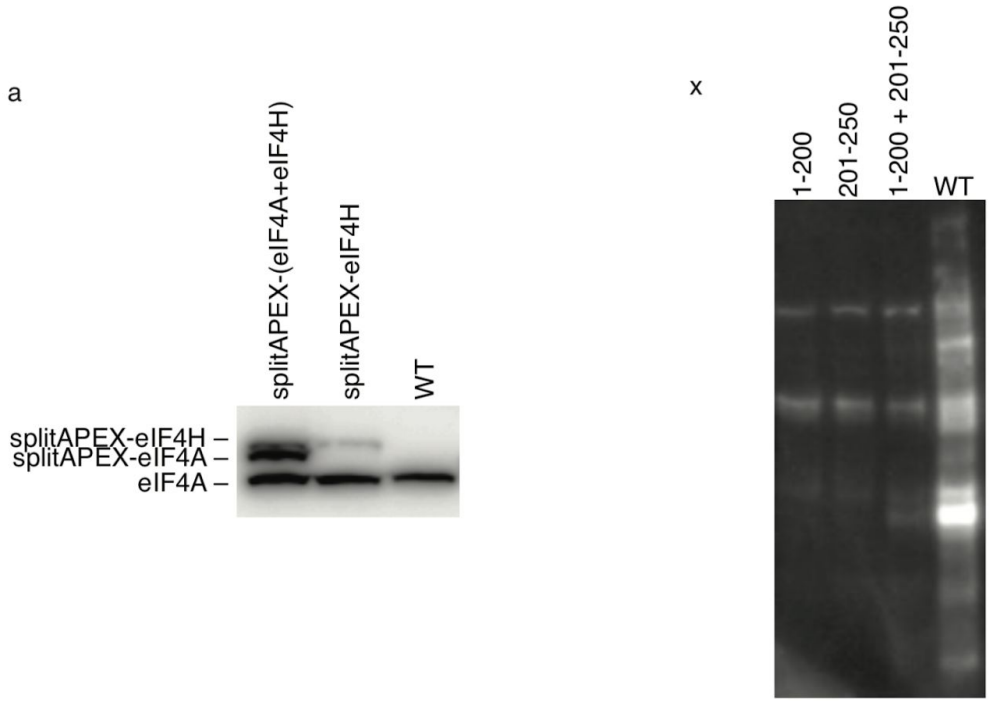
Under circumstances in which a researcher wants to investigate subtle spatial differences that coexist within the same cellular compartment, even modest amounts of background can make interpreting the signal of interest challenging. With this in mind, splitAPEX mutants were generated in order to reduce background levels. While the [1-200] fragment alone has no detectable activity (Fig S3D), chance encounters exist when expressing both, untethered halves in mammalian cells (Fig S4B). Previous work on split systems showed that disrupting electrostatic and helix stacking interactions being the most fruitful at diminishing background activity levels (Dixon et al. 2016) (Wehrman et al. 2005). With this in mind, we mutagenized residues that disrupted electrostatic and stacking interaction on the small APEX fragment, APEX[201-250] (Fig 4A). Several mutations reduced background levels of stochastic encounters, while maintaining WT levels of activity when matched mutations were introduced into full length APEX (Fig. 4B).

Figure 4.



(**a**) Mutagenesis of splitAPEX. In pink are the selected residues mutagenized in the small fragment. The small fragment is highlighted in darkgray, and the large fragment in lightgray. (**b**) Quantification of mutagenesis of splitAPEX. STREP-HRP quantification shows mutations that reduce or maintain wild-type activity levels. Blue dots indicate mutations that maintain wild-type activity levels while reducing chance stohastic encounters when expressing both halves of splitAPEX

Figure S4



(**a**). Western blot for splitAPEX-eIF4A, splitAPEX-eIF4H, and splitAPEX-eIF4B. (**x**) STREP-HRP assay for background levels of activity expressing both untethered halves of splitAPEX

Discussion

Here we present a conditional splitAPEX proximity labeling tool that can monitor the spatial landscape of binary interactions. The promise of splitAPEX is a refinement of the spatial landscape of macromolecules, as well as a better understanding of the shared spatial and temporal landscape of paired interactions. Since enzymatic activity of splitAPEX depends on the coincidence of two protein halves, the specificity should be proportional to the square of the localization of the components in question. We show the reconstitution of enzymatic activity using antiparallel leucine zippers in a protein fragment complementation assay. To showcase the utility of splitAPEX we have used it to probe translation initiation using the cofactors eIF4B, and eIF4H, conditioned on eIF4A, coupled with quantitative TMT mass spectrometry. This revealed eIF4H-eIF4A translation initiation complexes more proximal toward certain histone proteins, while eIF4B-eIF4A translation initiation complexes were statistically more likely to be

distal from these proteins. Conversely, PABPC4 appears to be distal from eIF4A-eIF4H translation complexes. This falls in line with the peculiarity of histone mRNAs, which are the only known cellular non polyadenylated mRNAs (Marzluff 2005), and are shuttled out of the nucleus through the interaction of a stem-loop binding protein and a 3' stem-loop on histone mRNAs(Marzluff, Wagner, and Duronio 2008), omitting any requirement for poly(A)-binding proteins. Differential proximity of certain protein folding chaperones, like CCT, also exist between these two types of translation complexes. This may highlight differences in protein folding requirements for the mRNAs these complexes associate with. mRNAs containing the RNA binding protein IGF2BP1 may implicate eIF4B-containing translation initiation complexes with IGF2BP1 biology, which is thought to stabilize c-Myc1 mRNA and has been shown to be misregulated in human breast cancers (Doyle et al. 2000).

Comparing these data to wild-type APEX-eIF4A, we find a strong depletion of stress granule forming proteins like VCP and YBX1 for both eIF4H-eIF4A, and eIF4B-eIF4A translation initiation complexes relative to eIF4A alone. This suggest that when eIF4A is not complexed with initiation factors it may be more proximal to stress granule forming proteins. Interestingly, YBX1, eIF4A, VCP, and RUV-like proteins are all "core" components of stress granules(Jain et al. 2016). Recent evidence has also suggested that stress granule "seeds" seem to exist prior to a stress response (Youn et al. 2018). Perhaps a way of establishing robust cellular remodeling in the face of stress is to pre-seed the cellular environment with macromolecular complexes ready to create fast, dynamic responses.

