Supporting Information: Characterizing the structure-function relationship of a naturally-occurring RNA thermometer

Sarai Meyer[1], Paul D. Carlson[1] and Julius B. Lucks[2*]

[1] Robert F. Smith School of Chemical and Biomolecular Engineering, Cornell University; 120 Olin Hall; Ithaca, NY 14853; USA

[2] Department of Chemical and Biological Engineering, Northwestern University; 2145 Sheridan Rd.; Evanston, IL 60208; USA

*Corresponding author: jblucks@northwestern.edu; Telephone: (847) 467-2943

Supplementary Figures

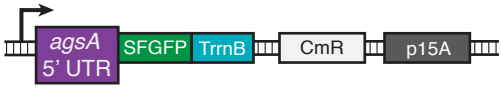

Figure S1. Plasmid architecture of the *agsA* testing constructs. The wild-type (WT) or mutant *agsA* 5' UTR (purple) is fused directly after the J23119 promoter (black arrow), followed by the fluorescent reporter SFGFP (green) and the rrnB intrinsic terminator (blue). The plasmid also harbors the chloramphenicol resistance selection marker (light gray) and the p15a origin of replication (dark gray).

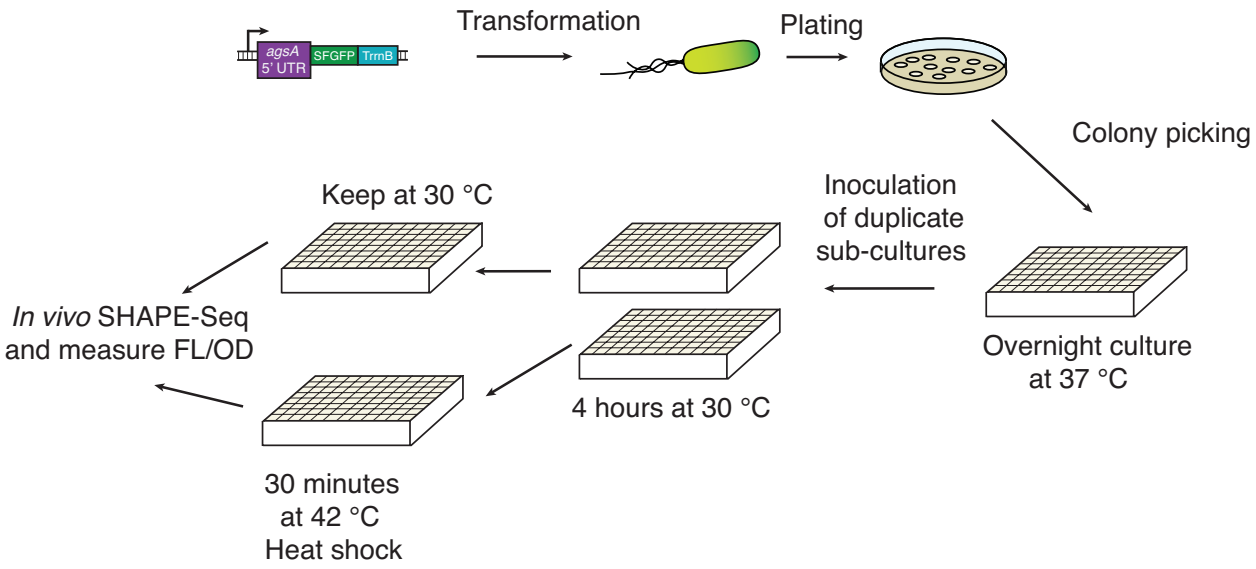

Figure S2. Overview of the experimental heat shock protocol for testing the *agsA* thermometer variants and control *in vivo*. The 5' untranslated region of the *agsA* thermometer (WT, G21C and A29C point mutants) was fused to the gene encoding superfolder fluorescent green protein (SFGFP) followed by the rrnB transcriptional terminator (TrrnB). The resulting constructs were transformed into TG1 *E. coli* and plated onto LB (lysogeny broth)/agar plates with chloramphenicol. Individual colonies were picked and grown overnight in 1 mL LB with chloramphenicol in culture blocks at 37 °C and 1000 rpm, then sub-cultured in 1.2 mL LB with chloramphenicol for 4 hours at 30 °C and 1000 rpm before transfer to 42 °C and 1000 rpm (heat shock condition) for 30 minutes or remaining at 30 °C (no heat shock condition). Cultures were then subjected to in-cell SHAPE-Seq and measurement of fluorescence (FL) and optical density (OD) as described in Experimental Procedures.

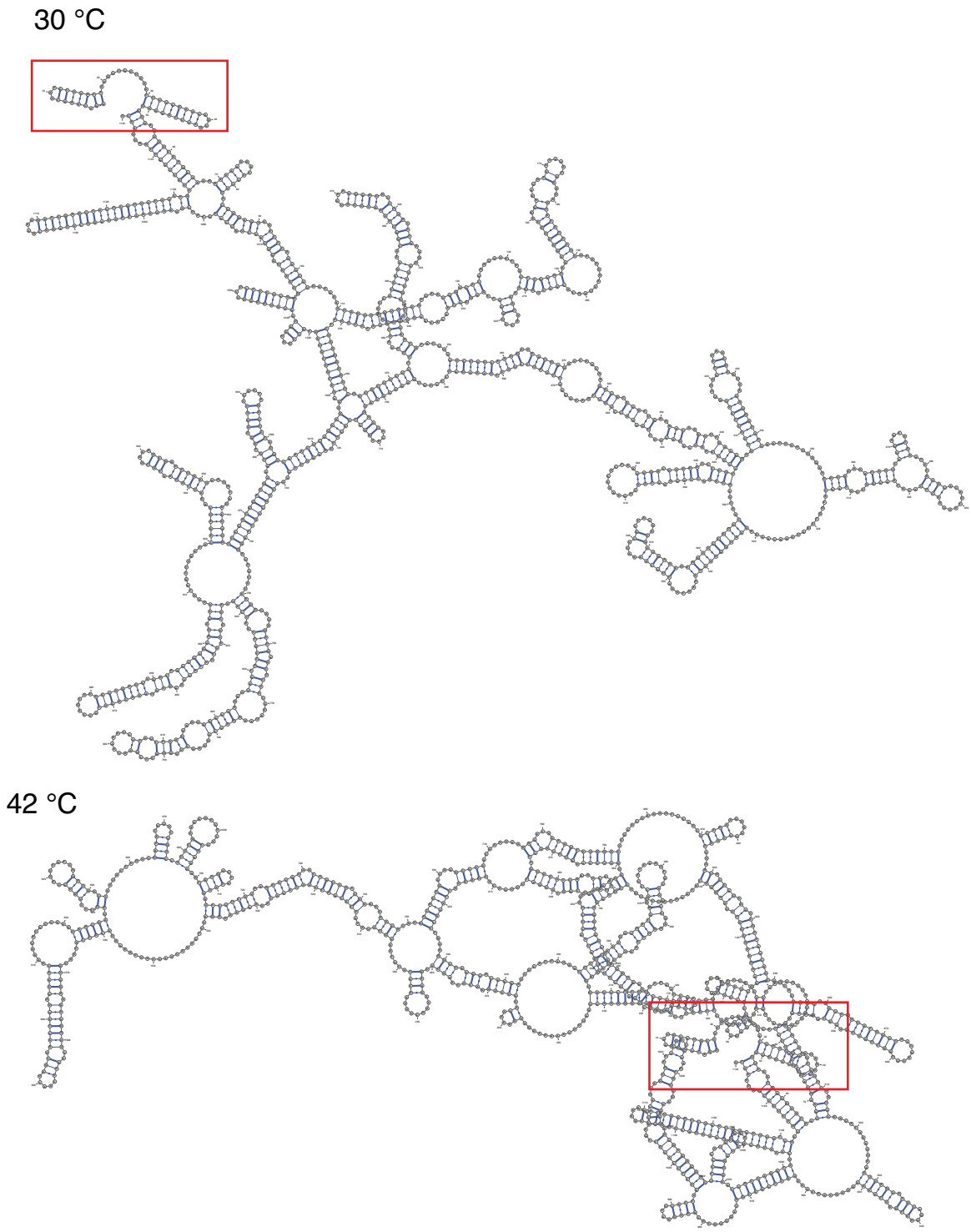

Figure S3.

