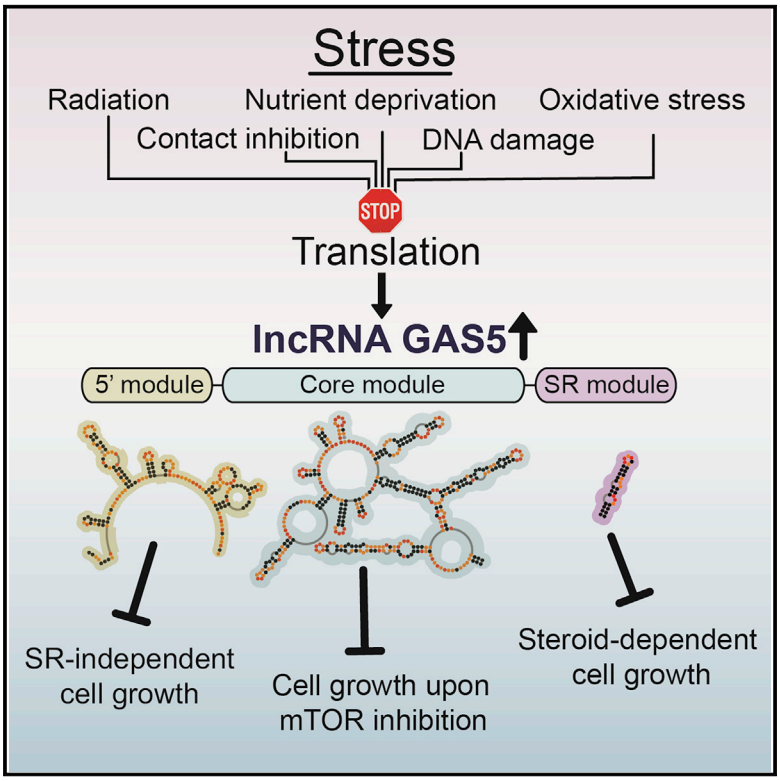
Cell Reports

Report

The lncRNA Growth Arrest Specific 5 Regulates Cell Survival via Distinct Structural Modules with Independent Functions

Graphical Abstract



Highlights

- The lncRNA GAS5 contains three main secondary structure modules
- Each module can independently reduce cell viability under different conditions
- Specific secondary structure elements are required for each module's activity

Authors

Filipp Frank, Nadieh Kavousi, Aikaterini Bountali, Eric B. Dammer, Mirna Mourtada-Maarabouni, Eric A. Ortlund

Correspondence

eortlund@emory.edu

In Brief

Frank et al. show that the lncRNA GAS5 uses a modular secondary structure architecture to regulate cell survival in response to varying cellular stress conditions.

Frank et al., 2020, Cell Reports *32*, 107933
July 21, 2020
https://doi.org/10.1016/j.celrep.2020.107933

CellPress

Cell Reports

CellPress
OPEN ACCESS

Report

The lncRNA Growth Arrest Specific 5 Regulates Cell Survival via Distinct Structural Modules with Independent Functions

Filipp Frank,[1] Nadieh Kavousi,[2] Aikaterini Bountali,[2] Eric B. Dammer,[1] Mirna Mourtada-Maarabouni,[2] and Eric A. Ortlund[1,3,*]

[1]Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322, USA
[2]Faculty of Natural Sciences, School of Life Sciences, Keele University, Keele ST5 5BG, UK
[3]Lead Contact
*Correspondence: eortlund@emory.edu
https://doi.org/10.1016/j.celrep.2020.107933

SUMMARY

There is increasing evidence that the architecture of long non-coding RNAs (lncRNAs)—just like that of proteins—is hierarchically organized into independently folding sub-modules with distinct functions. Studies characterizing the cellular activities of such modules, however, are rare. The lncRNA growth arrest specific 5 (GAS5) is a key regulator of cell survival in response to stress and nutrient availability. We use SHAPE-MaP to probe the structure of GAS5 and identify three separate structural modules that act independently in leukemic T cells. The 5′ terminal module with low secondary structure content affects basal survival and slows the cell cycle, whereas the highly structured core module mediates the effects of mammalian target of rapamycin (mTOR) inhibition on cell growth. These results highlight the central role of GAS5 in regulating cell survival and reveal how a single lncRNA transcript utilizes a modular structure-function relationship to respond to a variety of cellular stresses under various cellular conditions.

INTRODUCTION

Long non-coding RNAs (lncRNAs) represent a broad class of transcripts with critical roles in all aspects of cellular biology and a wide spectrum of molecular functions (Kopp and Mendell, 2018). They are aberrantly expressed in various cancers, are key players in tumor development and progression, and are linked to resistance against chemotherapy, demonstrating their potential as biomarkers and therapeutic targets (Arun et al., 2018). Determining the specific molecular function of individual lncRNAs, however, has proven difficult, and the mechanisms of most lncRNAs remain uncharacterized.

Structural studies of lncRNAs can provide valuable insight into their functional properties. Focused, in-depth secondary structure analyses of individual lncRNAs, for example, have identified modular architectures with distinct sub-domains (Hawkes et al., 2016; Ilik et al., 2013; Novikova et al., 2012; Smola et al., 2016; Somarowthu et al., 2015). Few studies, however, have tested if these sub-domains function independently in cells (Quinn et al., 2014; Uroda et al., 2019; Chillón and Pyle, 2016).

The growth arrest-specific 5 (*GAS5*) gene encodes a lncRNA that was identified in a subtraction cDNA library enriched for RNA sequences preferentially expressed in growth-arrested cells (Schneider et al., 1988). Later studies showed that it is required for normal growth arrest, slows down the cell cycle (Mourtada-Maarabouni et al., 2008), and controls apoptosis (Kino et al., 2010; Mourtada-Maarabouni et al., 2010; Pickard et al., 2013). In agreement with this role in inhibiting cell proliferation, GAS5 expression is reduced in numerous cancers, including breast, lung, gastric, pancreatic, bladder, and prostate cancer, as well as renal cell carcinoma, B cell lymphoma, and leukemia (Yu and Li, 2015).

The *GAS5* gene contains a number of intronic small nucleolar RNAs (snoRNAs) involved in ribosome biogenesis (Smith and Steitz, 1998). It is therefore highly transcribed in all tissue types and is one of the most highly expressed lncRNAs in the genome (Gibb et al., 2011). Like many other lncRNAs, the GAS5 RNA is spliced and polyadenylated, and it associates with ribosomes (Coccia et al., 1992; Renganathan et al., 2014). In actively dividing cells, GAS5 levels are kept low by non-sense mediated decay (NMD) (Mourtada-Maarabouni and Williams, 2013; Smith and Steitz, 1998), a process requiring active translation (Isken and Maquat, 2007). Upon exposure to stress such as density arrest or nutrient deprivation, translation is inhibited, and GAS5 levels increase (Coccia et al., 1992; Fleming et al., 1998). Under these conditions, GAS5 exerts its negative effects on cell proliferation and survival.

The mammalian target of rapamycin (mTOR) pathway is a central hub for regulating cell growth in response to intra- and extracellular signals, such as growth factors and mitogens or the availability of nutrients and ATP. Inhibition of mTOR by rapamycin results in inhibition of cell proliferation, and GAS5 is required for this effect in human immune cells (Mourtada-Maarabouni et al., 2010). These observations establish GAS5 as a central

cellular regulator responding to various stress signals by inhibiting cell growth. However, mechanistic insight into the function the GAS5 lncRNA is limited.

At its 3′ terminal end (nucleotides 546–566), GAS5 contains a predicted stem-loop structure that specifically interacts with steroid receptors (SRs) and blocks DNA-dependent steroid signaling (Hudson et al., 2014; Kino et al., 2010). In steroid-sensitive cancer cells, such as prostate cancers, this SR binding motif is responsible for GAS5 effects on cell growth (Hudson et al., 2014). This is not true in other cell types, however, where proliferation is not strongly dependent on SR signaling. Therefore, other regions in GAS5 must be active and use different mechanisms to regulate cell survival.

Here, we use Selective 2′ Hydroxyl Acylation analyzed by Primer Extension by Mutational Probing (SHAPE-MaP) chemical probing to analyze the secondary structure of GAS5 *in vitro* and *in cellulo*. We find that the secondary structure of endogenous GAS5 resembles that of *in vitro* transcribed GAS5 RNA. The molecule contains three separate structural modules: a 5′ module with low secondary structure content; a highly structured core module; and the SR binding module, which forms separate from the rest of the molecule close to its 3′ end. Functional studies in leukemic T cells show that the 5′ module mediates GAS5's role in inhibiting basal cell survival and slowing the cell cycle, whereas the core module is required for mediating the effects of mTOR inhibition. These results confirm that the GAS5 structural modules function independently in cells, and each module acts under different cellular conditions, likely using different molecular mechanisms.

RESULTS

In Vitro SHAPE-MaP Identifies Three Structural Modules within GAS5

We used SHAPE-MaP with 1-methyl-7-nitroisatoic anhydride (1M7) to probe the secondary structure of *in vitro* transcribed GAS5 lncRNA (ENST00000450589.5). Modified RNA was reverse transcribed and amplified for sequencing by two overlapping PCRs of 400 nucleotides in length each (Figure S1). We obtained sequencing read depths of ~100,000–200,000 per nucleotide and mutation rates of modified samples of ~0.1% (Figure S1). The resulting mutation profiles allowed us to determine accurate SHAPE reactivities for GAS5 nucleotides 1–599 (Figure 1A; SHAPE reactivities and pairing probabilities). SHAPE reactivities in the overlap between the two PCR products determined in separate analyses showed good agreement (Spearman R = 0.79; Figure S1). Independent experiments performed either on the same day (Spearman R = 0.88) or several months apart also showed good agreement (Spearman R = 0.81; Figure S1).

