Neuron type-specific mRNA translation programs provide a gateway for memory consolidation.

Mauricio M. Oliveira[1], Olivia Mosto[1#], Robert Carney[1#], Wendy Liu[1#], Emmanuel Makinde[1], Emily H. Lu[1], Karen S.A. Ruiz[2], Thomas Carew[1], Eric Klann[1$]

[1]Center for Neural Science, New York University, New York City, NY, US ZIP 10003

[2]Department of Molecular Biology, Massachusetts General Hospital, Harvard University, Boston, MA, US ZIP 02114

[#]Co-second authors

[$] Corresponding author: ek65@nyu.edu

Keywords: Protein synthesis, mRNA translation, neuron, memory consolidation, long-term memory

Summary

Long-term memory consolidation is a dynamic process that requires a heterogeneous ensemble of neurons, each with a highly specialized molecular environment. Considerable effort has been placed into understanding how the molecular mechanisms in specific neuron types are modified in memory, but these studies are often undertaken hours or days after training, when memory is already consolidated. Studies have shown that protein synthesis is elevated during the early stages of consolidation, but there is limited information as to how it impacts neuronal function. We hypothesize that mRNAs being translated during the early stages of consolidation could provide clues as to how diverse neurons involved in memory formation restructure their molecular architecture to support memory formation. Here, we generate a landscape of the translatome of three neuron types in the dorsal hippocampus during the first hour of contextual memory consolidation. Our results show that these neurons share a common backbone of readily translated mRNAs. However, excitatory neurons undergo deep reconfiguration of proteostatic control, whereas interneurons modify their synaptic transmission. We demonstrate that the translational control in these neurons requires GADD34, which promotes translation initiation. Finally, we show that the differential expression of mRNAs by neurons during early consolidation can be explained by features hard coded in the mRNA, suggesting ubiquitous mechanisms controlling activity-dependent neuronal translation. Altogether, our work uncovers previously unknown checkpoints controlling activity-dependent translation in neurons and provides a large, readily available resource for further investigations of memory formation in health and disease.

Introduction

The induction of a new memory involves the synchronous engagement of a heterogeneous population of neurons, promoting molecular changes that strengthen the connectivity of a neuronal ensemble in the brain[1]. These neurons can be either excitatory or inhibitory, depending on their specialized molecular signatures and electrophysiological properties. [2-4]. The molecular environment of a neuron enables it to quickly respond to convergent afferent stimuli from the ensemble and average them into a final output. Thus, the molecular framework of a neuron is likely the main driver of dynamic responses to stimuli by each cell type. However, we still lack a full understanding of rapid neuronal dynamics at the molecular level *in vivo*. Many studies have examined these molecular changes at distant time points (e.g. days to weeks), after the memory has been consolidated, but have not focused on how these processes start and develop. Furthermore, these studies generally rely on measurements of the transcriptome at the single-cell level, which is unlikely to capture true functional differences, given that transcript levels do not automatically correspond to protein abundance[5,6], and single-cell sequencing generates sparse data, with only a subset of enriched transcripts being incorporated in the final analysis. Conversely, studying the dynamic changes in the proteome is challenging, due to low technical sensitivity, especially when studying sparse neuronal subpopulations. The translatome, the intersection between transcriptome and proteome, provides a suitable alternative to dissect functional changes occurring in neurons during early stages of consolidation.

Long-term memory consolidation follows encoding and is characterized by a malleable memory, due to synaptic lability and enhanced protein synthesis, a transient feature which is critical for consolidation[7,8]. The full picture of mRNAs translated at such early stages of memory consolidation remain elusive, but translation disruption is capable of modifying memory formation[9,10], reinforcing it as a therapeutic target for memory-related disorders, such as post-traumatic stress disorder and Alzheimer's disease. Thus, we leveraged translating ribosome affinity purification (TRAP)[11] to examine threat conditioning-induced alterations in the translatome of three different neuron types of the dHPC: excitatory *Camk2a*$^+$ neurons, and inhibitory *Pvalb*$^+$ and *Sst*$^+$ neurons. We find that these neurons display different molecular dynamics and boost the translation of selective categories of mRNAs. Our results suggest the necessity of early translation of proteostatic control factors, particularly in excitatory neurons. In addition, we find that the translatome plasticity requires the translation initiation modulator GADD34, providing evidence for an upstream regulator of threat-induced modifications to the translatome. Finally, we find that rapid modifications of the translatome during consolidation are a function of cis elements encoded in the mRNA sequence. Altogether, our findings depict both common and specific changes of neuronal

responses during threat memory consolidation and pave the way to integrate molecular frameworks into the biology of neuronal ensemble formation and systems consolidation.

Results

The cell type-specific translatome reflects biological peculiarities linked to each neuronal type. We used the contextual threat conditioning paradigm, which is reliant on the dorsal hippocampus (dHPC), as a model to examine rapid molecular dynamics neurons in memory consolidation. To determine conditioning-dependent rapid modifications in the translatome of hippocampal neurons, we generated three double transgenic mouse lines, by crossing *Gt(ROSA)26Sor-eGFP::L10a* mouse line with the Cre recombinase-expressing driver lines *Pvalb.Cre, Camk2a.Cre* or *Sst.Cre*. The resulting mice had GFP-tagged ribosomes in a cell type-specific manner, enabling purification using anti-GFP-coated beads, a technique termed translating ribosome affinity purification (TRAP)[11]. To restrict our screening to the early protein synthesis-dependent stage of memory consolidation, mice underwent the threat conditioning training paradigm, and we collected dHPC samples either 15 or 60 minutes later (Fig 1a). Home cage mice ("HC", i.e. never introduced to the conditioning paradigm) and mice that went inside the box but were not shocked ("Box only"), were used as controls. The latter was demonstrated to not induce aversive memory formation to the context (Fig 1b).

We confirmed TRAP efficacy at isolating cell type-specific translatomes. Immunostaining confirmed the cell type-specific expression of eGFP::L10a (Fig 1c). To confirm the purity of our immunoprecipitated fraction, we compared it against the total fraction (i.e. all mRNAs in the dHPC). We found selective enrichment for excitatory markers in *Camk2a*$^+$ neurons, whereas inhibitory markers were enriched in *Pvalb*$^+$ and *Sst*$^+$ neurons. Glial markers were depleted in all purified fractions (Fig 1d). We further confirmed that our datasets reproduced similar neuron-specific translatome datasets found in the literature[12-14] (Ext. Fig 1a). Principal component analysis (PCA) revealed neuron type as the main source of variance (73% variance explained), with *Sst*$^+$ and *Pvalb*$^+$ translatomes being more similar to each other than to *Camk2a*$^+$ (Fig 1e). Indeed, only 30% of the translatome was shared among all three neuron types, and the intersection between the inhibitory neurons was ~2.5-fold bigger than between inhibitory and excitatory neurons (Fig 1f). Furthermore, fold changes of *Sst*$^+$ and *Pvalb*$^+$ translatome, when compared to *Camk2a*$^+$, have a correlation coefficient of 0.72 (Fig 1g; Suppl table 1).

We next used gene set enrichment analysis (GSEA) to identify cellular features that were reflected in the translatome (Fig 1h; Suppl table 2). Interestingly, we found that *Camk2a*$^+$ and *Sst*$^+$ neurons have

more mRNAs encoding translation modulators and ribosomal proteins, when compared to *Pvalb*$^{+}$ neurons, suggesting a differential modulation of translation in different neuron types. This agrees with earlier findings that enhancing *de novo* translation in these cells, but not *Pvalb*$^{+}$, enhances long-term memory retention[15]. We further found that *Pvalb*$^{+}$ neurons present enrichment of mRNAs encoding immune modulators and blood vessel biology, indicating that they may promote neurovascular coupling, a crucial biological process that enhances blood flow and tissue oxygenation in brain regions with high activity[16,17]. Overall, these data highlight the fidelity of our translatome datasets to biological peculiarities previously linked to each neuronal type.

