



Figure S2: RNA localization, fluorescent protein content and 3' UTR sequence controls.

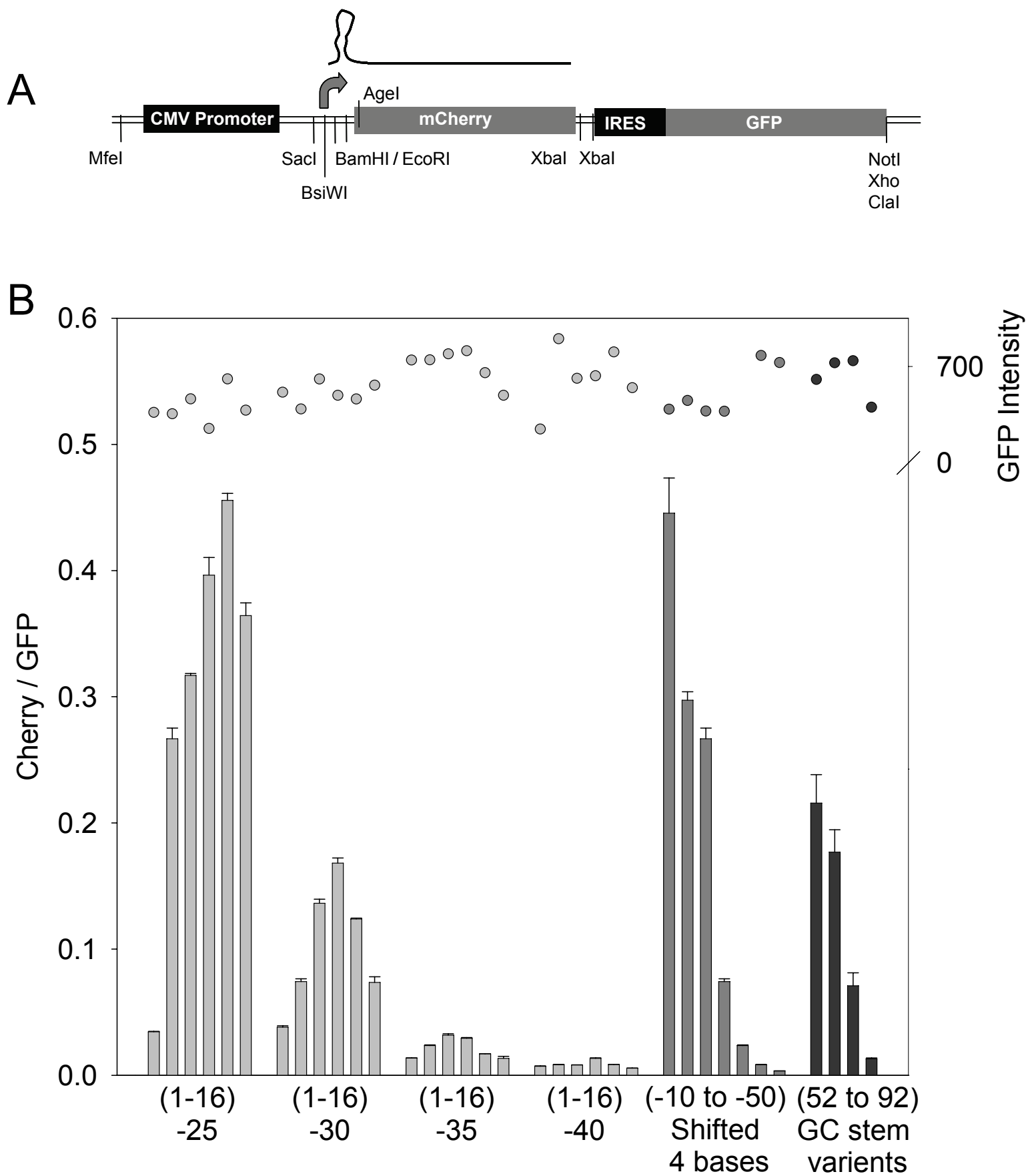


Figure S3: Translation efficiencies in CHO cells

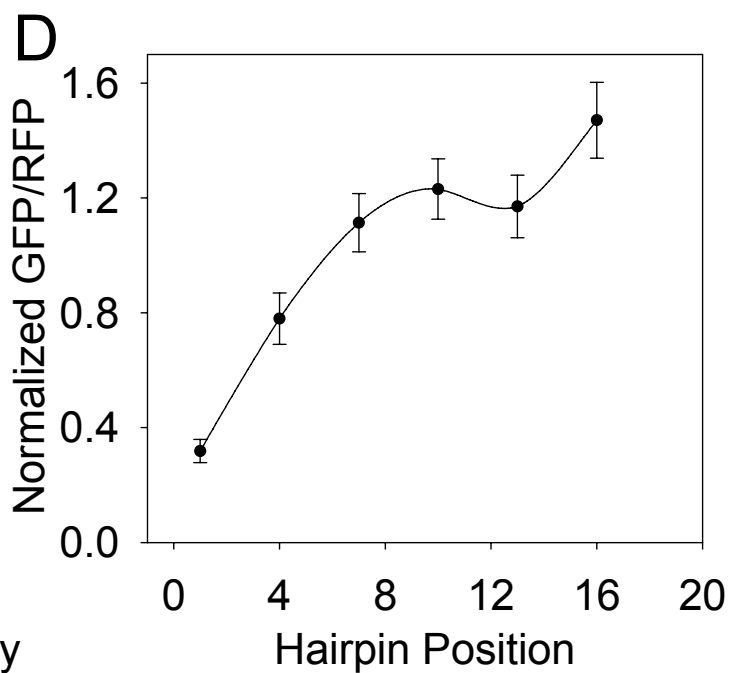