Under conditions in which subtle differences need to be teased out between paired interactions, minimizing background activity levels may be necessary to detect a signal over background. We have performed a targeted mutagenesis on the small splitAPEX fragment in search for mutants that minimize background levels of activity, while maintaining WT levels of activity in the full length APEX protein. Overall, these data demonstrate the utility of conditional spatial proximity labeling tools as a means to hone in on spatial and temporal differences between interacting factors in living cells.

Materials and Methods

General

HEK 293 Flp-In T-Rex cells (Invitrogen) were cultured and recombined according to manufacturer's instructions. Stable integrants of APEX-eIF4A were produced by co-transfection of these plasmids along with pOG44 by X-tremeGENE 9 (Roche) and selection using Hygromycin B. Transient transfections of splitAPEX leucine zipper components and splitAPEX eIF4A-eIF4H/eIF4B were generated using Nucleofection (Lonza 4D). TMT mass spec reagents (Thermo Fisher Scientific) were used for every quantitative mass spec experiment. Biotin-tyramide was purchased from Sigma-Aldrich.

APEX labeling

Plate cells in 15 cm dish and cultured for 3 days with 1 μg ml^{-1}. Cells were treated like in[1]. Briefly, Biotin-phenol containing (500 μM final) prewarmed DMEM media for 30 minutes prior to the start of the experiment. 1 mM final H_2O_2 was added to each dish for a total labeling time of 1 minute. Cells were gently agitated for 1 minute, and quenched with 2X quenching buffer (10 mM Trolox and 20 mM sodium ascorbate in DPBS). One 1X quenching solution wash was used to resuspend cells, which were then gently pelleted and lysed in lysis buffer containing 1X quenching buffer.

Biotinylated protein pulldown

Protein lysate were clarified by centrifugation for 10 min at 20,000g, 4 °C. Streptavidin beads (Pierce) were equilibrated with lysis buffer (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM MgCl2, and 1 mM DTT) containing 1% Triton X-100, 10 mM sodium ascorbate, 5 mM trolox ,and 0.1% SDS and EDTA-free complete protease inhibitor cocktail) for a total of two washes. Lysate was loaded at a volumetric ratio of 8:5 (streptavidin beads to lysate), and incubated at RT for 1 hr. Beads were washed twice with lysis buffer 1M KCl buffer, once with 2M urea, pH 8, and twice with lysis buffer (w/o detergent). Biotinylated proteins were eluted using 8M urea, pH 8 and boiling samples for 3 minutes at 98C.

Mass spec

100 μg of eluted proteins were brought up to 100 μl with 100mM TEAB. 5μl of 200mM TCEP and incubate sample at 55°C for 1 hour. Immediately before use, iodoacetamide with 100mM TEAB to make 375 mM iodoacetamide. 5 μl of the 375 mM iodoacetamide to the sample and incubate for 30 minutes protected from light at room temperature. Six volumes (or more) (~600μL) of pre-chilled (-20°C) acetone. Samples were allowed to precipitate overnight at

-20°C. Samples were pelleted dry and resuspended with 100 µl of 50mM TEAB. Immediately before use, 2.5 µl of trypsin (1 µg/µl) were used per 100 µg of protein overnight at 37°C. Peptides were quantified and normalized using a quantitative colorimetric peptide assay (Thermo Fisher Scientific). 41 µl of TMT labeling reagent (Thermo Fisher Scientific) were used for each 100 µl sample. Reactions were incubated for 1 hr at RT. 8 µl of 5% hydroxylamine was added to each sample and incubate for 15 minutes to quench the reaction. Samples were speed vacced dried or were frozen in -80C for storage. Samples were run on the UC Davis OrbiTrap mass spec using high pH fractionation, and analyzed using Proteome Discoverer. DESeq2 was used to perform quantitative ratiometric comparisons between splitAPEX-(eIF4H/B-eIF4A), splitAPEX-(leucine zipper), and full length APEX-GFP.

Western Blots

STREP-HRP (1:1000, Cell Signaling) was used for the activity blots. Anti-His (1:1000, Abcam), anti-flag (1:1000, Cell Signaling), and anti-GFP (1:10,000, Abcam) antibodies were used as primary antibodies. Chemiluminescence was induced by Pierce ECL Western Blotting Substrate (Thermo Scientific) and images were acquired by a FluorChem R imaging system (ProteinSimple).

Code availability

scripts to run the analyses mentioned above are available upon request.

Protein Purification

BL21 Star (DE3) Escherichia coli cells (Invitrogen) transformed with pHis-APEX[1-200], pHis-APEX[201-250]-Halo, pHis-APEX2 (WT), or pHis-APEX[1-200]-TEV-[201-250] in 1.5 L culture were cultivated to an absorbance at 600 nm, A600 nm, of 0.5 at 37 °C with 50 µg ml^{-1} kanamycin and then grown at 16 °C overnight with 1 mM IPTG. The cell pellets were resuspended in His buffer (20 mM HEPES-NaOH, pH 7.5, 500 mM NaCl, 10 mM imidazole, 10 mM β-mercaptoethanol) with 0.5% NP-40, sonicated, and centrifuged at 35,000g for 20 min. The supernatant was incubated with 1.5 ml bed volume of Ni-NTA Superflow (Qiagen) for 1 h. The beads were loaded on a gravity column and washed with His buffer containing 1 M NaCl. The proteins were eluted with 50 mM Na-phosphate buffer, pH 7.5, 500 mM NaCl, 100 mM Na2SO4, 250 mM imidazole, 10 mM β-mercaptoethanol. Samples were run through an FPLC

Ion exchange, mixed with 0.25 volumes of 80% glycerol, shock-frozen in liquid nitrogen, and stored at −80°C. All purification steps were performed at 4 °C.

***in vitro* activity assay**

Recombinant pHis-APEX[1-200]-TEV-[201-250] was treated overnight at 22°C with tobacco etch virus protease and enzymatic activity levels were compared to those of WT-APEX2 levels in a guaiacol assay as in[1]. *In vitro* APEX labeling reactions were catalyzed in the presence of absence of TEV Protease and samples were monitored for activity in a TECAN plate reader.

