Supplementary Figure 1: Gel source data for Fig 3D

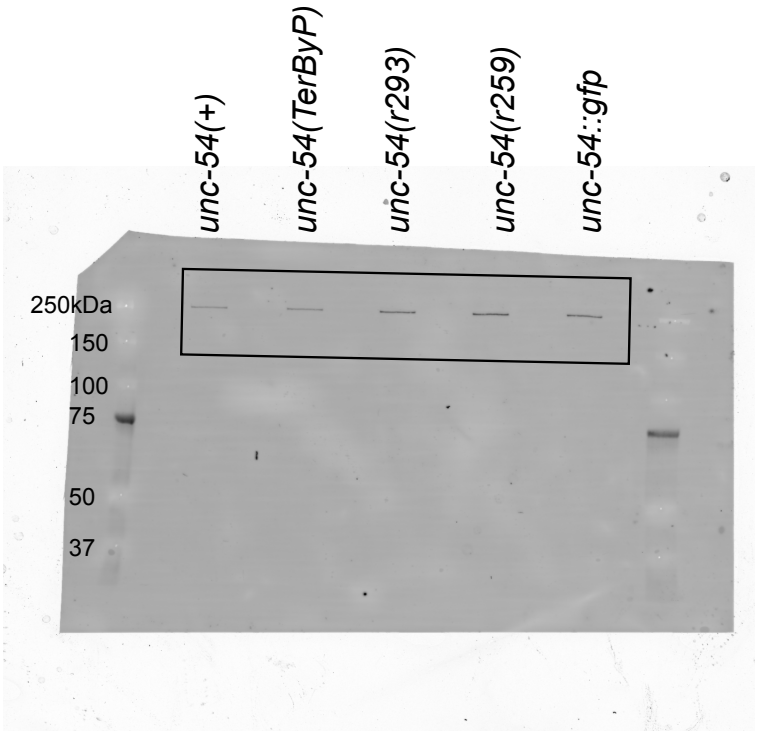
Supplementary Table 1: Strains used in this study

Supplementary Table 2: Plasmids used in this study

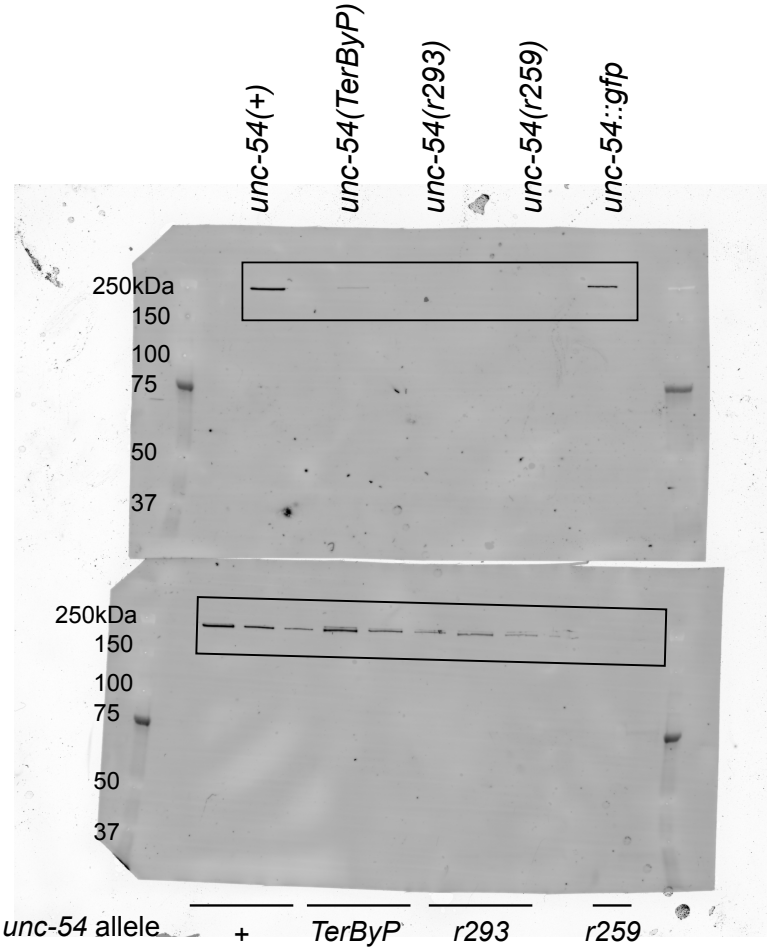
Supplementary Table 3: DNA oligos for rRNA digestion by RNaseH

Supplementary Information

We do not yet know the determinants of a translated 3'UTR sequence that confer a loss of protein. Our mutational analysis of *unc-54* and *tbb-2* readthrough regions identified translation of the sequences LFLLLS and NLFAF (respectively) as important for the loss of GFP. Non-synonymous substitutions of a similar sequence (VLFL) in the *rpl-14* 3'UTR did not restore GFP expression, suggesting some flexibility in the relationship between the C terminal extension and the loss of protein accumulation (Fig 2A); similar results with *unc-54* and *tbb-2* indicate that not all portions of the 3' peptide are essential for suppressing the accumulation of the 3' extended protein. Examining other 3'UTR sequences (*rps-17*, *myo-2*, *hlh-1*, *eef-1A.1*, *bar-1*, *daf-6*, *mut-16*, *alr-1*, *r74.6*, *rps-20*, *rps-30*, *unc-22*, and *unc-45*), or 3'UTRs genome-wide, revealed no well-defined peptide motif, and no obvious conservation signal (Extended Data 7). We did however note a preponderance of bulky and hydrophobic amino acids (F, L, I, Y), a property inherent in the underlying sense-strand nucleotide frequency bias of 3'UTRs in *C. elegans* (T>A>C>G) (Extended Data 8). To better understand the determinants of 3'UTR sequences detrimental for expression, we shuffled codons of the readthrough regions of *unc-54*, *tbb-2*, and *rpl-14(VLFL>RSCA)* (see diagram in Extended Data 4), and inserted the sequences downstream of GFP. Surprisingly, each of three shufflings conferred a substantial loss of GFP (Fig 2B). Randomizations of readthrough regions at the nucleotide level in humans also tended to confer low eGFP expression (Fig 4C). The ability of some shuffled sequences to reduce GFP expression is consistent with the lack of a well-defined motif (for comparatively degenerate phenomena, see[53,54]), suggesting instead that the negative effects of translated 3'UTRs on protein expression are a structural or highly degenerate sequence property. Consistent with the idea that readthrough peptides' effects may be mediated via their biophysical characteristics, we observed a significant negative relationship between hydrophobicity and GFP expression in both *C. elegans* (Spearman ρ=-0.77, p-value=6.8e-5, Extended Data 8D) and humans (Spearman ρ=-0.68, p-value=4.1e-5, Extended Data 9B). Similar trends were observed with the *unc-22* and *unc-54* loci in *C. elegans*, suggesting the effects are not peculiar to GFP (Extended Data 8E).



α MYO-3
"5-6" Ab



α UNC-54
"5-8" Ab