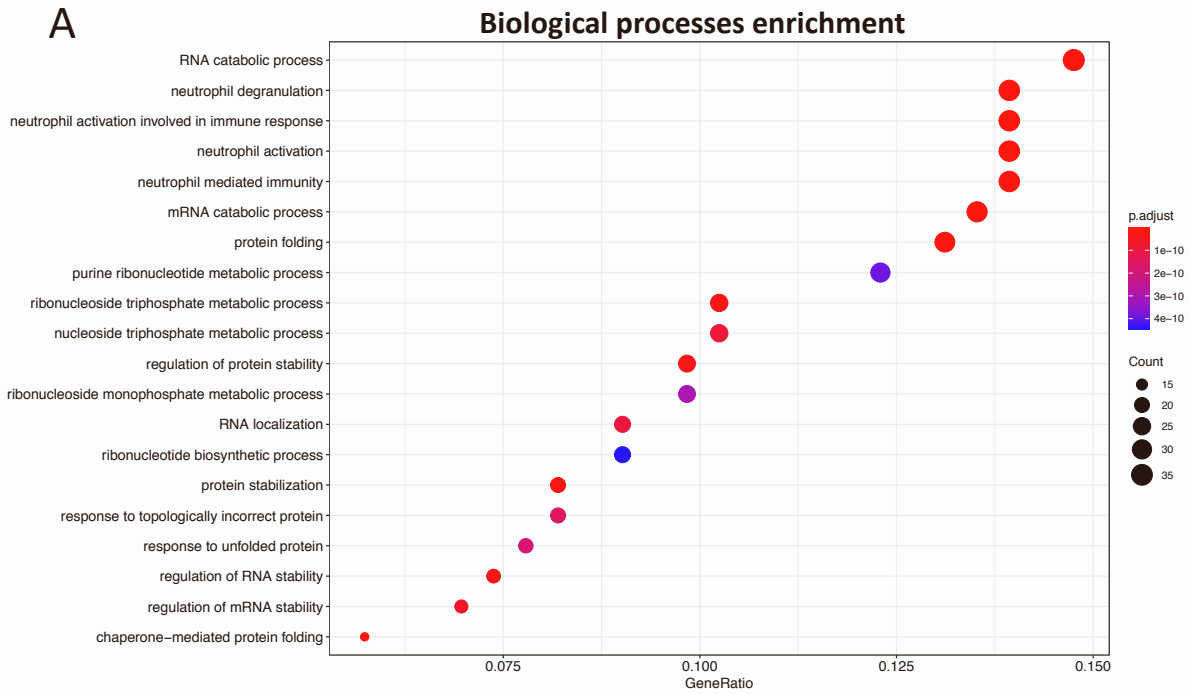
**iScience, Volume *25***

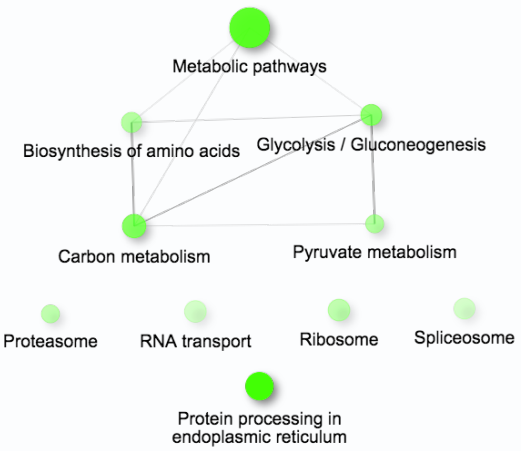
## Supplemental information

# SARS-CoV-2 nucleocapsid protein binds host mRNAs and attenuates stress granules to impair host stress response

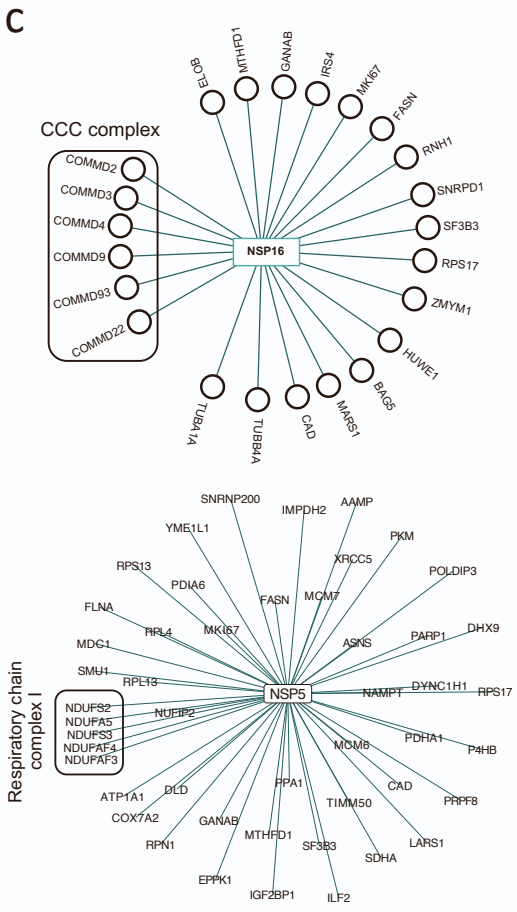
**Syed Nabeel-Shah, Hyunmin Lee, Nujhat Ahmed, Giovanni L. Burke, Shaghayegh Farhangmehr, Kanwal Ashraf, Shuye Pu, Ulrich Braunschweig, Guoqing Zhong, Hong Wei, Hua Tang, Jianyi Yang, Edyta Marcon, Benjamin J. Blencowe, Zhaolei Zhang, and Jack F. Greenblatt**

Figure S1




FDR <0.05




FDR <0.01

**Figure S1: SARS-COV-2 protein interactions, related to Figure 1. A:** GO enrichment analysis related to biological processes using cellular proteins identified as high confidence (FDR≤1%) interaction partners for 27 SARS-CoV-2 proteins. **B**: KEGG pathway enrichment analysis using cellular proteins identified as high confidence (FDR≤1%) interaction partners for the SARS-CoV-2 proteins. Darker nodes are more significantly enriched gene sets. Bigger nodes represent larger gene sets. Edge thickness represents overlap of genes. **C:** Network illustration of protein-protein interactions for selected SARS-CoV-2 bait proteins. Co-purifying subunits that belong to a specific protein-complex are grouped together.