Chapter 4: Conclusion and future direction

Proximity labeling serves as a useful approach for understanding the spatial and temporal organization of non-membrane bound organelles as well as protein-RNA complexes. A powerful aspect of the APEX system is the ability to couple protein proximity labeling with RNA proximity labeling, for a more comprehensive understanding of macromolecular organization. Coupling APEX-Seq with APEX-MS now opens the doors to questions such as, how variable are other RNA granule complexes, how much do these complexes change over time, and are there any unifying themes in terms of their respective constituents? Moreover, rheology-based questions could also be asked such as: are there different macromolecules that contribute to the material state of an RNA granule?

A potential use of APEX-Seq is in highly polarized cells such as gut epithelium, and in neurons where RNA localization and organization is most dramatic. In polarized cells, RNA localization is deeply tied to translational control. Quite surprisingly, in gut epithelial cells Moor et al., looked at the relationship between localization of enterocyte mRNAs and the proteins they encode, respectively, and found that they mostly anticorrelated with each other. For instance, the mRNAs encoding ribosomal proteins were localized basally, ribosomal proteins themselves were apically enriched (Gáspár and Ephrussi 2017). Interestingly, The authors performed an experiment in which mice received a meal prior to determining RNA localization patterns, and found that many mRNAs changed their localization patterns through active motor protein transport (Moor et al. 2017; Gáspár and Ephrussi 2017). Some of these RNAs appear to be change localization patterns and were translated upon feeding (Gáspár and Ephrussi 2017), which raises the question as to whether RNA localization is regulated and controlled through mTOR signalling and macronutrient detection.

The use of a conditional split APEX should in principle offer significantly higher specificity in proximity labeling of both protein and RNA. This added specificity lends itself nicely toward the understanding of what RNAs use unique translational programs. An example is the use of a conditional splitAPEX on the eIF4A and two accessory proteins, eIF4B, and eIF4H. Interestingly, eIF4H and eIF4B bind to eIF4A in a mutually exclusive manner (Rozovsky, Butterworth, and Moore 2008), implying that some mRNAs are likely composed of unique combinations of translation initiation complexes. Moreover, the spatial landscape can now be

monitored based on conditional interactions between two interacting proteins, or proteins that come in close proximity through interactions in macromolecular complexes. An example of conditional protein interactions is the DCC receptors, which interact directly with the ribosome and other translation initiation components in neurons (Tcherkezian et al. 2010). Perhaps, using a splitAPEX, the synaptic spatial proteome and transcriptome might be captured through the conditional labeling of DCC receptors with the 80S ribosome. Moreover, splitAPEX might be leveraged at the synapse to enrich biotinylated ribosomes that are either directly contacting or near synapses through DCC receptors.

A potential use of APEX-Seq arises in trying to gain a deeper appreciation for subcellular structure. As an RNA matures and makes its way through different regions of the cell, the RNA is expected to undergo various structural rearrangements. Although the implications of subcellular RNA structure are not yet well understood, APEX-Seq coupled with techniques like DMS-MaPseq(Zubradt et al. 2017, 2016), may offer a deeper appreciation in this area. Moreover, coupling techniques such as TimeLapse-Seq with APEX-Seq might offer insights as to what the contribution of RNA age is toward RNA localization (Schofield et al. 2018). Questions regarding RNA age offer additional sight into spatial subcellular control, since RNA turnover is likely not uniformly the same across all cellular compartments and organelles.

Acknowledgements

To all of the incredible people who have supported me, and continue to support me. To Nicholas Ingolia, my wonderful mentor whom I am thankful, and grateful for having the privilege of working with. To my mother, Ana Amores, for her countless sacrifices, optimism and support. To my father, Calixto Padron, for giving my family and I opportunities we would have never had otherwise. To my amazing partner, Lacey Kitch, for her support, love, and affection.

References

Adivarahan, Srivathsan, Nathan Livingston, Beth Nicholson, Samir Rahman, Bin Wu, Olivia S. Rissland, and Daniel Zenklusen. 2018. “Spatial Organization of Single mRNPs at Different Stages of the Gene Expression Pathway.” *Molecular Cell* 72 (4): 727–38.e5.

Alberta, Bruce, Julian Lewis, Keith Roberta, Alexander Johnson, Martin Raff, and Peter Walter. 2008. *Molecular Biology of the Cell*.

Bassell, Gary J., and Robert H. Singer. 2001. “Neuronal RNA Localization and the Cytoskeleton.” *Results and Problems in Cell Differentiation*. https://doi.org/10.1007/978-3-540-40025-7_3.

Berleth, T., M. Burri, G. Thoma, D. Bopp, S. Richstein, G. Frigerio, M. Noll, and C. Nüsslein-Volhard. 1988. “The Role of Localization of Bicoid RNA in Organizing the Anterior Pattern of the Drosophila Embryo.” *The EMBO Journal* 7 (6): 1749–56.

Brangwynne, Clifford P., Christian R. Eckmann, David S. Courson, Agata Rybarska, Carsten Hoege, Jöbin Gharakhani, Frank Jülicher, and Anthony A. Hyman. 2009. “Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/condensation.” *Science* 324 (5935): 1729–32.

Branon, Tess C., Justin A. Bosch, Ariana D. Sanchez, Namrata D. Udeshi, Tanya Svinkina, Steven A. Carr, Jessica L. Feldman, Norbert Perrimon, and Alice Y. Ting. 2017. “Directed Evolution of TurboID for Efficient Proximity Labeling in Living Cells and Organisms.” https://doi.org/10.1101/196980.

Buchan, J. Ross, and Roy Parker. 2009. “Eukaryotic Stress Granules: The Ins and Outs of Translation.” *Molecular Cell* 36 (6): 932–41.

Cayley, S., B. A. Lewis, H. J. Guttman, and M. T. Record Jr. 1991. “Characterization of the Cytoplasm of Escherichia Coli K-12 as a Function of External Osmolarity. Implications for Protein-DNA Interactions in Vivo.” *Journal of Molecular Biology* 222 (2): 281–300.

Chen, Yu, Yang Zhang, Yuchuan Wang, Liguo Zhang, Eva K. Brinkman, Stephen A. Adam, Robert Goldman, Bas van Steensel, Jian Ma, and Andrew S. Belmont. 2018. “Mapping 3D Genome Organization Relative to Nuclear Compartments Using TSA-Seq as a Cytological Ruler.” *The Journal of Cell Biology*, August. https://doi.org/10.1083/jcb.201807108.

