Comprehensive analysis of yeast +1 ribosomal frameshifting unveils a novel stimulator affirming two distinct frameshifting mechanisms

Darren A Fenton[1,#], Maria Bożko[2], Michał I. Świrski[2], Gary Loughran[1], Martina M Yordanova[1], Joanna Kufel[2], John F Atkins[1,3], Pavel V Baranov[1*]

[1]School of Biochemistry and Cell Biology, University College Cork, Ireland

[2]Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw, Warsaw, Poland

[3]School of Microbiology, University College Cork, Cork T12 K8AF, Ireland.

#current affiliation: Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA

*Corresponding author: p.baranov@ucc.ie

Abstract

Ribosomal frameshifting allows the synthesis of a protein from different reading frames. Generally, it involves a dissociation of the P-site tRNA from its codon and movement to the +1 codon setting the new frame for incoming tRNAs. This movement is stabilized by tRNA repairing with the +1 codon. However, in several occurrences in the yeast *Saccharomyces cerevisiae*, P-site tRNA repairing with the +1 codon is impossible. One hypothesis suggests that frameshifting involves P-site tRNA movement, with repairing not being essential. In the alternative model the A-site tRNA acceptance at the +1 codon occurs in the absence of P-site tRNA movement. Here we performed a comparative analysis of all known +1 frameshifting sites in *S. cerevisiae*. In addition, we discovered a novel gene requiring +1 frameshifting for its expression. We also identified a conserved RNA structure located upstream of the *ABP140* frameshifting site. This structure substantially increases frameshifting efficiency. Its location suggests that it is formed in mRNA exiting the ribosome, creating a pulling effect favouring positioning of the +1 codon in the P-site. We show that the stimulation is limited to frameshifting sites where P-site tRNA repairing is possible, supporting the existence of two distinct mechanisms of +1 ribosomal frameshifting.

Keywords: Translational Recoding, +1 Ribosomal Frameshifting, mRNA translation, Ty1, ABP140, OAZ1, EST3, YPL034W, YJR112W-A, YFS1, YFS2

Introduction

Translating ribosomes are capable of deviating from the “standard” decoding rules in response to specific mRNA signals. These deviations are collectively termed “recoding events” (1–3). Ribosomal frameshifting is a form of recoding where some ribosomes shift their reading frame at a specific site during translation, resulting in the creation of a trans-frame protein in addition to the standard translation product. While ribosomal frameshifting is common in mobile elements and especially in RNA viruses (4), it is extremely rare in chromosomal gene decoding with the exception of ciliates *Euplotes* where it is considered a feature of its genetic code (5). Normally, frameshifting occurs at a specific combination of (i) codons that allow tRNA repositioning with codons in a new frame and (ii) stimulatory signals that elevate its efficiency through a variety of molecular mechanisms.

The first example of a eukaryotic gene requiring frameshifting was discovered in the *Saccharomyces cerevisiae* transposable element Ty1 and was found to utilize efficient +1 ribosomal frameshifting (6–9). Strikingly, it was found that a minimal sequence of 7 nucleotides is sufficient to support a remarkably high (~40%) frameshifting efficiency in the absence of additional RNA stimulators in the vicinity of the frameshifting site (10). The sequence of the Ty1 frameshifting heptamer is CUU_A.GG_C where underscore indicates codon boundaries in the initial (0) reading frame and dot indicates codon boundaries in the new (+1) frame (this notation will be used thereafter). A remarkable feature of this frameshifting site is the high imbalance in the levels of tRNAs recognising the AGG (Arg) and GGC (Gly) codons. The *S. cerevisiae* genome contains only a single gene copy of Arg-tRNA (CCT), while the number of gene copies of Gly-tRNA (GCC) is up to 18 depending on the strain (11). This suggests a far greater concentration of tRNAs cognate for GGC over tRNAs cognate for AGG and subsequently faster decoding of the GGC codon in comparison with the AGG. Presumably, slow decoding of AGG allows more time for repairing of the Leu-tRNA in the P-site from CUU to UUA while fast decoding of GGC would prevent repairing mRNA in the reverse direction from UUA to CUU (10, 12). Subsequently, the same frameshift-inducing heptamer was identified in the genes *ABP140* (13) and *YPL034W*/*YFS1* (14). Here, we also report its utilization for the expression of *YJR112W-A*/*YFS2*. A different pattern, using the same P-site codon but different A-site codons (CUU_A.GU_U) was found to cause +1 frameshifting in *EST3* (15). Albeit different and less efficient, the mechanism of *EST3* frameshifting may be similar. In *S. cerevisiae*, AGU decoding requires wobble interactions and may be slow, while the number of copies for the cognate Val-tRNA is high reaching up to 21 (11).