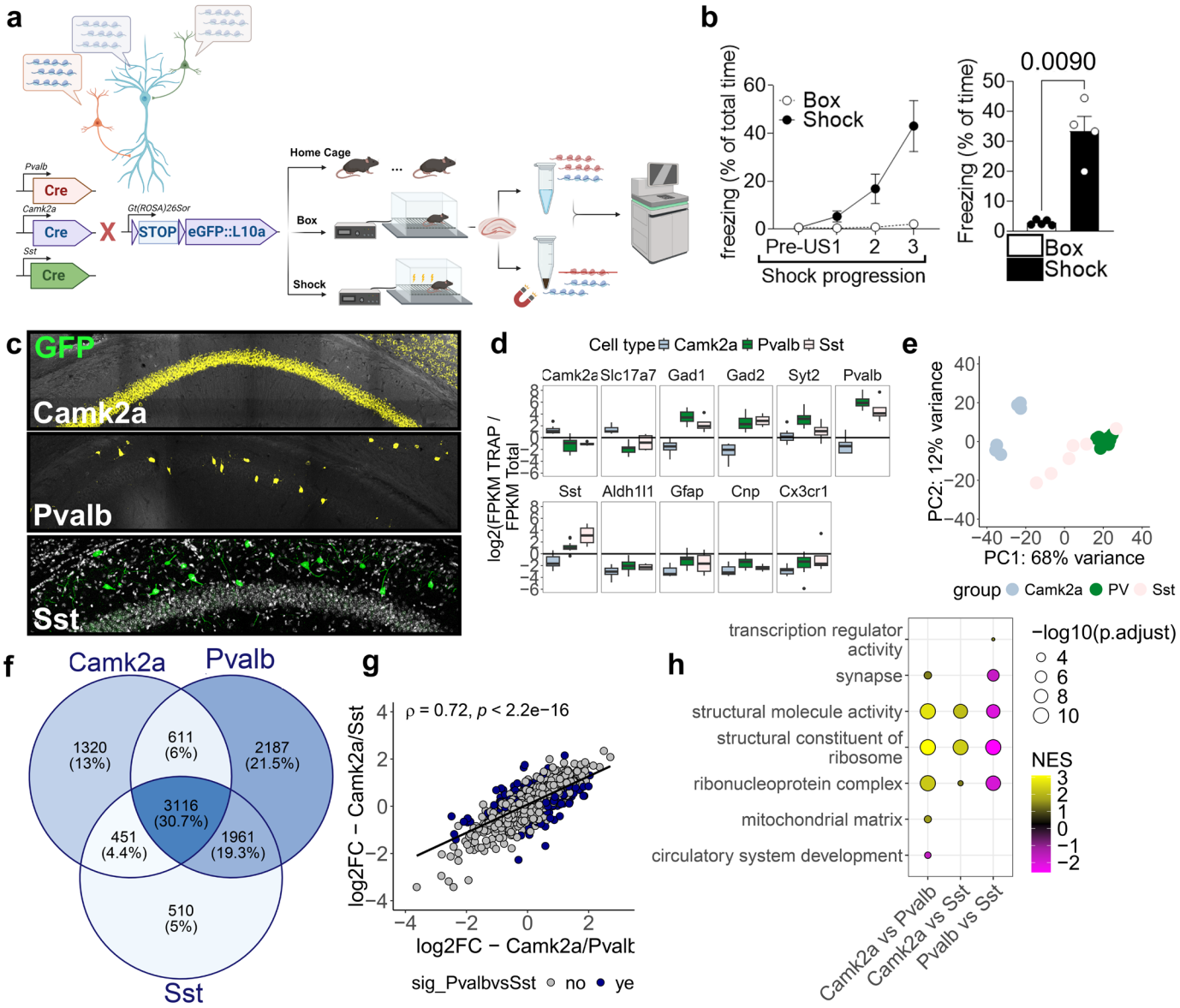

Figure 1. The cell type-specific translatome reflects biological function linked to each neuronal type. (**a**) Schematics of the experimental approach used to study cell type-specific dynamic changes in the translatome induced by threat conditioning. (**b**) Contextual threat conditioning experiment. White = control mice placed in the experimental box, but not shocked; Black = mice that were shocked in the experimental box. Left panel = Training session; Right panel = Testing session, 24h after training session. Unpaired *t* test with Welch's correction. $p = 0.0090$. $N = 4$-5 mice/group; Mean ± SEM; $F = 124.8$ (**c**) Immunofluorescence confirmation of cell type-specific expression of RPL10a-EGFP. Green = GFP; Blue = DAPI; (**d**) Enrichment of cell type-specific markers in the purified translatome of each cell type. *Camk2a* and *Slc17a7* = excitatory cell markers; *Gad1, Pvalb, Sst* = Inhibitory cell markers; *Gfap, Cnp, Cx3cr1* = Glial markers. The upper and lower whiskers represent the largest and smallest value within 1.5 times above or below the interquartile range, respectively. (**e**) PCA of baseline translatomes. Light-blue = *Camk2a*$^+$; dark green = *Pvalb*$^+$; Light-green = *Sst*$^+$. $N = 6$-9 mice/group. (**f**) Venn diagram of the mRNAs identified in each cell type-specific purification. Color intensity illustrates the size of the set. (**g**) Spearman correlation of Log_2FC of *Camk2a*$^+$ resting translatome versus *Pvalb*$^+$ (x axis) or *Sst*$^+$ (y axis). Spearman Rho = 0.8. (**h**) Gene set enrichment analysis of resting cell type-specific translatomes. Yellow means enriched in the first cell type in the comparison, magenta means enriched in the second cell type in the comparison. NES = normalized enrichment score.

Threat conditioning leads to rapid modifications to the translatome of neurons. We next examined conditioning-dependent rapid modifications to neuronal translatome. Differential expression analysis comparing the shock groups to their respective HC groups showed a similar magnitude of change to the translatome of *Camk2a*$^+$, *Pvalb*$^+$ and *Sst*$^+$ neurons at 15 min, with each neuron type ranging from ~2.5-9.5% differentially expressed genes (DEGs). Surprisingly, *Pvalb*$^+$ neurons showed the most dramatic change at 60 min, with 3622 DEGs. Intersection analysis between the DEGs found in each dataset showed a neuron-specific response to stimulus, with little overlap between datasets (Ext Fig 2a; Suppl. Table 3). However, we confirmed global expression of immediate early genes (IEGs), a classical marker of neuronal activation (Ext. Fig 2c). The cell type-dependent diversity of response led us to analyze intra-cell type differences individually. Comparison between home cage-normalized conditions (*Box vs Shock*) demonstrated robust correlation across neuron types, suggesting that similar molecular changes are triggered regardless of the presence of aversive stimulus, but less robust without it (Ext Fig 2c). Time-dependent changes in the translatome showed a consistent learning-associated response in *Pvalb*$^+$ and *Sst*$^+$ neurons ($R_{pvalb} = 0.34$, $p < 2.2^{-16}$; $R_{sst} = 0.14$, $p < 2.2^{-16}$; Fig 2a). On the other hand, *Camk2a*$^+$ translatome showed low correlation between the two timepoints, indicating a dynamic response ($R_{camk2a} = -0.27$, $p < 2.2^{-16}$) (Fig 2a). Box control mice did not display any notable correlation (Fig 2a, panels on the right). Overall, these results reinforce the notion that each type of neuron deploys specific mRNA programs to handle translational plasticity.

We used GSEA to gain a broad picture of physiological alterations that take place in different neuron types during consolidation. We found that neuron-specific physiological adaptation after conditioning (Fig 2b; Suppl Table 4). For instance, *Pvalb*$^+$ neurons produced mitochondrial proteins, in agreement with previous reports that demonstrate an increase in calcium transients in response to memory consolidation[18] . *Camk2a*$^+$ neurons, on the other hand, rapidly translate mRNAs encoding ribosomal proteins, which are then underrepresented at 60 minutes, suggesting ribosomal run-off of mRNAs without replenishment (Fig 2c). Furthermore, *Camk2a*$^+$ and *Sst*$^+$ neurons undergo notable reconfiguration of translation control, agreeing with previous studies that demonstrate that proteostatic control is essential for proper memory consolidation[19-24].

Given the role of synaptic plasticity for memory consolidation, we decided to investigate how conditioning affects the translatome of synaptic protein-encoding mRNAs (i.e. the synaptome). We took advantage of a recent synapse-specific proteome resource to identify all the proteins present in the synapses of *Camk2a*$^+$, *Pvalb*$^+$ and *Sst*$^+$ neurons in our datasets[25]. We found changes to the synaptome as early as 15 min, but they are more prominent at 60 min post-training (Ext Fig 3b). Notably, *Pvalb*$^+$ neurons undergo greater modification of the synaptome than *Camk2a*$^+$ neurons during consolidation,

with 73.2% and 17.1% of the synaptome-encoding mRNAs being differentially translated after conditioning, respectively (Ext Fig 3b). Furthermore, the *Camk2a*$^+$ synaptome undergoes a dynamic reshaping process, with little in common between the 15- and 60-min changes (Ext Fig 3b, connecting lines). On the other hand, *Pvalb*$^+$ neurons have a structured response that persists through the first hour of consolidation. We further noted that *Camk2a*$^+$ neurons upregulate multiple modulators of cytoskeleton dynamics, and the opposite occurs in *Pvalb*$^+$ neurons (Ext Fig 3c-d). On the other hand, *Pvalb*$^+$ neurons enhance the production of endoplasmic reticulum (ER) and mitochondria modulators (Ext Fig 3d). In addition, *Pvalb*$^+$ neurons rapidly change the translatome related to neurotransmission, both at the presynaptic vesicle release process and post-synaptic GABA receptor subunits. (Ext Fig 3d). Altogether, these data strongly support a prominent role for translation in the conditioning-dependent reconfiguration of structural plasticity and E/I balance in the dHPC during consolidation.

mRNA translation is regarded as a non-stochastic process, and proteins related to similar biological processes/protein complex often get produced at similar rates[25]. We thus decided to identify groups of mRNAs with correlated expression patterns across cell types and conditions tested. We created a gene network using weighted gene co-expression network analysis (WGCNA)[26]. The input dataset, consisting of 70 independent samples containing purified fractions from *home cage, box* and *shock* conditions for each cell type, evidenced a clear distinction between cell types (Ext Fig 4a-b). WGCNA identified 10 gene modules (Ext Fig 4c-d; Suppl. Table 5). ~3200 genes were not inserted in any module, being assigned to module "*grey*" (Ext Fig 4c). We examined the correlation between modules and each dataset feature (i.e. neuron type, time point and condition). Coherent with the PCA analysis, we found the strongest correlations of modules with cell types (Ext Fig 2e). However, we also found significant correlations with other traits. For instance, the magenta and purple modules correlate well (R = 0.37 and 0.36, respectively) with the time point of 60 minutes. On the other hand, the blue (R=-0.36) and green (R=-0.34) modules had better correlation with 15 minutes time point. We thus interrogated the biological features that were overrepresented in each module. We found that the black and green modules, with strong correlation with *Camk2a*$^+$ neurons (Fig 2e), contains multiple mRNAs related to "*RNP granules"*. On the other hand, the turquoise module, which presents strong correlation with interneurons, had mRNAs encoding synaptic signaling components (Fig 2f).

WGCNA presents the advantage of finding global patterns of conditioning-induced translatome shifts that fit in every cell type analyzed, but it fails to identify cell type-specific effects. We thus mapped the $\log_2$FC values found in the differential expression analysis to the WGCNA network (Fig 2d). We identified discrete patterns of altered translation across modules. For instance, we found a transient upregulation of the brown and yellow modules in *Camk2a*$^+$ 15 minutes after training. The same pattern

occurs at 60 min in *Pvalb*$^{+}$ neurons, with smaller differences seen in *Sst*$^{+}$. Furthermore, we found a select upregulation of the blue module that occurs only in excitatory neurons 60 min post-conditioning. Gene ontology of this subset of 336 mRNAs revealed that these mRNAs encode proteins mostly related to proteostasis (Fig 2e), such as ER homeostasis (e.g., *Tmx1, Ptges3*, *Emc2, Emc7*), ubiquitinylation (e.g., *Ube* family of genes and *Rnf139*) and protein folding (e.g., *Vapa, Pdzd8, Dnaja1, Dnaja2* and *Dnajb4*), indicating a cell type-specific feature where excitatory neurons must undergo a tight control of proteostasis to cope with increased translation during memory consolidation. Altogether, these analyses suggest that excitatory neurons recruit proteostatic control mediators during early stages of long-term memory consolidation, a feature that is not shared with inhibitory neurons.