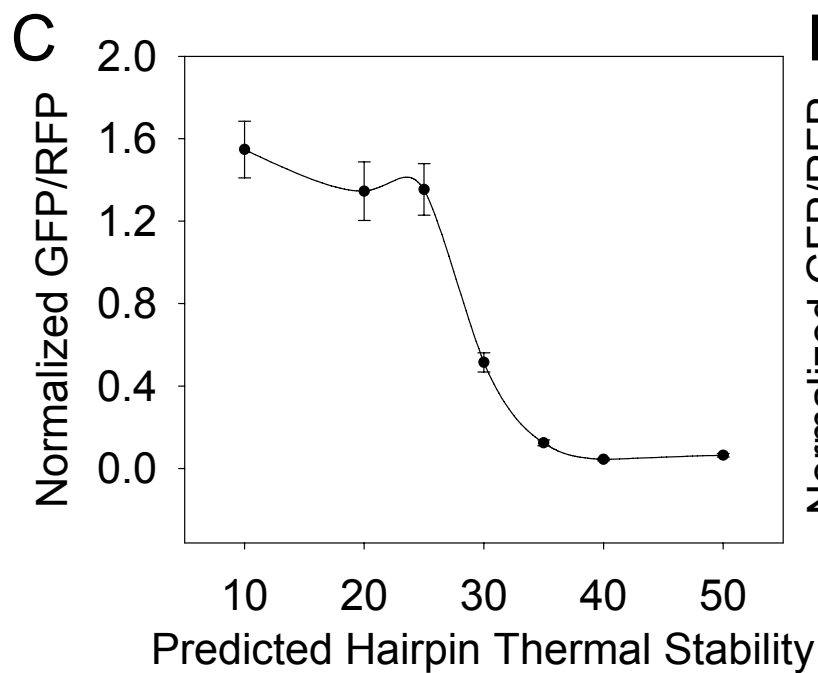
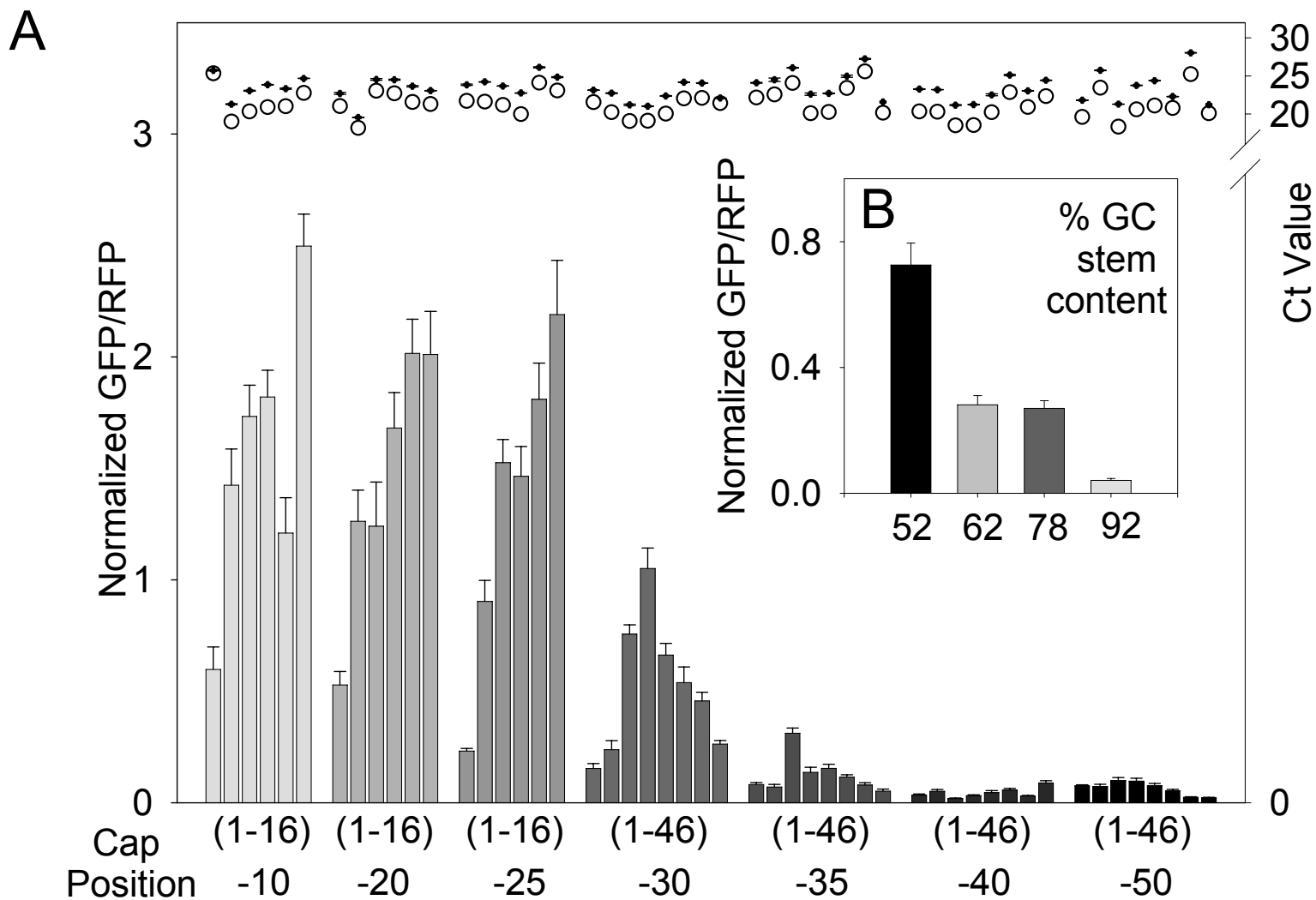