Choi-Rhee, Eunjoo, Howard Schulman, and John E. Cronan. 2004. “Promiscuous Protein Biotinylation by Escherichia Coli Biotin Protein Ligase.” *Protein Science: A Publication of the Protein Society* 13 (11): 3043–50.

Das, Sulagna, Robert H. Singer, and Young J. Yoon. 2019. “The Travels of mRNAs in Neurons: Do They Know Where They Are Going?” *Current Opinion in Neurobiology* 57 (February): 110–16.

De Benedetti, Arrigo, and Jeremy R. Graff. 2004. “eIF-4E Expression and Its Role in Malignancies and Metastases.” *Oncogene* 23 (18): 3189–99.

De Munter, Sofie, Janina Görnemann, Rita Derua, Bart Lesage, Junbin Qian, Ewald Heroes, Etienne Waelkens, Aleyde Van Eynde, Monique Beullens, and Mathieu Bollen. 2017. “Split-BioID: A Proximity Biotinylation Assay for Dimerization-Dependent Protein Interactions.” *FEBS Letters* 591 (2): 415–24.

Dine, Elliot, Agnieszka A. Gil, Giselle Uribe, Clifford P. Brangwynne, and Jared E. Toettcher. 2018. “Protein Phase Separation Provides Long-Term Memory of Transient Spatial Stimuli.” *Cell Systems* 6 (6): 655–63.e5.

Dixon, Andrew S., Marie K. Schwinn, Mary P. Hall, Kris Zimmerman, Paul Otto, Thomas H. Lubben, Braeden L. Butler, et al. 2016. “NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells.” *ACS Chemical Biology* 11 (2): 400–408.

Doyle, G. A., J. M. Bourdeau-Heller, S. Coulthard, L. F. Meisner, and J. Ross. 2000. “Amplification in Human Breast Cancer of a Gene Encoding a c-Myc mRNA-Binding Protein.” *Cancer Research* 60 (11): 2756–59.

Driever, Wolfgang, and Christiane Nüsslein-Volhard. 1988. “The Bicoid Protein Determines Position in the Drosophila Embryo in a Concentration-Dependent Manner.” *Cell*. https://doi.org/10.1016/0092-8674(88)90183-3.

Elden, Andrew C., Hyung-Jun Kim, Michael P. Hart, Alice S. Chen-Plotkin, Brian S. Johnson, Xiaodong Fang, Maria Armakola, et al. 2010. “Ataxin-2 Intermediate-Length Polyglutamine Expansions Are Associated with Increased Risk for ALS.” *Nature* 466 (7310): 1069–75.

Fei, Jingyi, and Cynthia M. Sharma. 2018. “RNA Localization in Bacteria.” *Microbiology Spectrum*. https://doi.org/10.1128/microbiolspec.rwr-0024-2018.

Ferrandon, D., I. Koch, E. Westhof, and C. Nüsslein-Volhard. 1997. “RNA-RNA Interaction Is Required for the Formation of Specific Bicoid mRNA 3’ UTR-STAUFEN Ribonucleoprotein Particles.” *The EMBO Journal* 16 (7): 1751–58.

Ferrandon, Dominique, Lisa Elphick, Christiane Nüsslein-Volhard, and Daniel St Johnston. 1994. “Staufen Protein Associates with the 3′UTR of Bicoid mRNA to Form Particles That Move in a Microtubule-Dependent Manner.” *Cell*.

https://doi.org/10.1016/0092-8674(94)90013-2.

Fischer, David S., Fabian J. Theis, and Nir Yosef. 2018. “Impulse Model-Based Differential Expression Analysis of Time Course Sequencing Data.” *Nucleic Acids Research*, August. https://doi.org/10.1093/nar/gky675.

Franzmann, Titus M., Marcus Jahnel, Andrei Pozniakovsky, Julia Mahamid, Alex S. Holehouse, Elisabeth Nüske, Doris Richter, et al. 2018. “Phase Separation of a Yeast Prion Protein Promotes Cellular Fitness.” *Science* 359 (6371). https://doi.org/10.1126/science.aao5654.

Gagnon, James A., and Kimberly L. Mowry. 2011. “Molecular Motors: Directing Traffic during RNA Localization.” *Critical Reviews in Biochemistry and Molecular Biology* 46 (3): 229–39.

Gáspár, Imre, and Anne Ephrussi. 2017. “RNA Localization Feeds Translation.” *Science* 357 (6357): 1235–36.

Georges, Amedee des, Vidya Dhote, Lauriane Kuhn, Christopher U. T. Hellen, Tatyana V. Pestova, Joachim Frank, and Yaser Hashem. 2015. “Structure of Mammalian eIF3 in the Context of the 43S Preinitiation Complex.” *Nature* 525 (7570): 491–95.

Ghosh, Indraneel, Andrew D. Hamilton, and Lynne Regan. 2000. “Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein.” *Journal of the American Chemical Society* 122 (23): 5658–59.

Graff, Jeremy R., Bruce W. Konicek, Thomas M. Vincent, Rebecca L. Lynch, David Monteith, Spring N. Weir, Phil Schwier, et al. 2007. “Therapeutic Suppression of Translation Initiation Factor eIF4E Expression Reduces Tumor Growth without Toxicity.” *The Journal of Clinical Investigation* 117 (9): 2638–48.

Han, Yisu, Tess Caroline Branon, Jeffrey D. Martell, Daniela Boassa, David Shechner, Mark H. Ellisman, and Alice Ting. 2019. “Directed Evolution of Split APEX2 Peroxidase.” *ACS Chemical Biology*, March. https://doi.org/10.1021/acschembio.8b00919.

Hashem, Yaser, Amedee des Georges, Vidya Dhote, Robert Langlois, Hstau Y. Liao, Robert A. Grassucci, Christopher U. T. Hellen, Tatyana V. Pestova, and Joachim Frank. 2013. “Structure of the Mammalian Ribosomal 43S Preinitiation Complex Bound to the Scanning Factor DHX29.” *Cell* 153 (5): 1108–19.

Hayat, M. A. 2007. *Microscopy, Immunohistochemistry, and Antigen Retrieval Methods: For Light and Electron Microscopy*. Springer Science & Business Media.