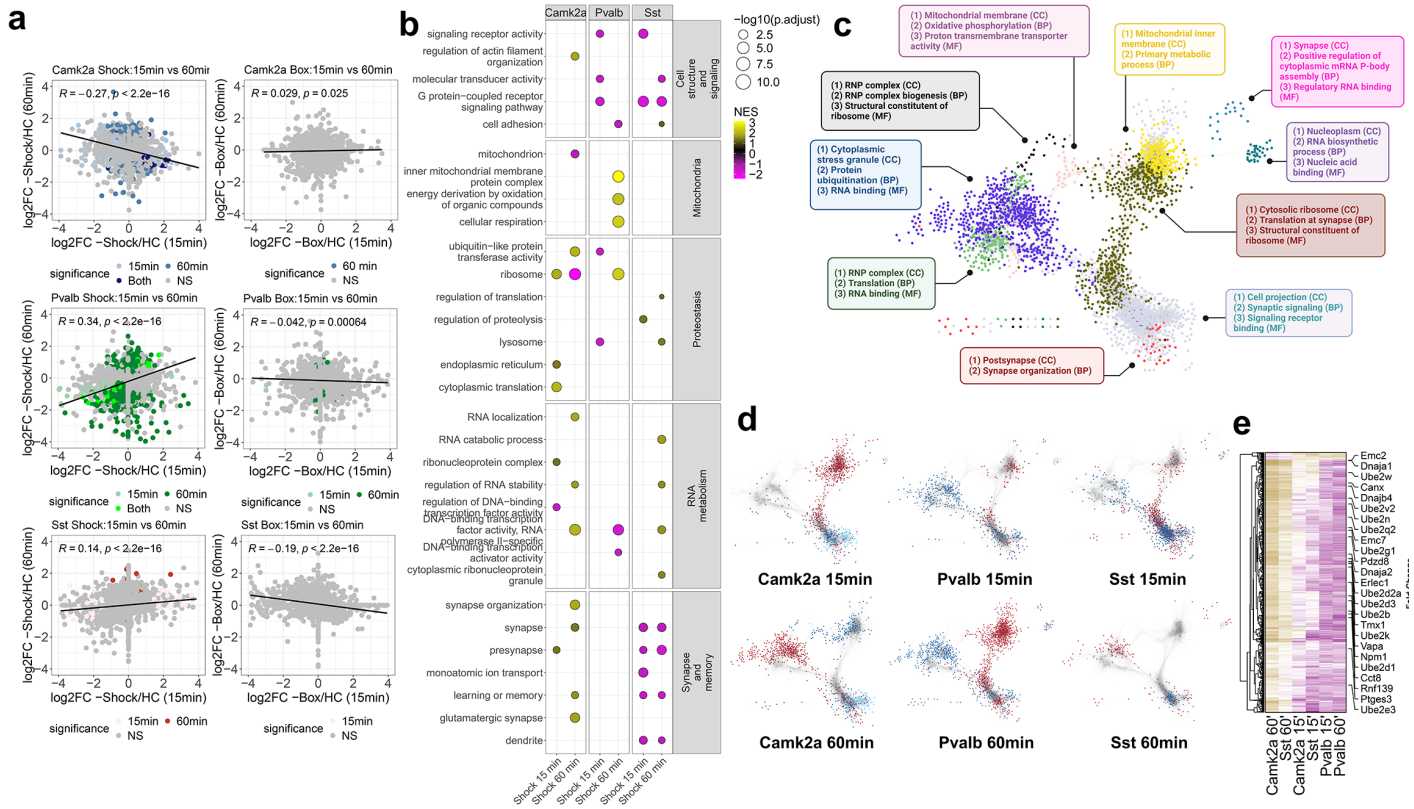

Figure 2. Threat conditioning leads to rapid modifications to the translatome of neurons. (**a**) Pearson's correlation of time-dependent mRNA translation in *shock* (left panels) or *box* conditions (right panels). Light colors = mRNAs differentially regulated exclusively at the 15 min time point; bright colors = mRNAs that are differentially regulated in both time points; dark colors = mRNAs that are differentially regulated in 60 min time point. Grey = mRNAs not significantly regulated. Top left corner = *R* represents Pearson's correlation coefficient; *p* is the significance of the correlation. (**b**) GSEA per dataset. NES = normalized enrichment score. Size of circles is mapped to -$\log_{10}$(FDR). Yellow = Upregulated terms; Magenta = Downregulated terms. (**c**) Representation of WGCNA network and the modules identified. Colors of modules are mapped to the network. Insets represent best overrepresented gene ontologies for each of the modules identified. (**d**) Map of the conditioning-induced fold in the gene network obtained through WGCNA. Red = upregulated (FC > 0.64), blue = downregulated (FC < -0.64). Top left = *Camk2a*$^+$ neurons 15 minutes after conditioning. Bottom left = *Camk2a*$^+$ neurons 60 minutes after conditioning. Top middle = *Pvalb*$^+$ neurons 15 minutes after conditioning. Bottom left = *Pvalb*$^+$ neurons 60 minutes after conditioning. Top left = *Sst*$^+$ neurons 15 minutes after conditioning. Bottom left = *Sst*$^+$ neurons 60 minutes after conditioning. (**e**) Heatmap of fold changes of all mRNAs clustered in the blue module. Magenta = downregulated genes; Gold = upregulated genes. Selected mRNAs with symbols highlighted are encompassed in GOs "*Proteasomal protein catabolic process*", "*Endoplasmic reticulum membrane"* and *"Unfolded protein binding*". Number of mice per group in the dataset used is listed in *Methods* section. Differential expression analysis and GSEA results were corrected for multiple comparisons using Benjamini-Hochberger *post-hoc* test.

GADD34 is a conditioning-induced modulator of translation and threat memory. The modifications in the nature of mRNAs being translated that occur during consolidation are accompanied by a sharp increase in *de novo* protein synthesis[8]. We recently found that this is controlled *in vitro* by the scaffolding protein GADD34, an effector of the Integrated Stress Response that is repurposed by neurons to localize phospho-eIF2α with protein phosphatase 1, resulting in eIF2α dephosphorylation and a subsequent increase in neuronal *de novo* translation (Fig. 3a)[27]. We found that GADD34 mRNA (*Ppp1r15a*) is significantly upregulated in the translatome of *Camk2a*$^+$, but not *Pvalb*$^+$ or *Sst*$^+$ neurons, 15 minutes following conditioning (Fig. 3b), in line with previous findings that suggest that the control of translation via eIF2α is required in *Camk2a*$^+$ neurons during consolidation[15]. The constitutive, non-inducible, counterpart of GADD34, CReP (*Ppp1r15b*), did not show any training-associated modifications (Fig. 3c). Thus, we reasoned that GADD34 plays a critical role in long-term memory formation. To assess this possibility, we crossed *Camk2a, Pvalb* or *Sst*.Cre driver lines with GADD34$^{fl/fl}$ mice, to achieve cell type-specific knockout of GADD34 (Fig. 3d). We then tested the memory of each mouse line using contextual threat conditioning. Notably, we found that cell type-specific knockout of GADD34 in *Camk2a*$^+$ neurons reduced long-term threat memory retention (Fig. 3d-e), consistent with our previous findings that GADD34 controls the translation of plasticity-related mRNAs in excitatory neurons[27]. Furthermore, overexpression of GADD34 in pyramidal neurons of the dHPC led to long-term memory enhancement, agreeing with earlier findings that this pathway bidirectionally controls memory formation[9,15,28,29] (Ext. Fig 5a-b). Enhanced threat memory also was observed when GADD34 was knocked out of *Pvalb*$^+$ (Fig. 3f) and *Sst*$^+$ (Fig. 3g) neurons, suggesting improper memory generalization, a feature characteristic of post-traumatic stress disorder[9,30].

We previously showed that GADD34 controls activity-dependent translation of plasticity-related mRNAs in neurons *in vitro*[27]. We thus hypothesized that a similar mechanism would occur *in vivo*. To test this hypothesis, we utilized the *Camk2a.Cre*-GADD34$^{fl/fl}$ mouse line, as this was the only neuron type to show an increased expression of GADD34 at early times post-conditioning, thus enabling rapid control of the neuronal translatome (Fig. 3b). We injected *Camk2a.Cre* (wild-type; WT) *Camk2a.Cre*-GADD34$^{fl/fl}$ (knockout; KO) mice with AAV9.*Camk2a*.TRAP virus in the dHPC (Fig. 3h). Mice were trained in contextual threat conditioning and their dHPC collected an hour after for TRAP. We confirmed efficacy of GADD34 knockout in excitatory neurons (Ext Fig 5d). We chose to examine changes in the dHPC translatome 60 min post-training as we hypothesized that the early (~15 min) increase in GADD34 translation would impact the later time point. Remarkably, we found that the KO mice displayed almost no DEGs, when compared against WT mice, which showed the typical response to threat conditioning (Fig 3i; Suppl. Table 6). Indeed, the conditioning-induced changes in translatome

were consistently smaller in KO mice (Fig. 3j), suggesting a global impairment of translation induced by conditioning. Failure in translation initiation is a hallmark of proteostatic disruption and cellular stress. Because proteostatic control reconfiguration is a feature of *Camk2a*$^{+}$ neurons early response to consolidation (Fig 2), we hypothesized that GADD34 knockout would abrogate this response. When we mapped the conditioning-induced fold change values found in KO mice in the co-variance gene network previously built, we found a failure to induce the expression of the *"blue*" module, which encompassed mRNAs encoding modulators of proteostatic control (Fig 3m). Overall, our data identifies a new mechanism employed by neurons early in memory consolidation, possibly to confer resilience to by-products (e.g., enhanced ER trafficking, reactive oxygen species, calcium dynamics) of the extreme anabolic situation induced by neuronal activity.

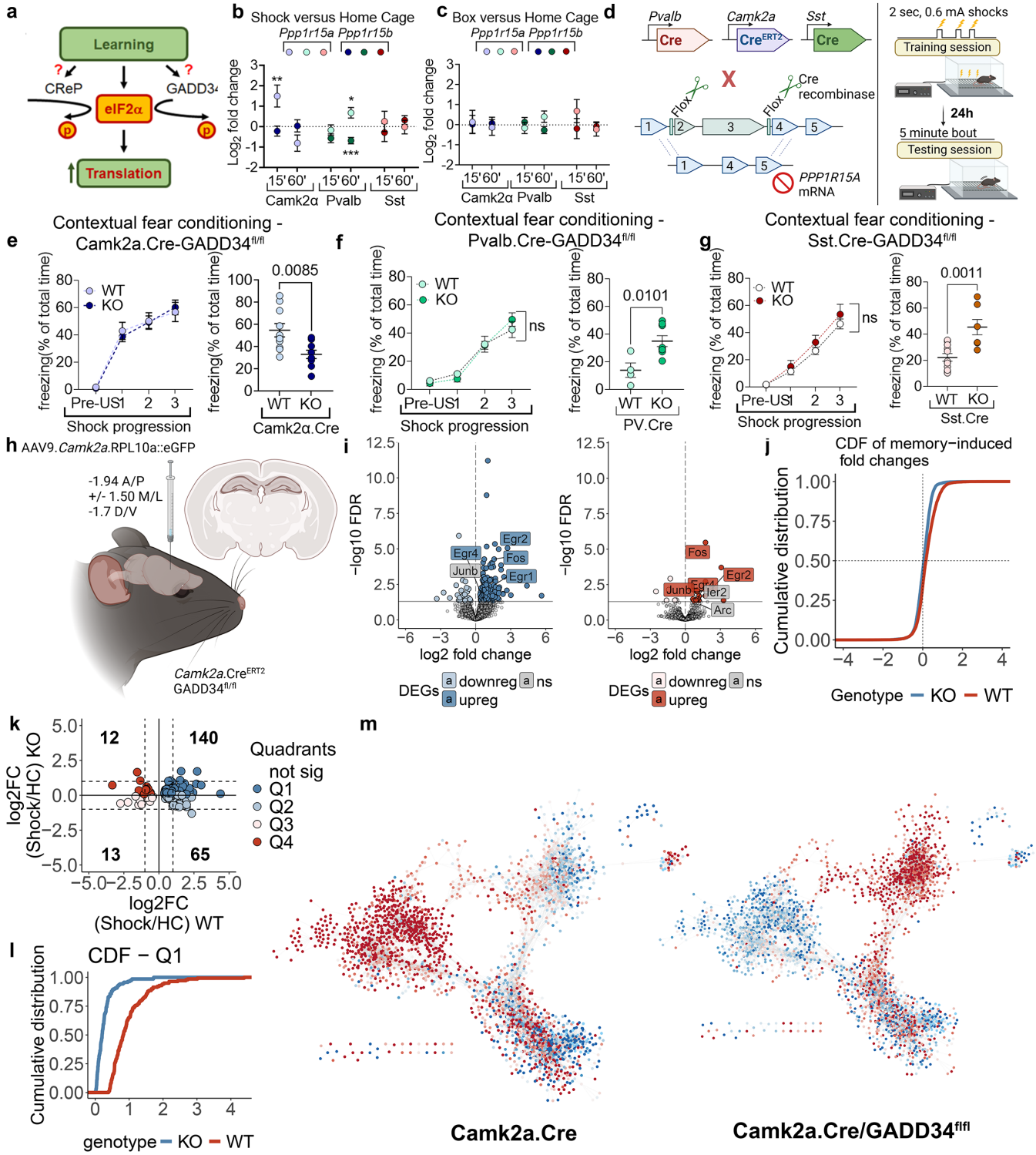

Figure 3. GADD34 is a conditioning-induced modulator of translation and threat memory. (**a**) Rationale for eIF2α dephosphorylation-mediated increase in *de novo* translation. (**b**) Log_2FC of *Ppp1r15a* (mRNA encoding GADD34) and *Ppp1r15b* (CReP) in the translatome of *Camk2a*$^+$, *Pvalb*$^+$ *or Sst*$^+$ neurons of mice that underwent contextual threat conditioning against home cage baseline. ** is p = 0.0054; * is p = 0.0076; *** is p = 0.00009. Number of mice per group in the dataset used is listed in *Methods* section. Mean ± SEM. (**c**) Log_2FC of *Ppp1r15a* and *Ppp1r15b* in the box control against home cage baseline. Number of mice per group in the dataset used is listed in *Methods* section. Mean ± SEM. (**d**) Schematics of contextual threat conditioning paradigm to test the cell type-specific relevance of GADD34 in long-term memory formation. (**e**) contextual threat conditioning of *Camk2a.Cre* ^{ERT2}*/GADD34*$^{fl/fl}$. Left panel = training session. Right panel = testing session. N = 9-10 mice/group; Mean ± SEM; F = 2.871. (**f**) contextual threat conditioning of *Pvalb.Cre/GADD34*$^{fl/fl}$. Left panel = training session. Right panel = testing session. N = 4-8 mice/group; Mean ± SEM.; F = 1.122.(**g**) contextual threat conditioning of *Sst.Cre/GADD34*$^{fl/fl}$. Left panel = training session. Right panel = testing session. N = 7-11 mice/group; Mean ± SEM; F = 2.319. (**h**) Experimental plan for TRAP-seq of excitatory neurons from the dHPC of *Camk2a.Cre*^{ERT2}*/GADD34*$^{fl/fl}$ mice.