A more different +1 frameshifting heptamer (GCG_A.GU_U) was identified in the Ty3 transposable element (16). While the A-site component of this pattern is identical to *EST3* frameshifting, the P-site codon (GCG) is different and seemingly does not support repairing in the +1 reading frame. Importantly, while the optimal Ala-tRNA (CGU) for GCG is present, it is rare. The same codon can be decoded by a near cognate isoacceptor tRNA with anticodon CGI which is apparently responsible for frameshifting since the oversupply of Ala-tRNA(CGU) inhibits frameshifting (17). The same P-site codon is used in the *S. cerevisiae* antizyme (*OAZ1*/*YPL052W*) frameshifting heptamer (GCG_U.GA_C) where frameshifting is driven by competition between termination at a stop codon in poor context with incorporation of a tRNA at the +1 codon. The

efficiency of antizyme frameshifting is sensitive to polyamine levels (18–20). The discovery of an unusual P-site codon in the Ty3 frameshifting heptamer prompted a subsequent investigation of the repertoire of P-site codons supporting +1 ribosomal frameshifting. It revealed several P-site codons supporting above-background levels of +1 frameshifting in A.GU_U or A.GG_C A-site contexts with little P-site repairing potential with the +1 codon (21). This led to the suggestion that certain tRNAs, when in the P site, possess properties that allow ribosomes to incorporate A-site tRNAs in the +1 frame without prior P-site tRNA slippage (22). To unify these observations into a single parsimonious mechanism, kinetic considerations have been used to argue that it is the dissociation of the P-site tRNA from its zero-frame codon that is the limiting step. Even short-lived unstable P-site tRNA interactions with the +1 codon may be sufficient for high efficiency frameshifting given the high imbalance between cognate in- and out-of-frame tRNAs in the A-site (12). However, until now, little experimental evidence has been provided to resolve this conundrum. It has remained unclear whether one or two different mechanisms are responsible for +1 ribosomal frameshifting in yeast at different P-site codons (Fig. 1).