Figure S2

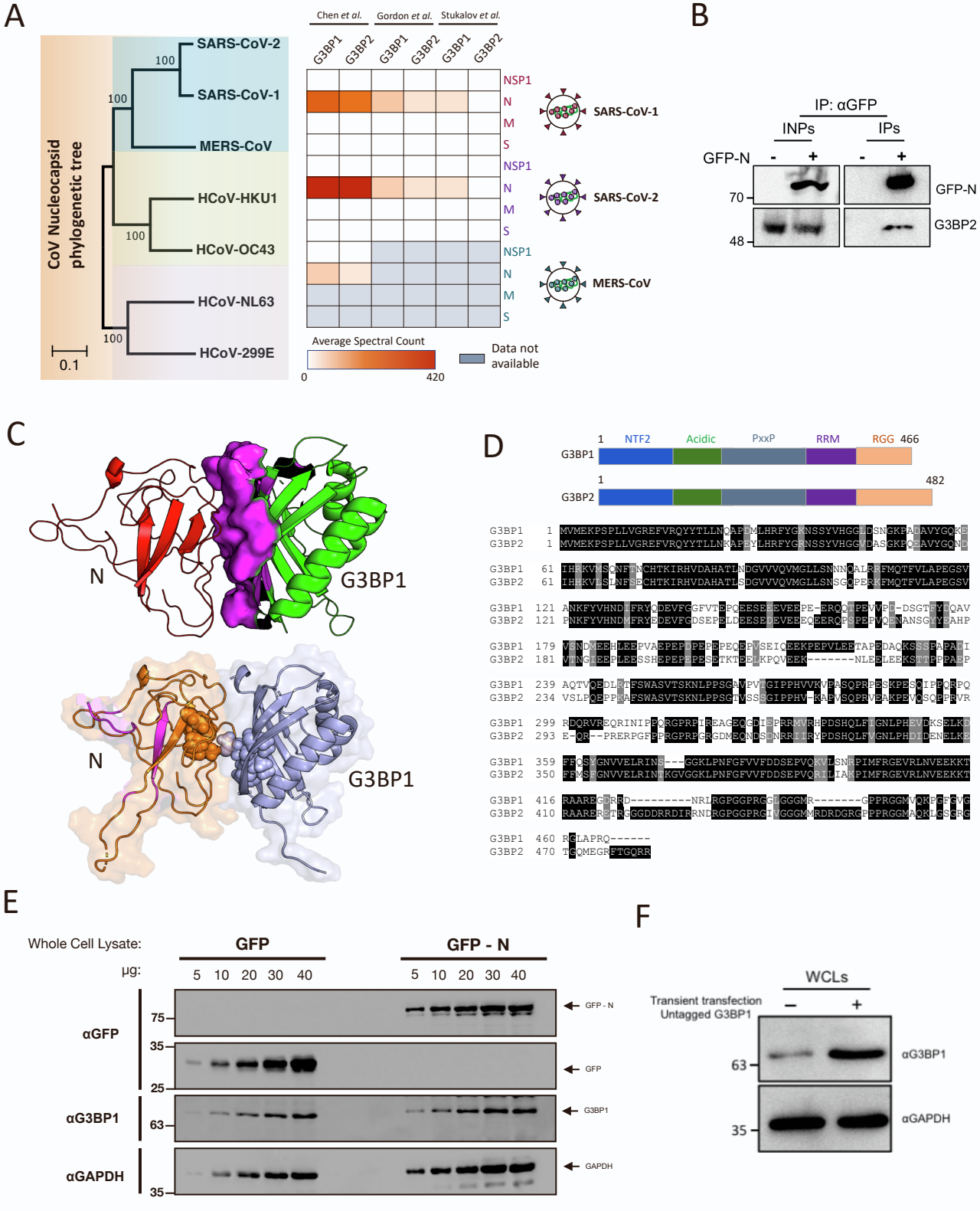

A
CoV Nucleocapsid phylogenetic tree
SARS-CoV-2
SARS-CoV-1
MERS-CoV
HCoV-HKU1
HCoV-OC43
HCoV-NL63
HCoV-299E
100
100
100
100
0.1
Chen et al.
Gordon et al.
Stukalov et al.
G3BP1
G3BP2
NSP1
N
M
S
SARS-CoV-1
SARS-CoV-2
MERS-CoV
Average Spectral Count
0
420
Data not available
B
IP: αGFP
INPs
IPs
GFP-N
- + - +
70
48
GFP-N
G3BP2
C
N
G3BP1
N
G3BP1
D
1
NTF2
Acidic
PxxP
RRM
RGG
466
G3BP1
1
482
G3BP2
E
Whole Cell Lysate:
GFP
GFP - N
μg:
5 10 20 30 40
5 10 20 30 40
αGFP
αG3BP1
αGAPDH
75
35
25
63
35
GFP - N
GFP
G3BP1
GAPDH
F
WCLs
Transient transfection
Untagged G3BP1
– +
63
35
αG3BP1
αGAPDH

**Figure S2: N specifically interacts with G3BP1 and G3BP2, related to Figures 1 and 2. A:** Left, Protein phylogeny of N across indicated coronaviruses. Different clusters are highlighted in colors. Tree topology represents the Neighbour Joining (NJ) estimations. Confidence values are indicated based on 1,000 bootstrap replicas. The scale bar indicates the number of substitutions per site. Right, Heatmap illustration of SARS-CoV-2 proteins' interaction with stress granule proteins G3BP1 and G3BP2 across three different studies as indicated; (Gordon et al. 2020 [PMID: 33060197]; Stukalov et al. 2021 [PMID: 33845483]; Chen et al. 2021 [PMID: 34232536]). **B:** Interaction of SARS-CoV-2 N protein with G3BP2 detected by co-immunoprecipitation experiments. IPs were performed with GFP antibody using whole cell lysates prepared from control HEK293 cells or cells expressing GFP-N. Cell lysates were treated with nuclease prior to IPs. Blots were probed with the indicated antibodies. INPs and IPs had different exposure time. **C: Top**: Cartoon illustration of protein docking studies. The complex structure model between G3BP1 (4FCJ_A, green) and N protein (6M3M_B, red) is shown. The interface areas are shown as a purple surface. A residue is defined as an interface residue if it is within 6 Å of its binding partner. The interface area of G3BP1-N was 780.73 $A^2$. The interface energy of the final model was -5.46, which indicates a reliable model according to the Rosetta document. **Bottom**: Alternative view of the complex structure model between G3BP1 (light blue cartoon) and the N protein (orange cartoon) is shown. The orange spheres and the light blue spheres are the predicted interface residues for N and G3BP1, respectively. The reported theoretical RNA-binding residues [PMID: 33264373] on N are highlighted in magenta. **D:** Top: Domain architecture of G3BP1 and G3BP2 proteins. Bottom: Multiple sequence alignment analysis using G3BP1 and G3BP2 protein sequences. Highly conserved residues are shaded in black. **E:** Western blotting analysis using whole cell extracts prepared from either GFP- or GFP-N-expressing cells. Blots were probed with indicated antibodies. **F:** Western blotting analysis using whole cell extracts prepared from cells transiently transfected with either G3BP1 or empty vector. Blots were probed with indicated antibodies.