N = 3-5 mice/group. Mean ± SEM. (**i**) Volcano plot illustrating the DEGs found in WT (left panel) and KO (right panel). Light blue = downregulated. Dark blue = upregulated. Grey = not significant. (**j**) cumulative distribution fraction of *shock vs home cage* log2 fold changes in WT (orange line) and KO (dark blue line). (**k**) Dot plot evidencing DEGs in WT that are differentially modulated in KO mice subjected to contextual threat conditioning. Q1-4 = Quadrant numbers. Dashed lines represent abs(Log_2FC) > 1. (**l**) Comparative cumulative distribution of genes belonging to Q1. (**m**) Map of the log_2FC values in the WGCNA network previously obtained. The WT network is similar to what was previously found. Red = upregulated (log_2FC > 0.35), blue = downregulated (log_2FC < -0.35).

AU-rich elements explain the conditioning-induced modifications to neuronal translatome. mRNA translation is controlled by multiple mechanisms, including cis- (i.e. features present in the mRNA sequence/structure) or trans-regulatory elements (i.e. elements that target mRNAs or ribosomes and control translation). Cis elements were previously shown to control mRNA stability, localization and translation, and are often found in the 3'UTR[31]. We thus interrogated whether cis elements in mRNAs would dictate rules guiding consolidation-induced mRNA translation. We gathered information regarding composition of the following mRNA cis-elements from literature: m6A enrichment, number of AU-rich elements (AREs) in the 3' UTR[32], number of RBPs targeting the 3' UTR, number of miRNAs targeting the transcript, length of the 3' UTR, and codon optimality[33].

To investigate whether mRNA regulatory elements were predictive of enrichment or depletion of an mRNA transcript in the translatome following conditioning, we leveraged classical machine learning techniques to build predictive models trained on seven key features of each transcript. Random forest (RF) models were trained separately for each training condition and time point because regulatory elements within an mRNA's sequence are static and are therefore differentially utilized by several hidden effectors whose regulatory valence is not captured by their dynamics (for a thorough characterization of each model's input, refer to Ext. Fig 6-8). Model performance was assessed by the area under curve (AUC) of their receiver operating characteristic (ROC) plots as well as validation accuracy metrics (Fig. 4b). Model performance was enhanced across datasets obtained from both *Pvalb*$^+$ and *Camk2a*+ neurons following conditioning and after longer post-training timescales, as the best model performance for both cell types occurred for the shock 60' group (shock 60' AUC: *Pvalb*$^+$ = 0.92, *Camk2a*$^+$ = 0.92) (Fig 4c-d). To determine which features most impacted the predictive power of the trained models, permutation importance tests were performed on each RF model. For those cell types, conditions, and time points where the respective trained models performed better than chance, permuting the number of AREs consistently resulted in the greatest decreases in validation accuracy. This result indicates that ARE count was the most predictive feature of up or downregulation for those models (Figure 4e, Ext Fig 6-8). This finding was confirmed again when RF models trained on single features from the *Pvalb*$^+$ and *Camk2a*$^+$ shock 60' datasets performed best when trained on the ARE feature set (ARE-shock 60' AUC: *Pvalb*$^+$ = 0.86, *Camk2a*$^+$ = 0.88) (Ext Fig 9). These results suggest that mRNA regulatory elements may be used to predict enrichment or depletion of mRNAs in the translatome post training, especially when predicting regulation valence following longer post-training

timescales. Additionally, the results show that AREs are principally informative features for our classical machine learning approach, and this importance may suggest that AREs play a distinct role in differential mRNA translation following conditioning.

The results obtained above suggest that cis-regulatory elements play a global role in directing reconfiguration of neuronal biological properties triggered by memory consolidation. To interrogate this, we first examined whether these mRNA elements could explain the concomitant translation of mRNAs related to particular biological properties targeted during consolidation. We isolated the identity of all mRNAs belonging to GOs significantly upregulated in previously performed GSEA (Fig 2) and collected information regarding each feature described above. We found that, in *Camk2a*$^{+}$ neurons, mRNAs with lower number of AREs in the 3'UTR are translated earlier, an opposite trend to what was observed in *Pvalb*$^{+}$ neurons (Fig 4d). Here, we found no feature driving notable changes in the translatome of *Sst*$^{+}$ neurons (Fig 4d, red dots and bars). Altogether, our results indicate a strong influence of hard-coded mRNA features in fast-paced changes in the translatome of neurons during memory consolidation.

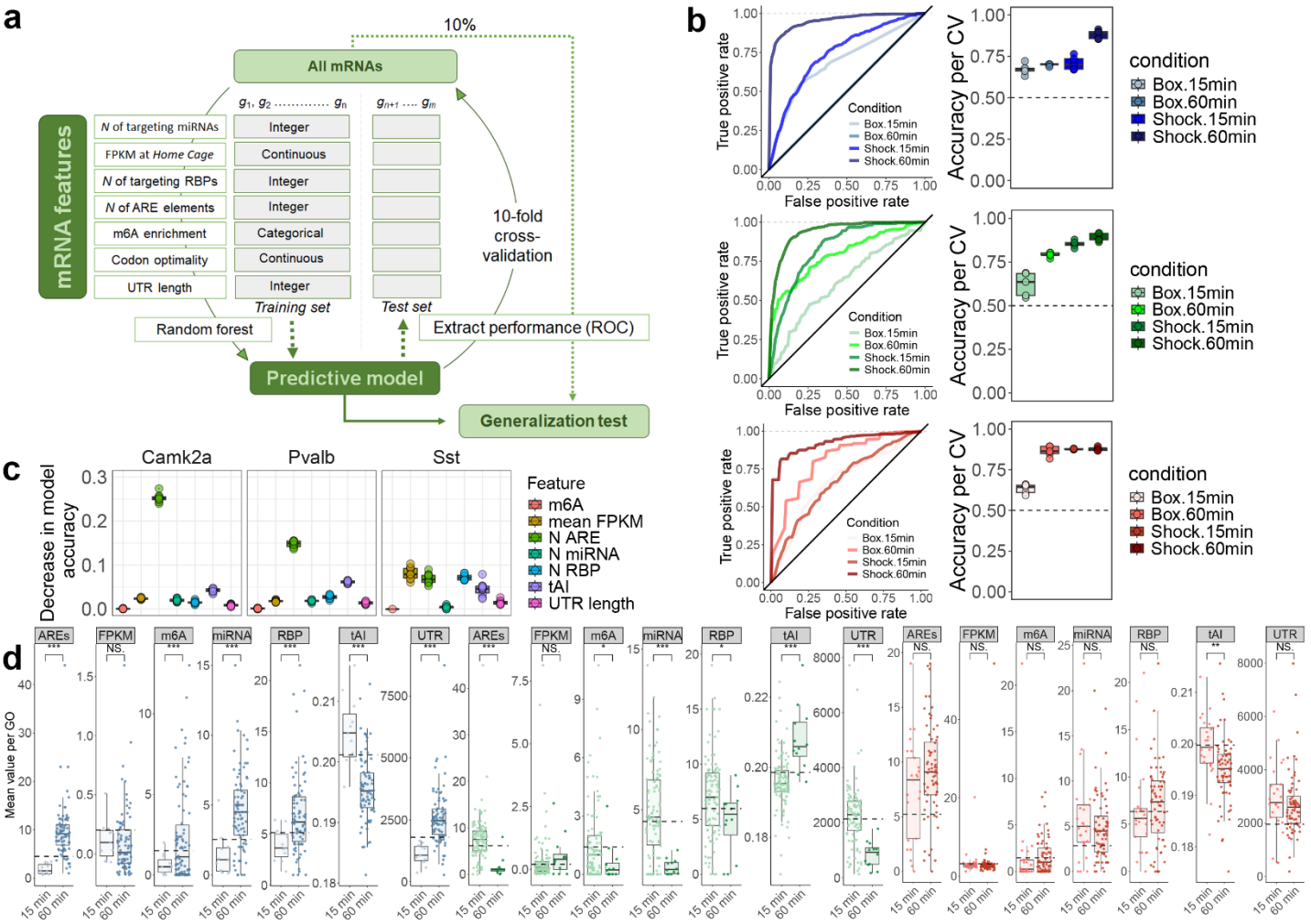

Figure 4. AU-rich elements explain the conditioning-induced modifications to neuronal translatome. (**a**) Sketch illustrating the design of the machine learning model. Adapted from ref[34]. (**b**) On the left: Receiver Operating Characteristics (ROC) curves on the left, categorized by condition tested. Diagonal line represents random classification performance (50% chance). On the right: Accuracy per cross validation performed (k = 10). Horizontal dashed line represents random classification performance. Blue shades graphs = *Camk2a*; Green shades graphs = *Pvalb*; Red shades graphs =*Sst*. (**c**) Permutation importance test (k = 10). (**d**) Average value of each mRNA feature per GO that was upregulated 15- or 60-min following training. Blue graphs = GOs upregulated in *Camk2a*$^+$ neurons; Green graphs = GOs upregulated in *Pvalb*$^+$ neurons; Red graphs = GOs upregulated in *Sst*$^+$ neurons. Light colors represent 15-minute time points, darker colors represent the 60-minute time point. Number of mice per group in the dataset used is listed in *Methods* section. NS = not significant; * = p<0.05; ** = p<0.01; *** = p<0.001.. Student's t test, unpaired. The upper and lower whiskers represent the largest and smallest value within 1.5 times above or below the interquartile range, respectively.

hNRNPD controls long-term memory formation. AREs are targeted by numerous RBPs, that can promote stabilization, translation or decay of the targeted mRNA[35]. Because of that, we examined the role RBPs played in the training-induced translatome modifications. To assess this, we used RBPMap[36] to identify unbiased map motifs at the 3'UTR that were targeted by 126 different RBPs (Suppl. Table 7). All the mRNAs targeted by a given RBP had their log_2FC values transformed into an arithmetic means, that was used as a coordinate for the centroid positioning on the cartesian plane. We found a notable correlation in the centroid positioning when we compared shock 60' and shock 15' groups (R_{Pvalb} = 0.98; R_{Camk2a} = -0.94; Ext. Fig 10; Suppl. Table 8). The box control group displayed good correlation values (R_{Pvalb} = -0.72; R_{Camk2a} = 0.45), but the effects were more mild and not significant when compared against the centroid for all genes (Ext Fig 10). To confirm that these results were reliable, we also collected data from the literature that experimentally defined mRNAs targeted by RBPs. We collected data from 47 different RBPs, 14 of which overlapped with the prediction assay. Similar analyses demonstrated the same correlational trends, thus reinforcing the reliability of our findings (Ext Fig 10; Suppl. Table 9-10).

