analytical chemistry

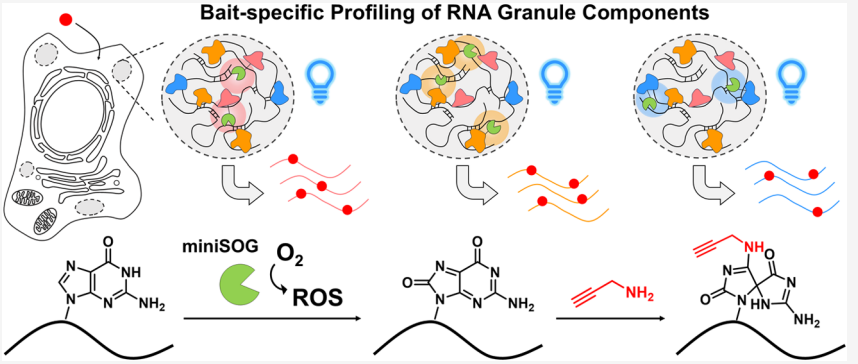
pubs.acs.org/ac Article

Spatially Resolved Multibait Mapping of Stress Granule and Processing Body Transcriptome

Ziqi Ren, Songrui Zhao, Wei Tang, and Peng Zou*

Cite This: *Anal. Chem.* 2025, 97, 12767−12775 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



ABSTRACT: Stress granules (SGs) are dynamic, membrane-less organelles that house complex RNA-protein networks. Although previous profiling methods have characterized SG RNAs as long, translation-repressed, and extensively epigenetically modified, it remains unclear whether these RNAs are evenly distributed within SGs. In this study, we genetically targeted the photocatalyst protein miniSOG to multiple SG core proteins, enabling the comprehensive CAP-seq profiling of SG-associated RNAs. Our results reveal that RNAs near different SG core proteins display heterogeneous distributions and distinct intrinsic features. We also employed CAP-seq to map RNAs associated with processing body (PB) marker protein DDX6 under both unstressed conditions and arsenite-induced stress. By comparing the transcriptomes proximal to SGs and PBs, our data suggest that m^{6}A modification may promote RNA localization to SGs, whereas higher AU content may facilitate mRNA targeting to PBs. These findings point to potential regulatory mechanisms that determine the subcellular localization of mRNAs within membrane-less organelles.

INTRODUCTION

The spatial organization of biological macromolecules within cells is often closely tied to their functions. This link is exemplified by membrane-less organelles, such as stress granules (SGs) and processing bodies (PBs), which form via liquid−liquid phase separation driven by interactions among proteins, RNAs, and other biomolecules.[1] By compartmentalizing specific cellular components, these organelles regulate critical processes including stress responses and RNA metabolism.[2] SGs rapidly assemble in response to various stress stimuli, such as oxidative stress, hyperosmolarity, starvation, and heat shock. They typically range in size from 100 nm to 1 μm^2, and contain core proteins including G3BP1, G3BP2, and TIA1.[3,4] Advanced imaging techniques (e.g., super-resolution and electron microscopy) have revealed a nonuniform distribution of proteins and RNAs within these granules.[4,5] In contrast, PBs are approximately 0.5 μm in diameter and harbor RNA-binding proteins involved in RNA decay (e.g., decapping enzymes, exonucleases, and helicases like DDX6[6−8]). Under stress conditions such as sodium arsenite exposure, SGs and PBs often become spatially connected,[6,9] with live-cell fluorescence imaging showing dynamic RNA shuttling between them within seconds.[10]

Understanding the functions of these membrane-less organelles requires a precise definition of their molecular composition. SG assembly involves both proteins (such as G3BP1) and RNAs. Due to their flexibility and abundant secondary structures, RNAs can condense at lower concentrations than intrinsically disordered proteins, driving phase separation, modulating the physical properties of droplets, and defining compartment identity independent of protein−protein interactions.[11] Recent findings that SG RNAs can form persistent assemblies without protein scaffolds underscore the importance of RNA in these processes.[12]

Received: March 17, 2025
Revised: June 3, 2025
Accepted: June 4, 2025
Published: June 10, 2025