Henkin, Tina. 2004. “Faculty of 1000 Evaluation for RNA Dynamics in Live Escherichia Coli Cells.” *F1000 - Post-Publication Peer Review of the Biomedical Literature*. https://doi.org/10.3410/f.1021598.246622.

Hinnebusch, Alan G. 2014. “The Scanning Mechanism of Eukaryotic Translation Initiation.” *Annual Review of Biochemistry* 83 (1): 779–812.

Hsieh, Andrew C., Yi Liu, Merritt P. Edlind, Nicholas T. Ingolia, Matthew R. Janes, Annie Sher, Evan Y. Shi, et al. 2012. “The Translational Landscape of mTOR Signalling Steers Cancer Initiation and Metastasis.” *Nature* 485 (7396): 55–61.

Hung, Victoria, Namrata D. Udeshi, Stephanie S. Lam, Ken H. Loh, Kurt J. Cox, Kayvon Pedram, Steven A. Carr, and Alice Y. Ting. 2016. “Spatially Resolved Proteomic Mapping in Living Cells with the Engineered Peroxidase APEX2.” *Nature Protocols* 11 (3): 456–75.

Hung, Victoria, Peng Zou, Hyun-Woo Rhee, Namrata D. Udeshi, Valentin Cracan, Tanya Svinkina, Steven A. Carr, Vamsi K. Mootha, and Alice Y. Ting. 2014. “Proteomic Mapping of the Human Mitochondrial Intermembrane Space in Live Cells via Ratiometric APEX Tagging.” *Molecular Cell* 55 (2): 332–41.

Hüttelmaier, Stefan, Daniel Zenklusen, Marcell Lederer, Jason Dictenberg, Mike Lorenz, Xiuhua Meng, Gary J. Bassell, John Condeelis, and Robert H. Singer. 2005. “Spatial Regulation of Beta-Actin Translation by Src-Dependent Phosphorylation of ZBP1.” *Nature* 438 (7067): 512–15.

Hyman, Anthony A., and Kai Simons. 2012. “Cell Biology. Beyond Oil and Water--Phase Transitions in Cells.” *Science* 337 (6098): 1047–49.

Iwasaki, Shintaro, Stephen N. Floor, and Nicholas T. Ingolia. 2016. “Rocaglates Convert DEAD-Box Protein eIF4A into a Sequence-Selective Translational Repressor.” *Nature* 534 (7608): 558–61.

Jain, Saumya, Joshua R. Wheeler, Robert W. Walters, Anurag Agrawal, Anthony Barsic, and Roy Parker. 2016. “ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure.” *Cell* 164 (3): 487–98.

Johannes, G., M. S. Carter, M. B. Eisen, P. O. Brown, and P. Sarnow. 1999. “Identification of Eukaryotic mRNAs That Are Translated at Reduced Cap Binding Complex eIF4F Concentrations Using a cDNA Microarray.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (23): 13118–23.

Kaewsapsak, Pornchai, David Michael Shechner, William Mallard, John L. Rinn, and Alice Y. Ting. 2017. “Live-Cell Mapping of Organelle-Associated RNAs via Proximity Biotinylation Combined with Protein-RNA Crosslinking.” *eLife* 6 (December). https://doi.org/10.7554/eLife.29224.

Katoh, Hiroshi, Toru Okamoto, Takasuke Fukuhara, Hiroto Kambara, Eiji Morita, Yoshio Mori,

Wataru Kamitani, and Yoshiharu Matsuura. 2013. “Japanese Encephalitis Virus Core Protein Inhibits Stress Granule Formation through an Interaction with Caprin-1 and Facilitates Viral Propagation.” *Journal of Virology* 87 (1): 489–502.

Kedersha, Nancy, Marc D. Panas, Christopher A. Achorn, Shawn Lyons, Sarah Tisdale, Tyler Hickman, Marshall Thomas, et al. 2016. “G3BP–Caprin1–USP10 Complexes Mediate Stress Granule Condensation and Associate with 40S Subunits.” *The Journal of Cell Biology* 212 (7): 845–60.

Khong, Anthony, Tyler Matheny, Saumya Jain, Sarah F. Mitchell, Joshua R. Wheeler, and Roy Parker. 2017. “The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules.” *Molecular Cell* 68 (4): 808–20.e5.

Kim, Dae In, Samuel C. Jensen, Kyle A. Noble, Birendra Kc, Kenneth H. Roux, Khatereh Motamedchaboki, and Kyle J. Roux. 2016. “An Improved Smaller Biotin Ligase for BioID Proximity Labeling.” *Molecular Biology of the Cell* 27 (8): 1188–96.

Kislauskis, Edward H., Xiao-Chun Zhu, and Robert H. Singer. 1997. “β-Actin Messenger RNA Localization and Protein Synthesis Augment Cell Motility.” *The Journal of Cell Biology*. https://doi.org/10.1083/jcb.136.6.1263.

Kozak, M. 1991. “An Analysis of Vertebrate mRNA Sequences: Intimations of Translational Control.” *The Journal of Cell Biology* 115 (4): 887–903.

Lad, Latesh, Martin Mewies, and Emma Lloyd Raven. 2002. “Substrate Binding and Catalytic Mechanism in Ascorbate Peroxidase: Evidence for Two Ascorbate Binding Sites.” *Biochemistry* 41 (46): 13774–81.

Lam, Stephanie S., Jeffrey D. Martell, Kimberli J. Kamer, Thomas J. Deerinck, Mark H. Ellisman, Vamsi K. Mootha, and Alice Y. Ting. 2015. “Directed Evolution of APEX2 for Electron Microscopy and Proximity Labeling.” *Nature Methods* 12 (1): 51–54.

Lee, Song-Yi, Myeong-Gyun Kang, Jong-Seok Park, Geunsik Lee, Alice Y. Ting, and Hyun-Woo Rhee. 2016. “APEX Fingerprinting Reveals the Subcellular Localization of Proteins of Interest.” *Cell Reports* 15 (8): 1837–47.

Lin, R. Y., J. C. Vera, R. S. Chaganti, and D. W. Golde. 1998. “Human Monocarboxylate Transporter 2 (MCT2) Is a High Affinity Pyruvate Transporter.” *The Journal of Biological Chemistry* 273 (44): 28959–65.