Because we found AREs to be the best predictors of differential translation of mRNAs, we restricted our analyses to RBPs that target AREs. We found that their targeted mRNAs were downregulated in *Pvalb*$^+$ neurons, whereas they were upregulated in *Camk2a*$^+$ neurons 1h following conditioning (Fig 5a). The modulation seemed to be ubiquitous to all ARE-targeting RBPs, consistent with the notion that these proteins compete with each other for the targets. Our analysis indicated that heterogeneous nuclear Ribonucleoprotein D (hnRNPD) displayed significant changes in mRNA translation of its targets across datasets (Suppl. Table 8), suggesting a universal role for this protein in memory consolidation. hnRNPD is primarily found in the nucleus, participating in alternative splicing[35]. It is also known to destabilize mRNAs via de-adenylation, resulting in decay[37], thereby granting it a major role in the regulation of the mRNA pool present in the cell. We hypothesized that its expression in neurons is essential to proper long-term memory formation. We perturbed expression of hnRNPD by either overexpressing or silencing it in each neuron type studied (Fig 5b). We found the most notable role for hnRNPD in Camk2a+ cells, where either increasing or decreasing its expression caused memory impairments (Fig5c-e). In *Pvalb*+ cells, we found that increasing, but not decreasing, hnRNPD expression caused memory impairment; Fig 5f-h). Finally, no clear effects were observed in *Sst*+ cells (Fig 5i-k). Altogether, these findings demonstrate that the dataset generated can be effectively used to discover new modulators of memory.

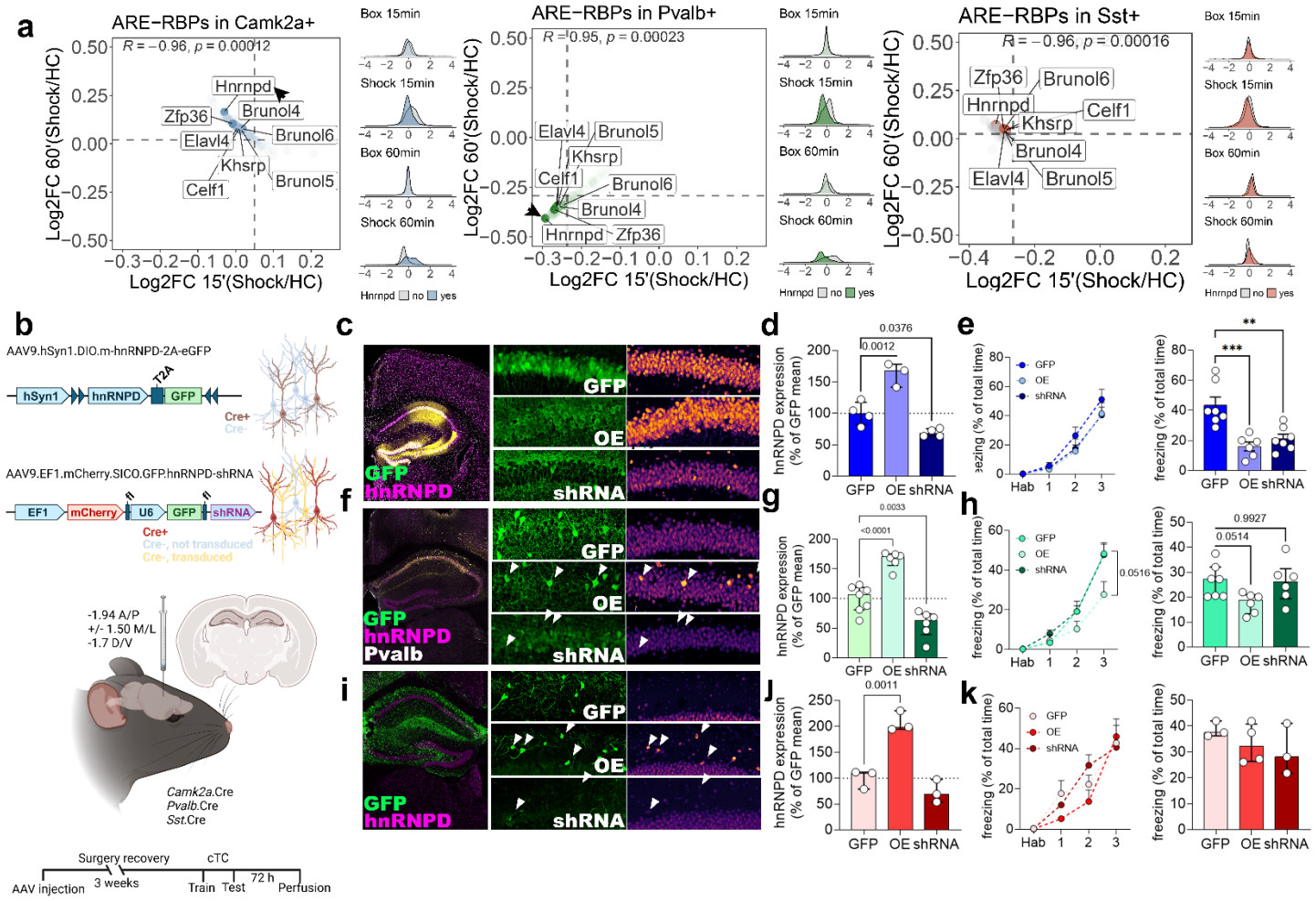

Figure 5. hNRNPD controls long-term memory formation. (a) Average fold changes of mRNAs predicted to be targets of ARE-RBPs in *Camk2a*$^+$, *Pvalb*$^+$ and *Sst*$^+$ neurons (names of the RBPs are highlighted inside each plot). Top left corner = *R* represents Pearson's correlation coefficient; *p* is the significance of the correlation. On the right of each plot, density plots show fold change distribution of targets of hnRNPD (colored) versus non-targets (in gray). Number of mice per group in the dataset used is listed in *Methods* section. (b) Schematics of bi-directional manipulation approach to target hnRNPD expression in a cell type-specific way. Green neurons = GFP$^+$ neurons; red neurons = mCherry$^+$ neurons; yellow neurons = GFP$^+$/mCherry$^+$ neurons. A/P = anterior/posterior. M/L = medial/lateral; D/V = dorsal/ventral coordinates. (c) On the left: representative images of the viral delivery to the dHPC. (d) Quantification of hnRNPD levels in AAV9.GFP-, OE- or shRNA-transduced neurons. Mean ± SEM. (e) contextual threat conditioning of mice injected with AAV9.GFP-, OE- or shRNA viruses. On the left, training session; On the right, test session . OE = overexpression. Hab = habituation (N=6-7 mice per group); Mean ± SEM; *F* for panel d = 0.7307; *F* for panel e = 0.6853. (f-h) Same as in c-e, but for *Pvalb*$^+$ neurons (N=6-7 mice per group); Mean ± SEM; *F* for panel g = 0.7307; *F* for panel h = 1.097. (i-k) Same as in c-e, but for *Sst*$^+$ neurons (N=3-4 mice per group); Mean ± SEM; *F* for panel j = 0.0421; *F* for panel k = 1.412. *P* values are demonstrated to the fourth digit above bar graphs in each panel. One-Way ANOVA corrected with Tukey's *post-hoc* test.

Discussion

The creation and stabilization of long-term memories require orchestrated communication between cells in different brain regions, as well as the local tuning of activity through the interaction of neighbor neurons in brain region microcircuits. Thus, each neuron type presents molecular and electrophysiological signatures that enable them to effectively play their specific role during memory consolidation. This molecular framework dictates how neurons dynamically respond to external cues, such as converging network stimuli, causing adaptations that are specific to each neuron type. Here we probed the alterations to the translational programs of three different neuron types induced by threat conditioning: *Camk2a*$^+$*, Pvalb*$^+$ and *Sst*$^+$ neurons. Using the translatome as a readout of molecular programs triggered during consolidation proved to be valuable, as it focuses on mRNAs bound in ribosomes that are likely undergoing active translation. This approach provides important advantages compared to current strategies to study RNA metabolism in a cell type-specific way, as it approximates the readout to the functional outcome of neuronal plasticity - modifications at the proteome level – while preserving the high depth of sequencing that isessential for identifying shifts in molecular programs that are characteristic of dynamic cellular behavior. To our knowledge, this is the first report on neuron-specific translational programs of the dHPC during memory consolidation. Our results demonstrate that the initial pattern of cellular adaptation is diverse, both in the nature of the cellular processes/structures altered and in the molecular mechanisms that guide the translation of the proper set of mRNAs.

One of the most noticeable features in the datasets that we collected was that *Camk2a*$^+$ neurons displayed a dynamic response within the first hour of consolidation, propelling translation plasticity as early as 15 minutes after conditioning by increasing the production of controllers of protein metabolism. This rapid shift in the production of proteostasis machinery is noteworthy, given that it is comparable to the speed of translation of immediate early genes[38]. Interneurons displayed a delayed proteostatic reconfiguration, instead prioritizing rapid changes in regulators of neurotransmission. One possible interpretation of such a phenomenon is that negative feedback systems tend to be tightly controlled, such as to produce phasic inhibition of increasing excitatory activity in the hippocampus[39]. Thus, it is possible that interneurons, after targeting proteins that control synaptic transmission, then respond with a ubiquitous reconfiguration of protein metabolism.

The changes in the nature of mRNAs being translated by neurons during consolidation are accompanied by an increase in translation, dependent on the dephosphorylation of the initiation factor eIF2α that is enabled by the scaffolding protein GADD34[8,9,15,22,27,29]. GADD34 promotes protein phosphatase 1-induced dephosphorylation eIF2α after acute cellular stress is resolved, restoring

protein synthesis to its homeostatic levels[40]. However, in neurons, we found that GADD34 is repurposed to regulate a rapid increase in protein synthesis. We suggest that GADD34 mediates the capacity of neurons for large increases in *de novo* translation, promoting translation of mRNAs related to protein folding, degradation and translation initiation. This finding is unexpected, given the canonical role assigned for GADD34 under stressful conditions[41]. Based on our findings, we hypothesize that the repurposing of GADD34 is necessary for neuronal resilience, not only promoting traditional neuronal plasticity, but also ensuring proper cellular adaptation to a stressful condition that accompanies plasticity. Indeed, pathways associated with buffering cellular stress, such as the unfolded protein response[24], chaperone production[42] and alternative translation initiation pathways[43], are consistently linked to neuronal plasticity and memory formation. Its impairment, on the other hand, plays a fundamental role in synaptic and cognitive failures observed in neurodegenerative diseases[44-46]. Overall, these findings reinforce the notion that neurons simultaneously engage both a large-scale anabolic response and a stress-resolving pathway during plasticity, a previously undescribed cellular state that warrants further investigation.

In addition to regulating increases in mRNA translation via upstream signaling molecules, such as GADD34, mRNAs carry cis-regulatory elements in their sequence that are recognized by the cell machinery, which ultimately decides the fate of that mRNA. These elements were previously shown to be crucially involved in subcellular localization of mRNAs in neurons[31], and to be involved in the translational control of specific mRNAs[47]. Here, we have demonstrated that cis-regulatory elements have a critical role in regulating mRNA dynamics during long-term memory consolidation. Our model is currently limited by knowledge of static features pertaining to the mRNA sequence. This can be solved via a deep characterization of cis- and trans- features occurring in specific types of neurons, which will provide insights into how the molecular diversity of neurons determines their role during memory consolidation. Furthermore, our model does not account for transcriptional modulation of gene expression, which has a clear effect on mRNA availability for translation. This stems from the lack of efficient methods to study rapid transcriptional fluctuations with cell type-specific precision in the brain in dynamic settings such as the one employed in this work. However, we were still able to identify and characterize a previously unknown modulator of long-term memory formation, hnRNPD. This protein is associated with post-transcriptional mRNA decay[37] and alternative splicing[35], and its microdeletion results in intellectual disability in humans[48]. It remains to be determined how hnRNPD with its two main post-transcriptional roles in mRNA modulation culminates in memory formation. Nevertheless, approaches such as those we have used in the current work will enable the identification of many other proteins that play key roles in cognitive tasks and how they are linked to brain disorders.