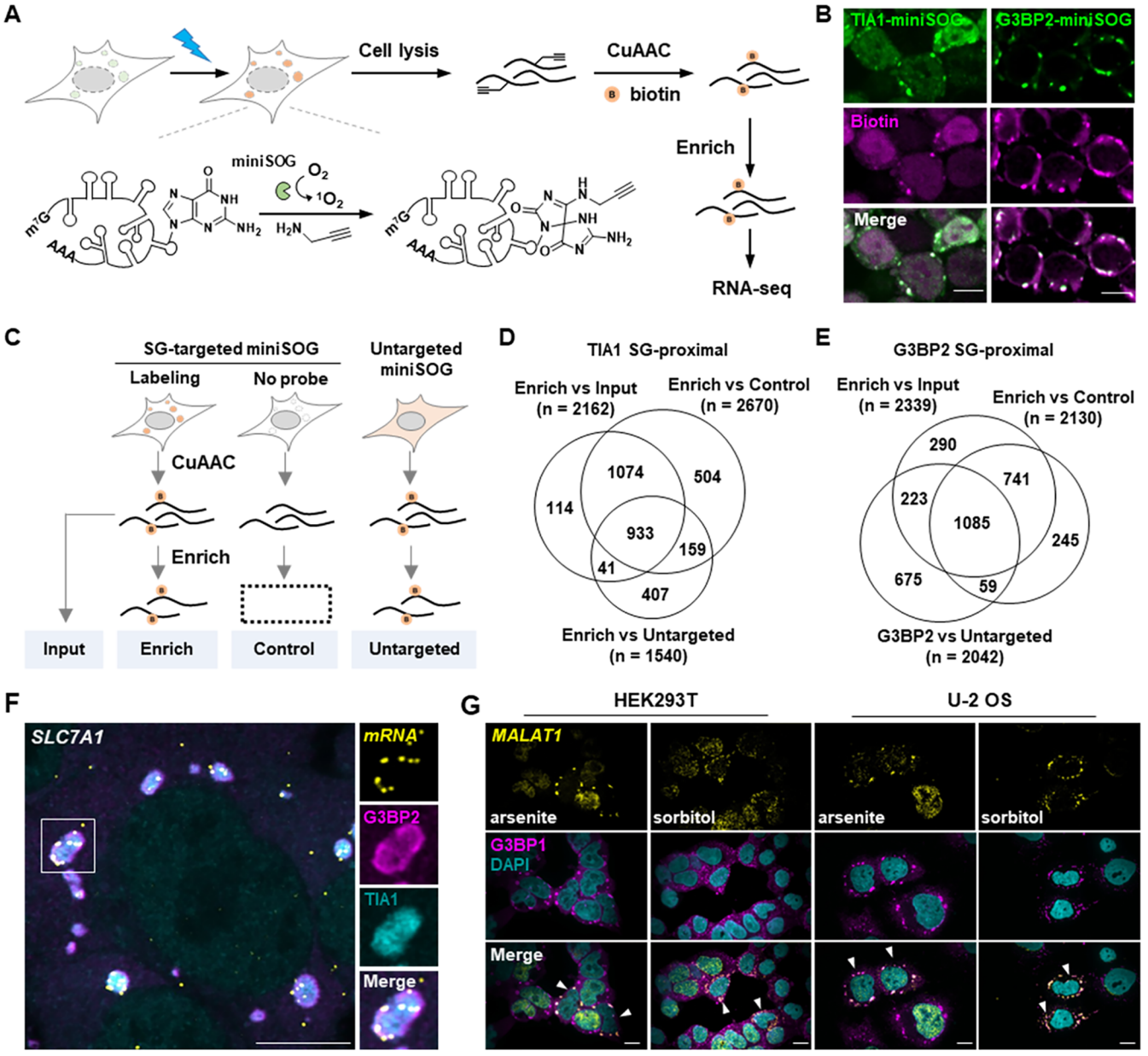


Figure 1. CAP-seq profiling of SG transcriptome with multiple baits. (A) Schematic workflow of SG CAP-seq. Cells were treated with 0.5 mM sodium arsenite for 60 min, labeled with 10 mM PA, and reacted with azide-conjugated biotin via copper-assisted alkyne−azide cycloaddition (CuAAC). Biotinylated RNAs were enriched for next-generation sequencing. (B) Confocal images of biotinylated HEK293T cells expressing TIA1-miniSOG and G3BP2-miniSOG (green). Biotinylation signal is detected with streptavidin-conjugated dye (magenta). (C) Workflow of Enrich, Input, Control, and Untargeted sample preparation. (D, E) Venn diagrams comparing RNAs enriched in arsenite-treated HEK293T expressing TIA1-miniSOG (D) and G3BP2-miniSOG (E). SG-proximal RNAs are defined as the overlap of: (1) Enrich vs Input; (2) Enrich vs Control; (3) Enrich vs Untargeted. (F, G) smFISH images of *SLC7A1* (F) and *MALAT1* (G) in HEK293T and U-2 OS cells under arsenite or sorbitol stimulation. Zoom-in views of the boxed region in (F) are shown on the right. For all fluorescence images, scale bars are 10 μm.

Traditional approaches, such as affinity purification combined with high-throughput sequencing, have profiled SG RNA content and revealed an enrichment of longer RNAs with reduced translation efficiency.[13,14] However, these methods risk losing weakly associated components and introducing contaminants, particularly given the dynamic nature of these organelles.[13,15] Proximity labeling techniques overcome these challenges by directly tagging RNAs in live cells, thereby enabling more accurate studies of RNA composition during the assembly and disassembly of SGs.[15,16] For example, APEX-seq has been used to characterize the SG transcriptome during heat shock,[16] while TRIBE-ID has demonstrated that G3BP1-associated RNAs correlate positively with transcript length and negatively with translation efficiency.[13,14,16]

Our recent work employed a photocatalytic proximity labeling method, CAP-seq,[17] to analyze the transcriptome proximal to G3BP1 in cultured mammalian cells under both stressed and basal conditions.[15] By employing a genetically encoded photosensitizer (miniSOG), CAP-seq mediates bait-proximal RNA oxidation and covalently captures RNAs with an amine probe bearing a click reaction handle (propargylamine, PA).[17] This approach revealed that SG-enriched RNAs correlate with increased length and AU content, yet exhibit

https://doi.org/10.1021/acs.analchem.5c01590
Anal. Chem. 2025, 97, 12767−12775

reduced translation efficiency,[15] and that m^{6}A modifications promote RNA targeting to SGs during arsenite-induced stress. CAP-seq has also proved valuable in characterizing RNAs during SG disassembly, uncovering AU-rich, translationally repressed RNA nanostructures that persist long after stress removal.[15] However, previous studies using CAP-seq have focused on a single SG component, leaving it unclear whether similar RNA composition patterns are shared by other SG core proteins. Moreover, the close spatial association between SGs and PBs raises questions about the extent of the transcriptomic overlap between these organelles.

In this study, we selected multiple SG core proteins and applied CAP-seq to resolve the subgranular transcriptome under sodium arsenite stress. We reveal heterogeneous RNA distributions within SGs and compare the transcriptomes of SGs and PBs, demonstrating that sequence features and m^{6}A modification levels influence RNA targeting these organelles under stress conditions.

EXPERIMENTAL SECTION

For reagents and experimental details, please refer to the Supporting Information.

Mammalian Cell Culture and RNA Labeling. HEK293T/17 cells were cultured in 6-well plates and co-transfected with target pLX304 plasmids and packing plasmids for homemade lentivirus. HEK293T cells were infected with lentivirus to generate cell lines expressing TIA1-miniSOG, G3BP2-miniSOG, miniSOG-DDX6, or TIA1-EGFP. Cell lines were characterized via immunofluorescence, detailed protocols are provided in the Supporting Information.

HEK293T cells expressing TIA1-miniSOG, G3BP2-miniSOG, miniSOG-DDX6, or untargeted miniSOG were cultured in 15 cm culture dishes until they reached ∼90% confluency. To induce stress, cells were treated with complete medium containing 0.5 mM sodium arsenite for 40 min at 37 °C, then the cells were rinsed once with 1× HBSS and the medium was replaced with HBSS containing 10 mM PA and 0.5 mM sodium arsenite. Cells were illuminated with a blue light-emitting diode (LED) (emission peak 465−475 nm, 24 mW/cm^2) for 15 min at room temperature. For RNA sample preparation, cells were washed with phosphate-buffered saline (PBS) and then lysed with TRIzol Reagent. RNA was extracted according to the manufacturer's instructions, treated with DNaseI at 37 °C for 30 min. Next, the RNA was incubated with click reagents consisting of 0.1 mM biotin-azide, 2 mM THPTA, 0.5 mM $CuSO_4$, and 5 mM sodium ascorbate for a 10 min CuAAC reaction. The RNA was purified using the RNA Clean & Concentrator kit and eluted with prewarmed nuclease-free water.

RNA Library Preparation and Data Analysis. RNA integrity was assessed using a Fragment Analyzer (Agilent). Approximately 1 μg of RNA was set aside as pre-enrichment samples (Input), while the remaining RNA was purified with Dynabeads MyOne Streptavidin C1 beads as the manufacturer's instructions indicated. Pre-enrichment RNAs (100 ng) and post-enrichment RNAs (5 μL, labeled sample or the negative control) were used for cDNA library construction with the NEBNext Ultra II RNA Library Prep Kit for Illumina according to the manufacturer's instructions. he final cDNA libraries were sequenced using 150 bp paired-end reads, with approximately 40 million reads per sample, on the Illumina HiSeq X Ten platform. The adaptors sequence in the reads were removed using Cutadapt (v.1.18)[18] and quality controlled with FastQC (v0.11.8) to ensure complete removal of adaptors. The cleaned reads were then mapped to the human genome assembly GRCh38 (hg38) using Hisat2 (v2.1.0)[19] with gene annotation (v87) downloaded from the Ensembl website. The mapped reads were counted by htseq-count (v0.7.2)[20] with the option '−stranded no'. Differential analysis was performed using the R package DESeq2 (v1.34.0)[21] to define the SG or DDX6-proximal data sets.