Lipshitz, Howard D., and Craig A. Smibert. 2000. “Mechanisms of RNA Localization and Translational Regulation.” *Current Opinion in Genetics & Development*. https://doi.org/10.1016/s0959-437x(00)00116-7.

Li, Yun R., Oliver D. King, James Shorter, and Aaron D. Gitler. 2013. “Stress Granules as Crucibles of ALS Pathogenesis.” *The Journal of Cell Biology* 201 (3): 361–72.

Llácer, Jose L., Tanweer Hussain, Laura Marler, Colin Echeverría Aitken, Anil Thakur, Jon R. Lorsch, Alan G. Hinnebusch, and V. Ramakrishnan. 2015. “Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex.” *Molecular Cell* 59 (3): 399–412.

MacDonald, P. M., A. Leask, and K. Kerr. 1995. “Exl Protein Specifically Binds BLE1, a Bicoid mRNA Localization Element, and Is Required for One Phase of Its Activity.” *Proceedings of the National Academy of Sciences of the United States of America* 92 (23): 10787–91.

Malka-Mahieu, Hélène, Michelle Newman, Laurent Désaubry, Caroline Robert, and Stéphan Vagner. 2016. “Molecular Pathways: The eIF4F Translation Initiation Complex—New Opportunities for Cancer Treatment.” *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 23 (1): 21–25.

Marintchev, Assen, Katherine A. Edmonds, Boriana Marintcheva, Elthea Hendrickson, Monika Oberer, Chikako Suzuki, Barbara Herdy, Nahum Sonenberg, and Gerhard Wagner. 2009. “Topology and Regulation of the Human eIF4A/4G/4H Helicase Complex in Translation Initiation.” *Cell* 136 (3): 447–60.

Markmiller, Sebastian, Sahar Soltanieh, Kari L. Server, Raymond Mak, Wenhao Jin, Mark Y. Fang, En-Ching Luo, et al. 2018. “Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules.” *Cell* 172 (3): 590–604.e13.

Martell, Jeffrey D., Thomas J. Deerinck, Stephanie S. Lam, Mark H. Ellisman, and Alice Y. Ting. 2017. “Electron Microscopy Using the Genetically Encoded APEX2 Tag in Cultured Mammalian Cells.” *Nature Protocols* 12 (9): 1792–1816.

Martell, Jeffrey D., Masahito Yamagata, Thomas J. Deerinck, Sébastien Phan, Carolyn G. Kwa, Mark H. Ellisman, Joshua R. Sanes, and Alice Y. Ting. 2016. “A Split Horseradish Peroxidase for the Detection of Intercellular Protein-Protein Interactions and Sensitive Visualization of Synapses.” *Nature Biotechnology* 34 (7): 774–80.

Marzluff, William F. 2005. “Metazoan Replication-Dependent Histone mRNAs: A Distinct Set of RNA Polymerase II Transcripts.” *Current Opinion in Cell Biology* 17 (3): 274–80.

Marzluff, William F., Eric J. Wagner, and Robert J. Duronio. 2008. “Metabolism and Regulation of Canonical Histone mRNAs: Life without a poly(A) Tail.” *Nature Reviews. Genetics* 9 (11): 843–54.

Mateju, Daniel, Titus M. Franzmann, Avinash Patel, Andrii Kopach, Edgar E. Boczek,

Shovamayee Maharana, Hyun O. Lee, Serena Carra, Anthony A. Hyman, and Simon Alberti. 2017. “An Aberrant Phase Transition of Stress Granules Triggered by Misfolded Protein and Prevented by Chaperone Function.” *The EMBO Journal* 36 (12): 1669–87.

Mazroui, Rachid, Rami Sukarieh, Marie-Eve Bordeleau, Randal J. Kaufman, Peter Northcote, Junichi Tanaka, Imed Gallouzi, and Jerry Pelletier. 2006. “Inhibition of Ribosome Recruitment Induces Stress Granule Formation Independently of Eukaryotic Initiation Factor 2alpha Phosphorylation.” *Molecular Biology of the Cell* 17 (10): 4212–19.

McMahon, Aoife C., Reazur Rahman, Hua Jin, James L. Shen, Allegra Fieldsend, Weifei Luo, and Michael Rosbash. 2016. “TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins.” *Cell* 165 (3): 742–53.

Medioni, Caroline, Kimberly Mowry, and Florence Besse. 2012. “Principles and Roles of mRNA Localization in Animal Development.” *Development* 139 (18): 3263–76.

Milo, Ron, and Rob Phillips. 2015. *Cell Biology by the Numbers*. Garland Science.

Molliex, Amandine, Jamshid Temirov, Jihun Lee, Maura Coughlin, Anderson P. Kanagaraj, Hong Joo Kim, Tanja Mittag, and J. Paul Taylor. 2015. “Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization.” *Cell* 163 (1): 123–33.

Montero Llopis, Paula, Audrey F. Jackson, Oleksii Sliusarenko, Ivan Surovtsev, Jennifer Heinritz, Thierry Emonet, and Christine Jacobs-Wagner. 2010. “Spatial Organization of the Flow of Genetic Information in Bacteria.” *Nature* 466 (7302): 77–81.

Moor, Andreas E., Matan Golan, Efi E. Massasa, Doron Lemze, Tomer Weizman, Rom Shenhav, Shaked Baydatch, et al. 2017. “Global mRNA Polarization Regulates Translation Efficiency in the Intestinal Epithelium.” *Science* 357 (6357): 1299–1303.

Morisaki, Tatsuya, Kenneth Lyon, Keith F. DeLuca, Jennifer G. DeLuca, Brian P. English, Zhengjian Zhang, Luke D. Lavis, et al. 2016. “Real-Time Quantification of Single RNA Translation Dynamics in Living Cells.” *Science* 352 (6292): 1425–29.

Myers, Samuel A., Jason Wright, Ryan Peckner, Brian T. Kalish, Feng Zhang, and Steven A. Carr. 2018. “Discovery of Proteins Associated with a Predefined Genomic Locus via dCas9–APEX-Mediated Proximity Labeling.” *Nature Methods* 15 (6): 437–39.

Nevo-Dinur, K., A. Nussbaum-Shochat, S. Ben-Yehuda, and O. Amster-Choder. 2011. “Translation-Independent Localization of mRNA in E. Coli.” *Science*. https://doi.org/10.1126/science.1195691.