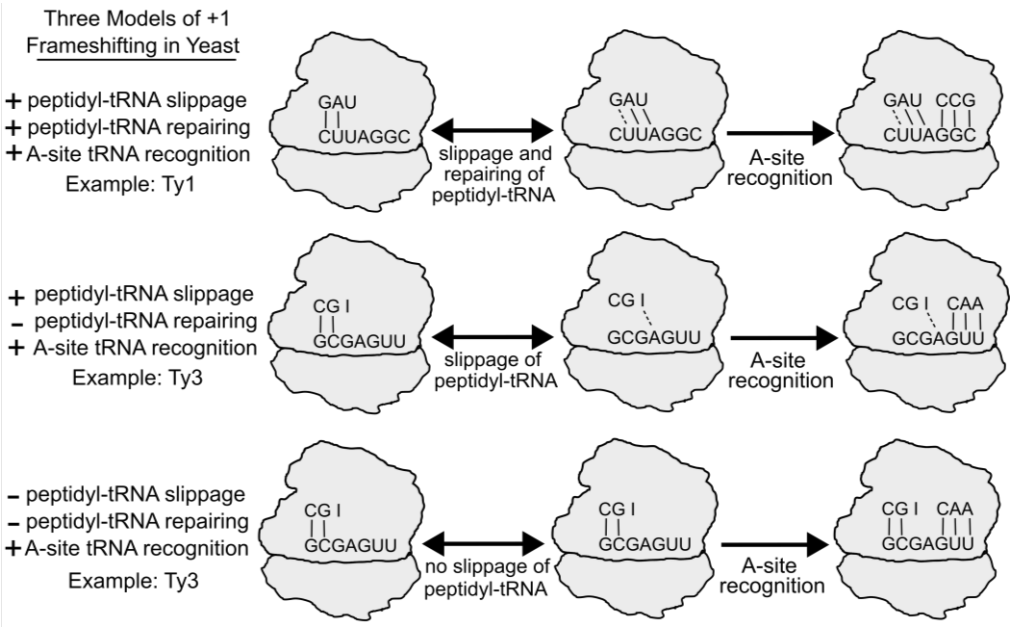

Figure 1. Models of +1 Frameshifting Mechanisms in *Saccharomyces cerevisiae*. Top model: Peptidyl-tRNA slippage and repairing to the overlapping UUA codon in the +1 frame is followed by A-site recognition of the tRNA decoding the +1-frame GGC codon. The middle and bottom depict two alternative mechanisms proposed to explain frameshifting when tRNA repairing with the +1 codon in the P-site could not occur. Middle model: Peptidyl-tRNA slippage, with no repairing, followed by A-site recognition of the tRNA decoding the +1-frame GUU codon. Bottom model: +1 frameshifting involves no slippage or repairing of the peptidyl-tRNA. Instead, frameshifting occurs due to direct recognition of tRNA to the +1 codon in the A-site.

In this study, our initial motivation was to carry out a systematic comparative study of naturally occurring +1 frameshifting sites, with a goal to provide a reliable reference for +1 frameshifting efficiencies in yeast. This was due, since prior frameshifting measurements were made in different research groups using different reporter systems that may be subject to various technical artifacts (23), and thus could not be directly compared. For this purpose, we have developed a new vector designed for studying recoding signals in yeast, based on the StopGo reporter system previously developed for mammalian cell lines (24). Using this

reporter system, we measured ribosomal frameshifting at all the above-mentioned cases with and without their surrounding mRNA context. The exploration of the contexts revealed the existence of a previously unknown +1 frameshifting stimulator within the *ABP140* mRNA, located upstream of the frameshifting site heptamer. This stimulator increases +1 frameshifting efficiency from 40% to 60%. Phylogenetic analysis of the corresponding sequence combined with site-directed mutagenesis strongly suggest that this stimulator is an RNA structure whose formation pushes the ribosome downstream into the +1 frame. We tested the effect of this stimulator on different frameshifting heptamers. This enabled us to discriminate frameshifting occurring due to P-site tRNA slippage or direct out-of-frame A-site tRNA recognition.

In addition to resolving a long-standing mechanistic conundrum, our study provides a reference to all *S. cerevisiae* +1 frameshifting cases reported so far, including a novel case we discovered while exploring available ribosome profiling data. Finally, ribosome profiling data was used as an orthogonal approach for estimating endogenous levels of ribosomal frameshifting.

Results

A new reporter to test recoding efficiencies in *S. cerevisiae.*

Measurements of frameshifting efficiencies are obtained using vectors that contain candidate frameshifting sequences fused between two expression reporters. Previously used reporters in yeast frameshifting studies included β-galactosidase, luciferase, β-galactosidase-luciferase fusions, and dual-luciferase assays (7, 25, 26). Luciferase assays are attractive as they allow sensitive readings, especially for cases where a low recoding signal needs to be measured. An in-frame control (i.e. both reporters in the same frame) is used to establish the "100% efficiency" signal, allowing comparison with out-of-frame reporters to determine frameshifting efficiency (23). However, the protein sequence encoded by the test sequence could affect the activities of one or both reporters, leading to alterations in measured activities and distorting accurate measurements of frameshifting (23). To mitigate this issue, the original dual luciferase reporter vector (pDLuc) developed for use in mammalian models (27, 28) was modified to incorporate StopGo/2A on both sides of the test sequence in pSGDLuc vector (24). The StopGo/2A peptide motif results in the synthesis of two separated protein products (encoded upstream and downstream of it) because the translating ribosome fails to form a peptide bond at a specific position (29) These motifs were originally found in viruses where they are responsible for the production of distinct products from the same translated ORF (30). Therefore, most reporter proteins produced from the in-frame and test sequence are identical, allowing a more accurate determination of the frameshifting efficiencies. Indeed the use of StopGo/2A-containg dual luciferase reporter led to identification of a false positive frameshifting case in a human gene due to limitations of the previous reporter (31).

We adapted the mammalian pSGDLuc reporter plasmid for use in *S. cerevisiae* with three modifications. 1). We optimised Renilla and Firefly luciferase coding sequences by introducing synonymous codon

substitutions based on yeast codon usage (Methods). 2). As the StopGo/2A motif used in the mammalian reporter (F2A, derived from foot-and-mouth disease virus) is only ~50% efficient in *S. cerevisiae*, we substituted it with a more efficient variant, E2A (derived from the equine rhinitis A virus) which is ~90% efficient (32). The new vector backbone encodes both leucine and G418 selection markers (Fig. 2A). 3). We modified a multiple cloning site to incorporate the NheI and NdeI restriction enzyme sites to allow cloning via T4 ligation or Gibson assembly reactions. The new vector termed pYSGDLuc (Fig. 2A) can be used for various types of recoding studies in yeast.