Detailed data analysis process and smFISH validation protocols can be found in the Supporting Information.

RESULTS AND DISCUSSION

CAP-Seq Profiling of SG-Proximal Transcriptome with Multiple Baits. To comprehensively investigate RNAs associated with stress granules (SGs), we selected two additional core proteins, TIA1 and G3BP2, as baits for genetically targeting the photocatalyst miniSOG. Using lentiviral transduction, we generated HEK293T cell lines stably expressing either TIA1-miniSOG or G3BP2-miniSOG. Immunofluorescence microscopy confirmed their expected basal localization: TIA1-miniSOG was partially in the nucleus, whereas G3BP2- miniSOG was diffusely localized throughout the cytoplasm (Figure S1). Following treatment with 0.5 mM sodium arsenite for 1 h, both fusion proteins translocated into cytoplasmic SGs and colocalized with established SG markers. Notably, some targeting was incomplete: TIA1-miniSOG retained residual nuclear localization, and G3BP2-miniSOG remained partially diffused in the cytoplasm (Figure S1), consistent with previous reports.[15,22] This observation underscores the necessity of an untargeted miniSOG cell line as a reference to account for background signals from non-SG-localized miniSOG.

We performed biological replicates of CAP-seq labeling in cells expressing TIA1-miniSOG, G3BP2-miniSOG, or untargeted miniSOG. In each experiment, arsenite-stressed cells were incubated with 10 mM alkyne-bearing propargylamine (PA) for 5 min before blue light illumination at 24 W/cm^2 for 15 min (Figure 1A).[15,22,23] Immunofluorescence microscopy revealed strong colocalization between the alkyne-labeled signals (detected via click chemistry with biotin-azide and streptavidin-fluorophore staining) and miniSOG expression (Figures 1B and S2). After cell lysis, RNA was extracted, subjected to a click reaction with biotin-azide, and affinity purified prior to next-generation sequencing analysis. Enrichment was highly reproducible across replicates, with Pearson correlation coefficients exceeding 0.97 (Figure S3). A negative control without PA yielded only minimal RNA enrichment, indicating minor nonspecific binding to streptavidin beads and emphasizing the importance of this control in data analysis. Differential expression analysis using DESeq2 was conducted in three comparisons: (1) enriched (Enrich) versus input (Input) RNAs for TIA1-/G3BP2-miniSOG samples, (2) enriched RNAs versus the negative control without PA (Control), and (3) enriched RNAs versus those labeled with untargeted miniSOG (Untargeted) (Figure 1C).

To define the TIA1-proximal SG transcriptome, we applied thresholds of $\log_2$FoldChange ($\log_2$FC) > 0.3 and adjusted p-value <0.05. This filtering yielded 2162 transcripts enriched over input (Enrich vs Input), 2670 enriched over the PA-negative control (Enrich vs Control), and 1540 enriched relative to untargeted miniSOG (Enrich vs Untargeted). Intersecting these three data sets produced a stringent list of 933 TIA1-proximal transcripts (the TIA1 SG-proximal data

https://doi.org/10.1021/acs.analchem.5c01590
Anal. Chem. 2025, 97, 12767−12775

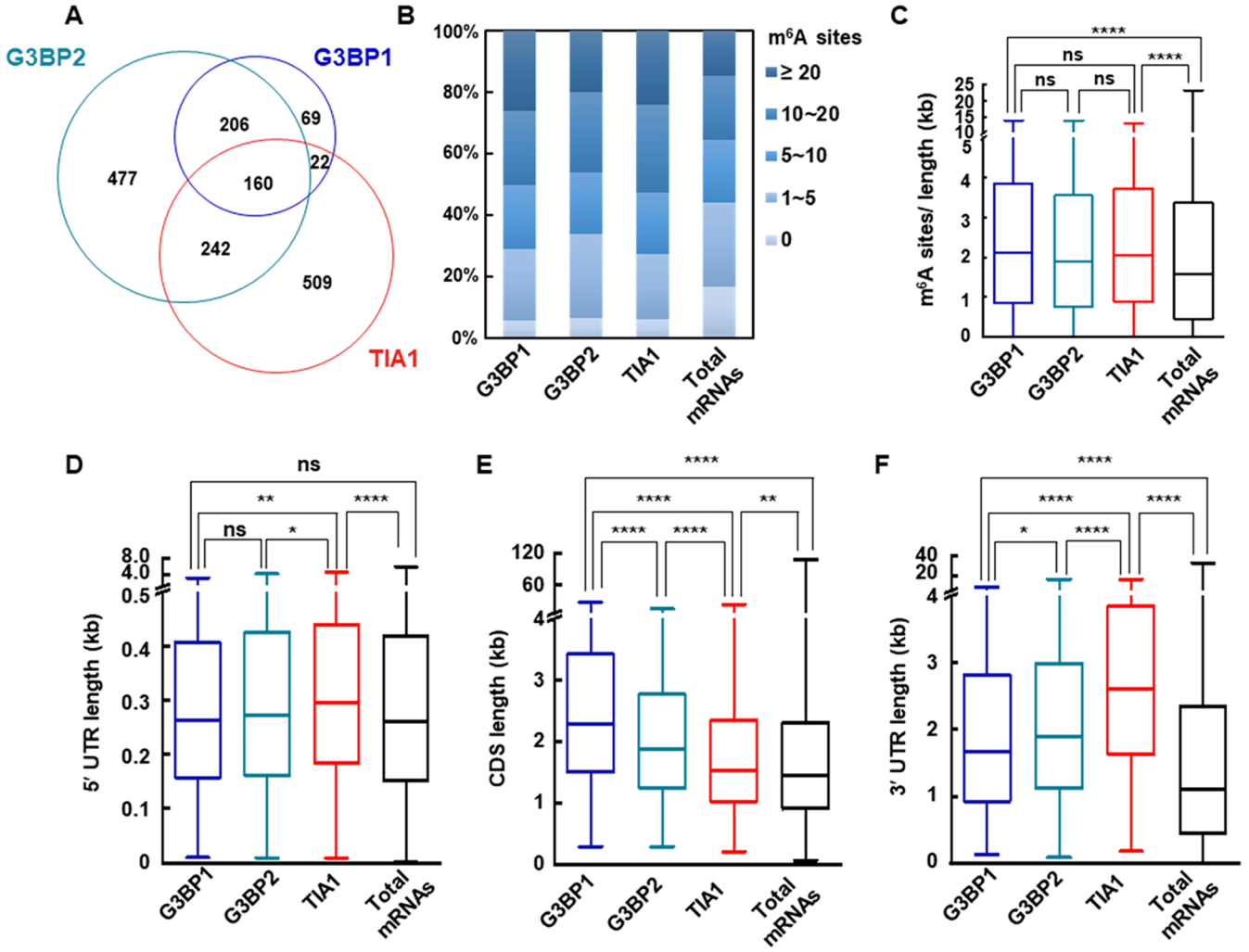


Figure 2. Comparison of RNAs proximal to different SG core proteins. (A) Venn diagram comparing SG-proximal RNAs captured by CAP-seq with G3BP1, TIA1, and G3BP2 as baits. (B) Distributions of m^{6}A sites[25] in SG-proximal RNAs and total mRNAs. (C) Comparison of m^{6}A densities (sites per kilobase) of mRNAs in SG-proximal data sets and total mRNAs. (D–F) Comparisons of 5′ UTR length (D), coding sequence (CDS) length (E), and 3′ UTR length (F) between SG-proximal mRNAs and total mRNAs. Boxes mark the first and third quartiles; the central line represents the median. Statistical significance was calculated with unpaired Mann–Whitney *U* test (two-sided). ns, not significant ($p > 0.05$); * $p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$.