Figure S3. Predicted minimum free energy *agsA* WT construct structures for full mRNA sequence *in vivo*. *In vivo* SHAPE-Seq restraints incorporated by RNAStructure `Fold`[1] covered 5' UTR (red box) and SFGFP leader sequence up until the site of reverse transcription primer binding. The remainder of the RNA was folded without SHAPE-Seq restraints.

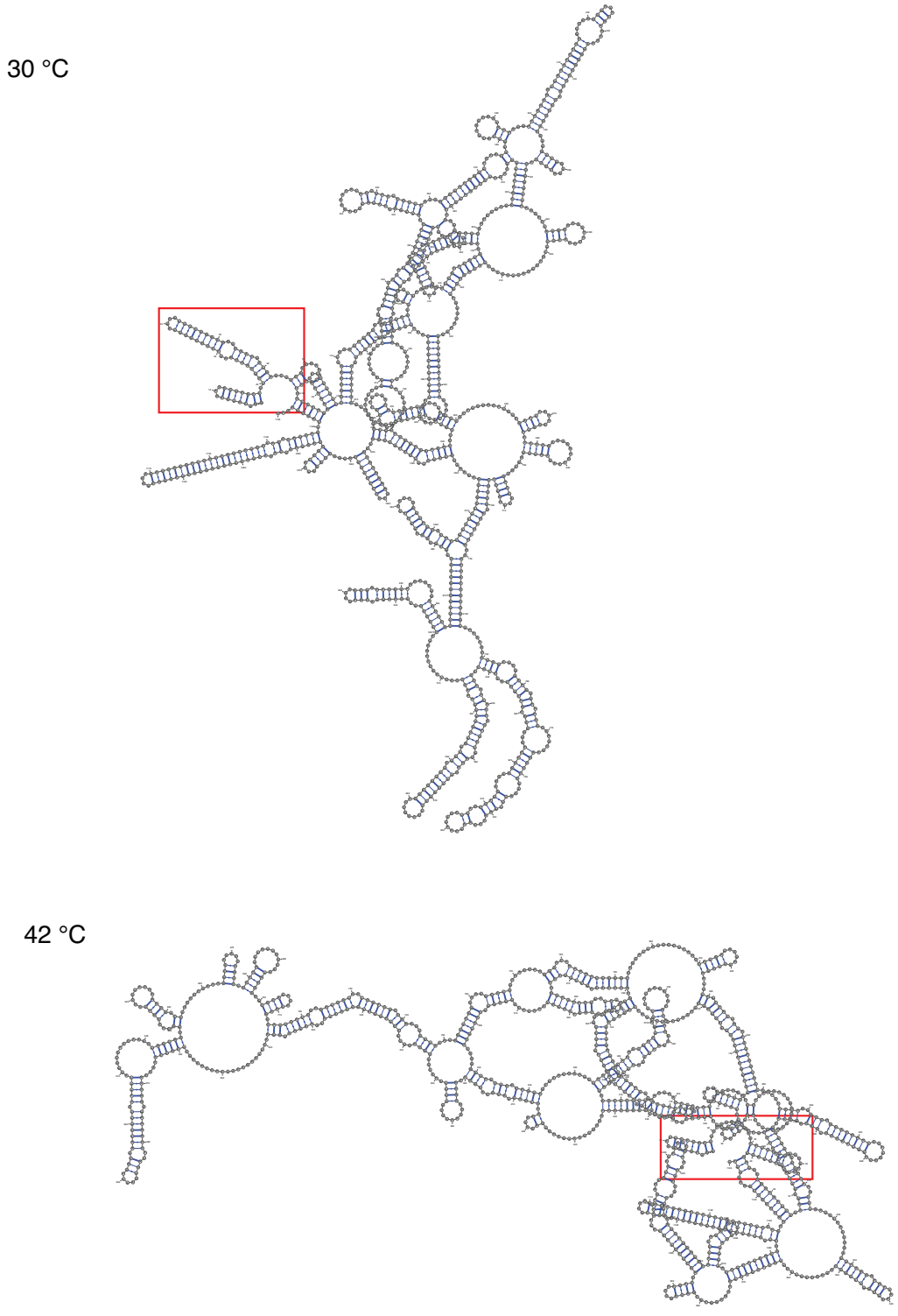

Figure S4.

Figure S4. Predicted *agsA* G21C construct structures for full mRNA sequence *in vivo*. *In vivo* SHAPE-Seq retraints incorporated by RNAStructure `Fold`[1] covered the 5' UTR (red box) and SFGFP leader sequence up until the site of reverse transcription primer binding. The rest of the RNA was folded without SHAPE-Seq restraints.