The SHAPE reactivity profile of GAS5 shows the typical pattern of defined regions of low and high reactivities corresponding to low and high nucleotide flexibility (Figure 1B). The resulting GAS5 structure model identified a modular architecture with three separate sub-modules (Figures 1B and 1C). Interestingly, the short hairpin containing the SR binding motif is formed as predicted previously (Hudson et al., 2014; Kino et al., 2010) and is located at the 3′ terminus as an independent structural module that is separated from the rest of the molecule. This is consistent with the evolutionary origin of this element. It arose in the haplorhine lineage from an Alu insertion at the GAS5 intron11/exon12 boundary (Hudson et al., 2014). Hence, it is not present in, for example, mice, and it appears to have been added as an independent structural and functional module late in evolution. This observation is reminiscent of the modular architecture and function of proteins, in which domains are combined in a mix-and-match fashion to create large, multi-domain proteins with diverse functions. In analogy, the addition of the SR binding motif expanded the functional repertoire of GAS5.

A second structural module at the 5′ terminus of GAS5 is characterized by high SHAPE reactivities reflecting low secondary structure content with mostly short stretches of base-paired regions (nucleotides 1–170; 78 out of 170 nucleotides predicted to be base paired, 46%). This module coincides with GAS5's putative open reading frame (nucleotides 31–183; Smith and Steitz, 1998), and the low structural content is consistent with efficient translation initiation as well as elongation and might explain the rapid and efficient degradation of the message by NMD. Low secondary structure content also facilitates binding to miR-21 and miR-135, the binding sites of which are located within this module (Hu et al., 2016; Song et al., 2014; Xue et al., 2016; Zhang et al., 2013). Finally, the single-stranded RNA stretches in this region are also ideal for interaction with other complementary nucleic acids such as genomic DNA or nascent RNAs. These are common targets for lncRNA regulation via the recruitment of protein factors such as chromatin regulatory proteins.

The central region, downstream of the open reading frame and upstream of the SR binding motif, constitutes the third structural module of GAS5. It has a relatively high secondary structure content (nucleotides 172–540; 250 out of 369 nucleotides predicted to be base paired, 68%). This structured core contains 12 helices (core helices cH1–cH12), 9 stem loops, 2 three-way junctions, and a multi-way junction connecting helices cH4 to cH11 (Figure 1C). The high secondary structure content suggests this module forms a compact core and may act as a scaffold to recruit protein interaction partners via distinct binding surfaces.

For some nucleotides in some of the core module helices (in particular, cH2, cH7, and cH10), we measured high SHAPE reactivities (Figures 1B and 1C). SHAPE experiments do not distinguish well between structural flexibility and the presence of multiple conformations in an RNA. Highly SHAPE-modified nucleotides within stems may suggest that these regions in GAS5 can adopt multiple stable conformations.

In Cellulo Structure Probing Is Similar to *In Vitro* Data

To assess the structure of GAS5 in its cellular environment, we probed endogenous GAS5 in HeLa cells (Figure 2). The resulting SHAPE profile of endogenous GAS5 shows a similar distribution of SHAPE reactivities as samples prepared *in vitro* (Figure 2A). The degree of correlation between *in cellulo* and *in vitro* data was high and similar to that for *in-cellulo*-measured SHAPE reactivities in the overlap of the two GAS5 amplicons used for high-throughput sequencing (Figure 2B). To assess potential local differences, we calculated windowed correlation coefficients over the entire sequence (Figure 2C). The results show that *in cellulo* data more strongly correlate with *in vitro* data than with a random distribution.

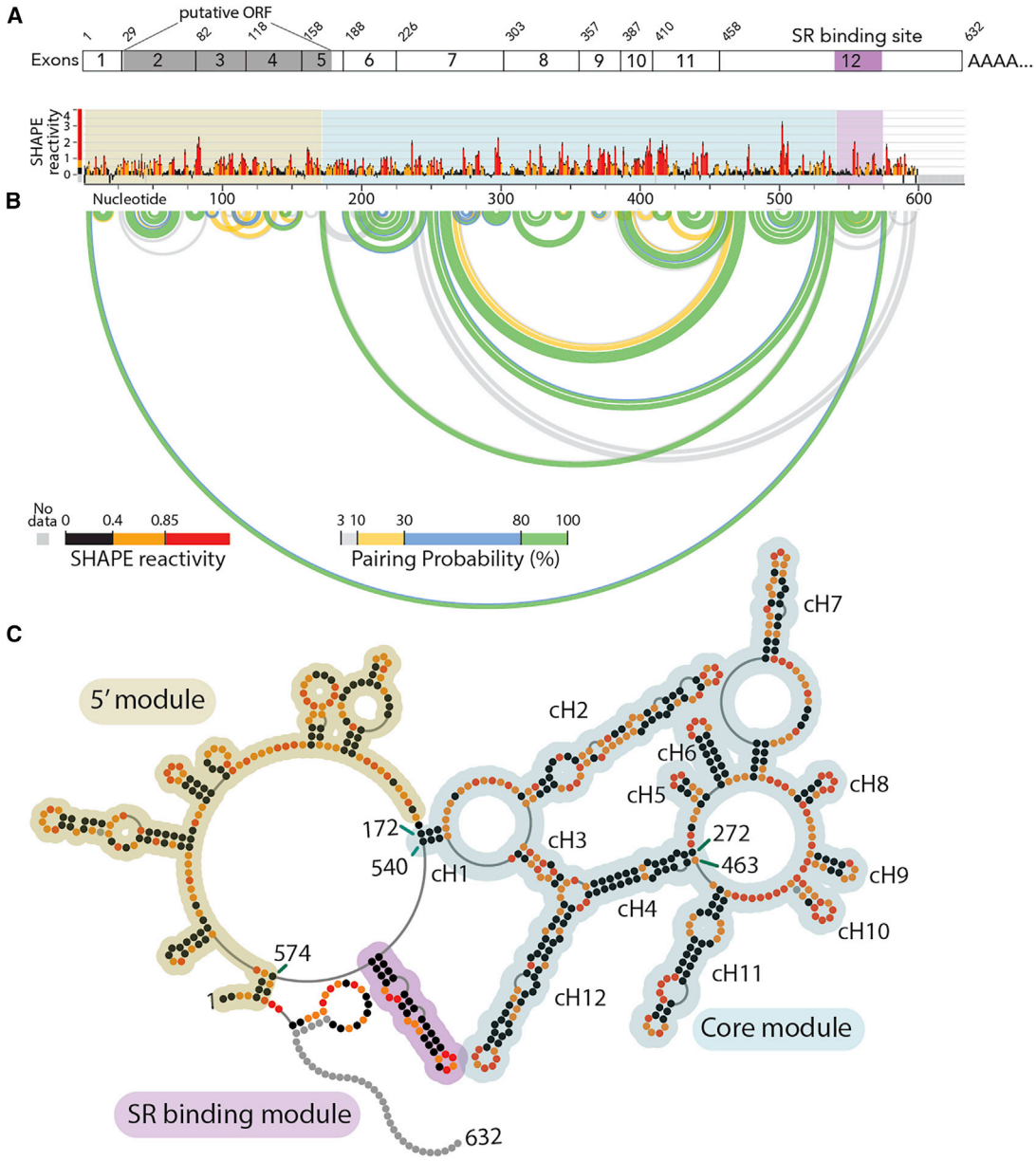


Figure 1. SHAPE-MaP of GAS5 RNA *In Vitro* Identifies a Modular Architecture

(A) Exon structure of GAS5 with the putative ORF and the SR binding motif highlighted.

(B) Nucleotide resolution SHAPE reactivities for *in vitro* transcribed and folded GAS5 are shown color coded for low, medium, and high reactivities. Below: base-pairing probabilities in GAS5. Arcs represent base pairs and are color coded by probability.

(C) Minimum free energy secondary structure model of GAS5 color coded for SHAPE reactivities. Structural modules are highlighted.

While these data show that *in vitro* transcribed GAS5 is highly similar to the endogenous lncRNA, local correlation between the two is weaker than between two independent *in vitro* datasets, in particular in the first ~450 nucleotides (Figure 2C).

The observed minor differences between endogenous and *in vitro* transcribed RNA may result from effects of heat denaturation and refolding on the *in vitro* sample leading to more compact structures (Dethoff and Weeks, 2019). Another source of difference is found in the cellular environment of endogenous GAS5. Interactions of GAS5 with proteins or other nucleic acids may affect SHAPE reactivities. To investigate these differences further, we applied the *deltaSHAPE* algorithm (Smola et al., 2015a) for the identification of protein binding sites from *in cellulo* SHAPE-MaP data (Figure S2). Comparison of *in cellulo* data with three independent *in vitro* datasets revealed localized sites of differences. The 5′ module contains regions with reduced SHAPE reactivities, which is usually associated with protein binding sites (Smola et al., 2016) and may also reflect the translation of the