set) (Figures 1D, S4 and Supporting Data 1). Similarly, for G3BP2-proximal RNAs, we identified 2339 (Enrich vs Input), 2130 (Enrich vs Control), and 2042 (Enrich vs Untargeted) transcripts, with 1085 overlapping across all comparisons to form the G3BP2 SG-proximal data set (Figures 1E, S4 and Supporting Data 1). Most of these enriched RNAs were mRNAs: 919 of 933 in the TIA1 data set and 1082 of 1085 in the G3BP2 data set, including transcripts such as *USP7*, *APLP2*, *GAS1*, and *BMS1*, which were previously validated by single-molecule fluorescence in situ hybridization (smFISH).[15] The remaining 3 non-mRNA transcripts in the G3BP2 data set are nonfunctional pseudogenes. In addition, the TIA1 data set contains two long noncoding RNAs (lncRNAs): *NORAD*, known for SG localization,[14] and *MALAT1*, a key component of nuclear speckles.[24]

To validate these data sets, we used smFISH imaging to examine the subcellular localization of selected RNA targets. We chose *SLC7A1*, which was specifically enriched in the TIA1 SG-proximal data set, and *C5orf30*, enriched in both the TIA1 and G3BP2 data sets. Using G3BP2 as a fluorescent marker, confocal imaging revealed significant SG enrichment of *SLC7A1* (52 ± 15%) and *C5orf30* (55 ± 15%) under arsenite stress, whereas *GNB1* was largely excluded from SGs (16 ± 7%) (Figures 1F and S5). These findings demonstrate that employing multiple bait proteins in CAP-seq effectively uncovers additional SG-proximal RNA candidates.

The detection of the lncRNA *MALAT1* in the TIA1 SG-proximal data set prompted further investigation into its subcellular localization using smFISH. Although *MALAT1* is well known for stabilizing nuclear speckles, its association with SGs has not been previously reported. Upon sodium arsenite stimulation, we observed recruitment of *MALAT1* to SGs in both HEK293T and U-2 OS cells (Figure 1G). To quantify this nucleocytoplasmic redistribution, we performed tiled wide-field imaging at 20× magnification, capturing fluorescence images from 5000 to 20,000 cells (Figure S6). Statistical analysis revealed that approximately 5% of HEK293T cells and 2% of U-2 OS cells displayed *MALAT1* localization within SGs. Interestingly, *MALAT1* was observed to colocalize with SGs in adjacent cell pairs, suggesting that cell cycle stage may influence SG transcriptome composition. Notably, CAP-seq targeting G3BP1 and G3BP2 under arsenite stress did not