Figure S3

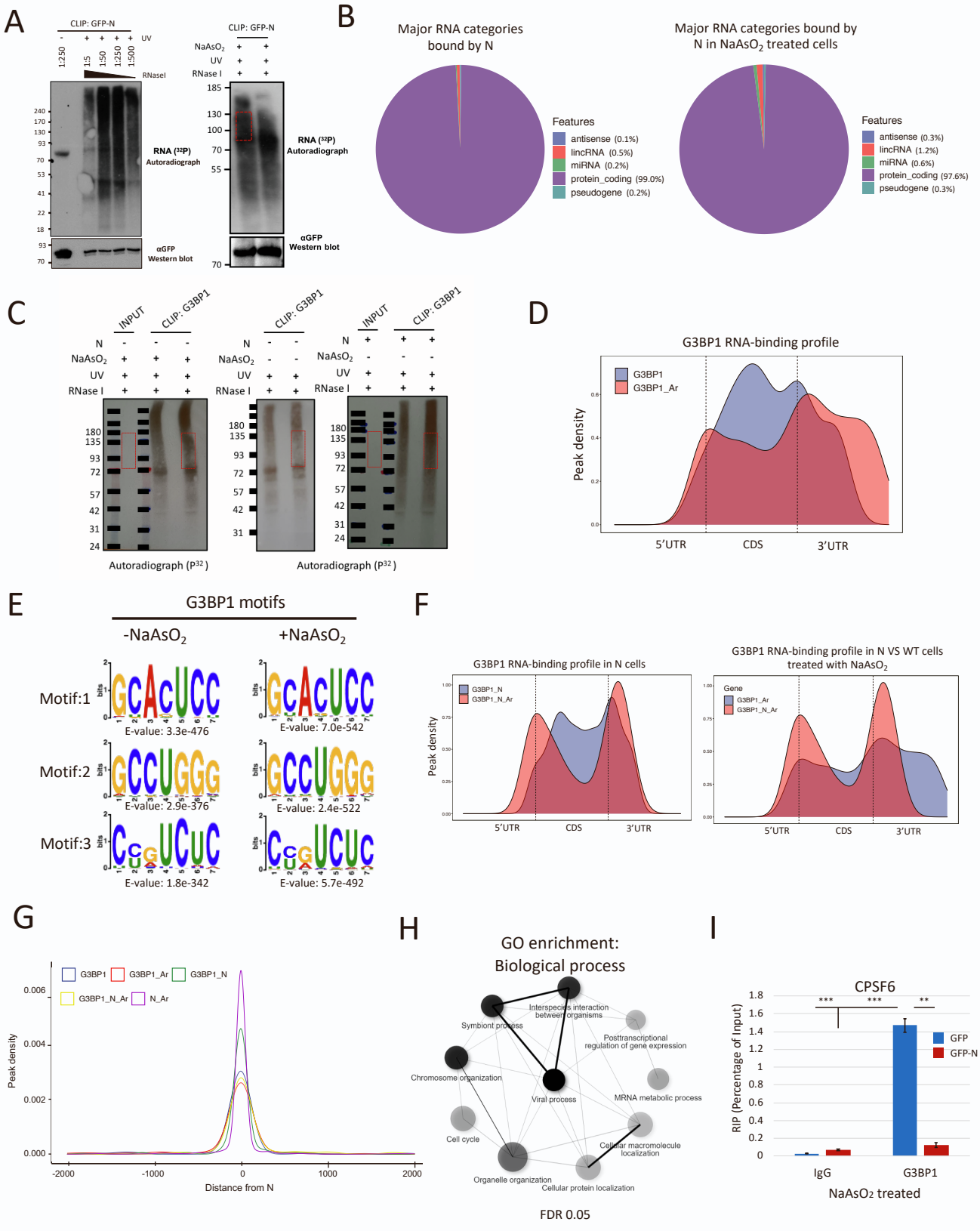