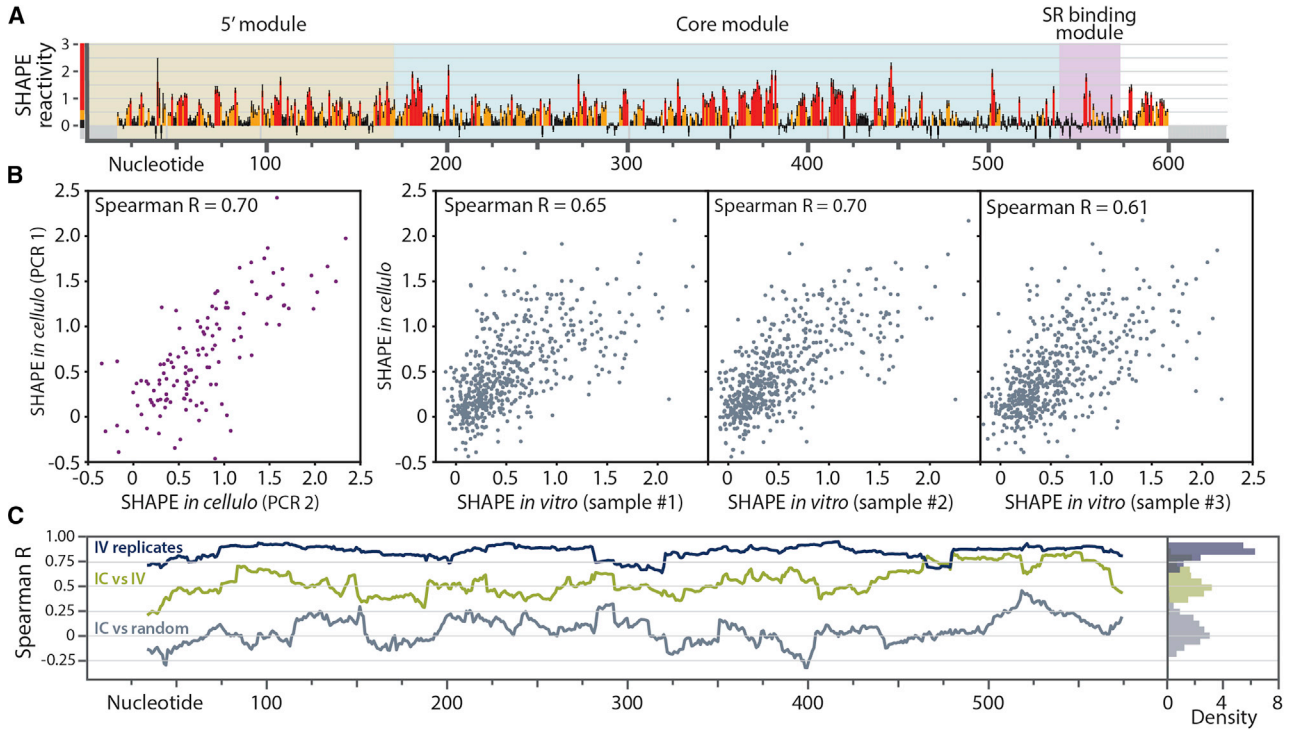


Figure 2. GAS5 SHAPE Reactivities Appear Similar *In Cellulo* and *In Vitro*

(A) SHAPE reactivities from probing endogenous GAS5 in HeLa cells.

(B) Correlation of SHAPE reactivities for individual nucleotides between different datasets. A similar degree of correlation is observed within the *in cellulo* data and between *in cellulo* and *in vitro* datasets. Left: the two amplicons used to generate double-stranded DNA (dsDNA) for sequencing had an overlap of 132 nucleotides (248–379). Shown here is the correlation between the SHAPE reactivities of those nucleotides when the two amplicons are processed separately. Right: correlation between the *in cellulo* data and three independent *in vitro* datasets.

(C) Windowed correlation between different SHAPE-MaP datasets (over a 40-nucleotide window). IV replicates, two replicates of *in vitro* data; IC versus IV, *in vitro* versus *in cellulo* data; random, *in cellulo* data versus randomly generated data with a similar distribution of values. Right: distribution of windowed correlation coefficients.

putative open reading frame by ribosomes. The core module comprises sites with enhanced cellular SHAPE reactivities, suggesting that minor structural changes in the structured core occur in endogenous GAS5 (Smola et al., 2016).

It is important to note, however, that the SHAPE differences observed here are smaller than those previously reported for very stable ribonucleoproteins (RNPs) (Smola et al., 2015a). The original analysis was developed using benchmark RNAs that are short (<300 nucleotides) and form highly stable ribonucleoprotein complexes (e.g., 5S rRNA and U1 small nuclear RNA [snRNA]). Smaller differences in the case of GAS5 may result from several factors, such as incomplete saturation of GAS5 with cellular protein binding partners under the conditions tested, transient protein binding, or differences between the cytoplasmic and nuclear populations of GAS5 (probed together in our experiments). All these factors will dilute any differences in SHAPE reactivities and may explain the observed data.

Apoptosis and Cell Cycle Arrest Are Mediated by the 5′ Module

The SR binding motif mediates most of the effects of GAS5 on the basal survival of steroid-driven cancer cells (Hudson et al., 2014). A single mutation in this motif that abrogates SR binding almost completely abolished its effects on cell survival (Hudson et al., 2014), consistent with this motif being an independent functional module. In immune cells, which are less dependent on SR signaling for growth (Mourtada-Maarabouni et al., 2008), the SR binding motif mutant only partially reduces the effects of GAS5. This suggests that other regions of the lncRNA play important roles in regulating cell survival in those cells.

To test if the 5′ terminal and core modules identified by SHAPE are functioning independently in cells, we overexpressed these RNAs in Jurkat and CEM-C7 leukemic T cells (Figures 3A and 3B). Overexpression of GAS5 induces robust negative effects on cell survival (Figure 3C). These effects are mediated by an increase in apoptosis and a slowing of the cell cycle, with an increased proportion of cells in G0/G1 and a corresponding decrease of cells in G2/M (Figures 3D–3F). Of the two GAS5 structural modules, overexpression of the core region did not cause any significant effect on cell viability, apoptosis, or cell cycle. In contrast, the 5′ terminal module significantly reduced cell viability, increased apoptosis, and completely phenocopied the effects on the cell cycle (Figures 3C–3F). The additional negative effects of full-length GAS5 on viable cell number and apoptosis

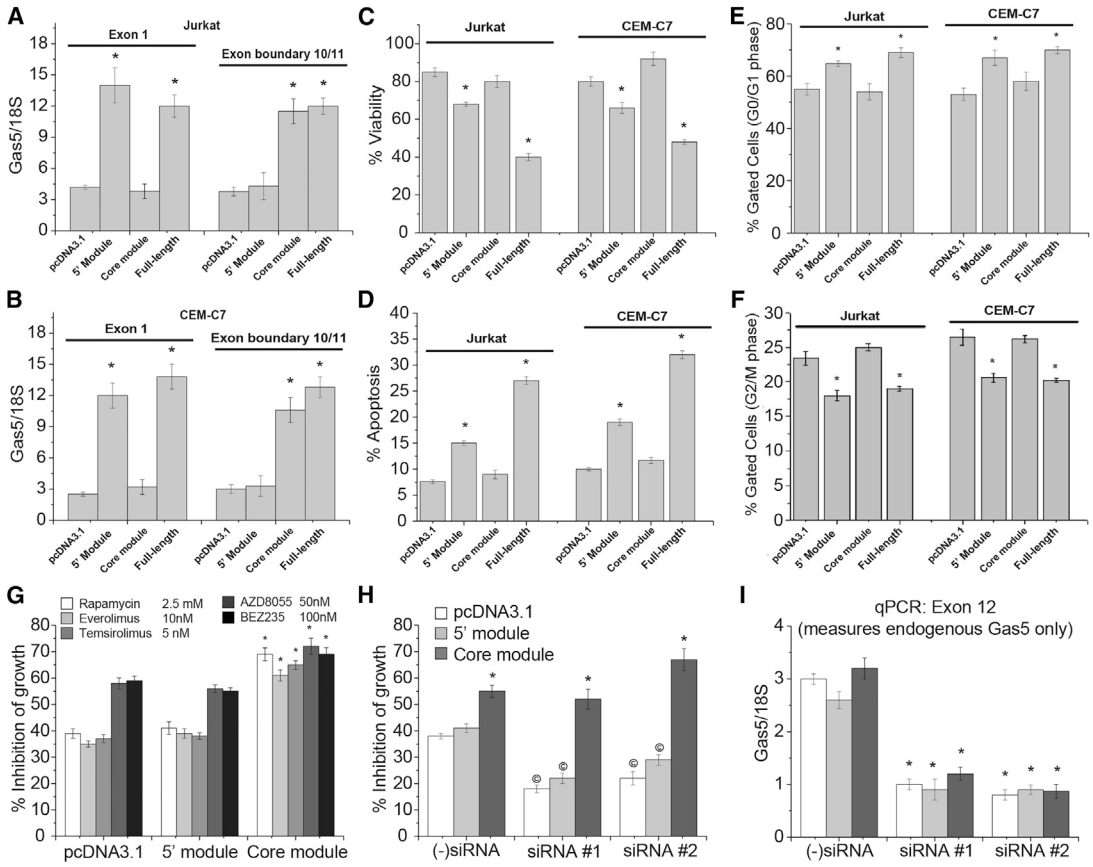


Figure 3. The GAS5 5′ and Core Modules Act Independently under Different Cellular Conditions

In (A)–(F), cells were transfected with GAS5 constructs, or empty pcDNA3 vector as control. After 24 h, cells were replated at a fixed density for analysis after a further 48 h. *p < 0.01 compared with cells transfected with vector alone (one-way ANOVA and Dunnet's post hoc analysis).

(A and B) Cellular GAS5 levels for Jurkat (A) and CEM-C7 cells (B) were determined by qRT-PCR at 24 h post-transfection using two different Taqman assays to distinguish the two modules: Exon 1 for the 5′ module and exon boundary 10/11 for the core module.

(C) Cell viability was determined by flow cytometry.

(D) Basal apoptosis level was measured by acridine orange staining.

(E and F) Cell cycle analysis was assessed by flow cytometry following nuclear propidium iodide staining.

(G) Jurkat cells were transfected with one of the GAS5 modules or empty vector. Cells were treated with mTOR inhibitors 24 h post-transfection. Cell growth was measured using MTS (3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium) after 72 h. Results are represented as the percentage of inhibition of cell proliferation relative to vehicle-treated cells. *p < 0.01 compared with cells transfected with vector alone (one-way ANOVA and Dunnet's post hoc analysis).

(H) Jurkat cells were transfected with specific GAS5 siRNAs (SI03652537 targeting exon 3 = siRNA #1; SI03652544 targeting exon 7 = siRNA #2) or negative control siRNA ((−)siRNA) and cultured at 37°C. After 48 h, cells were transfected with one of the GAS5 modules or empty vector and treated with rapamycin (2.5 mM) 24 h post-transfection. Cell growth was measured using MTS after 48 h. Results are represented as the percentage of inhibition of cell proliferation relative to vehicle-treated cells. *p < 0.01 compared with cells transfected with (−)siRNA; ©p < 0.01 compared with cells transfected with pcDNA3.1 (one-way ANOVA and Dunnet's post hoc analysis).