https://doi.org/10.1021/acs.analchem.5c01590
Anal. Chem. 2025, 97, 12767–12775

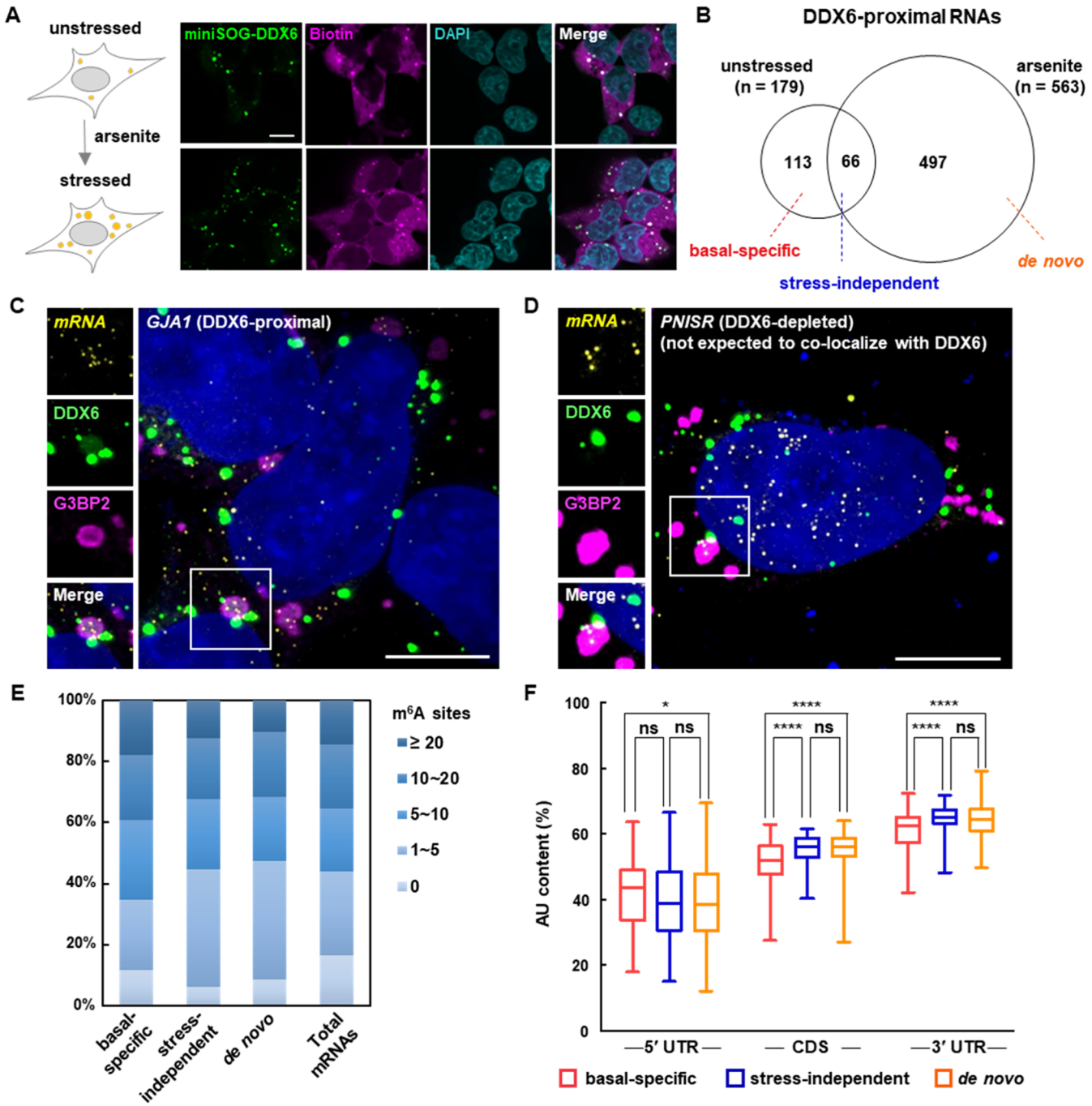


Figure 3. CAP-seq profiling of PB-proximal transcriptome in HEK293T cells. (A) Confocal images of HEK293T cells expressing miniSOG-DDX6 (green), before (top) and after (bottom) arsenite stress. Biotinylation signal is shown in magenta. (B) Venn diagram comparing PB-proximal RNAs under unstressed and arsenite-treated conditions. (C, D) smFISH images of enriched target *GJA1* (C) and depleted target *PNISR* (D) in HEK293T cells under arsenite stimulation. Zoom-in views of the boxed region are shown on the left. (E) Distributions of m^6A sites among basal-specific, stress-independent, and *de novo* PB-proximal RNAs. (F) Comparisons of AU content in the 5′ UTR, coding sequence (CDS), and 3′ UTR between PB-proximal mRNAs. Statistical significance was calculated with unpaired Mann−Whitney *U* test (two-sided). ns, not significant ($p > 0.05$); * $p < 0.05$; **** $p < 0.0001$. Scale bars: 10 μm.

show association of *MALAT1* with SGs, implying a preferential interaction between *MALAT1* and TIA1. Collectively, these findings underscore the utility of a multibait CAP-seq approach in unraveling the heterogeneous RNA composition within stress granules.

Comparison of RNAs Proximal to Different SG Core Proteins. In light of the heterogeneous model of the SG structure, an important question arises: How do CAP-seq data sets compare when different bait proteins are used? To answer this, we compared the RNA compositions of three SG-proximal data sets generated using G3BP1 (previously reported[15]), G3BP2, and TIA1 as bait proteins. Our analysis revealed that the G3BP1 and G3BP2 data sets were highly similar: 80% of the mRNAs in the G3BP1 data set were also found in the G3BP2 data set, whereas only 40% overlapped with the TIA1 data set (Figure 2A and Supporting Data 2).

https://doi.org/10.1021/acs.analchem.5c01590
Anal. Chem. 2025, 97, 12767−12775

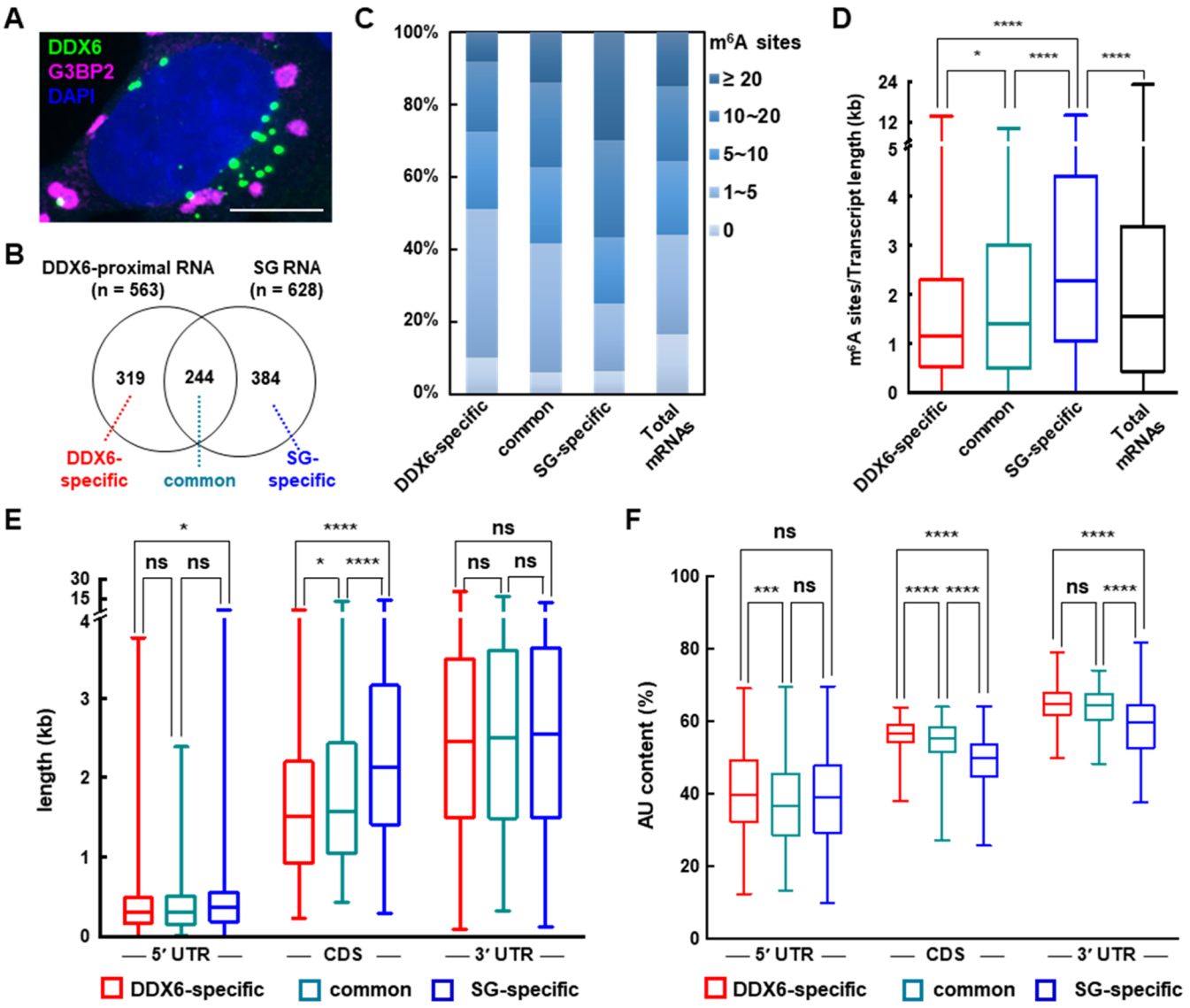


Figure 4. Comparison of DDX6- and SG-proximal RNAs in HEK293T cells under arsenite stress. (A) Representative immunofluorescence image of HEK293T cells under arsenite stress. PB marker DDX6 (green) and SG marker G3BP2 (magenta) are stained with antibodies. (B) Venn diagram comparing DDX6-proximal and SG CAP-seq RNAs captured in HEK293T cells under arsenite stress. (C) Distributions of m^6A sites among DDX6-specific mRNAs, SG-specific mRNAs, and mRNAs are common to both. (D) Comparison of the m^6A densities (sites per kilobase) of DDX6-specific and SG-specific data sets. (E, F) Comparisons of the lengths (E) and AU contents (F) of 5′ UTR, the coding sequence (CDS), and 3′ UTR between DDX6-specific mRNAs, SG-specific mRNAs, and mRNAs common to both. Statistical significance was calculated with unpaired Mann−Whitney *U* test (two-sided). ns, not significant ($p > 0.05$); * $p < 0.05$; *** $p < 0.001$; ****; $p < 0.0001$.