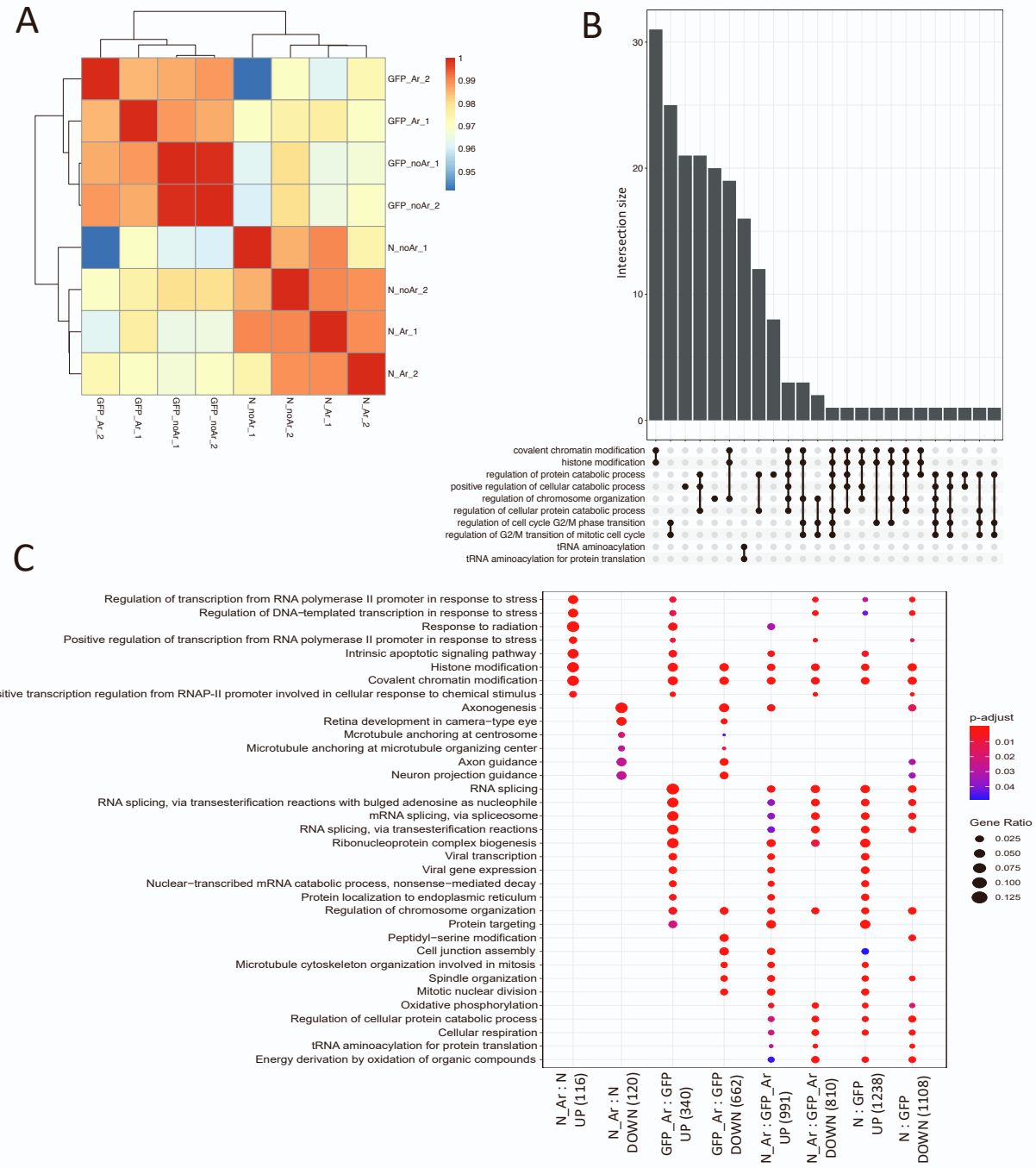