(I) qRT-PCRs for quantifying endogenous GAS5 levels. *p < 0.01 compared with cells transfected with (−)siRNA (one-way ANOVA and Dunnet's post hoc analysis). It is of note here that siRNAs #1 and #2 were designed to target GAS5 in the 5′ and core modules, respectively, so that overexpression of the modules could be selectively controlled. qRT-PCR results, however, showed that each siRNA targeting GAS5 reduced the extent of overexpression of both modules more so than did the control siRNA (Figure S3). This does not affect interpretation of the results since increased levels of the core module led to an increase in growth inhibition in all cases. Data represent mean ± SEM from four independent experiments.

may be explained by the activity of the SR binding motif (Hudson et al., 2014).

mTOR Signaling Is Mediated through the GAS5 Structured Core Domain

mTOR inhibition in leukemic T cells slows cell growth, and a substantial proportion of this effect is mediated through GAS5 (Mourtada-Maarabouni et al., 2010). To test if the individual modules could explain the role of GAS5 in the mTOR pathway, we measured their effects in the presence of a series of mTOR inhibitors in Jurkat cells (Figures 3G and 3H). Jurkat cells were transfected with the GAS5 modules, and 24 h post-transfection, cells were treated with mTOR inhibitors. Cell growth was monitored another 72 h later. Cells expressing the 5′ module did not

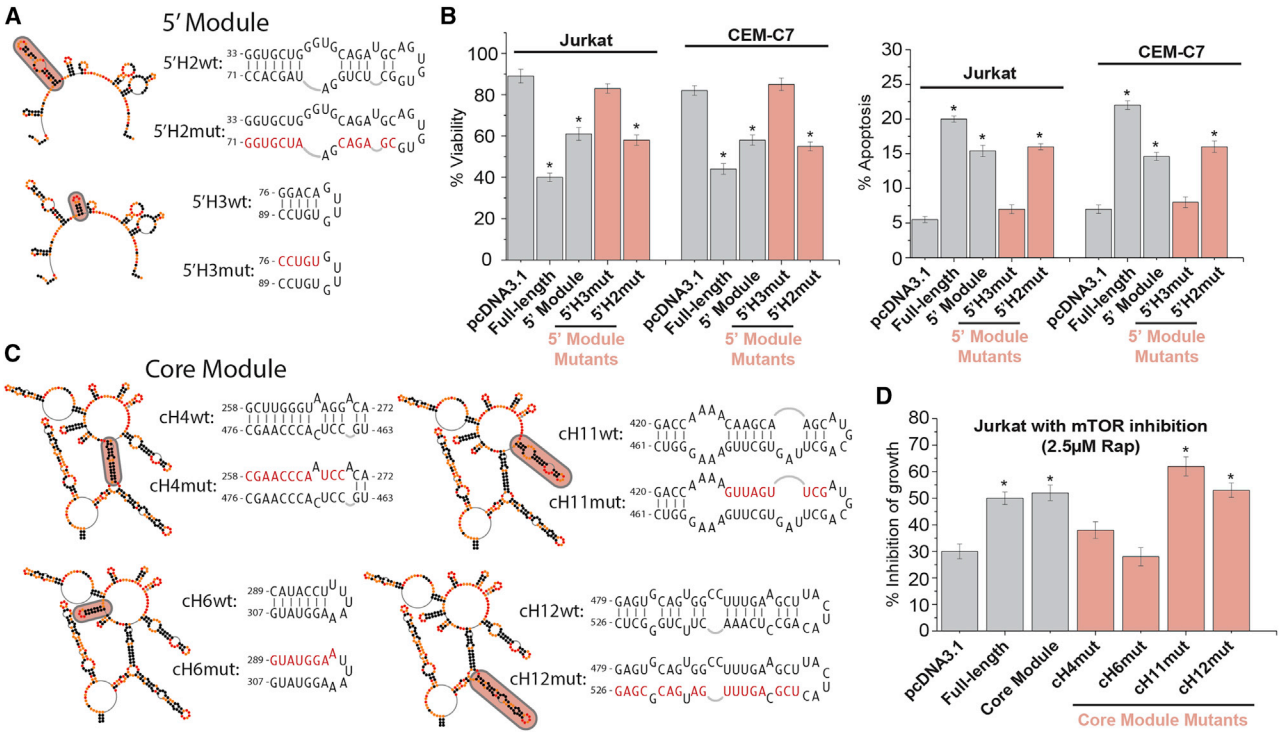


Figure 4. Distinct Secondary Structures within the GAS5 Modules Are Required for Function in Cells

(A) Design of mutations to disrupt helices 5′H2 and 5′H3 within the 5′ module.

(B) Effects of mutations in distinct secondary structures on the activity of the 5′ module in Jurkat and CEM-C7 cells.

(C) Design of mutations to disrupt helices cH4, cH6, cH11, and cH12 within the core module.

(D) Effects of mutations in distinct secondary structures on the activity of the core module in Jurkat cells in the presence of mTOR inhibitors. Data represent mean ± SEM from three independent experiments. *p < 0.01 compared with cells transfected with pcDNA3.1 (one-way ANOVA and Dunnet's post hoc analysis).

experience any increased inhibition of growth upon mTOR inhibition, compared to cells transfected with the empty vector (Figure 3G). The core module, however, significantly enhanced the effect of mTOR inhibition on cell growth with all mTOR inhibitors tested (Figure 3G).

To further support these observations, we knocked down endogenous GAS5 using small interfering RNAs (siRNAs) and re-introduced the GAS5 modules. Cell growth was then monitored after exposure to rapamycin. siRNAs efficiently knocked down endogenous GAS5 (Figure 3I). Subsequent overexpression of the 5′ module did not have any effect, whereas overexpression of the core module consistently increased the inhibition of growth in response to mTOR inhibition (Figure 3H).

The results of these experiments confirm that the core module acts when mTOR is inhibited, whereas the 5′ module does not affect cell growth under these conditions. Together, these experiments show that both the 5′ module and the core module have the capacity to regulate cell growth in leukemic T cells, but they do so under different cellular conditions.

Distinct Structural Elements Are Required for GAS5 Module Functions

To test whether individual structural features present in the 5′ or core modules are required for function, we disrupted a set of well-defined secondary structures in each GAS5 module. This strategy has been successfully applied in other studies to determine the precise location of functional RNA elements (Uroda et al., 2019; Chillón and Pyle, 2016). The effects of mutations were measured in leukemic T cells under normal growth conditions for the 5′ module (Figures 4A and 4B) and in the presence of mTOR inhibitors for the core module (Figures 4C and 4D).

The 5′ module contains six secondary structures of varying length, ranging from 2 to 13 base pairs. We introduced mutations to disrupt two well-defined structures with low SHAPE reactivities (5′H2 and 5′H3; Figure 5A). Expression of these constructs in Jurkat or CEM-C7 cells showed that helix 5′H2 is not required for proper function of the 5′ module (Figure 4B). No effect was observed on cell viability or levels of apoptosis when compared to the wild-type 5′ module. Disruption of helix 5′H3, however, abolished effects on viability and apoptosis completely. Thus, while the 5′ module is mostly unstructured, it contains a small hairpin structure (5′H3) that is responsible for its negative effects on cell survival.

The highly structured core module was tested in Jurkat cells under conditions of mTOR inhibition. We chose four stable helices (cH4, cH6, cH11, and cH12) and introduced mutations to disrupt these helices and probe their function in cells. Mutation of helices cH11 and cH12 did not have an effect on the function of the core module. Core module constructs containing mutations in helices cH4 and cH6, however, were not able to reduce

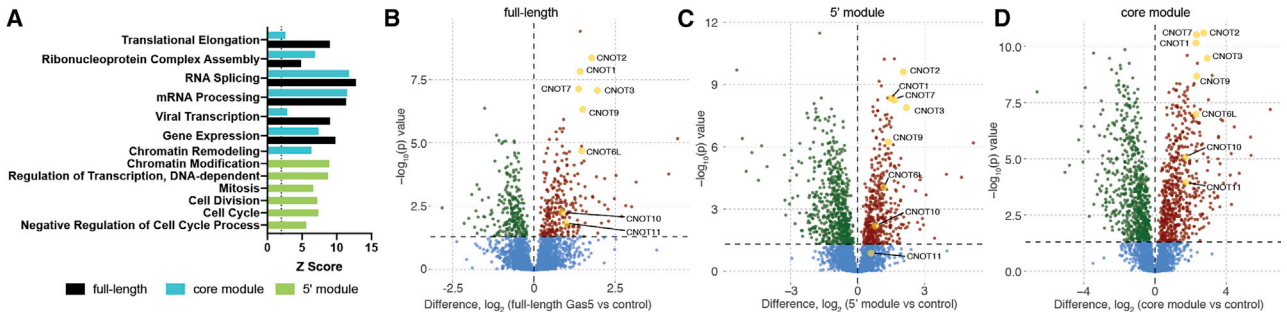


Figure 5. Proteomics Analysis of GAS5 Interactions

(A) GO analysis of proteins significantly enriched in GAS5 samples over control. Shown are the GO terms (biological process) that have a *Z* score of at least 5 for at least one sample. *Z* scores below 2 are not shown.

(B–D) Volcano plots of pairwise comparisons of proteins detected in pull-downs with the full-length GAS5 (B), the 5′ module (C), and the core module (D) versus control. Proteins enriched in the GAS5 pulldowns are shown in red, proteins enriched in the controls are shown in green, and proteins with $p < 0.05$ are shown in blue. The CCR4-NOT complex proteins are highlighted as large yellow circles.

cell growth beyond the level of the control (Figure 4D). Helix cH4 is formed by long-range interactions between nucleotides 258–272 and 463–476 (Figure 4C). It is located at the center of the core module and bridges the three-way junction of helices cH3, cH4, and cH12 with the multi-way junction connecting helices cH4 to cH11. Disruption of this structure may have effects on the integrity of the core module as a whole, which could explain why it is required for function. Helix cH6 is part of a short, seven-base-pair hairpin structure. Abolishing cH6 ablates the core module's effect on cell growth, suggesting that it is responsible for mediating its activity directly.