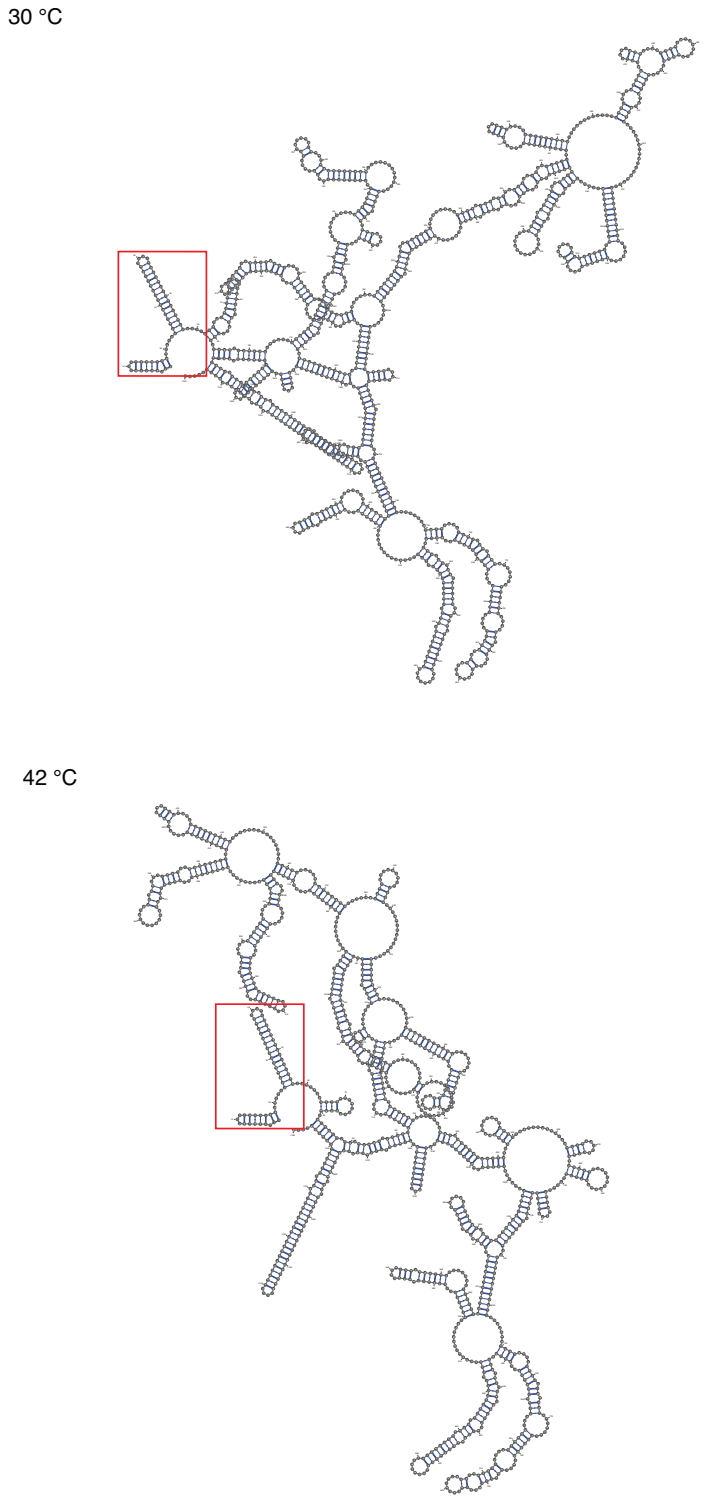

Figure S5.

Figure S5. Predicted *agsA* A29C construct structures for full mRNA sequence *in vivo*. *In vivo* SHAPE-Seq restraints incorporated by RNAStructure `Fold`[1] covered the 5' UTR (red box) and SFGFP leader sequence up until the site of reverse transcription primer binding. The remainder of the RNA was folded without SHAPE-Seq restraints.

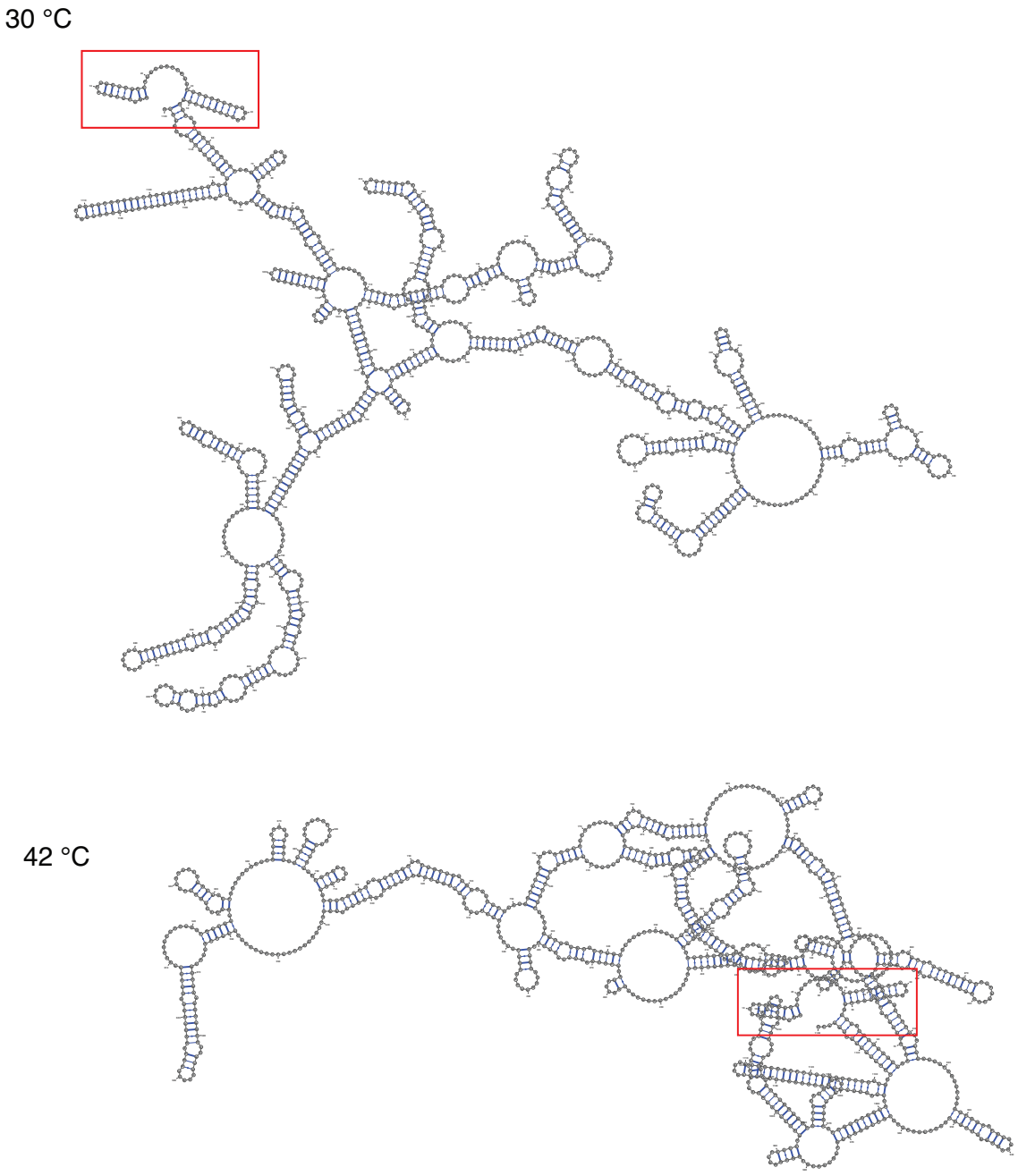

Figure S6. Predicted *agsA* WT construct structures for full mRNA sequence *in vitro* in PURExpress without ribosomes. *In vitro* SHAPE-Seq restraints incorporated by RNAStructure `Fold`[1] covered the 5' UTR (red box) and SFGFP leader sequence up until the site of reverse transcription primer binding. The remainder of the RNA was folded without SHAPE-Seq restraints.