A
CLIP: GFP-N
UV
RNaseI
1:250
1:5
1:50
1:250
1:500
RNA (32P) Autoradiograph
αGFP Western blot
NaAsO2
RNase I
B
Major RNA categories bound by N
Major RNA categories bound by N in NaAsO2 treated cells
Features
antisense (0.1%)
lincRNA (0.5%)
miRNA (0.2%)
protein_coding (99.0%)
pseudogene (0.2%)
antisense (0.3%)
lincRNA (1.2%)
miRNA (0.6%)
protein_coding (97.6%)
pseudogene (0.3%)
C
INPUT
CLIP: G3BP1
N
NaAsO2
UV
RNase I
Autoradiograph (P32)
D
G3BP1 RNA-binding profile
G3BP1
G3BP1_Ar
Peak density
5'UTR
CDS
3'UTR
E
G3BP1 motifs
-NaAsO2
+NaAsO2
Motif:1
Motif:2
Motif:3
E-value: 3.3e-476
E-value: 7.0e-542
E-value: 2.9e-376
E-value: 2.4e-522
E-value: 1.8e-342
E-value: 5.7e-492
F
G3BP1 RNA-binding profile in N cells
G3BP1 RNA-binding profile in N VS WT cells treated with NaAsO2
G3BP1_N
G3BP1_N_Ar
Gene
G
G3BP1
G3BP1_Ar
G3BP1_N
G3BP1_N_Ar
N_Ar
Peak density
Distance from N
H
GO enrichment:
Biological process
Interspecies interaction between organisms
Symbiont process
Posttranscriptional regulation of gene expression
Chromosome organization
Viral process
MRNA metabolic process
Cell cycle
Cellular macromolecule localization
Organelle organization
Cellular protein localization
FDR 0.05
I
CPSF6
GFP
GFP-N
RIP (Percentage of Input)
IgG
G3BP1
NaAsO2 treated