While it is possible that mutations of individual secondary structures may have broader effects on the overall structure of the GAS5 modules, these experiments strongly support the notion that the GAS5 structural modules we identified using SHAPE-MaP contain defined secondary structure elements that are required for their negative effects on basal cell survival of leukemic T cells.

The GAS5 Protein Interactome

The presence of defined secondary structures required for the function of GAS5 and its modules suggests that specific proteins recognize these structures and mediate the downstream effects of GAS5 on cell survival. To identify such proteins, we performed pull-down experiments using biotinylated, *in vitro* transcribed RNA and Jurkat cell lysates. Pull-downs were conducted with full-length GAS5, the 5′ module, the core module, or the complete reverse complement of GAS5 as a negative control. Proteins associating with these RNAs were then quantified using mass spectrometry (Figure 5).

We found 273, 448, and 619 proteins significantly enriched over controls in the full-length, 5′, and core module samples, respectively. Gene Ontology (GO) analyses of the proteins associating with the GAS5 samples identified enriched RNA-related processes (mRNA processing, splicing, ribonucleoprotein complex assembly, translation elongation) for the full-length and core modules. GO analysis of enriched proteins shows that the 5′ module associates with proteins involved in the cell cycle, which is in agreement with our cellular assays showing that the 5′ module, but not the core module, regulates cell survival through changes in the cell cycle (Figure 5A). Full-length GAS5, however, did not enrich for these proteins. Individual pulldowns were not performed with equimolar amounts of biotinylated RNA, but rather with an equal mass (3 μg). As a result, the shorter 5′ module sample (167 nucleotides) contained approximately 4 times more molecules compared to full-length GAS5. This may explain a greater enrichment of cell cycle proteins for the 5′ module.

A number of the most highly enriched proteins belong to the CCR4-NOT (Carbon Catabolite Repression—Negative On TATA-less) complex (CNOT1, CNOT2, CNOT3, CNOT6L, CNOT7, CNOT9, CNOT10, and CNOT11; Figures 5B–5D). This complex is enriched in all GAS5 samples, including the 5′ and core modules, suggesting that multiple parts of GAS5 are able to associate with it. Interaction with this complex and modulation of its activity would provide GAS5 with a broad platform to affect cell survival at multiple steps in gene expression. The CCR4-NOT complex is involved in almost all aspects of gene expression, including chromatin modification and transcription as well as mRNA processing, export, deadenylation, and translation (Collart, 2016).

In summary, the proteomics analysis identified groups of proteins that specifically interact with GAS5 and its structural modules. Differences between proteins associating with the different modules allude to their different activities under varying cellular conditions. These results may serve as a resource for further investigation into the molecular mechanisms employed by GAS5 to regulate cell survival.

DISCUSSION

Their flexibility in size and structure and their unique ability to form specific complexes with proteins, DNA, and other cellular RNAs make lncRNAs ideal molecular scaffolds to bring together different biomolecules in order to regulate and organize cellular processes (Deveson et al., 2017). A modular structural architecture is an important benefit, allowing lncRNAs to assemble distinct, functionally independent effector complexes on different parts of their structure and therefore diversify their functional toolkit.

GAS5 is a lncRNA with antiproliferative effects that is downregulated in numerous cancers. A short hairpin structure at its 3′ terminus acts as a decoy for SRs and inhibits their signaling (Hudson et al., 2014; Kino et al., 2010). Beyond that, not much is known about how GAS5 reduces cell growth and viability and which elements are required for these effects.

To further characterize GAS5 function, we determined its secondary structure *in vitro* and *in cellulo* using chemical probing experiments (Figures 1 and 2). The results show that GAS5 contains three structural modules with distinct sequence and structure: a 5′ module with low secondary structure content, a highly structured core module, and the previously predicted SR binding short hairpin motif.

Highly modular structures have been observed for other lncRNAs (Hawkes et al., 2016; Novikova et al., 2012; Somarowthu et al., 2015; Tsai et al., 2010), and the individual modules are thought to be functionally important. The SR binding hairpin of GAS5, for instance, has previously been shown to act independently of the rest of the molecule via interaction with SRs (Hudson et al., 2014; Kino et al., 2010). Here, using functional assays in leukemic T cells, we show that the 5′ and the core modules also function independently and do so in different cellular contexts. Under normal growth conditions in leukemic T cells, which have little dependence on SR signaling, the 5′ module is responsible for many of the effects of GAS5 on cell viability, apoptosis, and the cell cycle (Figure 3). The core module has no effect under those conditions. When cells were treated with mTOR inhibitors, however, we find that the core module is active, and the 5′ module has no effects (Figure 3). Under these conditions, the core module recapitulates the previously observed function of GAS5 in mediating the inhibition of growth upon shutting down the mTOR signaling pathway (Mourtada-Maarabouni et al., 2010). Mutations within specific RNA secondary structure elements in the 5′ and core modules disrupt their effects in cells, showing that distinct structural features are critical for each module's function. Thus, GAS5 uses its modular structural organization to assemble distinct functional complexes on different parts of its sequence, each modulating cell growth independent of the other. This allows GAS5 to reliably respond to cellular stress in different cell types and under different external conditions.

How did these structural modules of GAS5 arise? While there is currently no evidence that GAS5 encodes a functional protein product, it is possible that when the GAS5 gene first appeared in evolution, its ORF initially produced a functional protein. GAS5 belongs to the family of 5′ terminal oligopyrimidine (5′ TOP) genes (Smith and Steitz, 1998), which include all ribosomal proteins as well as translation elongation factors. This suggests that its initial function was related to translation. Over time, the transcript may have gained beneficial non-coding RNA functions in its 3′ untranslated region. Subsequently, with evolutionary pressure low on the ORF itself, it lost its coding function, and GAS5 changed into a lncRNA gene. Here, we show that the 5′ module, which covers the putative ORF, functions independently at the RNA level and regulates cell survival. This shows that the ORF not only lost its coding function, but also gained novel, non-coding functions. Finally, the SR binding module at the GAS5 3′ end was added more recently in evolution via an Alu insertion in the haplorrhine lineage (Hudson et al., 2014). As a result, human GAS5 now encodes a lncRNA with at least three structural modules, each with a separate, independent function.

The translation of 5′ TOP genes is selectively inhibited under conditions of nutrient deprivation. It is, thus, conceivable that the non-coding functions of GAS5 evolved as part of the adaptive response to nutrient starvation, which involves the slowing of cell growth as a means to enhance survival. Cells respond to various kinds of stress with transcript-specific—such as for the 5′ TOP family—or general post-transcriptional mechanisms that reprogram translation to favor survival (Ivanov et al., 2011). By regulating GAS5 using NMD, a process that requires active translation, cells ensure that any global negative effects on translation as a response to stress increase the levels of GAS5 so that it can exert its antiproliferative effects. This establishes GAS5 as a key effector in the cellular stress response.

Our understanding of the mechanisms employed by GAS5 to regulate cell survival is still very limited. The GAS5 structural modules identified here likely assemble into complexes with other cellular components (proteins, DNA, and RNA) that mediate GAS5's effects. We used a proteomics approach to identify proteins that specifically interact with GAS5 and its structural modules. The results show strong interaction of GAS5 with the CCR4-NOT machinery, suggesting that it may modulate gene expression through this large multi-component complex. However, further mechanistic explorations will be necessary to substantiate these results and to gain further insight into GAS5's biological mechanism.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Preparation of RNA
  - Synthesis of 1-methyl-7-nitroisatoic anhydride
  - Modification of *in vitro* transcribed GAS5 RNA with 1M7
  - Modification of cellular RNA with 1M7
  - Reverse transcription of 1M7 modified RNA
  - Amplification of cDNA for SHAPE-MaP
  - Library preparation and high-throughput sequencing
  - SHAPE data analysis
  - Effects of GAS5 constructs on the basal survival of Jurkat and CEM-C7 leukemic T cells
  - Effects of GAS5 constructs on cell growth in the presence of mTOR inhibitors
  - RNA-pulldown
  - On bead digestion
  - LC-MS/MS analysis
  - Protein Identification and Quantification with MaxQuant

- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Assays for determination of cell viability, apoptosis, survival, and cell cycle
  - Mass Spectrometry Quantitative Data Processing and Statistics

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2020.107933.

ACKNOWLEDGMENTS

This work was supported by a W.M. Keck Foundation Medical Research grant. High-throughput sequencing was performed at the Yerkes Non-Human Primate Genomics Core at Emory. We thank the Kevin Weeks lab and Steven Busan for making ShapeMapper and other SHAPE analysis software available and for advice on data analysis. We thank Jeffrey A. Kohn for help with the synthesis of 1M7. This study was supported in part by the Emory Integrated Proteomics Core (EIPC), which is subsidized by the Emory University School of Medicine and is one of the Emory Integrated Core Facilities. Additional support was provided by the Georgia Clinical & Translational Science Alliance of the National Institutes of Health under award UL1TR002378. The content is solely the responsibility of the authors and does not necessarily reflect the official views of the National Institutes of Health.

AUTHOR CONTRIBUTIONS

F.F. performed all SHAPE experiments and analysis and RNA pulldown experiments. M.M.-M., N.K., and A.B. performed cellular assays in leukemic T cells. E.B.D. performed the proteomics data analysis. F.F., M.M.-M., and E.A.O. designed experiments. F.F. and E.A.O. wrote the manuscript. All authors reviewed the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: September 18, 2019
Revised: April 17, 2020
Accepted: June 22, 2020
Published: July 21, 2020

REFERENCES

Arun, G., Diermeier, S.D., and Spector, D.L. (2018). Therapeutic Targeting of Long Non-Coding RNAs in Cancer. Trends Mol. Med. *24*, 257–277.

Chillón, I., and Pyle, A.M. (2016). Inverted repeat Alu elements in the human lincRNA-p21 adopt a conserved secondary structure that regulates RNA function. Nucleic Acids Res. *44*, 9462–9471.

Coccia, E.M., Cicala, C., Charlesworth, A., Ciccarelli, C., Rossi, G.B., Philipson, L., and Sorrentino, V. (1992). Regulation and expression of a growth arrest-specific gene (gas5) during growth, differentiation, and development. Mol. Cell. Biol. *12*, 3514–3521.