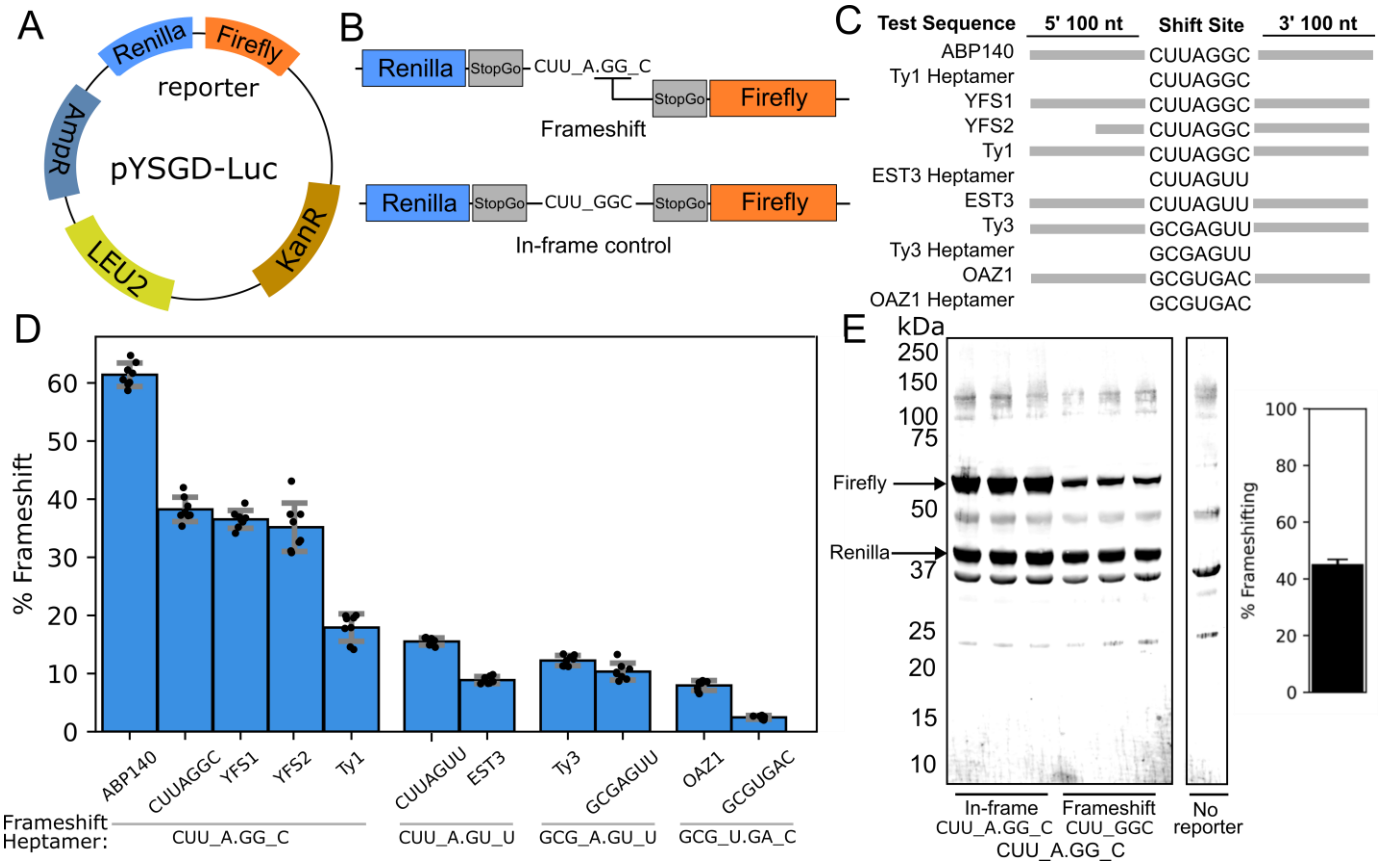

Figure 2. Uniform assessment of frameshifting efficiency at known frameshifting cases. A, A schematic representation of the pYSGDLuc. *LEU2* represents the Leucine biosynthesis gene and AmpR and KanR represent ampicillin (for *E. coli* cloning) and kanamycin resistance (for yeast G418 selection) genes. **B,** Schematic of frameshifting reporters for testing the Ty1 heptamer sequence, including WT and mutant in-frame control sequences. **C,** Representation of natural frameshifting cassettes tested in this study, grey bars indicate the length of tested sequences flanking heptamers. **D,** Frameshifting efficiencies determined with luciferase assays. Cases are clustered together based on the frameshifting heptamer (for example, *ABP140*, *YFS1, YFS2* and Ty1 frameshifting at CUU_A.GG_C). Error bars represent standard deviation (n=8). **E,** Western Blotting of replicate cell extracts derived from cultures expressing the Ty1 heptamer (CUU_A.GG_C) reporters, using anti-Renilla and anti-Firefly antibodies. Densitometry analyses estimate ~45% frameshifting.

Comparative analysis of S. cerevisiae frameshifting cassettes using pYSGDLuc and publicly available ribosome profiling data.

Using pYSGDLuc, we performed a comparative study of all known naturally occurring cases of efficient +1 frameshifting in *S. cerevisiae*. Measuring these cases in the same reporter system under the same environments (e.g. yeast strain, media, equipment and reagents), allows objective comparison of frameshifting efficiencies, that are not distorted by the differences in environmental parameters. Each tested

case included both an in-frame control and frameshifting test sequence as recommended by newly developed guidelines (23), see Fig. 2B. For all cases, we tested the minimum heptamer sequence (e.g. CUU_A.GG_C) required for frameshifting that includes the P-site and A-site codons before and after frameshifting (Fig. 2C). This provided the basal rate of frameshifting at corresponding heptamers independent of their natural surrounding nucleotide context. To assess the effect of the frameshifting heptamer within the native mRNA context, we included a ~100 nt flanking region. The frameshifting site for *YSF1* is located close (~42 nt) to the 5' end of the mRNA transcript, therefore, a shorter upstream test sequence was included. The frameshifting efficiencies calculated as a ratio of luciferases ratios between test sequences and controls are shown in Figure 2D as percentages.