In summary, we have examined the plasticity of the translatome induced by threat conditioning in three different neuron types involved in memory consolidation. Our results indicate a clear pattern of response by each neuron, distinguishable by specific biological features rapidly modified to adapt in response to an afferent input. However, we find that the translatome modifications are governed by elements hard coded in the mRNA sequence, regardless of the cell type examined. Overall, we anticipate this study to provide novel insights into the changes to the molecular environment of neurons induced in response to learning-related stimuli, with broad implications in the collective fields of the molecular, systems, and computational neuroscience.

Methods

Mice. All animal protocols were reviewed and approved by the New York University Animal Care and Use Committee (protocol # 2021-1139). Mice were provided with food and water *ad libitum* and maintained in a 12h-12h light-dark cycle at New York University at a stable temperature (78oF) and humidity (~50%). Male and female mice were equally distributed across experimental groups and used between the ages of 3-5 months. B6;129S6-Tg(Camk2a-cre/ERT2)1Aibs/J (Camk2a.Cre^{ERT2}), B6.129P2-Pvalb$^{tm1(cre)Arbr}$/J (Pvalb.Cre), Sst$^{tm2.1(cre)Zjh}$/J (Sst.IRES.Cre), B6;129S4-Gt(ROSA)26Sor$^{tm9(EGFP/Rpl10a)Amc}$/J (TRAP^{fl}) and B6.Cg-Tg(Camk2a-cre)T29-1Stl/J (T29.Cre) mouse lines were obtained from Jackson laboratories. B6.129P2-Ppp1r15a$^{tm1.1Ajf}$/Mmnc (GADD34^{fl}) was obtained from Mutant Mouse Resource & Research Centers (MMRRC). Every mouse strain was backcrossed with C57Bl/6J mice from Jackson laboratories every 5 weeks, to ensure minimal background variation. T29.Cre, Pvalb.Cre or Sst.Cre were crossed with TRAP^{fl} mice to obtain Camk2a.Cre-TRAP, Pvalb.Cre-TRAP or Sst.Cre-TRAP mouse lines. Heterozygous GADD34$^{fl/+}$-Heterozygous Camk2a.Cre^{ERT2} were crossed with heterozygous GADD34$^{fl/+}$ mice to obtain Camk2a.CreERT2.GADD34$^{fl/fl}$. Heterozygous GADD34$^{fl/+}$- Homozygous Pvalb.Cre were crossed with heterozygous GADD34$^{fl/+}$ mice to obtain Pvalb.Cre.GADD34$^{fl/fl}$. Heterozygous GADD34$^{fl/+}$-Homozygous Sst.IRES.Cre were crossed with heterozygous GADD34$^{fl/+}$ mice to obtain Sst.IRES.Cre.GADD34$^{fl/fl}$.

Behavior. All behavior was conducted during the light cycle. Mice were randomly assigned for the order of testing. All data were collected by experimenters' blind to mouse genotype. To habituate mice, over the course of two days mice were placed in the threat conditioning room for 1hr then each mouse was handled for 2min each day. For simple contextual threat conditioning, mice were placed in the context, which consisted of a metal rod floor and white house light, first for 298s and then presented with three series 2s, 0.6mA foot shocks. The first and second inter-trial intervals were 128s and 98s respectively, and after the last shock, mice remained in the chamber for an additional 170s. This training protocol lasted in total 700s. LTM was tested 24h after training by placing mice back in the same context for 300s. Freezing behavior was automatically measured using Freeze Frame 4 (ActiMetrics).

Translating ribosome affinity purification (TRAP). TRAP was performed as before[11,27]. In brief, we separate the procedure into three steps: (1) anti-GFP-coated magnetic bead preparation; (2) sample collection and ribosome purification; (3) RNA extraction. All steps were performed in RNAse-free condition. Total number of mice per experimental group is listed below:

<u>*Camk2a* 15 min dataset</u>

. Home cage = 3 mice

. Box only = 3 mice

. Shock = 4 mice

Camk2a 60 min dataset

. Home cage = 4 mice

. Box only = 4 mice

. Shock = 5 mice

Pvalb 15 min dataset

. Home cage = 4 mice

. Box only = 4 mice

. Shock = 5 mice

Pvalb 60 min dataset

. Home cage = 6 mice

. Box only = 4 mice

. Shock = 6 mice

Sst 15 min dataset

. Home cage = 4 mice

. Box only = 3 mice

. Shock = 5 mice

Sst 60 min dataset

. Home cage = 2 mice

. Box only = 2 mice

. Shock = 2 mice

Below we list each step of the TRAP purification process:

Anti-GFP-coated magnetic bead preparation: 300 μl of MyOne Streptavidin beads per reaction was collected, washed once with PBS, and incubated with 120 μl of recombinant protein L suspension (1 μg/ml) and 180 μl of PBS for 35 minutes/RT/rotating. Beads were washed 5x with 1 ml PBS + 3% IgG-free bovine serum albumin solution and incubated with anti 50 μg of each anti-GFP Htz antibodies for 1h/RT/rotating. Beads were then washed 3x with cold low-salt buffer (80 mM HEPES-KOH, 600 mM KCl, 40 mM $MgCl_2$, 1% NP-40, 1x Halt protease/phosphatase inhibitor, 0.5 mM DTT, 100 ug/ml Cycloheximide) and resuspended in 200 μl low-salt buffer per reaction. Beads were stored at 4oC until time of use.

Sample collection: After contextual threat conditioning training session, mice were returned to their home cage for 15 minutes or 60 minutes. Mice were quickly euthanized via cervical dislocation, and the brain removed and rapidly transferred to ice-cold PBS + 100 μg/ml Cycloheximide. The dorsal hippocampus was micro-dissected and transferred to an ice-cold pestle container containing 1 ml of tissue lysis buffer (80 mM HEPES KOH, 600 mM KCl, 40 mM MgCl2, 1% NP-40, 1x Halt protease/phosphatase inhibitor, 0.5 mM DTT, 100 μg/ml Cycloheximide, RNAsin). Tissue was homogenized with 12 strokes of a motorized pestle and the suspension transferred to an Eppendorf tube. The homogenate was cleared via centrifugation (2000 *g*/10 min/4°C) and the supernatant saved. The supernatant was added with 1/9th v:v of 300 mM DHPC, mixed by inversion and incubated for 5 minutes on ice. Samples were then centrifuged at 20.000 *g*/10 min/4°C and supernatant collected. 1/10th of the volume was saved for total lysate sequencing, and the remaining volume was added with 200 μl of anti-GFP-coated magnetic beads and incubated overnight/rotating/4°C. On the next day, samples were briefly spun to recover all beads and washed 4x with 1ml ice-cold high-salt buffer (80 mM HEPES-KOH, 1.4M KCl, 40 mM MgCl2, 1% NP-40, 1x Halt protease/phosphatase inhibitor, 0.5 mM DTT, 100 μg/ml Cycloheximide, RNAsin). Beads were then resuspended in 350 μl of room temperature RLT buffer added with 10% β-mercaptoethanol (RNEasy plus kit). 350 ul of RLT buffer + BME was further added to total lysate samples.

RNA purification. RNA purification was done using RNEasy plus kit (Qiagen). This kit was chosen due to the extra purification step to eliminate gDNA contamination. RNA extraction was performed in accordance with the suggested protocol, with following changes: (1) 50 μl of elution buffer was used, to maximize recovery; (2) two sequential elution steps were performed. RNA quality was measured using Agilent's ScreenTape for global assessment of RIN, and for detailed quantitative and qualitative assessment Agilent's Bioanalyzer was used.

Library preparation and sequencing. The SMART-Seq HT Kit (Takara) was utilized to generate complementary DNA (cDNA) from 1 ng of total RNA input. The synthesis process included 13 PCR cycles. The quality and concentration of the stock RNA were assessed using an Agilent Pico Chip on the Bioanalyzer system, ensuring high-quality RNA inputs. cDNA was quantified using the Invitrogen Quant-iT system and subsequently diluted to a concentration of 0.3 ng/µL for library preparation. Libraries were prepared using the Nextera XT DNA Library Preparation Kit, employing a total of 12 PCR cycles to amplify and uniquely barcode the samples. The quality of the generated libraries was assessed using the Agilent High Sensitivity TapeStation to verify fragment sizes and integrity. Additionally, the concentration of libraries was measured with the Invitrogen Quant-iT system to ensure accurate normalization prior to pooling. Libraries were pooled in equimolar ratios and sequenced on the Illumina NovaSeq X+ platform using paired-end 100-cycle reads. The sequencing data was processed using standard bioinformatics pipelines.

Alignment to the genome and feature count. Sequencing adapters and low-quality bases were trimmed using *Cutadapt* version 1.8.2 using the following arguments (-q 20 -O 1 -a CTGTCTCTTATA)[49]. Trimmed reads were then aligned to the mm9 mouse genome assembly using STAR version 2.7.7a and raw gene counts were obtained using *-quantmode* argument[50]. Total feature counts were obtained using *featureCounts*[51]. Feature counts were extracted and converted into one single raw count matrix, where each row was a gene, and each column was a sample. When samples were obtained from different cohorts of mice, we used ComBat-Seq[52] to batch-correct.

Bioinformatics.

<u>Quality control and filtering of enriched genes.</u> The datasets were subjected to a series of filtering steps and data clearing to ensure strict data analysis. One of the noticeable features across datasets was the increase in translation of mRNAs traditionally associated with glial cells when mice were conditioned, such as markers *Cx3cr1* and *Aldh1l1*. Because these are known to be low or not expressed by neurons, we decided to filter our datasets to select truly neuron-enriched transcripts in the translatome. To achieve this in an unbiased manner, we used the following approach to filter out genes that were not “truly enriched” in the purified fraction:

$$\frac{\frac{CPM_{X_{IP}}}{Length_X}}{\frac{CPM_{X_{TL}}}{Length_X}} \geq 1.3 \; in \geq 50\% \; of \; home \; cage \; samples$$

Where $\frac{CPM_X}{Length_X}$ is the FPKM per gene, IP is the counts in the purified fraction, and TL is the counts in the total fraction (*total lysate*). We used these values to select samples that had over 2-fold enrichment of cell type markers. Samples that did not have this enrichment were excluded from further analysis. The cell type markers used were:

Camk2a$^+$ *neurons = Camk2a, Slc17a6, Slc17a7*

Pvalb$^+$ *neurons = Gad1, Pvalb*

Sst$^+$ *neurons = Gad1, Sst*

Glial markers (to evaluate depletion) = *Aldh1l1, Gfap, Cnp, Cx3cr1*

Differential expression analysis and Gene Set Enrichment Analysis (GSEA). Identification of DEGs was made using DESeq2 package, from R[53]. Briefly, crude count matrices were batch corrected using ComBat-seq[52]. The resulting matrix served as input to DEseq2. Transcripts with less than 20 counts in more than half of the samples were filtered out. Datasets were then normalized by the size of library, and variance stabilized using the *vst*() function from DEseq2 package. DEG identification was done by filtering genes that had *p* adjusted value < 0.05 in comparisons between conditions. After a full list of DEGs were obtained, genes categorized as enriched per cell type (see above) were filtered in and kept for future analysis. We opted for doing the filtering at this later stage to ensure the rigor of adjusting the *P* value in the whole list of genes. For intra-cellular analyses, genes identified as enriched in that particular cell type were kept. For inter-cellular analyses, such as the WGCNA (see below), a full list, containing all genes enriched in all neuron types, was obtained and used as input. A complete list of all DEGs from all datasets is available in Supplementary table 3. These lists of DEGs were independently subjected to GSEA analysis using the *clusterProfiler* package from R, and only sets larger than 150 genes were considered in the analyses[54]. Because *Pvalb+* neurons showed no GSEA groups with positive modulation following 15 min shocks, we re-did the analysis including smaller GO sets (> 30 genes), which were used to generate panel 4d.