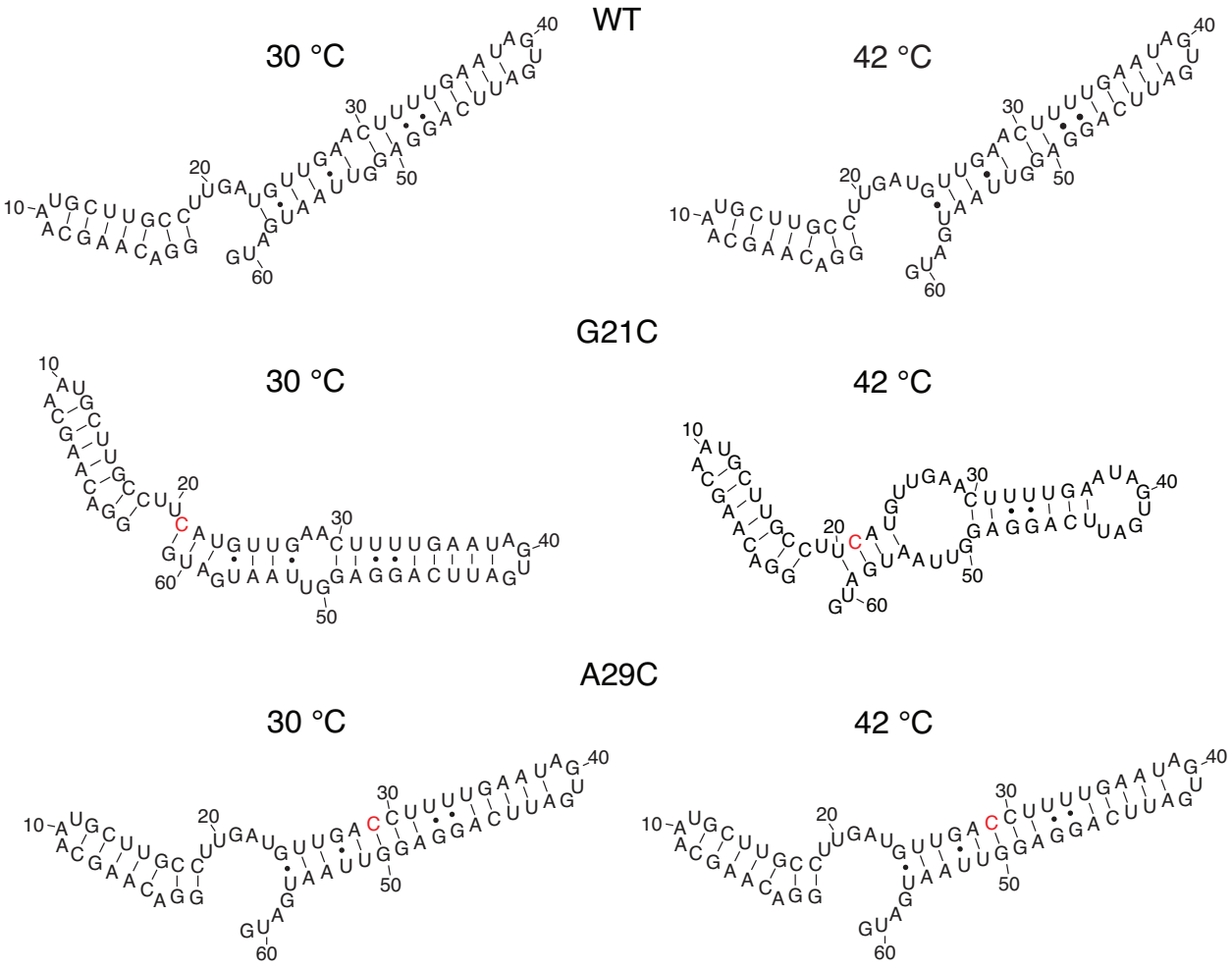

Figure S7. Unrestrained minimum free energy predictions of *agsA* construct 5' UTR structure through the start codon. Predicted *agsA* variant structures at 30 °C and 42 °C in the absence of SHAPE reactivity restraints, as predicted by RNAStructure `Fold`[1]. Point-mutations are highlighted in red.

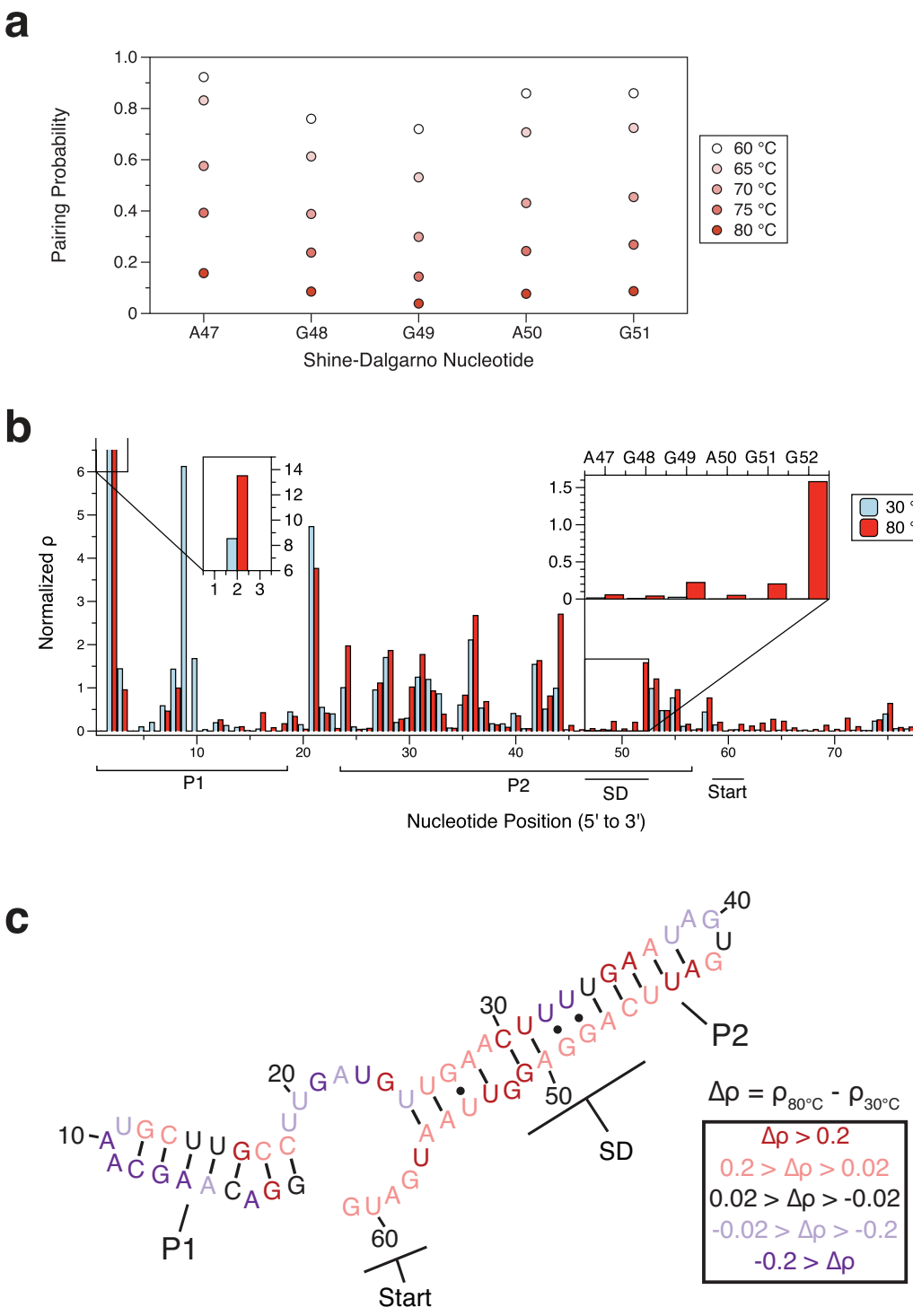
a
Pairing Probability
1.0
0.8
0.6
0.4
0.2
0
A47
G48
G49
A50
G51
Shine-Dalgarno Nucleotide
60 °C
65 °C
70 °C
75 °C
80 °C
b
Normalized ρ
A47 G48 G49 A50 G51 G52
30 °C
80 °C
P1
P2
SD
Start
Nucleotide Position (5' to 3')
c
P1
P2
SD
Start
Δρ = ρ80°C - ρ30°C
Δρ > 0.2
0.2 > Δρ > 0.02
0.02 > Δρ > -0.02
-0.02 > Δρ > -0.2
-0.2 > Δρ

Figure S8

Figure S8. (a) Predicted pairing-probability plot for the *agsA* WT thermometer Shine-Dalgarno sequence in the P2 hairpin at 60-80 °C to determine temperature at which Shine-Dalgarno sequence could show increased reactivity to SHAPE probing. Probabilities determined based on the partition function calculated using RNAStructure. **(b)** SHAPE-Seq normalized reactivity (ϱ) plots for *agsA* WT from *in vitro* experiments in folding buffer performed at 30 °C and 80 °C. Inset on the left shows the full scale of reactivities at nucelotide position 2, while the inset on the right shows the reactivities of the SD sequence. **(c)** Reactivity differences between 80 °C and 30 °C (Δϱ) overlayed with the predicted *agsA* 5' UTR structure.