Nicastro, Giuseppe, Adela M. Candel, Michael Uhl, Alain Oregioni, David Hollingworth, Rolf

Backofen, Stephen R. Martin, and Andres Ramos. 2017. “Mechanism of β-Actin mRNA Recognition by ZBP1.” *Cell Reports* 18 (5): 1187–99.

Niedzwiecka, Anna, Joseph Marcotrigiano, Janusz Stepinski, Marzena Jankowska-Anyszka, Aleksandra Wyslouch-Cieszynska, Michal Dadlez, Anne-Claude Gingras, et al. 2002. “Biophysical Studies of eIF4E Cap-Binding Protein: Recognition of mRNA 5′ Cap Structure and Synthetic Fragments of eIF4G and 4E-BP1 Proteins.” *Journal of Molecular Biology* 319 (3): 615–35.

Oh, Sekyung, Ryan A. Flynn, Stephen N. Floor, James Purzner, Lance Martin, Brian T. Do, Simone Schubert, et al. 2016. “Medulloblastoma-Associated DDX3 Variant Selectively Alters the Translational Response to Stress.” *Oncotarget* 7 (19): 28169–82.

Oostdyk, Luke T., Leonard Shank, Kasey Jividen, Natalia Dworak, Nicholas E. Sherman, and Bryce M. Paschal. 2019. “Towards Improving Proximity Labeling by the Biotin Ligase BirA.” *Methods* 157 (March): 66–79.

Özeş, Ali R., Kateryna Feoktistova, Brian C. Avanzino, and Christopher S. Fraser. 2011. “Duplex Unwinding and ATPase Activities of the DEAD-Box Helicase eIF4A Are Coupled by eIF4G and eIF4B.” *Journal of Molecular Biology* 412 (4): 674–87.

Panas, Marc D., Pavel Ivanov, and Paul Anderson. 2016. “Mechanistic Insights into Mammalian Stress Granule Dynamics.” *The Journal of Cell Biology* 215 (3): 313–23.

Parakh, Sonam, and Julie D. Atkin. 2016. “Protein Folding Alterations in Amyotrophic Lateral Sclerosis.” *Brain Research* 1648 (Pt B): 633–49.

Patel, Avinash, Liliana Malinovska, Shambaditya Saha, Jie Wang, Simon Alberti, Yamuna Krishnan, and Anthony A. Hyman. 2017. “ATP as a Biological Hydrotrope.” *Science* 356 (6339): 753–56.

Pelletier, Jerry, Jeremy Graff, Davide Ruggero, and Nahum Sonenberg. 2015. “Targeting the eIF4F Translation Initiation Complex: A Critical Nexus for Cancer Development.” *Cancer Research* 75 (2): 250–63.

Pilhofer, Martin, Marko Pavlekovic, Natuschka M. Lee, Wolfgang Ludwig, and Karl-Heinz Schleifer. 2009. “Fluorescence in Situ Hybridization for Intracellular Localization of nifH mRNA.” *Systematic and Applied Microbiology* 32 (3): 186–92.

Pisareva, Vera P., Andrey V. Pisarev, Anton A. Komar, Christopher U. T. Hellen, and Tatyana V. Pestova. 2008. “Translation Initiation on Mammalian mRNAs with Structured 5’UTRs Requires DExH-Box Protein DHX29.” *Cell* 135 (7): 1237–50.

Pokrywka, N. J., and E. C. Stephenson. 1991. “Microtubules Mediate the Localization of Bicoid

RNA during Drosophila Oogenesis." *Development* 113 (1): 55–66.

Protter, David S. W., and Roy Parker. 2016. "Principles and Properties of Stress Granules." *Trends in Cell Biology* 26 (9): 668–79.

Ramaswami, Mani, J. Paul Taylor, and Roy Parker. 2013. "Altered Ribostasis: RNA-Protein Granules in Degenerative Disorders." *Cell* 154 (4): 727–36.

Rauniyar, Navin, and John R. Yates 3rd. 2014. "Isobaric Labeling-Based Relative Quantification in Shotgun Proteomics." *Journal of Proteome Research* 13 (12): 5293–5309.

Rees, Johanna S., Xue-Wen Li, Sarah Perrett, Kathryn S. Lilley, and Antony P. Jackson. 2015. "Protein Neighbors and Proximity Proteomics." *Molecular & Cellular Proteomics: MCP* 14 (11): 2848–56.

Rhee, Hyun-Woo, Peng Zou, Namrata D. Udeshi, Jeffrey D. Martell, Vamsi K. Mootha, Steven A. Carr, and Alice Y. Ting. 2013. "Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging." *Science* 339 (6125): 1328–31.

Roux, Kyle J., Dae In Kim, Manfred Raida, and Brian Burke. 2012. "A Promiscuous Biotin Ligase Fusion Protein Identifies Proximal and Interacting Proteins in Mammalian Cells." *The Journal of Cell Biology* 196 (6): 801–10.

Rozovsky, Nadja, Aimee C. Butterworth, and Melissa J. Moore. 2008. "Interactions between eIF4AI and Its Accessory Factors eIF4B and eIF4H." *RNA* 14 (10): 2136–48.

Schnorrer, F., K. Bohmann, and C. Nüsslein-Volhard. 2000. "The Molecular Motor Dynein Is Involved in Targeting Swallow and Bicoid RNA to the Anterior Pole of Drosophila Oocytes." *Nature Cell Biology* 2 (4): 185–90.

Schofield, Jeremy A., Erin E. Duffy, Lea Kiefer, Meaghan C. Sullivan, and Matthew D. Simon. 2018. "TimeLapse-Seq: Adding a Temporal Dimension to RNA Sequencing through Nucleoside Recoding." *Nature Methods* 15 (3): 221–25.

Schopp, Isabel M., and Julien Béthune. 2018. "Split-BioID - Proteomic Analysis of Context-Specific Protein Complexes in Their Native Cellular Environment." *Journal of Visualized Experiments: JoVE*, no. 134 (April). https://doi.org/10.3791/57479.

Schopp, Isabel Myriam, Cinthia Claudia Amaya Ramirez, Jerneja Debeljak, Elisa Kreibich, Merle Skribbe, Klemens Wild, and Julien Béthune. 2017. "Split-BioID a Conditional Proteomics Approach to Monitor the Composition of Spatiotemporally Defined Protein Complexes." *Nature Communications* 8 (June): 15690.