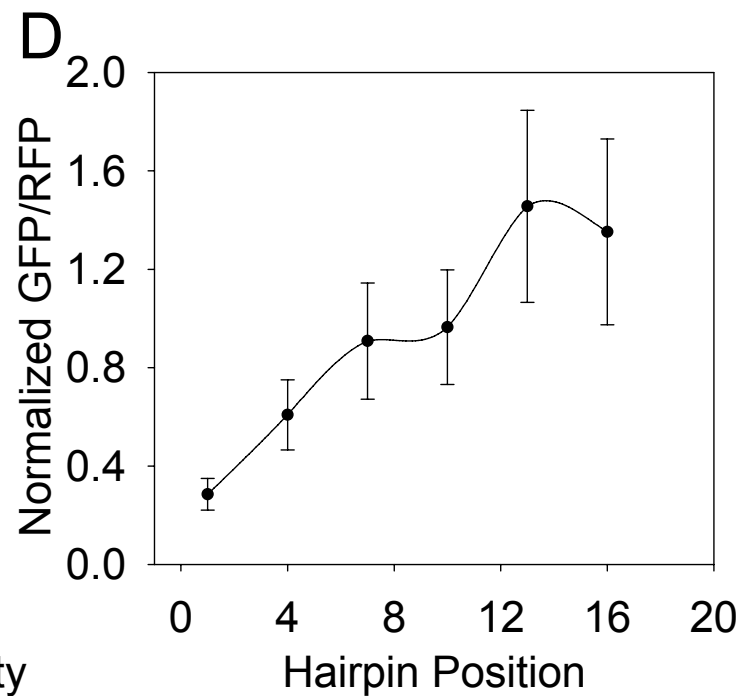
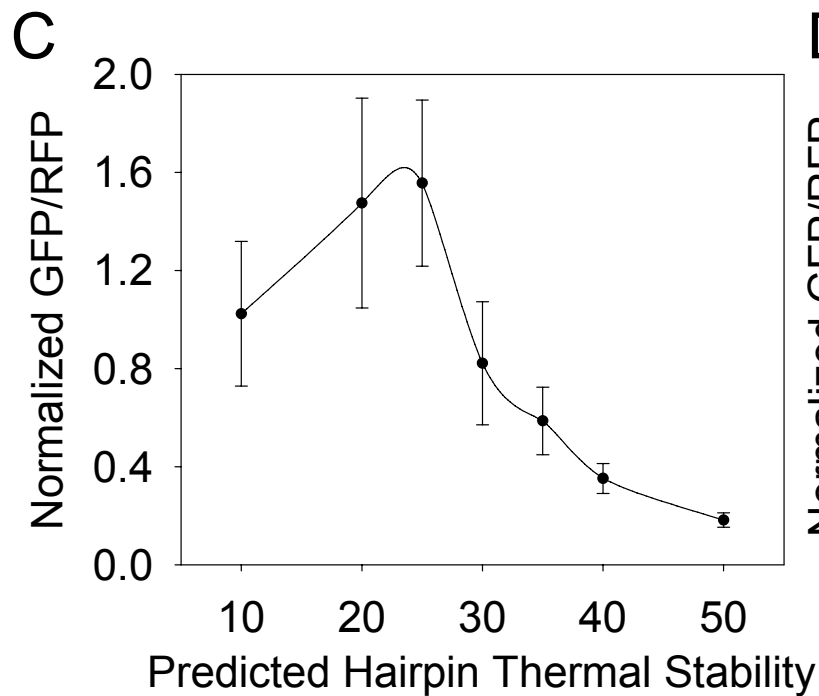
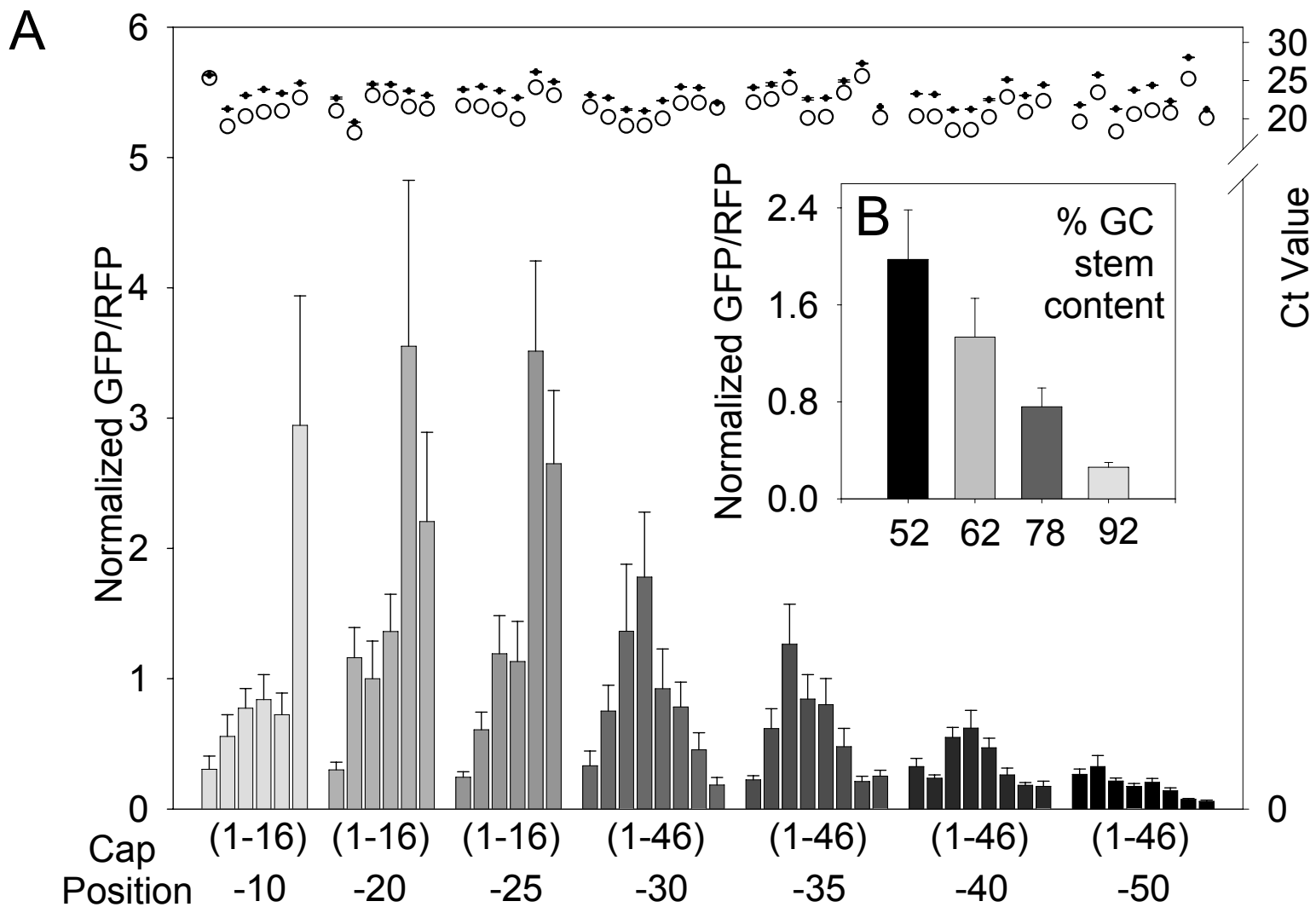