Collart, M.A. (2016). The Ccr4-Not complex is a key regulator of eukaryotic gene expression. Wiley Interdiscip. Rev. RNA *7*, 438–454.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. *10*, 1794–1805.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Mol. Cell. Proteomics *13*, 2513–2526.

Dethoff, E.A., and Weeks, K.M. (2019). Effects of Refolding on Large-Scale RNA Structure. Biochemistry *58*, 3069–3077.

Deveson, I.W., Hardwick, S.A., Mercer, T.R., and Mattick, J.S. (2017). The Dimensions, Dynamics, and Relevance of the Mammalian Noncoding Transcriptome. Trends Genet. *33*, 464–478.

Fleming, J.V., Hay, S.M., Harries, D.N., and Rees, W.D. (1998). Effects of nutrient deprivation and differentiation on the expression of growth-arrest genes (gas and gadd) in F9 embryonal carcinoma cells. Biochem. J. *330*, 573–579.

Gibb, E.A., Vucic, E.A., Enfield, K.S., Stewart, G.L., Lonergan, K.M., Kennett, J.Y., Becker-Santos, D.D., MacAulay, C.E., Lam, S., Brown, C.J., and Lam, W.L. (2011). Human cancer long non-coding RNA transcriptomes. PLoS ONE *6*, e25915.

Hawkes, E.J., Hennelly, S.P., Novikova, I.V., Irwin, J.A., Dean, C., and Sanbonmatsu, K.Y. (2016). COOLAIR Antisense RNAs Form Evolutionarily Conserved Elaborate Secondary Structures. Cell Rep. *16*, 3087–3096.

Hu, L., Ye, H., Huang, G., Luo, F., Liu, Y., Liu, Y., Yang, X., Shen, J., Liu, Q., and Zhang, J. (2016). Long noncoding RNA GAS5 suppresses the migration and invasion of hepatocellular carcinoma cells via miR-21. Tumour Biol. *37*, 2691–2702.

Hudson, W.H., Pickard, M.R., de Vera, I.M.S., Kuiper, E.G., Mourtada-Maarabouni, M., Conn, G.L., Kojetin, D.J., Williams, G.T., and Ortlund, E.A. (2014). Conserved sequence-specific lincRNA-steroid receptor interactions drive transcriptional repression and direct cell fate. Nat. Commun. *5*, 5395.

Ilik, I.A., Quinn, J.J., Georgiev, P., Tavares-Cadete, F., Maticzka, D., Toscano, S., Wan, Y., Spitale, R.C., Luscombe, N., Backofen, R., et al. (2013). Tandem stem-loops in roX RNAs act together to mediate X chromosome dosage compensation in Drosophila. Mol. Cell *51*, 156–173.

Isken, O., and Maquat, L.E. (2007). Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function. Genes Dev. *21*, 1833–1856.

Ivanov, P., Kedersha, N., and Anderson, P. (2011). Stress puts TIA on TOP. Genes Dev. *25*, 2119–2124.

Kino, T., Hurt, D.E., Ichijo, T., Nader, N., and Chrousos, G.P. (2010). Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. Sci. Signal. *3*, ra8.

Kopp, F., and Mendell, J.T. (2018). Functional Classification and Experimental Dissection of Long Noncoding RNAs. Cell *172*, 393–407.

Mohammed, H.N., Pickard, M.R., and Mourtada-Maarabouni, M. (2016). The protein phosphatase 4 - PEA15 axis regulates the survival of breast cancer cells. Cell. Signal. *28*, 1389–1400.

Mourtada-Maarabouni, M., and Williams, G.T. (2013). Growth arrest on inhibition of nonsense-mediated decay is mediated by noncoding RNA GAS5. BioMed Res. Int. *2013*, 358015–358019.

Mourtada-Maarabouni, M., Hedge, V.L., Kirkham, L., Farzaneh, F., and Williams, G.T. (2008). Growth arrest in human T-cells is controlled by the non-coding RNA growth-arrest-specific transcript 5 (GAS5). J. Cell Sci. *121*, 939–946.

Mourtada-Maarabouni, M., Hasan, A.M., Farzaneh, F., and Williams, G.T. (2010). Inhibition of human T-cell proliferation by mammalian target of rapamycin (mTOR) antagonists requires noncoding RNA growth-arrest-specific transcript 5 (GAS5). Mol. Pharmacol. *78*, 19–28.

Novikova, I.V., Hennelly, S.P., and Sanbonmatsu, K.Y. (2012). Structural architecture of the human long non-coding RNA, steroid receptor RNA activator. Nucleic Acids Res. *40*, 5034–5051.

Panda, A.C., Martindale, J.L., and Gorospe, M. (2016). Affinity pulldown of biotinylated RNA for detection of protein-RNA complexes. Bio Protoc. *6*, 1–10.

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. *47*, D442–D450.

Pickard, M.R., Mourtada-Maarabouni, M., and Williams, G.T. (2013). Long non-coding RNA GAS5 regulates apoptosis in prostate cancer cell lines. Biochim. Biophys. Acta *1832*, 1613–1623.

Quinn, J.J., Ilik, I.A., Qu, K., Georgiev, P., Chu, C., Akhtar, A., and Chang, H.Y. (2014). Revealing long noncoding RNA architecture and functions using

domain-specific chromatin isolation by RNA purification. Nat. Biotechnol. *32*, 933–940.

Renganathan, A., Kresoja-Rakic, J., Echeverry, N., Ziltener, G., Vrugt, B., Opitz, I., Stahel, R.A., and Felley-Bosco, E. (2014). GAS5 long non-coding RNA in malignant pleural mesothelioma. Mol. Cancer *13*, 119.

Schneider, C., King, R.M., and Philipson, L. (1988). Genes specifically expressed at growth arrest of mammalian cells. Cell *54*, 787–793.

Siegfried, N.A., Busan, S., Rice, G.M., Nelson, J.A.E., and Weeks, K.M. (2014). RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). Nat. Methods *11*, 959–965.

Smith, C.M., and Steitz, J.A. (1998). Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5′-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. Mol. Cell. Biol. *18*, 6897–6909.

Smola, M.J., Calabrese, J.M., and Weeks, K.M. (2015a). Detection of RNA-Protein Interactions in Living Cells with SHAPE. Biochemistry *54*, 6867–6875.

Smola, M.J., Rice, G.M., Busan, S., Siegfried, N.A., and Weeks, K.M. (2015b). Selective 2′-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. Nat. Protoc. *10*, 1643–1669.

Smola, M.J., Christy, T.W., Inoue, K., Nicholson, C.O., Friedersdorf, M., Keene, J.D., Lee, D.M., Calabrese, J.M., and Weeks, K.M. (2016). SHAPE reveals transcript-wide interactions, complex structural domains, and protein interactions across the Xist lncRNA in living cells. Proc. Natl. Acad. Sci. USA *113*, 10322–10327.

Somarowthu, S., Legiewicz, M., Chillón, I., Marcia, M., Liu, F., and Pyle, A.M. (2015). HOTAIR forms an intricate and modular secondary structure. Mol. Cell *58*, 353–361.

Song, J., Ahn, C., Chun, C.H., and Jin, E.J. (2014). A long non-coding RNA, GAS5, plays a critical role in the regulation of miR-21 during osteoarthritis. J. Orthop. Res. *32*, 1628–1635.

Tsai, M.-C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., and Chang, H.Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science *329*, 689–693.

Turner, R., Shefer, K., and Ares, M., Jr. (2013). Safer one-pot synthesis of the 'SHAPE' reagent 1-methyl-7-nitroisatoic anhydride (1m7). RNA *19*, 1857–1863.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat. Methods *13*, 731–740.

Uroda, T., Anastasakou, E., Rossi, A., Teulon, J.-M., Pellequer, J.-L., Annibale, P., Pessey, O., Inga, A., Chillón, I., and Marcia, M. (2019). Conserved Pseudoknots in lncRNA MEG3 Are Essential for Stimulation of the p53 Pathway. Mol. Cell *75*, 982–995.e9.

Xu, Z.Z., and Mathews, D.H. (2016). Experiment-Assisted Secondary Structure Prediction with RNAstructure. Methods Mol. Biol. *1490*, 163–176.

Xue, D., Zhou, C., Lu, H., Xu, R., Xu, X., and He, X. (2016). LncRNA GAS5 inhibits proliferation and progression of prostate cancer by targeting miR-103 through AKT/mTOR signaling pathway. Tumour Biol., 1–11.

Yu, X., and Li, Z. (2015). Long non-coding RNA growth arrest-specific transcript 5 in tumor biology. Oncol. Lett. *10*, 1953–1958.