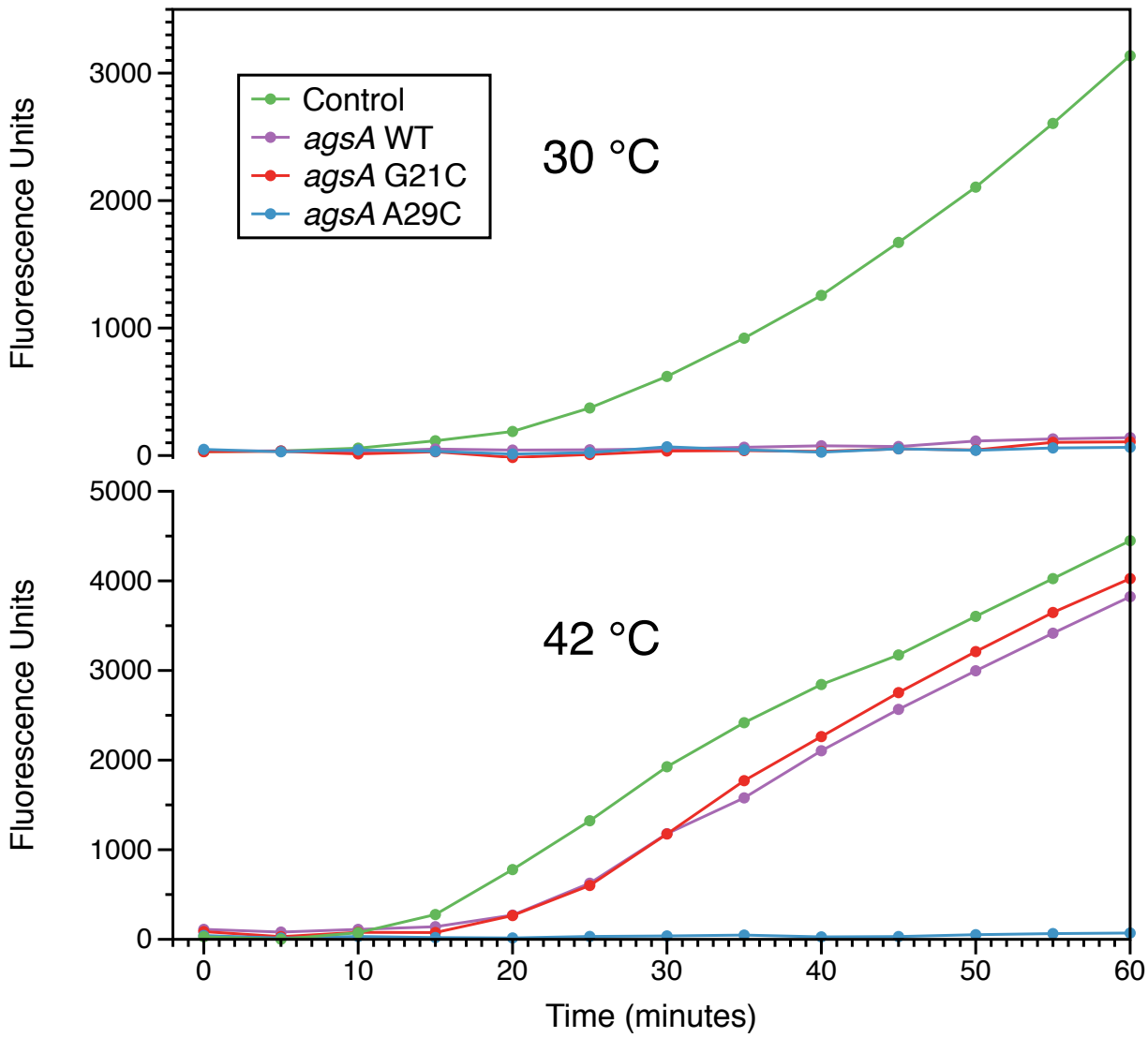

Figure S9. PURExpress testing of *agsA* constructs from plasmid DNA templates. (DNA templates consisted of a constitutive promoter followed by *agsA* 5' UTR, SFGFP gene, and intrinsic transcriptional terminator as depicted in Figure S1). Fluorescence traces demonstrate successful thermometer activation in the context of the PURExpress cell-free transcription and translation system.

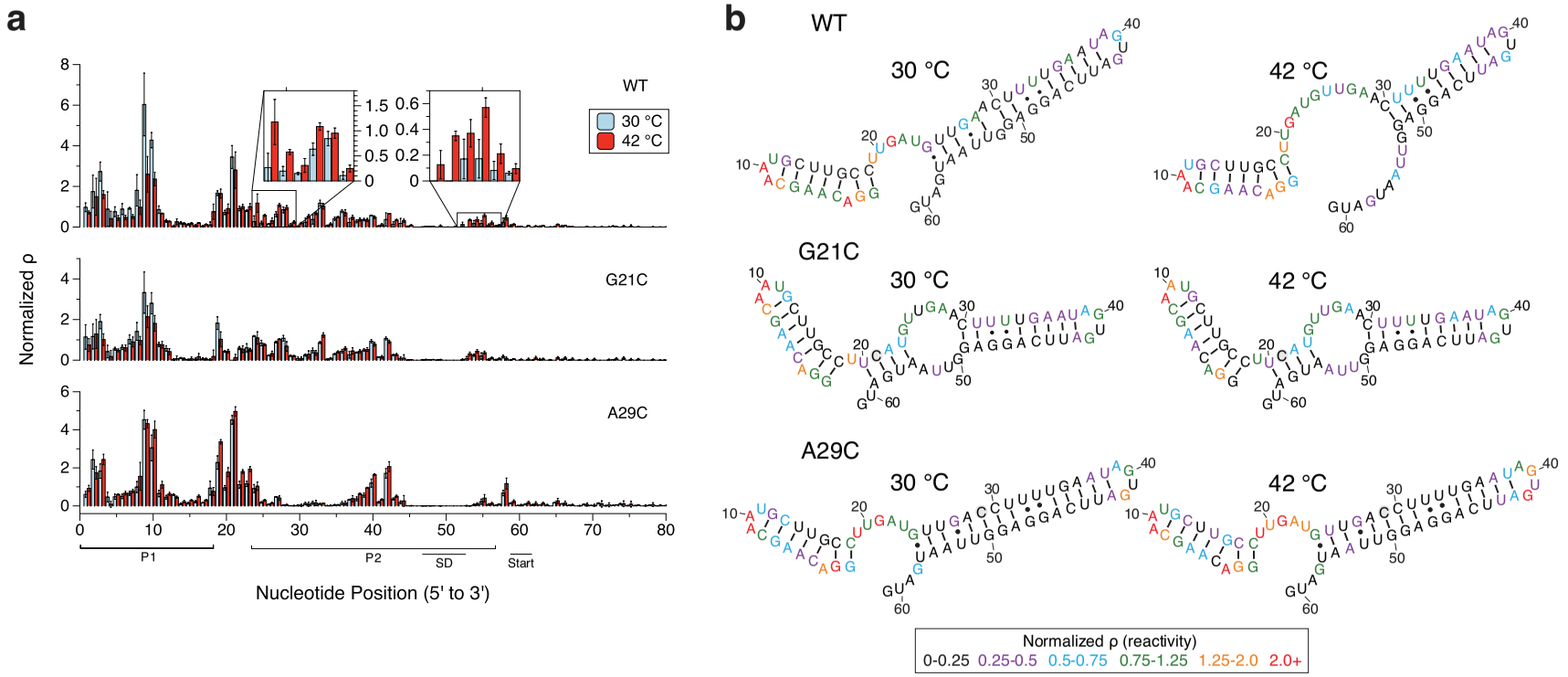
a
WT
30 °C
42 °C
G21C
A29C
Normalized ρ
P1
P2
SD
Start
Nucleotide Position (5' to 3')
b
WT
30 °C
42 °C
G21C
A29C
Normalized ρ (reactivity)
0-0.25 0.25-0.5 0.5-0.75 0.75-1.25 1.25-2.0 2.0+