This close similarity likely reflects the ability of G3BP1 and G3BP2 to form heterodimers via their N-terminal domains, resulting in their closer colocalization within SGs compared to TIA1. Among the 182 RNAs detected in both G3BP1 and TIA1 SG data sets (i.e., the 40% overlapped RNA in Figure 2A), approximately one-third (62 out of 182, or 34%) belong to G3BP1 pre-existing RNAs (i.e., RNAs that are proximal to G3BP1 in both basal and stress conditions).[15] In contrast, for the 275 G3BP1-specific RNAs, a higher proportion (52%, or 149 of 275) overlapped with G3BP1 pre-existing RNAs. This difference suggests that the heterogeneous distribution of SG RNAs may be partially inherited from the RNA-protein interactions present under basal conditions.

Previous studies have shown a positive correlation between RNA methylation, particularly m^6A modification, and RNA enrichment in SGs.[26−28] m^6A-binding proteins, such as YTHDF1/2/3, are recruited to SGs during stress (e.g., sodium arsenite exposure).[27,29] Using the average number of m^6A sites per transcript as a proxy for methylation, we observed that RNAs in the G3BP1 (13.3 sites), G3BP2 (11.6 sites), and TIA1 (13.3 sites) data sets have consistently higher methylation levels compared to the total transcriptome (9.4 sites) (Figure 2B). Moreover, the m^6A site density, defined as the number of sites per kilobase, was significantly higher in the SG-proximal data sets than in the total mRNA pool (Figure 2C).

Additionally, RNAs associated with all three bait proteins tended to have longer coding sequences (CDS) and untranslated regions (UTRs) than total RNAs (Figure 2D), which is consistent with the known positive correlation between RNA length and methylation level.[28,30] Despite this overall similarity, differences in CDS (Figure 2E) and 3′ UTR

https://doi.org/10.1021/acs.analchem.5c01590
Anal. Chem. 2025, 97, 12767−12775

lengths (Figure 2F) among the data sets suggest that mRNA distribution within SGs may vary depending on the specific bait protein used. Notably, while translation efficiency did not differ significantly among RNAs associated with the three bait proteins, those proximal to G3BP1 and G3BP2 exhibited substantially lower translation efficiencies compared with the total mRNA pool (Figure S7). In summary, our findings indicate that RNAs proximal to different SG core proteins display distinct intrinsic features, reflecting the heterogeneity of SG composition and organization.

Profiling PB-Proximal Transcriptome in HEK293T via CAP-seq. To profile the transcriptome of processing bodies (PBs), we created a HEK293T cell line that stably expresses miniSOG fused to the N-terminus of PB protein DDX6. Using the same CAP-seq labeling protocol as for SG bait proteins, we treated cells with 10 mM PA and exposed them to blue light for 15 min. Immunofluorescence microscopy confirmed that the miniSOG-DDX6 fusion protein and its associated biotinylation signal colocalized with PB puncta under both basal conditions and following sodium arsenite-induced stress (0.5 mM for 1 h), with a notable increase in PB formation during stress (Figure 3A). Sequencing data from biological replicates demonstrated high reproducibility, with Pearson's correlation coefficients exceeding 0.98 for enriched RNAs (Figure S8).

In unstressed cells, DESeq2 analysis (using a cutoff of $\log_2$FC > 0.3 and adjusted p-value <0.05) identified 3817 RNAs in the Enrich vs Input comparison, 3231 in the Enrich vs Control comparison, and 228 in the Enrich vs Untargeted comparison. Because DDX6 is only partially localized in PBs under basal conditions,[31] including an untargeted miniSOG control was essential to subtract nonspecific cytoplasmic labeling. The lower RNA count in the Enrich vs Untargeted comparison underscores the importance of this control. By intersecting these three data sets, we defined a basal PB-proximal transcriptome of 179 RNAs, including the lncRNA *NORAD* ($\log_2$FC = 0.54 in Enrich vs Untargeted). A similar approach identified 563 DDX6-proximal RNAs under stress (Figure S9 and Supporting Data 1).

Combining the basal and stressed DDX6-proximal transcriptomes yielded a total of 676 RNAs. Of these, 479 RNAs (71%) were recruited *de novo* to DDX6 under stress, 66 RNAs (10%) were consistently present under both conditions (stress-independent), and 113 RNAs (19%) were basal-specific. This shift in the RNA composition highlights the dynamic nature of DDX6-proximal RNAs upon stress induction (Figure 3B). For instance, smFISH imaging revealed that upon sodium arsenite stimulation, the DDX6-enriched target *GJA1* colocalized with both PBs and SGs (Figures 3C and S11). This overlap is expected because DDX6 signal partially coincides with SG markers, as noted in previous proteomic studies.[4,8,32] In contrast, the DDX6-depleted target *PNISR* was found in SGs but not in PBs (Figure 3D). *PNISR* demonstrated significantly weaker colocalization with DDX6 compared to *GJA1* (Figure S11), suggesting distinct RNA preferences between SG and PB marker proteins. These findings confirm the effectiveness of CAP-seq in profiling stress-related RNA candidates in PBs.

YTHDF2, an m^{6}A-binding protein, colocalizes with the PB marker protein DCP1A, indicating that m^{6}A-modified RNAs might be selectively recruited to PBs.[33] In our data set, RNAs localized to PBs under basal conditions showed slightly higher m^{6}A modification levels (9.7 sites per transcript) compared to the total transcriptome (9.4) (Figure 3E). Under arsenite stress, however, *de novo*-recruited PB-proximal RNAs had lower m^{6}A levels (7.90) than basal-specific RNAs (10.2), which contrasts with the trend observed in SGs, where m^{6}A levels generally increase under stress.[15] This suggests that m^{6}A modifications may play a less prominent role in RNA recruitment to PBs than to SGs.

Comparison of the DDX6-proximal transcriptomes under basal and stressed conditions revealed no significant differences in RNA lengths, including CDS and UTRs (Figure S10), suggesting that the RNA length does not influence stress-induced PB RNA recruitment. However, stress-recruited mRNAs displayed significantly higher AU content in both CDS and 3′ UTR compared to the basal PB transcriptome (Figure 3F), indicating that AU-rich elements may facilitate stress-induced RNA binding by DDX6.

Comparison of DDX6 and SG-Proximal Transcriptomes under Arsenite Stress. Under arsenite stimulation, DDX6 was observed adjacent to the SGs in the cytoplasm (Figure 4A). Given this proximity, we compared our CAP-seq data sets for SGs and DDX6. We defined SG-associated RNAs (SG RNAs) as those identified by at least two SG core proteins under stress, yielding a data set of 628 RNAs. Of these, 244 overlapped with the DDX6-proximal RNA data set under stress (defined as "common") (Figure 4B). SG-specific RNAs exhibited 58% higher levels of m^{6}A modification than DDX6-specific (i.e., PB-specific) RNAs (12.7 vs 8.0 sites per transcript) (Figure 4C) and a significantly greater m^{6}A density (Figure 4D and Supporting Data 2). This finding suggests that m^{6}A modifications play a more prominent role in the directing of RNAs to SGs.