Shin, Yongdae, Joel Berry, Nicole Pannucci, Mikko P. Haataja, Jared E. Toettcher, and Clifford P. Brangwynne. 2017. "Spatiotemporal Control of Intracellular Phase Transitions Using

Light-Activated optoDroplets." *Cell*. https://doi.org/10.1016/j.cell.2016.11.054.

Song, Tingting, Yi Zheng, Yarong Wang, Zachary Katz, Xin Liu, Shaoying Chen, Robert H. Singer, and Wei Gu. 2015. "Specific Interaction of KIF11 with ZBP1 Regulates the Transport of β-Actin mRNA and Cell Motility." *Journal of Cell Science* 128 (5): 1001–10.

Steensel, B. van, and S. Henikoff. 2000. "Identification of in Vivo DNA Targets of Chromatin Proteins Using Tethered Dam Methyltransferase." *Nature Biotechnology* 18 (4): 424–28.

Suter, Beat. 2018. "RNA Localization and Transport." *Biochimica et Biophysica Acta, Gene Regulatory Mechanisms* 1861 (10): 938–51.

Tcherkezian, Joseph, Perry A. Brittis, Franziska Thomas, Philippe P. Roux, and John G. Flanagan. 2010. "Transmembrane Receptor DCC Associates with Protein Synthesis Machinery and Regulates Translation." *Cell* 141 (4): 632–44.

Theurkauf, W. E., and T. I. Hazelrigg. 1998. "In Vivo Analyses of Cytoplasmic Transport and Cytoskeletal Organization during Drosophila Oogenesis: Characterization of a Multi-Step Anterior Localization Pathway." *Development* 125 (18): 3655–66.

Thoreen, Carson C., Lynne Chantranupong, Heather R. Keys, Tim Wang, Nathanael S. Gray, and David M. Sabatini. 2012. "A Unifying Model for mTORC1-Mediated Regulation of mRNA Translation." *Nature* 485 (7396): 109–13.

Toombs, James A., Michelina Petri, Kacy R. Paul, Grace Y. Kan, Asa Ben-Hur, and Eric D. Ross. 2012. "De Novo Design of Synthetic Prion Domains." *Proceedings of the National Academy of Sciences of the United States of America* 109 (17): 6519–24.

Vainberg, I. E., S. A. Lewis, H. Rommelaere, C. Ampe, J. Vandekerckhove, H. L. Klein, and N. J. Cowan. 1998. "Prefoldin, a Chaperone That Delivers Unfolded Proteins to Cytosolic Chaperonin." *Cell* 93 (5): 863–73.

Varnaitė, Renata, and Stuart A. MacNeill. 2016. "Meet the Neighbors: Mapping Local Protein Interactomes by Proximity-Dependent Labeling with BioID." *Proteomics* 16 (19): 2503–18.

Vogel, Maartje J., Lars Guelen, Elzo de Wit, Daniel Peric-Hupkes, Martin Lodén, Wendy Talhout, Marike Feenstra, Ben Abbas, Anne-Kathrin Classen, and Bas van Steensel. 2006. "Human Heterochromatin Proteins Form Large Domains Containing KRAB-ZNF Genes." *Genome Research* 16 (12): 1493–1504.

Wang, Chong, Boran Han, Ruobo Zhou, and Xiaowei Zhuang. 2016. "Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells." *Cell* 165 (4): 990–1001.

Wang, Jennifer T., and Geraldine Seydoux. 2014. "P Granules." *Current Biology: CB* 24 (14): R637–38.

Wehrman, Tom S., Clayton L. Casipit, Nevin M. Gewertz, and Helen M. Blau. 2005. “Enzymatic Detection of Protein Translocation.” *Nature Methods* 2 (7): 521–27.

Weil, Timothy T. 2014. “mRNA Localization in the Drosophila Germline.” *RNA Biology* 11 (8): 1010–18.

Wheeler, Joshua R., Saumya Jain, Anthony Khong, and Roy Parker. 2017. “Isolation of Yeast and Mammalian Stress Granule Cores.” *Methods* 126 (August): 12–17.

Wheeler, Joshua R., Tyler Matheny, Saumya Jain, Robert Abrisch, and Roy Parker. 2016. “Distinct Stages in Stress Granule Assembly and Disassembly.” *eLife* 5 (September). https://doi.org/10.7554/eLife.18413.

Wolke, Uta, Gilbert Weidinger, Marion Köprunner, and Erez Raz. 2002. “Multiple Levels of Posttranscriptional Control Lead to Germ Line-Specific Gene Expression in the Zebrafish.” *Current Biology*. https://doi.org/10.1016/s0960-9822(02)00679-6.

Wu, Bin, Adina R. Buxbaum, Zachary B. Katz, Young J. Yoon, and Robert H. Singer. 2015. “Quantifying Protein-mRNA Interactions in Single Live Cells.” *Cell* 162 (1): 211–20.

Wu, Bin, Carolina Eliscovich, Young J. Yoon, and Robert H. Singer. 2016. “Translation Dynamics of Single mRNAs in Live Cells and Neurons.” *Science* 352 (6292): 1430–35.

Yan, Xiaowei, Tim A. Hoek, Ronald D. Vale, and Marvin E. Tanenbaum. 2016. “Dynamics of Translation of Single mRNA Molecules In Vivo.” *Cell* 165 (4): 976–89.

Youn, Ji-Young, Wade H. Dunham, Seo Jung Hong, James D. R. Knight, Mikhail Bashkurov, Ginny I. Chen, Halil Bagci, et al. 2018. “High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies.” *Molecular Cell* 69 (3): 517–32.e11.

Zubradt, Meghan, Paromita Gupta, Sitara Persad, Alan M. Lambowitz, Jonathan S. Weissman, and Silvi Rouskin. 2017. “DMS-MaPseq for Genome-Wide or Targeted RNA Structure Probing in Vivo.” *Nature Methods* 14 (1): 75–82.

Zubradt, Meghan, Meghan Zubradt, Paromita Gupta, Sitara Persad, Alan M. Lambowitz, Jonathan S. Weissman, and Silvi Rouskin. 2016. “Target-Specific DMS-MaPseq for in Vivo RNA Structure Determination.” *Protocol Exchange*. https://doi.org/10.1038/protex.2016.069.