Figure S10

Figure S10. SHAPE-Seq characterization of the *agsA* thermometer *in vitro* in buffer. **(a)** Normalized SHAPE-Seq reactivity plots for *agsA* WT, G21C, and A29C variants from triplicate experiments *in vitro* performed on mRNAs equilibrated at 30 °C and 42 °C. Error bars represent sample standard deviation. **(b)** Experimentally-derived minimum free energy (MFE) RNA structures as predicted using the RNAStructure `Fold`[1] program with the reactivity values from (a) taken as pseudoenergy folding restraints. Nucleotides are color-coded according to normalized SHAPE reactivity (ϱ), with red indicating high reactivity and black low reactivity. Point-mutated nucleotides are shaded in grey.

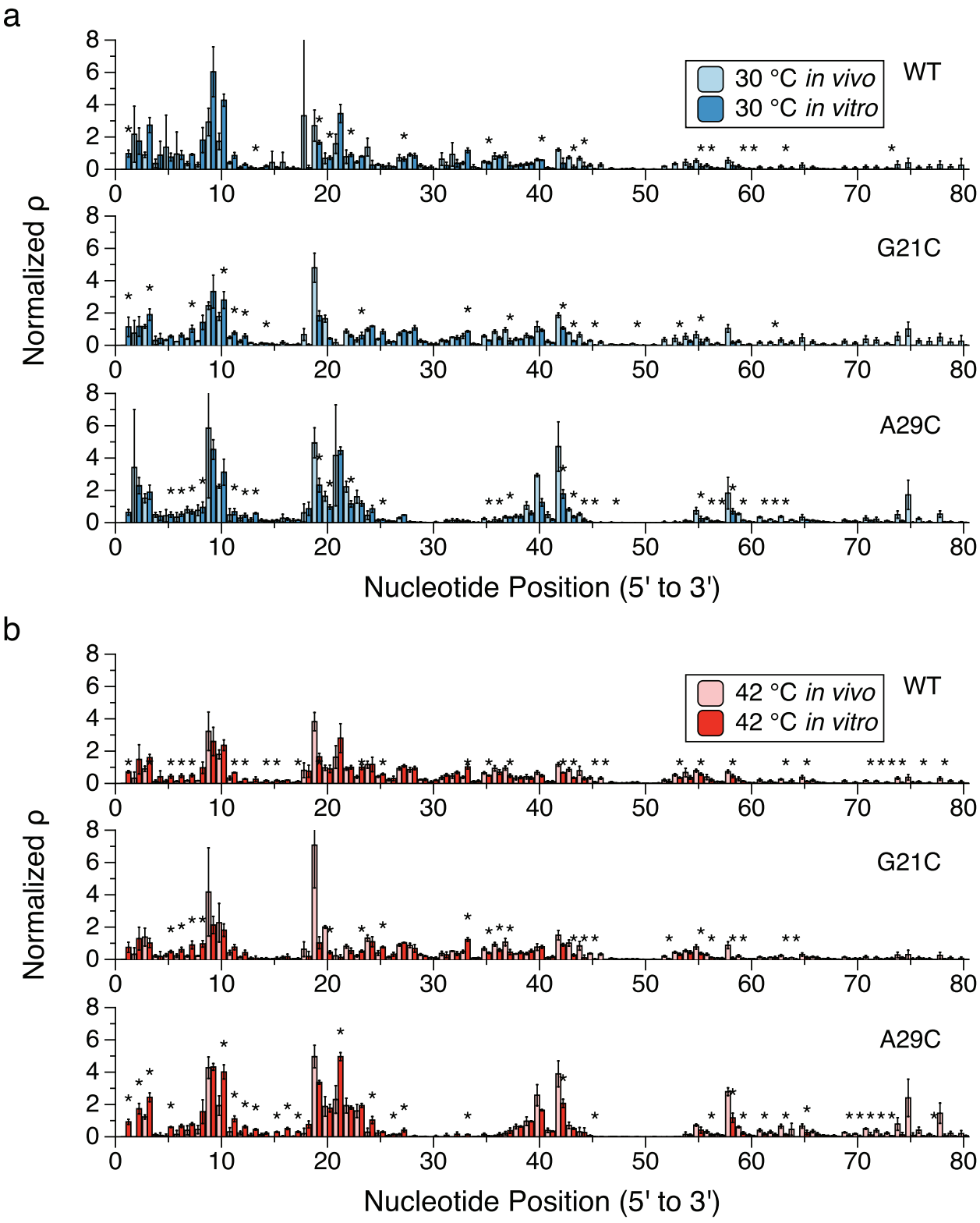
a
30 °C in vivo
30 °C in vitro
WT
G21C
A29C
Normalized ρ
Nucleotide Position (5' to 3')
b
42 °C in vivo
42 °C in vitro
WT
G21C
A29C
Normalized ρ
Nucleotide Position (5' to 3')

Figure S11

Figure S11. The effect of active translation on mRNA SHAPE-Seq reactivity. Reactivity comparisons for WT a*gsA* construct between *in vivo* and *in vitro* in buffer at 30 °C **(a)** and 42 °C **(b)**. Stars denote statistically significant reactivity differences, as determined using two-sided heteroscedastic Welch's t-tests. Error bars represent sample standard deviation.