**Figure S3: iCLIP analysis of N and G3BP1, related to Figures 3 and 4. A:** Left, Autoradiographs of immunopurified $^{32}$P-labeled N-RNA complexes after partial RNase I digestion at various dilutions. N-expressing HEK293 cells were UV-crosslinked and GFP-N was immunoprecipitated using anti-GFP antibody. Purified RNA-protein complexes were resolved on 4-12% Bis-Tris gels after radiolabeling the RNA and transferred to nitrocellulose membranes. Band in non-UV lane indicates radiolabeled N. Right, $NaAsO_2$-treated N-expressing HEK293 cells were used for CLIP-autoradiography. Red box indicates the region of membrane used to extract RNA **B:** Pie chart representing the distribution of N iCLIP peaks across various RNA categories in non-$NaAsO_2$- and $NaAsO_2$-treated cells. Legends are provided. **C:** Autoradiographs of immunopurified $^{32}$P-labeled G3BP1-RNA complexes after partial RNase I digestion. 2% input samples were loaded in parallel to CLIP samples. CLIP was performed using anti-G3BP1 antibody. Red boxes indicate the regions of membrane used to extract RNA. **D:** Standardized metaplot profile showing the normalized peak density of G3BP1 iCLIP in non-$NaAsO_2$- and $NaAsO_2$-treated wildtype cells. CDS represents the coding sequence. **E:** Enriched sequence motifs found in G3BP1 iCLIP-seq peaks. E-value represents the significance of the motif against randomly assorted control sequences. **F:** Left: Standardized metaplot profile showing the normalized peak density of G3BP1 iCLIP in non-$NaAsO_2$- and $NaAsO_2$-treated N-expressing cells. Right: Standardized metaplot profile showing the normalized peak density of G3BP1 iCLIP in $NaAsO_2$-treated N-expressing and wildtype cells. CDS represents the coding sequence. **G:** Average iCLIP peak density of G3BP1 around N-binding sites. iCLIP-seq samples across different conditions are color coded and legend is provided. **H:** GO enrichment analysis related to biological processes using N-specific G3BP1 target genes ($Q \leq 0.05$). Darker nodes are more significantly enriched gene sets. Bigger nodes represent larger gene sets. Edge thickness represents overlap of genes. **I:** Bar graph representation of G3BP1 RIP-qPCR experiments in $NaAsO_2$-treated cells expressing either GFP-N or GFP showing the G3BP1 RNA-binding to *CPSF6* transcript. The experiments were performed in biological triplicates, and p-values were calculated using the student's t-test ($^{***}p \leq 0.001$, $^{**}p \leq 0.01$). Error bars represent standard error of mean (SEM).

Figure S4




**Supplementary Figure 4: RNA-seq analysis of N-expressing cells, related to Figure 5.** **A:** Heatmap illustration of correlation between RNA-seq biological replicates. Pearson correlation of co-efficient values for the heatmap are shown on the right. **B:** GO enrichment analysis using genes that were significantly downregulated in cells expressing GFP-N in comparison to cells expressing GFP. Only the top 10 most significant terms related to biological processes are shown (Q≤0.05). **C:** GO enrichment analysis using genes that were significantly differentially expressed across the indicated comparisons. p-values and gene counts for each GO term are provided.