In addition, SG-specific RNAs tended to be longer than DDX6-specific RNAs (Figure S12), particularly in the CDS (Figure 4E). Although there was no significant difference in translation efficiencies between the two data sets (Figure S12), DDX6-specific RNAs showed significantly higher AU content in both CDS and 3′ UTR regions compared to SG-specific RNAs (Figure 4F). This indicates that AU-rich elements may facilitate RNA targeting of DDX6 under stress. Together, these analyses suggest that while m^{6}A modifications may preferentially drive RNA localization to SGs, AU-rich sequence features could be key for targeting mRNAs proximal to DDX6. These findings highlight potential regulatory roles for m^{6}A and AU content in determining mRNA distribution within membraneless organelles.

CONCLUSIONS

In this study, we used TIA1 and G3BP2 as baits to explore the SG transcriptome. Together with our previous G3BP1 data set,[15] these results provide a more comprehensive analysis of SG structure and composition. Under basal conditions (i.e., without sodium arsenite), TIA1 is partly localized in the nucleus. However, upon sodium arsenite stimulation, complexes of intracellular mRNA and translation initiation factors rapidly recruit both G3BP1 and TIA1, triggering SG assembly. Although G3BP1 and G3BP2 share 80% sequence similarity, they perform distinct functions.[34−37] For instance, SG formation is disrupted only when both G3BP1 and G3BP2 are knocked out.[3] Moreover, G3BP1 influences mTORC1 phosphorylation, whereas G3BP2 does not.[35] In the Wnt3a signaling pathway, G3BP2 acts as a positive regulator while G3BP1 functions as a negative regulator.[37] Additionally, G3BP2 contains a higher proportion of arginine residues in its RNA-binding domain and is longer than G3BP1, potentially

https://doi.org/10.1021/acs.analchem.5c01590
Anal. Chem. 2025, 97, 12767−12775

enhancing its capacity for liquid−liquid phase separation and binding to a broader range of RNAs.

It is important to note that our transcriptomic profiling captures RNAs that are proximal to a specific SG core protein rather than those directly interacting with the bait. The labeling radius of CAP-seq is primarily determined by the diffusion distance of singlet oxygen, which is about 70 nm in aqueous solution[38] and much larger than a typical protein (4−5 nm). Thus, RNAs indirectly associated with the core via adaptor proteins can also be labeled. For example, our data show that m^{6}A levels are significantly higher in the G3BP1/2 SG transcriptomes compared to the total transcriptome, a finding that might seem at odds with previous reports suggesting that G3BPs preferentially bind nonmethylated RNA.[39] One possible explanation is that G3BPs recruit m^{6}A-modified RNAs indirectly through interactions with m^{6}A-binding proteins. Further investigation into the dynamic changes in the SG proteome during assembly is needed to validate this hypothesis. Recent developments in proximity labeling methods with smaller labeling radii, such as μMAP (~4 nm),[40] might provide more precise information on direct RNA-protein interactions and offer a multiscale perspective on SG structure. Additionally, it is important to consider that only a subpopulation of the bait protein localizes to the SG. Therefore, employing an untargeted miniSOG can help subtract background labeling from the cytoplasmic bait population. Consequently, the cutoff ratio for defining SG RNAs should be carefully determined to ensure an accurate interpretation of the data.

Our results also indicate that the differences in m^{6}A modification levels between SGs and DDX6-proximal RNAs may arise from the distinct distributions of m^{6}A-binding proteins and chaperones associated with these membrane-less organelle core proteins. The localization of m^{6}A-binding proteins, such as YTHDF1/2/3, can influence where m^{6}A-modified RNAs accumulate.[26,27] The interplay between YTHDF proteins and other SG components likely causes mRNAs with higher m^{6}A modification levels to preferentially localize to SGs under stress.

Interestingly, our CAP-seq data reveal that lncRNA *MALAT1* is proximal to TIA1 following sodium arsenite stimulation. Although previous PAR-CLIP studies have shown interactions between *MALAT1* and SG proteins like G3BP1/2 and TIA1,[41,42] *MALAT1* is not typically considered a core SG component. We confirmed its subcellular localization with smFISH imaging, which showed *MALAT1* associated with SGs in approximately 5% of cells. This phenomenon was also observed in U-2 OS cells under different stress conditions (e.g., sorbitol treatment). Although the proportion is low, it is notable that cells with SG-associated *MALAT1* often appear in pairs, reminiscent of daughter cells shortly after mitosis. A recent study found that UV-induced DHX9 SGs are highly correlated with the cell cycle,[43] suggesting that stress-induced nuclear export and SG recruitment of *MALAT1* may be linked to specific mitotic stages. Further investigation is warranted to explore this potential correlation.

Future studies could expand upon our findings by mapping SG RNAs across various stress conditions and cell types. In our previous research, we analyzed G3BP1-associated RNAs under two distinct stressors: arsenite-induced oxidative stress and sorbitol-induced osmotic stress.[15] Notably, mRNAs uniquely enriched in SGs during sorbitol stress exhibited significantly shorter CDS, longer 3′ UTR, and fewer m^{6}A sites. To further elucidate stress-specific SG RNA profiles, future investigations should employ additional bait proteins beyond G3BP1. Similarly, most transcriptome-wide SG RNA studies are conducted in common immortalized cell lines, such as HEK293T, HeLa, and U-2 OS, under artificial stress conditions like heat shock, oxidative stress, or osmotic stress. Expanding research to include diverse cell lines or primary cells subjected to physiological stresses (e.g., amino acid deprivation) or pathological mutations could uncover universal features of SG components and facilitate the translation of omics findings into clinical applications.

■ ASSOCIATED CONTENT

Data Availability Statement

The sequencing data reported in this paper have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus (accession code: GSE271045).

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.5c01590.

Details in reagents information; detailed experimental methods and supporting figures on immunofluorescences; smFISH and sequencing results (PDF)

All DESeq2 results (XLSX)

Feature analysis summary of all transcripts of interest (XLSX)

Sequence of primers used in smFISH validation (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Peng Zou − *College of Chemistry and Molecular Engineering, Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, PKU-IDG/McGovern Institute for Brain Research, Beijing Advanced Center of RNA Biology (BEACON), Peking University, Beijing 100871, China; Academy for Advanced Interdisciplinary Studies, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China; Chinese Institute for Brain Research (CIBR), Beijing 102206, China;* orcid.org/0000-0002-9798-5242; Email: zoupeng@pku.edu.cn

Authors

Ziqi Ren − *College of Chemistry and Molecular Engineering, Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, PKU-IDG/McGovern Institute for Brain Research, Beijing Advanced Center of RNA Biology (BEACON), Peking University, Beijing 100871, China*

Songrui Zhao − *Academy for Advanced Interdisciplinary Studies, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China*

Wei Tang − *Academy for Advanced Interdisciplinary Studies, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.5c01590

Author Contributions

Z.R., S.Z., T.W., and P.Z. designed experiments; Z.R., S.Z., and T.W. performed experiments; Z.R. and P.Z. analyzed data; Z.R., S. Z., and P.Z. wrote the paper with inputs from all authors.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Ministry of Science and Technology (2022YFA1304700), the National Natural Science Foundation of China (32088101), and Beijing National Laboratory for Molecular Sciences (BNLMS-CXXM-202403). P.Z. is sponsored by Bayer Investigator Award. National Center for Protein Sciences at Peking University in Beijing, China, for assistance with Fragment Analyzer 12.