Zhang, Z., Zhu, Z., Watabe, K., Zhang, X., Bai, C., Xu, M., Wu, F., and Mo, Y.-Y. (2013). Negative regulation of lncRNA GAS5 by miR-21. Cell Death Differ. *20*, 1558–1568.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
DNase I	New Endland Biolabs	Cat#M0303S
Biotin-11-CTP	Sigma Aldrich	Cat#4739205001
Superscipt II reverse transcriptase	Invitrogen	Cat#M1302
Microspin™ G-25 columns	GE Healthcare	Cat#27532501
Q5® High-Fidelity DNA polymerase	New England Biolabs	Cat#M0491S
AmpureXP beads	Beckman Coulter	Cat#A63880
Lysyl Endopeptidase	Wako Chemicals	Cat#12505061
Kapa Biosystems Hyperprep Plus DNA kit	Roche	Cat#07962380001
Streptavidin Magnetic Beads	New England Biolabs	Cat#S1420S
Trypsin	Promega	V5280
Rapamycin	Sigma-Aldrich	Cat#553210
Everolimus	Sigma-Aldrich	Cat#SML2282
Temsirolimus	Sigma-Aldrich	Cat#PZ0020
AZD8055	Selleckchem	Cat#S1555
BEZ235	Selleckchem	Cat#S1009
CellTiter 96® AQ_{ueous} One Solution Cell Proliferation Assay (MTS)	Promega	Cat#G3582
Muse Cell Cycle Reagent	Merck Millipore	Cat#MCH100106
Acridine orange	Thermo Fisher Scientific	Cat#A1301
Random primers	Invitrogen	Cat#48190-011
RNaseOUT recombinant ribonuclease inhibitor	Invitrogen	Cat#10777019)
SensiFAST Probe Hi-ROX mix	Bioline	Cat#BIO-82020
Hs03671981_s1	Thermo Fisher Scientific	Cat#4331182
Hs03464472_m1	Thermo Fisher Scientific	Cat#4331182
Muse® Annexin V & Dead Cell Kit	Merck Millipore	Cat#MCH100105
Muse® Count & Viability Kit	Merck Millipore	Cat#MCH100102
Sperimidine	Sigma-Aldrich	Cat#S2626
Silencer Negative Control No. 1 siRNA	Thermo Fisher Scientific	Cat#AM4611
GAS5 siRNA#2	QIAGEN	Cat#SI03652544
GAS5 siRNA#1	QIAGEN	Cat#SI03652537
Critical Commercial Assays		
RNeasy Mini Kit	QIAGEN	Cat#74104
PureLink™ PCR purification kit	Invitrogen	Cat#K310001
Muse annexin V and dead cell assay kit	Merck Millipore	Cat#MCH100105
Cell Count and Viability Kit	Merck Millipore	Cat#MCH100102
Direct-zol RNA MiniPrep kit	ZYMO RESEARCH	Cat#R2050
Omniscript® RT kit	QIAGEN	Cat#205111
Mycoalert Mycoplasma Detection Kit	Lonza	Cat#LT07-318
Deposited Data		
Gas5 SHAPE data (*in vitro* data 1)	This Paper	https://doi.org/10.17632/98cwbscjvn.1
Gas5 SHAPE data (*in vitro* data 2)	This Paper	https://doi.org/10.17632/98cwbscjvn.1
Gas5 SHAPE data (in cellulo)	This Paper	https://doi.org/10.17632/98cwbscjvn.1

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Gas5 pull-down mass spectrometry proteomics data	This Paper	PXD018584
Experimental Models: Cell Lines		
CCRF-CEM cell line (Subclone C7KM1)	ATCC	Cat#CCL-119
Jurkat, Clone E6-1	ATCC	Cat#TIB-152
Oligonucleotides		
Human GAS5 Exon 12: Forward primer (CTTCTGGGCTCAAGTGATCCT)-	Applied Biosystems	CTTCTGGGCTCAAGTGATCCT
Gas5_1-15_FW	IDT	TTTCGAGGTAGGAGTC
Human GAS5 Exon 12 Reverse primer	Applied Biosystems	TTGTGCCATGAGACTCCATCAG
pcDNA3_Gas5_FW	IDT	GCTAGCGTTTAAACTTAAGCTTGG
Gas5_400_RE	IDT	GATAACAGGTCTGCCTGCAT
Gas5_227_FW	IDT	TCCTGGTAACGTTTTTATCC
RT01	IDT	GGATTGCAAAAATT TATTAAAATTGGAGACA
Recombinant DNA		
pcDNA3.1: Gas5 (1-167) 5′H3mut	This paper	N/A
pcDNA3.1: Gas5 (163-545) cH4mut	This paper	N/A
pcDNA3.1: Gas5 (163-545) cH6mut	This paper	N/A
pcDNA3.1: Gas5 (163-545) cH11mut	This paper	N/A
pcDNA3.1: Gas5 (163-545) cH12mut	This paper	N/A
pcDNA3.1: Gas5	Hudson et al., 2014	N/A
pcDNA3.1: Gas5 (1-167)	This paper	N/A
pcDNA3.1: Gas5 (163-545)	This paper	N/A
pcDNA3.1: Gas5 (1-167) 5′H2mut	This paper	N/A
Software and Algorithms		
ShapeMapper 2	Siegfried et al., 2014	https://weekslab.com/software/
DeltaSHAPE	Smola et al., 2015a	https://weekslab.com/software/
RNAstructure	Xu and Mathews, 2016	https://rna.urmc.rochester.edu/RNAstructure.html
MaxQuant v1.6.01	MaxQuant	https://www.maxquant.org
Thermo Foundation 2.0	Thermo Fisher	https://www.thermofisher.com/us/en/home/industrial/mass-spectrometry/mass-spectrometry-software.html

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Eric Ortlund (eortlun@emory.edu).

Materials Availability

Plasmids generated in this study are available upon request.

Data and Code Availability

Raw sequencing data used for SHAPE analyses were deposited at Mendeley Data: https://data.mendeley.com/datasets/98cwbscjvn/draft?a=c5ba08ab-66e3-4e4f-9f4a-25de8676cbe4

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD018584.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Jurkat cell line (Clone E6-1) and apoptosis-sensitive cloned CEM-C7 CKM1 cell line (CCRF-CEM) were obtained from ATCC. Cells were maintained in RPMI-1640 medium (Sigma) supplemented with 10% heat inactivated fetal calf serum (FCS; Hyclone), 2 mM L-glutamine and 200 μg/ml gentamycin (Sigma), at 37°C in a 5% CO2 humidified incubator. Both cell lines were mycoplasma-free as tested with the Mycoalert Mycoplasma Detection Kit (Lonza, Slough, UK).

METHOD DETAILS

Preparation of RNA

Full-length GAS5 RNA was generated by *in vitro* transcription using recombinant T7 RNA polymerase and pcDNA3 containing GAS5 (NR_002578.2 bases 1–631). The *in vitro* transcription reaction was carried out in 50 mM Tris-HCl, pH 7.5, 25 mM $MgCl_2$, 5 mM DTT and 2 mM spermidine with 2 mM of each NTP. Recombinant T7 RNA polymerase (0.1 mg/ml) was added and the reaction was incubated for 4 h at 37°C. Sample was DNaseI treated, extracted with phenol/chloroform/IAA, and the RNA was PAGE-purified. Biotinylated RNA for pull-down experiments was prepared under the same conditions with additional 0.5mM biotin-CTP present in the transcription reaction.

Synthesis of 1-methyl-7-nitroisatoic anhydride

1M7 was synthesized as described previously (Turner et al., 2013).

Modification of *in vitro* transcribed GAS5 RNA with 1M7

SHAPE-MaP was carried out according to published protocols (Smola et al., 2015b). For each reaction 500ng RNA was prepared in 12 μL water, heated to 95 C for 2minutes, and placed on ice immediately. 6 μL of 3.3x folding buffer (333 mM HEPES, pH8.0, 333 mM NaCl, 33 mM $MgCl_2$) was added and the sample was incubated at 37 C for 20 minutes to fold. 9 μL of sample was transferred to a tube containing either 1 μL 100mM 1M7 in DMSO or plain DMSO as the no reagent control. After 75 s incubation at 37C the sample was placed on ice and 40 μL water was added.

For the denaturing control reaction 250ng RNA was prepared in 3 μl water. 3 μL formamide and 1μL 10x denaturing control buffer (500 mM HEPES pH 8.0, 40 mM EDTA) was added and the sample was incubated at 95 C for 2 minutes to denature the RNA. The hot sample was then transferred to a new tube containing 1 μL 100 mM 1M7, mixed well, and immediately incubated at 95 C for one minute. Sample was then placed on ice and 40 μL water was added.

Finally, all samples were purified using the RNeasy Mini Kit (QIAGEN) according to the manufacturer's instructions. The elution volume was 150 μL.

Modification of cellular RNA with 1M7

Cells were plated in 6-well plates at 0.8×10^6 cells/well and grown for 4 days to reach density arrest under which endogenous GAS5 levels are increased(Smith and Steitz, 1998). Cells were washed with 1 mL of PBS, then 900 μL fresh growth medium and 100 μL 100 mM 1M7 (or plain DMSO) were added, and the solution was immediately mixed by swirling the culture plate. Samples were incubated for 5 minutes at 37 C, after which total RNA was extracted using TRIzol™ (Invitrogen).

Reverse transcription of 1M7 modified RNA

1 μL of primer RT01 at 2 μM was added to 10 μL of purified, modified (or control) RNA (or 3 μg of RNA modified *in cellulo*) and the samples were incubated at 65 C for 5 minutes. After cooling on ice 8 μL 2.5x SHAPE MaP buffer (125 mM Tris (pH 8.0), 187.5 mM KCl, 15 mM $MnCl_2$, 25 mM DTT and 1.25 mM dNTPs) was added and the samples were incubated at 42 C for 5 minutes. Finally, 1 μL of Superscipt II reverse transcriptase (Invitrogen) was added and the samples were incubated at 42 C for 3 hours. Reverse transcriptase was inactivated by incubation 70 C for 15 minutes and samples were purified using Microspin™ G-25 columns (GE Healthcare).

Amplification of cDNA for SHAPE-MaP

Two overlapping amplicons of GAS5 of ~400 nt length were generated by PCR using Q5® High-Fidelity DNA polymerase (NEB). Primers are listed in the Key Resources Table. 50μL PCR reactions were prepared and purified using a PureLink™ PCR purification kit (Invitrogen).

Library preparation and high-throughput sequencing

PCR amplicons were cleaned using the AmpureXP beads (Beckman Coulter). The DNA was quantified using a fluorometric based method and 10 ng was used for library preparation using the Kapa Biosystems Hyperprep Plus DNA kit. The amplicons were sequenced on an Illumina Miseq instrument in a paired-end 300 cycle format.

The raw sequencing data is available at: https://data.mendeley.com/datasets/98cwbscjvn/draft?a=c5ba08ab-66e3-4e4f-9f4a-25de8676cbe4

SHAPE data analysis

SHAPE reactivity data was analyzed using the *ShapeMapper 2* software, the arc plot was generated using the *SuperFold* software (Siegfried et al., 2014), and SHAPE differences were calculated using the *DeltaSHAPE* software (Smola et al., 2015a) (all can found be found at https://weekslab.com). GAS5 secondary structure was modeled using the *RNAstructure* software (https://rna.urmc.rochester.edu/RNAstructure.html) (Xu and Mathews, 2016) and *in vitro* SHAPE reactivity data.