Weighted gene co-expression network analysis (WGCNA). WGCNA was performed using the *WGCNA* R package[26]. Briefly, a variance-stabilized (vst) data frame containing all conditions from all cell types was fed into the pipeline. For *Camk2a* 15 min time point, we found a significant reduction in the total amount of represented genes (~1000 less genes than in other datasets). We thus imputed those values using classification and regression trees method, from the *mice* R package, in the pre-vst dataset. No noticeable difference in the raw count distribution was observed, suggesting that the imputation did not

significantly alter the nature of our dataset. To design a scale-free network, the premise behind WGCNA, we first calculated the power for soft-thresholding, which uses lasso regularization to select significant connections in the network. A power of 20 was found to reach the threshold of 0.8 in the scale-free index, suggested by the designers of WGCNA as optimal for producing scale-free networks.

The categorization of genes in co-expression modules was done using the *blockwiseModules()* function, using power = 20 and minimum module size of 30. We generated a signed network, to separate genes that had positive versus negative correlations. 10 modules were identified and assigned colors to them. The *grey* module, however, contained genes that were not assigned to any particular module, and was therefore excluded from functional analyses below. The gene ontology (GO) analysis for each module identified was performed using the *anRichment* package in R, with a threshold of 0.05 and Bonferroni correction. We manually curated GO terms per module to display in representative figure in panel 2c.

To design a network that could be exported and visualized in cytoscape[55], we first generated a weighted network using the function *TOMsimilarityFromExpr()* from *WGCNA* package. We then exported all the most significant connections (threshold = 0.1) from all modules but *"grey*" using the *exportNetworkToCytoscape()* function. The edges file was imported into cytoscape and matched with the nodes file, containing the information regarding gene-module membership. Modules were colored in the network with colors matching their names. To map log2FC values obtained in the differential expression analysis into the network, we matched the log2FC of each gene with the corresponding network node, and mapped both color (upregulated = *red*, downregulated = *blue*) and size (below threshold = *size 5*, above threshold = *size 30*). Genes that were in the third quartile of log_2FC (i.e. 25% of most differentially expressed genes) were considered above threshold.

<u>Filtering synaptic protein-encoding mRNAs</u>. Enriched proteome datasets were retrieved from Oostrum and colleagues. We then did two sequential filtering procedures: first filtered hippocampal proteomes, then for each of our specific cell types (*Camk2a+, Pvalb+* or *Sst+*). We subsequently filtered proteins that were significantly enriched over fraction *p2* (whole synaptosome fraction). The resulting list of genes was used as dictionary to filter only synaptome-encoding genes from our datasets containing differential expression analysis.

<u>Prediction of RBP targets:</u> 3'UTR from all transcripts identified in each sequencing were retrieved using biomaRt. The longest variant of 3'UTR was selected by transcript. The 3'UTRs were submitted as query to RBPMap[36], using the following parameters: *-db_motifs 'all' -stringency 'high' -conservation 'off'*. Results were extracted into a data frame using scripts built *in house* (Supplementary table 7). For

centroid plots, the arithmetic means of all mRNAs predicted as targets to a given RBP were calculated for each time point and plotted in a two-dimensional space.

Machine learning model. In order to begin our machine learning analysis, we constructed cell-type specific datasets consisting of 7 features of transcripts identified from our TRAP-seq analysis: AU-rich elements (ARE), codon optimality (tAI), UTR length, total number of known regulatory RNA binding proteins (RBPs), total counts of m6A modification, total number of known regulatory microRNAs (miRNA), and mean FPKM. These feature sets were obtained from different sources:

. ARE counts in the 3' UTR were retrieved from previous work that quantified them using regular expression in the longest transcript version[32].

.Codon optimality was retrieved from http://stadium.pmrc.re.kr/[33]. Codon optimality is represented as the geometric mean of the relative optimality of codons from the coding sequence of each mRNA.

. UTR lengths were retrieved using biomaRt database (R).

. Total number of known regulatory RNA binding proteins were retrieved using scripts built *in house* to identify motifs in the longest version of the 3'UTR of each transcript that are targeted by RBPs included in the RBPMap[36] database (see above).

. total counts of m6A modifications per transcript, per cell type, were obtained from a study that generated a single-cell database of the epitranscriptome[56].

. total number of regulatory microRNAs were obtained for each cell type. Specifically, we first retrieved a list of microRNAs expressed in each neuron type from previous literature[57]. We then used miRWalk[58] to identify microRNA-binding regions in the longest isoform of the 3'UTR of each transcript identified per neuron type. Finally, we used scripts built *in house* to filter the binding sites to keep only the ones targeted by microRNAs expressed in the specific neuron type.

. mean FPKM values were derived from the translatome dataset of *home cage* mice and represent the average ribosomal occupancy of each mRNA in resting conditions.

Furthermore, the dataset was filtered to select most notable fold changes relative to *home cage* (log2FC > 0.5 or log2FC < -0.5). Those transcripts which were deemed to be significantly up or downregulated in comparison to home-cage mice were used for training. This selection thereby narrows the scope of the subsequent analysis as identifying those features of the mRNA which predict the valence of

regulation of those transcripts known to be differentially regulated during TC. Using this dataset, Gradient-boosted decision tree ensemble models (XGBoost v2.0.3) were trained using 5-fold cross validation (CV). Per CV fold, feature columns were scaled to ensure they were distributed such that $\mu = 0$ and $\sigma = 1$. After scaling, outlier detection was performed using the IQR with bounds determined by the standard multiplicative factor 1.5, with values determined to be outliers removed and subsequently imputed using KNN imputation with K = 5 (scikit-learn v1.2.1). Hyperparameters were tuned using a Bayesian optimization strategy (scikit-optimize v0.9.0) by scoring optimal estimators using the accuracy of the trained estimator on the held-out validation folds. Due to the static nature of mRNA sequence regulatory features, optimal models were trained and selected per condition and per cell type, allowing for cross condition/cell type comparison. Optimal model topologies across conditions were compared using the mean area under the receiver operating characteristic (AUC-ROC) curve as computed per fold using the held-out validation data. Feature importances were determined using two methods. Firstly, untrained models with optimal hyperparameters were trained on single features of the dataset and performance across features was compared using the AUC computed per fold using the held-out validation data. Additionally, permutation importance tests were performed using 10 permutation sets per feature and importance was assessed by the mean reduction of model accuracy over these sets.

Viral injection. Mice aged 10-12 weeks were anesthetized with isoflurane (induction: 5%, maintenance: 2% in O2 vol/vol, via nose cone) and placed in a stereotaxic frame. Bilateral dorsal hippocampus (−1.94 mm AP, ±1.50 mm ML, and −1.5 mm DV) was stereotactically targeted with a 10ul Hamilton syringe. The syringe was lowered to the 0.2mm pocket below the target site and remained there for 5 min before moving to the target and beginning the injection. After the injection, the needle stayed in the target site for 5 min before it was withdrawn. A StereoDrive microinjector (Microinjection Robot, StereoDrive) was used to deliver 500 nl of viral solution (2.1-2.8 vg/ml titer) at a rate of 0.2ul/min. *Camk2a*.Cre, *Pvalb*.Cre, and *Sst*.Cre transgenic mouse lines were injected with either AAV9.hSyn1.DIO.m-hnRNPD.2A.eGFP or AAV9.EF1.mCherry.DIO.U6.GFP.hnRNPD-shRNA (Vector Biolabs). *Camk2a.Cre*^{ERT2}*/GADD34*^{flfl} mice were injected with AAV9.CAG.FLEX.RPL10a::eGFP (Addgene). Mice were given 3 mg kg−1 ketoprofen for 3 days post-op as an analgesic after surgery and allowed to recover for at least 2 weeks before behavioral experiments.

Immunohistochemistry and imaging quantification. Mice were anesthetized with isoflurane and perfused trans-cardial with PBS to clear blood and then with 4% PFA in PBS (~40 ml per mouse). Brain was removed and post-fixed in 4% PFA in PBS overnight/4oC. Brains were mounted in 3% agarose and sectioned using a vibratome Leica VT1200S. Around 40 sections of 40 μm thickness containing

the dorsal hippocampus area were collected and kept free-floating in PBS added with sodium azide. For immunohistochemistry, sections were permeabilized with PBS + 0.5% Triton X-100 solution (15min/RT/agitation). Sections were then blocked in 5% normal goat serum (NGS) + PBS + 0.1% Triton X-100 solution (60 min/RT/agitation). Brain sections were then transferred to blocking buffer added with primary antibodies (for concentrations, please refer to antibodies and reagents section) and incubated overnight/4oC/agitation. They were subsequently washed 3x with PBS + 0.1% Triton X-100 (10 min/agitation/RT) and incubated with blocking buffer added with secondary antibodies (1:500, 90 min/agitation/RT). Sections were washed 3x again with PBS + 0.1% Triton X-100, transferred to PBS and mounted in glass slides. After sections were dry, Prolong was added to the slides and covered with a rectangular coverslip. Imaging was performed in an upright Leica SP8 system, with 20x magnification objective. To measure the efficacy of hnRNPD overexpression and silencing systems, we imaged 4 CA1-containing fields per mouse (2 sections, 2 images per section, one CA1-containing image per hippocampus). Quantification of hnRNPD intensity was done using Fiji macros built in house. Briefly, cell type-specific markers were used to generate a mask. That mask was used to select the region-of-interest in the channel containing hnRNPD staining. The raw integrated density value was obtained and normalized against the total area selected. For better visualization in the graph (Fig 5c-e), GFP-injected signal was averaged, and the resulting value was determined as 100%. All the points were then re-calculated to obtain a relative value against the average of control.

Tamoxifen preparation and administration. Tamoxifen was prepared and administered as suggested by Jackson laboratories. In brief, tamoxifen powder was diluted in corn oil to the final concentration of 20 mg/ml, under vortex agitation for 4h/RT. Solution was stored at 4°C for up to a week. Mice were administered with tamoxifen solution intraperitoneally for 5 days, 100 μl per day (~ 75 mg/kg).

Statistical analysis. Mice were included as independent biological variables in all the experiments performed in this work. No replicated measures were obtained from the same mouse and treated as independent samples. Graphs were designed using *ggplot2* package from R or GraphPad Prism. The exact test and post-test corrections are described in figure legends, but as a general rule we used two-tailed unpaired *t* test for comparisons between two groups, and One-Way ANOVA for comparisons that involved more than one group, with all conditions independent from one another. One-Way ANOVA was corrected using Tukey *post-hoc* test. Prior to statistical analysis, data was tested for normality using Q-Q and residual plots.

References

1 Josselyn, S. A. & Tonegawa, S. Memory engrams: Recalling the past and imagining the future. *Science* **367**, doi:10.1126/science.aaw4325 (2020).