Figure S4: Translation efficiencies in MEF cells



*Figure S1: Comprehensive list of insert sequences for GR1, GR2, and Qc3c\_AgeI and GR3 vectors.* A) Insert sequences 70 bases in length placed in vector GR1. Sequences in bold represent hairpins with predicted thermal stabilities of -10.5, -20.5, -25.7, -30.3, -34.5, -39.4 and -50.2kcal/mol with +1, +4, +7, +10, +13 or +16 positions. Sequences were normalized to control insert 70 bases long composed of CAA repeats. B) Insert sequences 50 bases long placed in vector GR1. Sequences in bold represent hairpins with GC stem percentages of 52, 62, 78 and 92%. Each hairpin was designed with predicted thermal stability of -30 +/- 1kcal/mol and +4 position. Sequences were normalized to control insert 50 bases long composed of CAA repeats. C) Insert sequences 63 bases long placed in vector GR2. Vector GR2 was adapted from GR1 to place inserts 31 bases 3' of the transcription start site. Sequences in bold represent hairpins with predicted thermal stabilities of -30.3, -34.5, -39.4 and -50.2kcal/mol each placed at +31 or +46 positions. Sequences were normalized to control insert 63 bases long composed of CAA repeats. D) Natural 5'UTRs were placed in vector AgeI\_Qc3c or GR3 for *in vitro* or live cell analysis respectively.