Mutation rates of the *in cellulo* modified sample was lower than *in vitro* samples (0.06% median rate versus 0.09%, respectively; Figures S2 and S3), most likely owing to limited diffusion of 1M7 into the cells. Mutation rates of denaturing controls for *in cellulo* samples (generated by RNA extraction and exposure to 1M7 under denaturing conditions similar to *in vitro* samples) were too low and not usable. Instead we used denaturing control data from *in vitro* SHAPE samples for *in cellulo* samples as well.

The resulting quality data generated from *in cellulo* samples was sufficient for analysis as evidenced by small relative standard errors of SHAPE reactivities (Figure S2) and good agreement in the region of overlap between the two separate PCRs reactions employed to amplify the GAS5 cDNA: the two amplicons overlap between nucleotides 248 and 379 and SHAPE reactivities calculated independently for each amplicon correlate well in that region (Spearman R = 0.70; Figure 2C).

Effects of GAS5 constructs on the basal survival of Jurkat and CEM-C7 leukemic T cells

Cells were transfected with GAS5 constructs, or empty pcDNA3 vector as control. After 24 h, cells were replated at a fixed density for analysis after a further 48 h. Cellular GAS5 levels for Jurkat and CEM-C7 cells were determined by RT-qPCR at 24h post transfection using two different Taqman assays, Hs03671981_s1 (exon 1) and Hs03464472_m1 (exon boundary 10-11, position 409).

Apoptosis was routinely determined by assessment of nuclear morphology by fluorescence microscopy after staining with acridine orange (25 μg/ml); cells containing condensed or fragmented chromatin were scored as apoptotic. Apoptosis level was also determined by flow cytometry using a Muse annexin V and dead cell assay kit, according to the manufacturer's protocol (Merck Millipore, Cat # MCH100105). Representative images of Annexin V stained Jurkat cells is shown in Figure S4.

Cell viability was determined using a commercial Cell Count and Viability Kit (Merck Millipore; Cat # MCH 100102) and a Muse flow cytometer (Merck Millipore, Darmstadt, Germany) (Mohammed et al., 2016).

Cell cycle analysis was carried using nuclear propidium iodide (PI)-staining procedure and flow cytometry (Mohammed et al., 2016). Transfected cells were harvested 24 hours post-transfection and re-plated in fresh medium at 2 × 10^5 cells/well in 6-well plates. Following incubation for 48 hours, 1 million cells were washed in PBS before re-suspending the pellet in 200 μL PBS. Cells were then fixed in 1 mL ice cold (70% ethanol / 30% PBS) and stored at –20°C for at least 3 hours prior to cell cycle analysis. Fixed cells were centrifuged for 5 minutes at 1500 rpm. The supernatant was discarded and the cell pellet was re-suspended in 200 μL of Muse Cell Cycle Reagent (Merck Millipore # MCH100106). Cells were incubated for 30 minutes in the dark and data acquisition was carried out using the Millipore Muse cell analyzer (Mohammed et al., 2016).

Effects of GAS5 constructs on cell growth in the presence of mTOR inhibitors

Jurkat cells were transfected with one of the constructs or empty vector. Cells were treated with the mTOR inhibitors 24 post transfection. Cell growth was measured using MTS after 48 and 72 hours. Results are represented as the percentage inhibition of cell proliferation relative to vehicle-treated cells.

For experiments with knock-down, Jurkat cells were transfected with specific GAS5 siRNAs (SI03652537 targeting exon 3 = siRNA 1; SI03652544 targeting exon 7 = siRNA2) or negative control siRNA ((-)siRNA) and cultured at 37°C. After 48h, cells were transfected with one of the constructs or empty vector and treated with Rapamycin (2.5 mM) 24h post transfection. Cell growth was measured using MTS after 48 hours. Results are represented as the percentage inhibition of cell proliferation relative to vehicle-treated cells.

RNA-pulldown

The RNA pulldown experiments were carried out described in Panda et al. (2016). For each individual pulldown sample approximately 2 million Jurkat cells were required. Each sample was done in quadruplicates from four different passages of cells. Cells were centrifuged at 500 g for 5 minutes, washed with PBS once and resuspended in lysis buffer (10 mM HEPES (pH7.5), 100 mM KoAc, 10 mM MgoAc, 1% NP-40. 2.5 mM DTT, cOmplete Protease Inhibitor Cocktail (Roche)). Cells were snap frozen in liquid nitrogen, then thawed, and incubated for 40 minutes on a nutating platform at 4 C. Samples were then spun down at 15,000 g for 15 minutes at 4 C. Protein concentrations were measured and each pulldown was done with 1mg of total protein. Biotinylated RNA was prepared using recombinant T7 RNA Polymerase and buffers (see above). 3 μg of RNA (full-length GAS5, nucleotides 1-167, nucleotides 163-545, or the complete GAS5 reverse complement sequence) was incubated with jurkat cell lysates (1 mg total protein) for three hours on ice.

In the meantime, beads were prepared as follows. 50 μL of streptavidin magnetic beads (NEB) were used per pulldown reaction. Beads were washed three times with lysis buffer (no protease inhibitors), incubated with 1mg BSA per 200 μL beads for one hour, and washed four times with lysis buffer. Beads were then added to the RNA and jurkat cell lysates and incubated at 4 C overnight on a nutating platform. Beads were then washed four times with lysis buffer (without NP-40) and finally resuspended in 50 μL lysis buffer (without NP-40) for mass spectrometry analysis.

On bead digestion

After removing the supernatant from the IP samples, the bead solutions were resuspended in 200ul of 50 mM NH4HCO3. The samples were reduced with 1mM dithiothreitol (DTT) for 30 minutes at room temperature. This was followed by alkylation with 5mM iodoacetimide (IAA) for 30 minutes in the dark. The samples were then digested with 1:100 (w/w) lysyl endopeptidase (Wako) for 2 hours at room temperature. Trypsin (Promega) was added at 1:50 (w/w) and digestion was allowed to proceed overnight. The resulting peptides solutions were desalted with a Sep-Pak C18 column (Waters) and dried under vacuum.

LC-MS/MS analysis

Dried peptides were reconstituted in 10 μL of loading buffer (0.1% formic acid, 0.03% TFA, 1% acetonitrile). The sample (2μL) was loaded onto and eluted from a self-packed C18 fused silica column (25 cm x 75 μM internal diameter (ID); New Objective, Woburn, MA) driven by a Easy nLC 1200 coupled to a Fusion mass spectrometer (ThermoFisher Scientific, San Jose, CA). Elution was performed over a 140-minute gradient at a rate of 200 nl/min with buffer B ranging from 3% to 99% (buffer A: 0.1% formic acid in water, buffer B: 0.1% formic in 80% acetonitrile). The mass spectrometer cycle was programmed to collect at the top speed for 3 s cycles with higher-energy collision dissociation (HCD) fragmentation. The MS scans (350-1500 m/z range, 200,000 AGC, 50 ms maximum ion time) were collected at a resolution of 120,000 at m/z 200 in profile mode while the HCD MS/MS spectra (1.2 m/z isolation width, 30% collision energy, 10,000 AGC target, 35 maximum ion time) were detected in the ion trap. Dynamic exclusion was set to exclude previous sequenced precursor ions for 20 s within a 10 ppm window. Precursor ions with +1, and +7 or higher charge states were excluded from sequencing.

Protein Identification and Quantification with MaxQuant

RAW data was analyzed using MaxQuant v1.6.01 with Thermo Foundation 2.0 for RAW file reading capability. The search engine Andromeda was used to build and search a concatenated target-decoy UniProt Knowledgebase (UniProtKB) containing both Swiss-Prot and TrEMBL human reference protein sequences (90,307target sequences downloaded April 21, 2015), plus 245 contaminant proteins included as a parameter for Andromeda search within MaxQuant (Cox et al., 2011). Methionine oxidation (+15.9949 Da), and protein N-terminal acetylation (+42.0106 Da) were included as variable modifications (up to 5 allowed per peptide); cysteine was assigned a fixed carbamidomethyl modification (+57.0215 Da). Only fully tryptic peptides were considered with up to 2 missed cleavages in the database search. A precursor mass tolerance of ± 20 ppm and a 0.6 Da product ion tolerance was used. Other search settings included a maximum peptide mass of 4,600 Da, a minimum peptide length of 7 residues and match between runs. The false discovery rate (FDR) for peptide spectral matches, proteins, and site decoy fraction were all set to 1 percent. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PXD018584.

QUANTIFICATION AND STATISTICAL ANALYSIS

Assays for determination of cell viability, apoptosis, survival, and cell cycle

Details about statistical analyses of these experiments can be found in the legend of Figures 3 and 4.

Mass Spectrometry Quantitative Data Processing and Statistics

MaxQuant grouped protein quantification based on the MaxLFQ normalization algorithm (Cox et al., 2014) was extracted from MaxQuant output for 3,108 non-decoy proteins. Of these, 3,026 proteins were quantified in at least 3 out of 4 technical replicate LC-MS/MS run RAW data for at least one of the 4 different RNA baits. This filtering is important because imputation according to an informative missingness assumption, or missing not at random (MNAR) values remain for these well-quantified proteins with consistent non-missing quantification. LFQ normalized data for all but the outlier sample, determined later, had a Pearson correlation of at least 0.98 to unnormalized summed intensity data. Missing value imputation was performed in R according to the algorithm for LFQ intensity imputation described by Tyanova et al., with each sample's missing values imputed with a random distribution within 0.3 SD of the population mean minus 1.8 SD of that sample's population of nonmissing quantitations (Tyanova et al., 2016). Then outlier detection was performed using a 3SD cutoff for the z-scored connectivity measure calculated by the R WGCNA package fundamentalNetworkConcepts function on an adjacency matrix calculated for the imputed LFQ data iteratively until no outliers could be found; one outlier was removed (one replicate of the full-length GAS5 sample). One-way ANOVA with Tukey post hoc test for significance of differences between pairs of grouped samples was performed in R using base functions and TukeyHSD. Then volcano plots using Tukey p values were plotted in R using ggplot2 and plotly packages using an in-house script.