■ REFERENCES

(1) Protter, D. S. W.; Parker, R. *Trends Cell Biol.* **2016**, *26* (9), 668−679.
(2) Panas, M. D.; Ivanov, P.; Anderson, P. *J. Cell Biol.* **2016**, *215* (3), 313−323.
(3) Yang, P.; Mathieu, C.; Kolaitis, R. M.; Zhang, P.; et al. *Cell* **2020**, *181* (2), 325−345.
(4) Jain, S.; Wheeler, J. R.; Walters, R. W.; Agrawal, A.; et al. *Cell* **2016**, *164* (3), 487−498.
(5) Souquere, S.; Mollet, S.; Kress, M.; Dautry, F.; et al. *J. Cell Sci.* **2009**, *122* (20), 3619−3626.
(6) Standart, N.; Weil, D. *Trends Genet.* **2018**, *34* (8), 612−626.
(7) Luo, Y.; Na, Z.; Slavoff, S. A. *Biochemistry* **2018**, *57* (17), 2424−2431.
(8) Marmor-Kollet, H.; Siany, A.; Kedersha, N.; Knafo, N.; et al. *Mol. Cell* **2020**, *80* (5), 876−891.
(9) Tian, S.; Curnutte, H. A.; Trcek, T. *Molecules* **2020**, *25* (14), No. 3130.
(10) Wilbertz, J. H.; Voigt, F.; Horvathova, I.; Roth, G.; et al. *Mol. Cell* **2019**, *73* (5), 946−958.e7.
(11) Campos-Melo, D.; Hawley, Z. C. E.; Droppelmann, C. A.; Strong, M. J. *Front. Cell Dev. Biol.* **2021**, *9*, No. 621779.
(12) Parker, D. M.; Tauber, D.; Parker, R. *Mol. Cell* **2025**, *85* (3), 571−584.e7.
(13) Matheny, T.; Rao, B. S.; Parker, R. *Mol. Cell. Biol.* **2019**, *39*, No. e00313.
(14) Khong, A.; Matheny, T.; Jain, S.; Mitchell, S. F.; et al. *Mol. Cell* **2017**, *68* (4), 808−820.e5.
(15) Ren, Z.; Tang, W.; Peng, L.; Zou, P. *Nat. Commun.* **2023**, *14* (1), No. 7390.
(16) Padrón, A.; Iwasaki, S.; Ingolia, N. T. *Mol. Cell* **2019**, *75* (4), 875−887.e5.
(17) Wang, P.; Tang, W.; Li, Z.; Zou, Z.; et al. *Nat. Chem. Biol.* **2019**, *15* (11), 1110−1119.
(18) Martin, M.Cutadapt removes adapter sequences from high-throughput sequencing reads *EMBnet. j.* 2011; Vol. *17* .
(19) Kim, D.; Langmead, B.; Salzberg, S. L. *Nat. Methods* **2015**, *12* (4), 357−360.
(20) Anders, S.; Pyl, P. T.; Huber, W. *Bioinformatics* **2015**, *31* (2), 166−169.
(21) Love, M. I.; Huber, W.; Anders, S. *Genome Biol.* **2014**, *15* (12), No. 550.
(22) Wheeler, J. R.; Jain, S.; Khong, A.; Parker, R. *Methods* **2017**, *126*, 12−17.
(23) Seo, K. W.; Kleiner, R. E. *Nat. Chem. Biol.* **2023**, *19*, 1361−1371.
(24) Hutchinson, J. N.; Ensminger, A. W.; Clemson, C. M.; Lynch, C. R.; et al. *Bmc Genomics* **2007**, *8*, No. 39.
(25) Liu, C.; Sun, H.; Yi, Y.; Shen, W.; et al. *Nat. Biotechnol.* **2023**, *41* (3), 355−366.
(26) Ries, R. J.; Zaccara, S.; Klein, P.; Olarerin-George, A.; et al. *Nature* **2019**, *571* (7765), 424−428.
(27) Fu, Y.; Zhuang, X. *Nat. Chem. Biol.* **2020**, *16* (9), 955−963.
(28) Ries, R. J.; Pickering, B. F.; Poh, H. X.; Namkoong, S.; Jaffrey, S. R. *Nat. Struct. Mol. Biol.* **2023**, *30* (10), 1525−1535.
(29) Markmiller, S.; Soltanieh, S.; Server, K. L.; Mak, R.; et al. *Cell* **2018**, *172* (3), 590−604.
(30) Meyer, K. D.; Saletore, Y.; Zumbo, P.; Elemento, O.; et al. *Cell* **2012**, *149* (7), 1635−1646.
(31) Huang, J.-H.; Ku, W.-C.; Chen, Y.-C.; Chang, Y.-L.; Chu, C. Y. *Sci. Rep.* **2017**, *7* (1), No. 42853.
(32) Youn, J.-Y.; Dunham, W. H.; Hong, S. J.; Knight, J. D. R.; et al. *Mol. Cell* **2018**, *69* (3), 517−532.e11.
(33) Wang, X.; Lu, Z.; Gomez, A.; Hon, G. C.; et al. *Nature* **2014**, *505* (7481), 117−120.
(34) Sidibé, H.; Dubinski, A.; Velde, C. V. *J. Neurochem.* **2021**, *157* (4), 944−962.
(35) Prentzell, M. T.; Rehbein, U.; Sandoval, M. C.; De Meulemeester, A. S.; et al. *Cell* **2021**, *184* (3), 655−674.e27.
(36) Matsuki, H.; Takahashi, M.; Higuchi, M.; Makokha, G. N.; et al. *Genes Cells* **2013**, *18* (2), 135−146.
(37) Bikkavilli, R. K.; Malbon, C. C. *J. Cell Sci.* **2012**, *125* (10), 2446−2456.
(38) Makhijani, K.; To, T.-L.; Ruiz-González, R.; Lafaye, C.; et al. *Cell Chem. Biol.* **2017**, *24* (1), 110−119.
(39) Edupuganti, R. R.; Geiger, S.; Lindeboom, R. G. H.; Shi, H.; et al. *Nat. Struct. Mol. Biol.* **2017**, *24* (10), 870−878.
(40) Pan, C.; Knutson, S. D.; Huth, S. W.; MacMillan, D. W. C. *Nat. Chem. Biol.* **2025**, *21* (4), 490−500.
(41) Meyer, C.; Garzia, A.; Mazzola, M.; Gerstberger, S.; et al. *Mol. Cell* **2018**, *69* (4), 622−635.e6.
(42) Edupuganti, R. R.; Geiger, S.; Lindeboom, R. G. H.; Shi, H.; et al. *Nat. Struct. Mol. Biol.* **2017**, *24* (10), 870−878.
(43) Zhou, Y.; Panhale, A.; Shvedunova, M.; Balan, M.; et al. *Cell* **2024**, *187* (7), 1701−1718.e28.

https://doi.org/10.1021/acs.analchem.5c01590
Anal. Chem. 2025, 97, 12767−12775