2 Sun, W. *et al.* Spatial transcriptomics reveal neuron-astrocyte synergy in long-term memory. *Nature* **627**, 374-381, doi:10.1038/s41586-023-07011-6 (2024).

3 Bahl, E. *et al.* Using deep learning to quantify neuronal activation from single-cell and spatial transcriptomic data. *Nat Commun* **15**, 779, doi:10.1038/s41467-023-44503-5 (2024).

4 Vanrobaeys, Y. *et al.* Mapping the spatial transcriptomic signature of the hippocampus during memory consolidation. *Nat Commun* **14**, 6100, doi:10.1038/s41467-023-41715-7 (2023).

5 Schwanhausser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* **473**, 337-342, doi:10.1038/nature10098 (2011).

6 Fair, B. *et al.* Global impact of unproductive splicing on human gene expression. *Nat Genet*, doi:10.1038/s41588-024-01872-x (2024).

7 Alberini, C. M. & Chen, D. Y. Memory enhancement: consolidation, reconsolidation and insulin-like growth factor 2. *Trends Neurosci* **35**, 274-283, doi:10.1016/j.tins.2011.12.007 (2012).

8 Shrestha, P. & Klann, E. Spatiotemporally resolved protein synthesis as a molecular framework for memory consolidation. *Trends Neurosci* **45**, 297-311, doi:10.1016/j.tins.2022.01.004 (2022).

9 Shrestha, P. *et al.* Cell-type-specific drug-inducible protein synthesis inhibition demonstrates that memory consolidation requires rapid neuronal translation. *Nat Neurosci* **23**, 281-292, doi:10.1038/s41593-019-0568-z (2020).

10 Shrestha, P. *et al.* Amygdala inhibitory neurons as loci for translation in emotional memories. *Nature* **586**, 407-411, doi:10.1038/s41586-020-2793-8 (2020).

11 Heiman, M., Kulicke, R., Fenster, R. J., Greengard, P. & Heintz, N. Cell type-specific mRNA purification by translating ribosome affinity purification (TRAP). *Nat Protoc* **9**, 1282-1291, doi:10.1038/nprot.2014.085 (2014).

12 Thomson, S. R. *et al.* Cell-Type-Specific Translation Profiling Reveals a Novel Strategy for Treating Fragile X Syndrome. *Neuron* **95**, 550-563 e555, doi:10.1016/j.neuron.2017.07.013 (2017).

13 Furlanis, E., Traunmuller, L., Fucile, G. & Scheiffele, P. Landscape of ribosome-engaged transcript isoforms reveals extensive neuronal-cell-class-specific alternative splicing programs. *Nat Neurosci* **22**, 1709-1717, doi:10.1038/s41593-019-0465-5 (2019).

14 Bhatti, D. L. *et al.* Molecular and Cellular Adaptations in Hippocampal Parvalbumin Neurons Mediate Behavioral Responses to Chronic Social Stress. *Front Mol Neurosci* **15**, 898851, doi:10.3389/fnmol.2022.898851 (2022).

15 Sharma, V. *et al.* eIF2alpha controls memory consolidation via excitatory and somatostatin neurons. *Nature* **586**, 412-416, doi:10.1038/s41586-020-2805-8 (2020).

16 Iadecola, C. Neurovascular regulation in the normal brain and in Alzheimer's disease. *Nat Rev Neurosci* **5**, 347-360, doi:10.1038/nrn1387 (2004).

17 Attwell, D. *et al.* Glial and neuronal control of brain blood flow. *Nature* **468**, 232-243, doi:10.1038/nature09613 (2010).

18 Chamberland, S. & Topolnik, L. Inhibitory control of hippocampal inhibitory neurons. *Front Neurosci* **6**, 165, doi:10.3389/fnins.2012.00165 (2012).

19 Moon, S. L., Sonenberg, N. & Parker, R. Neuronal Regulation of eIF2alpha Function in Health and Neurological Disorders. *Trends Mol Med* **24**, 575-589, doi:10.1016/j.molmed.2018.04.001 (2018).

20 Costa-Mattioli, M., Sossin, W. S., Klann, E. & Sonenberg, N. Translational control of long-lasting synaptic plasticity and memory. *Neuron* **61**, 10-26, doi:10.1016/j.neuron.2008.10.055 (2009).

21 Banko, J. L. *et al.* Behavioral alterations in mice lacking the translation repressor 4E-BP2. *Neurobiol Learn Mem* **87**, 248-256, doi:10.1016/j.nlm.2006.08.012 (2007).

22 Costa-Mattioli, M. *et al.* Translational control of hippocampal synaptic plasticity and memory by the eIF2alpha kinase GCN2. *Nature* **436**, 1166-1173, doi:10.1038/nature03897 (2005).

23 Pandey, K., Yu, X. W., Steinmetz, A. & Alberini, C. M. Autophagy coupled to translation is required for long-term memory. *Autophagy* **17**, 1614-1635, doi:10.1080/15548627.2020.1775393 (2021).

24 Martinez, G., Khatiwada, S., Costa-Mattioli, M. & Hetz, C. ER Proteostasis Control of Neuronal Physiology and Synaptic Function. *Trends Neurosci* **41**, 610-624, doi:10.1016/j.tins.2018.05.009 (2018).

25 van Oostrum, M. *et al*. The proteomic landscape of synaptic diversity across brain regions and cell types. *Cell* **186**, 5411-5427 e5423, doi:10.1016/j.cell.2023.09.028 (2023).

26 Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559, doi:10.1186/1471-2105-9-559 (2008).

27 Oliveira, M. M. *et al*. The integrated stress response effector GADD34 is repurposed by neurons to promote stimulus-induced translation. *Cell Rep* **43**, 113670, doi:10.1016/j.celrep.2023.113670 (2024).

28 Sidrauski, C. *et al*. Pharmacological brake-release of mRNA translation enhances cognitive memory. *Elife* **2**, e00498, doi:10.7554/eLife.00498 (2013).

29 Costa-Mattioli, M. *et al*. eIF2alpha phosphorylation bidirectionally regulates the switch from short- to long-term synaptic plasticity and memory. *Cell* **129**, 195-206, doi:10.1016/j.cell.2007.01.050 (2007).

30 Dong, X., Wang, Y., Liu, Y. & Li, Y. Fear generalization modulated by shock intensity and protein synthesis inhibitor. *Psychopharmacology (Berl)*, doi:10.1007/s00213-024-06662-1 (2024).

31 Chekulaeva, M. Mechanistic insights into the basis of widespread RNA localization. *Nat Cell Biol* **26**, 1037-1046, doi:10.1038/s41556-024-01444-5 (2024).

32 Loedige, I. *et al*. mRNA stability and m(6)A are major determinants of subcellular mRNA localization in neurons. *Mol Cell* **83**, 2709-2725 e2710, doi:10.1016/j.molcel.2023.06.021 (2023).

33 Yoon, J., Chung, Y. J. & Lee, M. STADIUM: Species-Specific tRNA Adaptive Index Compendium. *Genomics Inform* **16**, e28, doi:10.5808/GI.2018.16.4.e28 (2018).

34 Jung, J. *et al*. mRNA transport, translation, and decay in adult mammalian central nervous system axons. *Neuron* **111**, 650-668 e654, doi:10.1016/j.neuron.2022.12.015 (2023).

35 Sidali, A. *et al*. AU-Rich Element RNA Binding Proteins: At the Crossroads of Post-Transcriptional Regulation and Genome Integrity. *Int J Mol Sci* **23**, doi:10.3390/ijms23010096 (2021).

36 Paz, I., Kosti, I., Ares, M., Jr., Cline, M. & Mandel-Gutfreund, Y. RBPmap: a web server for mapping binding sites of RNA-binding proteins. *Nucleic Acids Res* **42**, W361-367, doi:10.1093/nar/gku406 (2014).

37 Gouble, A. *et al*. A new player in oncogenesis: AUF1/hnRNPD overexpression leads to tumorigenesis in transgenic mice. *Cancer Res* **62**, 1489-1495 (2002).

38 Clayton, D. F. *et al*. The role of the genome in experience-dependent plasticity: Extending the analogy of the genomic action potential. *Proc Natl Acad Sci U S A* **117**, 23252-23260, doi:10.1073/pnas.1820837116 (2020).

39 Farrant, M. & Nusser, Z. Variations on an inhibitory theme: phasic and tonic activation of GABA(A) receptors. *Nat Rev Neurosci* **6**, 215-229, doi:10.1038/nrn1625 (2005).

40 Costa-Mattioli, M. & Walter, P. The integrated stress response: From mechanism to disease. *Science* **368**, doi:10.1126/science.aat5314 (2020).

41 Pakos-Zebrucka, K. *et al*. The integrated stress response. *EMBO Rep* **17**, 1374-1395, doi:10.15252/embr.201642195 (2016).

42 Chatterjee, S. *et al*. Endoplasmic reticulum chaperone genes encode effectors of long-term memory. *Sci Adv* **8**, eabm6063, doi:10.1126/sciadv.abm6063 (2022).

43 Hacisuleyman, E. *et al*. Neuronal activity rapidly reprograms dendritic translation via eIF4G2:uORF binding. *Nat Neurosci* **27**, 822-835, doi:10.1038/s41593-024-01615-5 (2024).

44 Lourenco, M. V. *et al*. TNF-alpha mediates PKR-dependent memory impairment and brain IRS-1 inhibition induced by Alzheimer's beta-amyloid oligomers in mice and monkeys. *Cell Metab* **18**, 831-843, doi:10.1016/j.cmet.2013.11.002 (2013).

45 Moreno, J. A. *et al*. Sustained translational repression by eIF2alpha-P mediates prion neurodegeneration. *Nature* **485**, 507-511, doi:10.1038/nature11058 (2012).

46 Freeman, O. J. & Mallucci, G. R. The UPR and synaptic dysfunction in neurodegeneration. *Brain Res* **1648**, 530-537, doi:10.1016/j.brainres.2016.03.029 (2016).

47 Das, S., Vera, M., Gandin, V., Singer, R. H. & Tutucci, E. Intracellular mRNA transport and localized translation. *Nat Rev Mol Cell Biol* **22**, 483-504, doi:10.1038/s41580-021-00356-8 (2021).

48 Gillentine, M. A. *et al*. Rare deleterious mutations of HNRNP genes result in shared neurodevelopmental disorders. *Genome Med* **13**, 63, doi:10.1186/s13073-021-00870-6 (2021).

49 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10-12, doi:https://doi.org/10.14806/ej.17.1.200 (2011).

50 Dobin, A. *et al*. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).

51 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930, doi:10.1093/bioinformatics/btt656 (2014).

52 Zhang, Y., Parmigiani, G. & Johnson, W. E. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genom Bioinform* **2**, lqaa078, doi:10.1093/nargab/lqaa078 (2020).

53 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

54 Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287, doi:10.1089/omi.2011.0118 (2012).

55 Shannon, P. *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504, doi:10.1101/gr.1239303 (2003).

56 Tegowski, M., Prater, A. K., Holley, C. L. & Meyer, K. D. Single-cell m(6)A profiling in the mouse brain uncovers cell type-specific RNA methylomes and age-dependent differential methylation. *Nat Neurosci* **27**, 2512-2520, doi:10.1038/s41593-024-01768-3 (2024).

57 He, M. *et al*. Cell-type-based analysis of microRNA profiles in the mouse brain. *Neuron* **73**, 35-48, doi:10.1016/j.neuron.2011.11.010 (2012).

58 Sticht, C., De La Torre, C., Parveen, A. & Gretz, N. miRWalk: An online resource for prediction of microRNA binding sites. *PLoS One* **13**, e0206239, doi:10.1371/journal.pone.0206239 (2018).