Supplementary Tables

Table S1. Plasmids used in this study.
Color code: yellow-promoter, ---insertion sequence, green-SFGFP light blue-terminator, dark blue-chloramphenicol resistance marker, and grey-origin of replication (p15a)

Plasmid number	Description	Sequence
n/a	p15a backbone	gaattctaaagatcttgacagctagctcagtcctaggtataatactagt---------- ATGagcaaaggagaagaacttttcactggagttgtcccaattcttgttgaattagatggtgatgtt aatgggcacaaattttctgtccgtggagagggtgaaggtgatgctacaaacggaaaactcaccct taaatttatttgcactactggaaaactacctgttccgtggccaacacttgtcactactctgacctatgg tgttcaatgcttttcccgttatccggatcacatgaaacggcatgacttttcaagagtgccatgcccg aaggttatgtacaggaacgcactatatctttcaaagatgacgggacctacaagacgcgtgctgaa gtcaagtttgaaggtgataccttgttaatcgtatcgagttaaagggtattgattttaaagaagatgg aaacattcttggacacaaactcgagtacaactttaactcacacaatgtatacatcacggcagacaa acaaaagaatggaatcaaagctaacttcaaaattcgccacaacgttgaagatggttccgttcaact agcagaccattatcaacaaaatactccaattggcgatggccctgtccttttaccagacaaccattac ctgtcgacacaatctgtcctttcgaaagatcccaacgaaaagcgtgaccacatggtccttcttgagt ttgtaactgctgctgggattacacatggcatggatgagctctacaaataaggatctgaagcttggg cccgaacaaaaactcatctcagaagaggatctgaatagcgccgtcgaccatcatcatcatcatca ttgagtttaaacggtctccagcttggctgttttggcggatgagagaagattttcagcctgatacagat taaatcagaacgcagaagcggtctgataaaacagaatttgcctggcggcagtagcgcggtggtc ccacctgaccccatgccgaactcagaagtgaaacgccgtagcgccgatggtagtgtggggtctc cccatgcgagagtagggaactgccaggcatcaaataaaacgaaaggctcagtcgaaagactgg gccttttcgttttatctgttgtttgtcggtgaactggatccttactcgagtctagactgcagttgatcgggg cacgtaagaggttccaactttcaccataatgaaataagatcactaccgggcgtatttttgagttatc gagattttcaggagctaaggaagctaaaatggagaaaaaaatcactggatataccaccgttgatat atcccaatggcatcgtaaagaacattttgaggcatttcagtcagttgctcaatgtacctataaccag accgttcagctggatattacggccttttaaagaccgtaaagaaaaataagcacaagttttatccgg cctttattcacattcttgcccgcctgatgaatgctcatccggaatttcgtatggcaatgaaagacggt gagctggtgatatgggatagtgttcaccttgttacaccgttttccatgagcaaactgaaacgtttc atcgctctggagtgaataccacgacgatttccggcagtttctacacatatattcgcaagatgtggcg tgttacggtgaaaacctggcctatttccctaaagggtttattgagaatatgtttttcgtctcagccaat ccctgggtgagtttcaccagttttgatttaaacgtggccaatatggacaacttcttcgcccccgtttc accatgggcaaatattatacgcaaggcgacaaggtgctgatgccgctggcgattcaggttcatca tgccgtttgtgatggcttccatgtcggcagaatgcttaatgaattacaacagtactgcgatgagtgg cagggcggggcgtaatttgatatcgagctcgcttggactcctgttgatagatccagtaatgacctc agaactccatctggatttgttcagaacgctcggttgccgccgggcgtttttttattggtgagaatcca agcctccgatcaacgtctcattttcgccaaaagttggcccagggcttcccggtatcaacagggac accaggatttatttattctgcgaagtgatcttccgtcacaggtatttattcggcgcaaagtgcgtcgg gtgatgctgccaacttactgatttagtgtatgatggtgtttttgaggtgctccagtggcttctgtttctat cagctgtccctcctgttcagctactgacggggtggtgcgtaacggcaaaagcaccgccggacat cagcgctagcggagtgtatactggcttactatgttggcactgatgagggtgtcagtgaagtgcttc atgtggcaggagaaaaaaggctgcaccggtgcgtcagcagaatatgtgatacaggatatattcc

		gcttcctcgctcactgactcgctacgctcggtcgttcgactgcggcgagcggaaatggcttacga acggggcggagatttcctggaagatgccaggaagatacttaacagggaagtgagagggccgc ggcaaagccgtttttccataggctccgccccctgacaagcatcacgaaatctgacgctcaaatc agtggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggcggctccctcgt gcgctctcctgttcctgcctttcggtttaccggtgtcattccgctgttatggccgcgtttgtctcattcc acgcctgacactcagttccgggtaggcagttcgctccaagctggactgtatgcacgaacccccc gttcagtccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggaaagacatgca aaagcaccactggcagcagccactggtaattgatttagaggagttagtcttgaagtcatgcgccg gttaaggctaaactgaaaggacaagttttggtgactgcgctcctccaagccagttacctcggttca aagagttggtagctcagagaaccttcgaaaaaccgccctgcaaggcggtttttcgttttcagagc aagagattacgcgcagaccaaaacgatctcaagaagatcatcttattaatcagataaaatatttcta gatttcagtgcaatttatctcttcaaatgtagcacctgaagtcagccccatacgatataagttgtaatt ctcatgtttgacagcttatcatcgataagcttccgatggcgcgccgagaggctttacactttatgctt ccggct
JBL 3048	WT *agsA* in p15a backbone	Insertion sequence: ggacaagcaatgcttgccttgatgttgaacttttgaatagtgattcaggaggttaatg
JBL 3057	*agsA* G21C mutant in p15a backbone	Insertion sequence: ggacaagcaatgcttgccttCatgttgaacttttgaatagtgattcaggaggttaatg
JBL 3070	*agsA* A29C mutant in p15a backbone	Insertion sequence: ggacaagcaatgcttgccttgatgttgaCcttttgaatagtgattcaggaggttaatg
JBL 001	empty backbone control	gaattctaaagatcctaactcgagtaatgagaagcttgggcccgaacaaaaactcatctcagaag aggatctgaatagcgccgtcgaccatcatcatcatcatcattgagtttaaacggtctccagcttggc tgttttggcggatgagagaagattttcagcctgatacagattaaatcagaacgcagaagcggtctg ataaaacagaatttgcctggcggcagtagcgcggtggtcccacctgaccccatgccgaactcag aagtgaaacgccgtagcgccgatggtagtgtggggtctccccatgcgagagtagggaactgcc aggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgttttatctgttgtttgtcggt gaactggatccttactcgagtctagactgcagttgatcggcacgtaagaggttccaactttcacc ataatgaaataagatcactaccgggcgtatttttgagttatcgagattttcaggagctaaggaagct aaaatggagaaaaaaatcactggatataccaccgttgatatatcccaatggcatcgtaaagaacat tttgaggcatttcagtcagttgctcaatgtacctataaccagaccgttcagctggatattacggcctt ttaaagaccgtaaagaaaaataagcacaagttttatccggcctttattcacattcttgcccgcctgat gaatgctcatccggaatttcgtatggcaatgaaagacggtgagctggtgatatgggatagtgttca cccttgttacaccgttttccatgagcaaactgaaacgttttcatcgctctggagtgaataccacgac gatttccggcagtttctacacatatattcgcaagatgtggcgtgttacggtgaaaacctggcctattt ccctaaagggtttattgagaatatgttttcgtctcagccaatccctgggtgagtttcaccagttttgat ttaaacgtggccaatatggacaacttcttcgcccccgttttcaccatgggcaaatattatacgcaag gcgacaaggtgctgatgccgctggcgattcaggttcatcatgccgtttgtgatggcttccatgtcg gcagaatgcttaatgaattacaacagtactgcgatgagtggcagggcggggcgtaatttgatatc gagctcgcttggactcctgttgatagatccagtaatgacctcagaactccatctggatttgttcaga acgctcggttgccgccgggcgttttttattggtgagaatccaagcctccgatcaacgtctcattttcg ccaaaagttggcccagggcttcccggtatcaacagggacaccaggatttatttattctgcgaagtg