*Figure S2: RNA localization, fluorescent protein content and 3'UTR sequence controls.*

A) General map of vector RG1. Original *SacI* and *BamHI* (or *BsiWI* and *BamHI*) were used to place hairpin constructs at the 5' UTR of a CMV driven mCherry fluorescent protein. An original *AgeI* was placed as the 2nd and 3rd amino acids for full length 5'UTR cloning. At the 3' UTR of the same transcript is an IRES driven GFP fluorescent protein. B) Hairpin position, thermal stability and GC stem content affect translation efficiency with RG1 vector. Constructs were transfected into Cos7 cells 24 hours prior to

fluorescence microscopy measurements. Average red to green ratios were computed after subtracting a background region from the images. All values were normalized to control CAA construct HPO. Bars labeled (1-16) refer to hairpins placed at positions +1, +4, +7, +10, +13 and +16 from left to right for the -25, -30, -35 and -40 kcal/mol hairpins. Bars labeled (-10 to -50) refer to the -10, -20, -25, -30, -35, -40 and -50 kcal/mol hairpins at +4 positions. Bars labeled (52 to 92) refer to hairpins with GC stem content percentages of 52, 62, 78 and 92 while maintaining a +4 position and a predicted thermal stability near -30 kcal/mol. Error bars represent the standard error of the mean for 20 fields. Each set is shown with increasing cap-to-hairpin distances for the -25, -30, -35 and -40 kcal/mol hairpins, increasing thermal stabilities (each at +4) and increasing GC stem content. Values at the top are IRES driven GFP values.

*Figure S3 – Translation efficiencies in Chinese Hamster Ovary (CHO) cells.* Library constructs were transfected into CHO cells and analyzed with fluorescence microscopy. Error bars represent the standard error of the mean for 20 fields. All values were normalized to their respective control CAA value. A) Each hairpin set is shown with increasing thermal stabilities (-10kcal/mol to -50kcal/mol) from left to right. Bars labeled (1-16) refer to hairpins placed at positions +1, +4, +7, +10, +13 and +16 from left to right for the -10, -20, and -25 kcal/mol hairpins. Bars labeled (1-46) refer to hairpins placed at positions +1, +4, +7, +10, +13, +16, +31 and +46 from left to right for the -30, -35, -40 and -50 kcal/mol hairpins. B) The percent stem GC content increases from left to right. C) Positions +1 to +16 for hairpins with predicted thermal stabilities between -10 to -25kcal/mol and positions +1 to +46 for hairpins with predicted thermal stabilities

between -30 to -50kcal/mol were averaged. Error bars are the standard error of the mean.

D) Average distance effect on translation efficiency. Each position for hairpins with predicted thermal stabilities between -10 to -35kcal/mol was averaged. Error bars are the standard error of the mean.

*Figure S4 – Translation efficiencies in Mouse Embryonic Fibroblast (MEF) cells.*

Library constructs were transfected into MEF cells and analyzed with fluorescence microscopy. Error bars represent the standard error of the mean for 20 fields. All values were normalized to their respective control CAA value. A) Each hairpin set is shown with increasing thermal stabilities (-10kcal/mol to -50kcal/mol) from left to right. Bars labeled (1-16) refer to hairpins placed at positions +1, +4, +7, +10, +13 and +16 from left to right for the -10, -20, and -25 kcal/mol hairpins. Bars labeled (1-46) refer to hairpins placed at positions +1, +4, +7, +10, +13, +16, +31 and +46 from left to right for the -30, -35, -40 and -50 kcal/mol hairpins. B) The percent stem GC content increases from left to right. C) Positions +1 to +16 for hairpins with predicted thermal stabilities between -10 to -25kcal/mol and positions +1 to +46 for hairpins with predicted thermal stabilities between -30 to -50kcal/mol were averaged. Error bars are the standard error of the mean. D) Average distance effect on translation efficiency. Each position for hairpins with predicted thermal stabilities between -10 to -35kcal/mol was averaged. Error bars are the standard error of the mean.