Figure S5

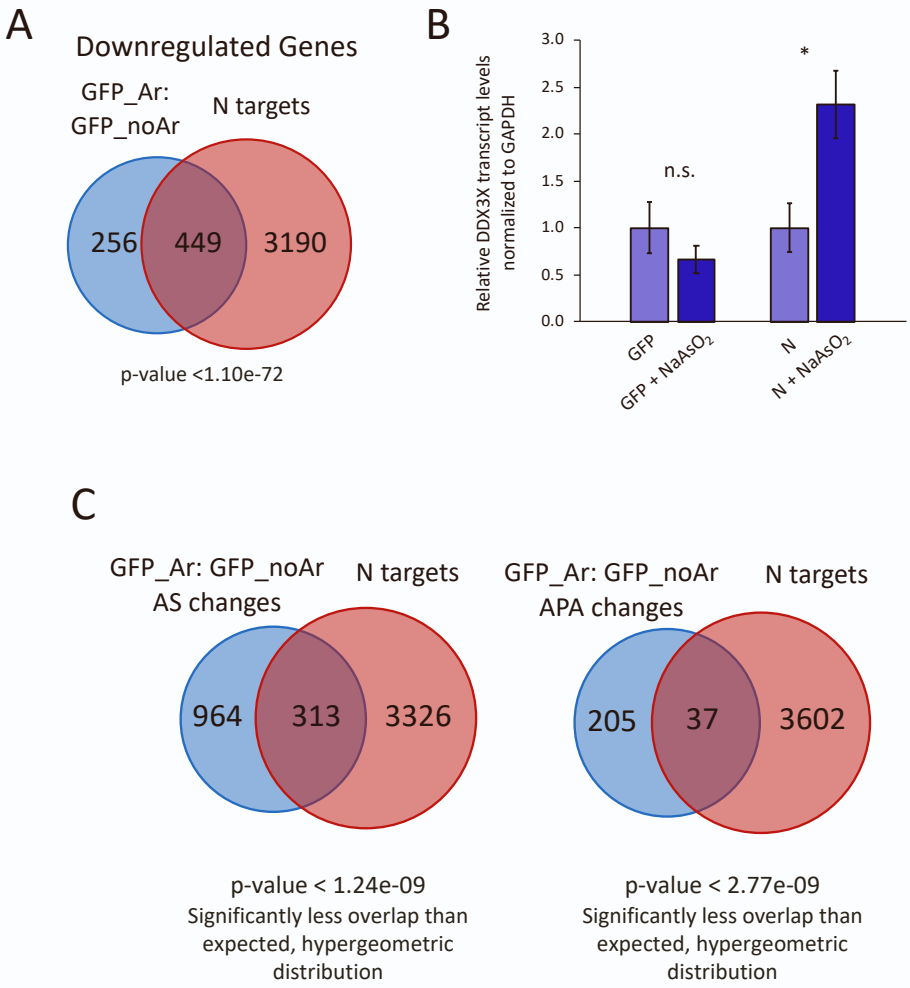



**Figure S5: N reverses posttranscriptional changes induced in response to stress, related to Figure 6. A:** Venn diagram showing the overlap of stress induced differential genes in wildtype GFP-expressing cells with those identified as N targets through iCLIP-seq. P-values were calculated using hypergeometric test. **B:** Bar graphs representing RT-qPCR results to examine the differential expression of selected targets in $NaAsO_2$-treated N-expressing and wildtype cells. The experiments were performed in biological triplicates, and p-value were calculated using the student's t-test ($^{***}p \leq 0.001$, $^{**}p \leq 0.01$, $^{*}p \leq 0.05$, n.s.: non-significant). Error bars represent SEM. **C:** Venn diagram showing the overlap of stress induced AS (left) and APA (right) changing genes with those identified as N targets. P-values indicate the significant underrepresentation of AS and APA genes in N targets (hypergeometric distribution).