		atcttccgtcacaggtatttattcggcgcaaagtgcgtcgggtgatgctgccaacttactgatttagt gtatgatggtgtttttgaggtgctccagtggcttctgtttctatcagctgtccctcctgttcagctactg acggggtggtgcgtaacggcaaaagcaccgccggacatcagcgctagcggagtgtatactgg cttactatgttggcactgatgagggtgtcagtgaagtgcttcatgtggcaggagaaaaaaggctg caccggtgcgtcagcagaatatgtgatacaggatatattccgcttcctcgctcactgactcgctac gctcggtcgttcgactgcggcgagcggaaatggcttacgaacggggcggagatttcctggaag atgccaggaagatacttaacagggaagtgagagggccgcggcaaagccgtttttccataggctc cgccccccctgacaagcatcacgaaatctgacgctcaaatcagtggtggcgaaacccgacagga ctataaagataccaggcgtttcccccctggcggctccctcgtgcgctctcctgttcctgcctttcggtt taccggtgtcattccgctgttatggccgcgtttgtctcattccacgcctgacactcagttccgggtag gcagttcgctccaagctggactgtatgcacgaacccccccgttcagtccgaccgctgcgccttatc cggtaactatcgtcttgagtccaacccggaaagacatgcaaaagcaccactggcagcagccact ggtaattgatttagaggagttagtcttgaagtcatgcgccggttaaggctaaactgaaaggacaag ttttggtgactgcgctcctccaagccagttacctcggttcaaagagttggtagctcagagaaccttc gaaaaaccgccctgcaaggcggtttttcgttttcagagcaagagattacgcgcagaccaaaacg atctcaagaagatcatcttattaatcagataaaatatttctagatttcagtgcaatttatctcttcaaatg tagcacctgaagtcagccccatacgatataagttgtaattctcatgtttgacagcttatcatcgataa gcttccgatggcgcgccgagaggctttacactttatgcttccggct
JBL 2348	gene expression control plasmid	gaattctaaagatcttttgacagctagctcagtcctaggtataatactagtgggagaaagaagaaga gatgcggaagagagaaatataacacaacaatacgtatatttatatcttccgcaaagaagaggaga ggaagaatgaaatatacgagacaacaaacaataaacagcATGagcaaaggagaagaacttttt cactggagttgtcccaattcttgttgaattagatggtgatgttaatgggcacaaattttctgtccgtgg agagggtgaaggtgatgctacaaacggaaaactcacccttaaatttatttgcactactggaaaact acctgttccgtggccaacacttgtcactactctgacctatggtgttcaatgcttttcccgttatccgga tcacatgaaacggcatgacttttcaagagtgccatgcccgaaggttatgtacaggaacgcactat atctttcaaagatgacgggacctacaagacgcgtgctgaagtcaagtttgaaggtgatacccttgt taatcgtatcgagttaaagggtattgattttaaagaagatggaaacattcttggacacaaactcgag tacaactttaactcacacaatgtatacatcacggcagacaaacaaaagaatggaatcaaagctaa cttcaaaattcgccacaacgttgaagatggttccgttcaactagcagaccattatcaacaaaatact ccaattggcgatggccctgtccttttaccagacaaccattacctgtcgacacaatctgtcctttcgaa agatcccaacgaaaagcgtgaccacatggtccttcttgagtttgtaactgctgctgggattacacat ggcatggatgagctctacaaataaggatcttagcataacccccttggggcctctaaacgggtcttga ggggttttttgtggatccttactcgagtctagactgcagttgatcgggcacgtaagaggttccaactt tcaccataatgaaataagatcactaccgggcgtatttttgagttatcgagattttcaggagctaagg aagctaaaatggagaaaaaaatcactggatataccaccgttgatatatcccaatggcatcgtaaag aacattttgaggcatttcagtcagttgctcaatgtacctataaccagaccgttcagctggatattacg gccttttaaagaccgtaaagaaaaataagcacaagttttatccggcctttattcacattcttgcccg cctgatgaatgctcatccggaatttcgtatggcaatgaaagacggtgagctggtgatatgggatag tgttcacccttgttacaccgttttccatgagcaaactgaaacgttttcatcgctctggagtgaatacca cgacgatttccggcagtttctacacatatattcgcaagatgtggcgtgttacggtgaaaacctggc ctatttccctaaagggtttattgagaatatgtttttcgtctcagccaatccctgggtgagtttcaccagt tttgatttaaacgtggccaatatggacaacttcttcgcccccgttttcaccatgggcaaatattatacg caaggcgacaaggtgctgatgccgctggcgattcaggttcatcatgccgtttgtgatggcttccat gtcggcagaatgcttaatgaattacaacagtactgcgatgagtggcagggcggggcgtaatttga tatcgagctcgcttggactcctgttgatagatccagtaatgacctcagaactccatctggatttgttc agaacgctcggttgccgccgggcgtttttttattggtgagaatccaagcctccgatcaacgtctcatt

		ttcgccaaaagttggcccagggcttcccggtatcaacagggacaccaggatttatttattctgcga agtgatcttccgtcacaggtatttattcggcgcaaagtgcgtcgggtgatgctgccaacttactgat ttagtgtatgatggtgtttttgaggtgctccagtggcttctgtttctatcagctgtccctcctgttcagct actgacggggtggtgcgtaacggcaaaagcaccgccggacatcagcgctagcggagtgtata ctggcttactatgttggcactgatgagggtgtcagtgaagtgcttcatgtggcaggagaaaaaag gctgcaccggtgcgtcagcagaatatgtgatacaggatatattccgcttcctcgctcactgactcg ctacgctcggtcgttcgactgcggcgagcggaaatggcttacgaacggggcggagatttcctgg aagatgccaggaagatacttaacagggaagtgagagggccgcggcaaagccgtttttccatag gctccgcccccctgacaagcatcacgaaatctgacgctcaaatcagtggtggcgaaacccgac aggactataaagataccaggcgtttcccccctggcggctccctcgtgcgctctcctgttcctgccttt cggtttaccggtgtcattccgctgttatggccgcgtttgtctcattccacgcctgacactcagttccg ggtaggcagttcgctccaagctggactgtatgcacgaaccccccgttcagtccgaccgctgcgc cttatccggtaactatcgtcttgagtccaacccggaaagacatgcaaaagcaccactggcagcag ccactggtaattgatttagaggagttagtcttgaagtcatgcgccggttaaggctaaactgaaagg acaagttttggtgactgcgctcctccaagccagttacctcggttcaaagagttggtagctcagaga accttcgaaaaaccgccctgcaaggcggtttttttcgttttcagagcaagagattacgcgcagacca aaacgatctcaagaagatcatcttattaatcagataaaatatttctagatttcagtgcaatttatctcttc aaatgtagcacctgaagtcagccccatacgatataagttgtaattctcatgtttgacagcttatcatc gataagcttccgatggcgcgccgagaggctttacactttatgcttccggct

Works Cited

(1) Reuter, J. S., and Mathews, D. H. (2010) RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics 11*, 129.