## Protocol S5

### Cloning of vector RG1

Vector RG1 was cloned by modifying sites in the 3' and 5' UTRs of vector Qc3c. We modified the 5' UTR, by adding an original *Bsi*WI between *Sac*I and *Bam*HI sites cloning with oligos O and P using *Sac*I and *Bam*HI sites. We modified the 3' UTR by adding *Nhe*I, *Not*I and *Cla*I sites as well as swapping the order of *Xba*I and *Xho*I sites with oligos Q and R using *Xho*I and *Xba*I site. Once complete, we exchanged the GFP with an mCherry having a 5' proximal *Age*I cloning site. As a result, we amplified mCherry with primers S and T that added an *Eco*RI site 5' of the Kozak sequence and an *Age*I site as the 2nd and 3rd amino acids of mCherry. This PCR product was digested with *Eco*RI and *Xba*I sites and cloned into modified Qc3c vector. We then inserted an IRES driven GFP by amplifying an IRES-GFP sequence from an Invitrogen vector with primers U and V. This PCR product was digested with *Xba*I and *Not*I and inserted into the Cherry Qc3c vector, named RG1.

```
O CGTACGGCTAACTAGGCCAG
P GATCCTGGGCCTAGTTAGCCGTACGAGCT
Q TCGAATCTAGAGCTAGCGGCCGCTCGAGAATCGATA
R CTAGTATCGATTCTCGAGGCGCCGCTAGCTCTAGAT
S GATCCGAATTGCGCCACCATGACCGGTGTGAGCAAGGGCGAGGAGG
T GCATTTAGGTGACACTATAG
U CTAGCAATTAGCGTACCAGCACATCTAGAC
V GATCGTCTAGATGTGCTGGTACGCTAATTG
W GATCAATTGCAACTCGAGCAATGATCACAATCGATCAAGCTTAGC
X CGGCTAAGCTTGATCGATTTGTGATCATTGCTCGAGTTGCAATT
```

### Analysis of results with RG1 constructs

We desired to rule out alternative explanations of our findings. First, could trends be dependent upon how well RNAs were exported from the cell? For example, if highly structured RNAs were less efficiently transported from the nucleus to the cytoplasm, then observed effects



would be due to RNA localization, not translatability. Secondly, could trends be repeated with a different reporter protein and 3' UTR? To assess independence of the results from reporter gene, 3' UTR and nuclear export, we designed vector RG1 shown in figure S2A. The reporter protein was changed from an EGFP to mCherry fluorescent protein to address the question if results could be repeated with a different reporter. To address the questions of mRNA export, we cloned an internal ribosome entry site (IRES) in the 3' UTR of the mCherry construct allowing normalized protein expression from the same transcript as the modulated protein. If mRNAs were exported from the nucleus to the cytoplasm at different efficiencies, than control GFP levels would change at the same rate. If so, there would be no difference in red to green ratios between samples. Finally, the 3' IRES sequence extends the length of the 3' UTR addressing the question of 3' UTR sequence.

As a result, we cloned 36 of the original 57 constructs including the -25, -30, -35 and -40 kcal/mol hairpin sets at positions +1, +4, +7, +10, +13 and +16, 3 additional hairpins at +4 (-10, -20, and -50 kcal/mol), all 4 GC stem constructs (52, 62, 78 and 92%) and control CAA construct. Figure S2B shows the results for these control experiments. In summary, we observed the same trends as with GR constructs: the closer hairpin structures were placed to the cap, the more they inhibited translation; increased hairpin thermal stability inhibited translation; and increased GC content inhibited translation. Trends with RG -30, -35 and -40 kcal/mol hairpins sets appeared almost identical to GR constructs in figure 3A. Similar to QPCR results with GR vectors, we did not observe any trends in translation levels from the IRES-GFP control shown in the top margin of figure S2B, showing values were due to changes in the modulated mCherry mRNA. Furthermore, the variability in GFP expression illustrates the necessity of including an internal control.

We noticed some differences in changing construct parameters. First, the relative amount of translation compared with control CAA was less. GR constructs reached ratios of 10, while RG1 constructs reached values of 1. In addition, we noticed less dramatic changes in translation with respect to cap to hairpin distance. The -25kcal/mol hairpin set resembled the -10 and -20kcal/mol sets in figure 3 with an 8 fold enhancement between +1 and +4. Between +4 and +16 translation increased slightly. GC stem content trends with RG constructs appeared similar as well, however the RNA stem with 62% GC content was closer to the 52% stem as opposed to the 78% stem (compare Figure S2 B to Figure 4). We conclude the reporter protein, 3' UTR sequence and nuclear export have minimal effects on translation efficiency trends observed with GR constructs.