Figure S6

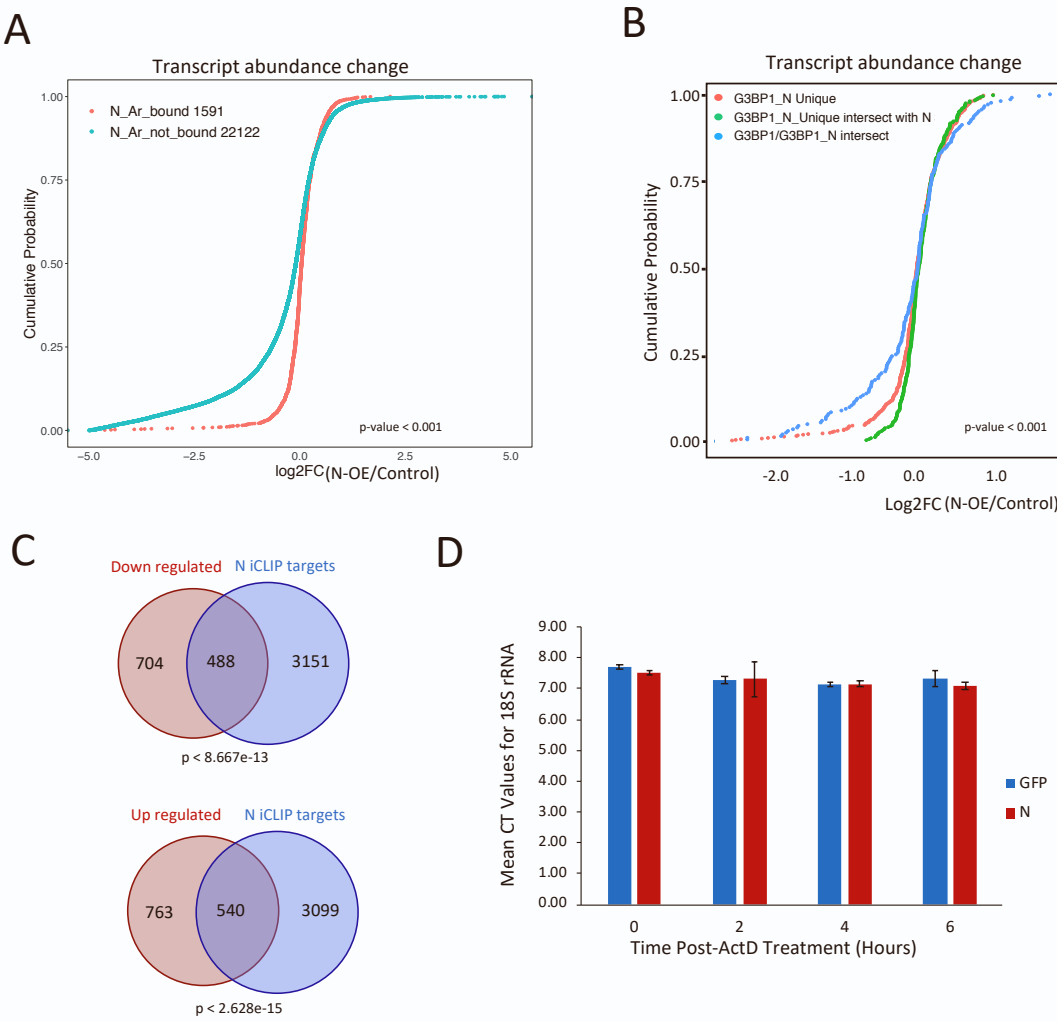



**Figure S6: N modulates the stability of target mRNAs, related to Figure 6. A:** Cumulative distribution analysis in abundance N iCLIP target mRNAs after N expression in $NaAsO_2$-treated HEK293 cells. P-value < 0.001 (Kruskal–Wallis test). **B:** Cumulative distribution analysis in abundance of G3BP1 target mRNAs after N overexpression in HEK293 cells. G3BP1 targets were identified by iCLIP-seq in N-expressing cells. Note: 'G3BP1_N Unique' represent those targets that were bound by G3BP1 only in N-expressing cells; 'G3BP1_N_Unique intersect with N' targets were obtained by overlapping 'G3BP1_N Unique' with N bound transcripts as defined by iCLIP-seq of N; 'G3BP1/G3BP1_N intersect' represent the common targets of G3BP1 in wildtype and N-expressing cells. Legend is provided. P-value < 0.001 (Kruskal–Wallis test). **C:** Venn diagram showing the overlap of differential genes in N-expressing cells with those identified as N targets through iCLIP-seq. P-values were calculated using hypergeometric test. **D:** Bar graph showing the RT-qPCR mean CT values for 18S rRNA transcripts across indicated time points after Actinomycin D treatment. Error bars represent standard error of mean (n=3).

Figure S7

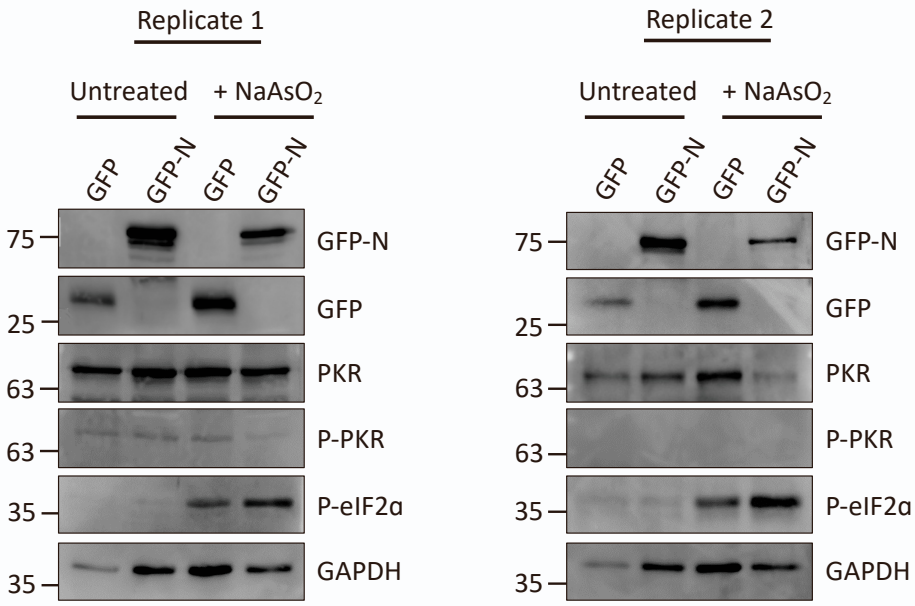



**Figure S7: Effect of N on phosphorylation levels of PKR and eIF2α in non-$NaAsO_2$- and $NaAsO_2$-treated cells. Related to Figure 2.** Western blotting analysis using whole cell extracts prepared from either GFP- or GFP-N-expressing non-NaAsO2- and NaAsO2-treated cells. Blots were probed with the indicated antibodies.