To test if there may be unexpected products produced from these reporters, we carried out western blotting to confirm product sizes for Renilla and Firefly reporter proteins (Fig. 2E, left). Densitometry analyses estimated similar frameshifting efficiencies as the luciferase assays (Fig. 2E, right).

The measured differences between four frameshifting heptamer-only reporters appeared to be consistent with previous reports, CUU_A.GG_C ~40% > CUU_A.GU_U ~15% > GCG_A.GU_U ~10% > GCG_U.GA_C ~3% (Fig. 2D). Testing frameshifting efficiencies at these heptamers demonstrated a certain degree of context dependence. The inclusion of natural sequence contexts of Ty3, *YFS1* and *YFS2* (Fig. 2C) did not alter basal levels of frameshifting at their corresponding frameshifting heptamers. However, frameshifting in the *ABP140* context was increased 1.5-fold to ~60% and unsurprisingly, near 3-fold in *OAZ1* natural context to ~8%. The Ty1 and *EST3* contexts had opposite effects, with TY1 reducing frameshifting efficiency 2-fold to ~20% and 1.5-fold to 10% in *EST3*.

Inclusion of the 200 nt surrounding sequence into pYSGDLuc cassette does not entirely recreate the parameters of natural frameshifting at least for two reasons. First, sequence elements at large distances from the frameshifting site may alter its efficiencies (33), second the efficiency of frameshifting may depend on parameters of mRNA translation not directly linked to the surrounding sequence, such as ribosome loading (34). Therefore, to assess the natural levels of frameshifting we used publicly available ribosome profiling data and inferred frameshifting efficiency from a change in ribosome footprint density downstream of frameshifting sites, similarly to previous work on viral frameshifting (35), see Methods. A global aggregate of these ribosome profiling data from different studies for *ABP140*, *EST3* and *OAZ1* is presented in Figure 3A. For most cases, we found a strong concordance between frameshifting efficiencies measured with these two methods (Fig. 3B), suggesting that our pYSGD-Luc reporter provides accurate measurements of frameshifting efficiencies and the absence of relevant context outside of the 200 nt proximal window. We note that for *YFS1* and *YJR112W-A/YFS2*, the short length of coding region upstream of frameshifting sites limits the accuracy of this method in comparison with *ABP140*, *EST3* and *OAZ1*. Nonetheless, for *ABP140*, *EST3* and *OAZ1*, we find a strong concordance in frameshifting efficiencies between luciferase assays and ribosome profiling.

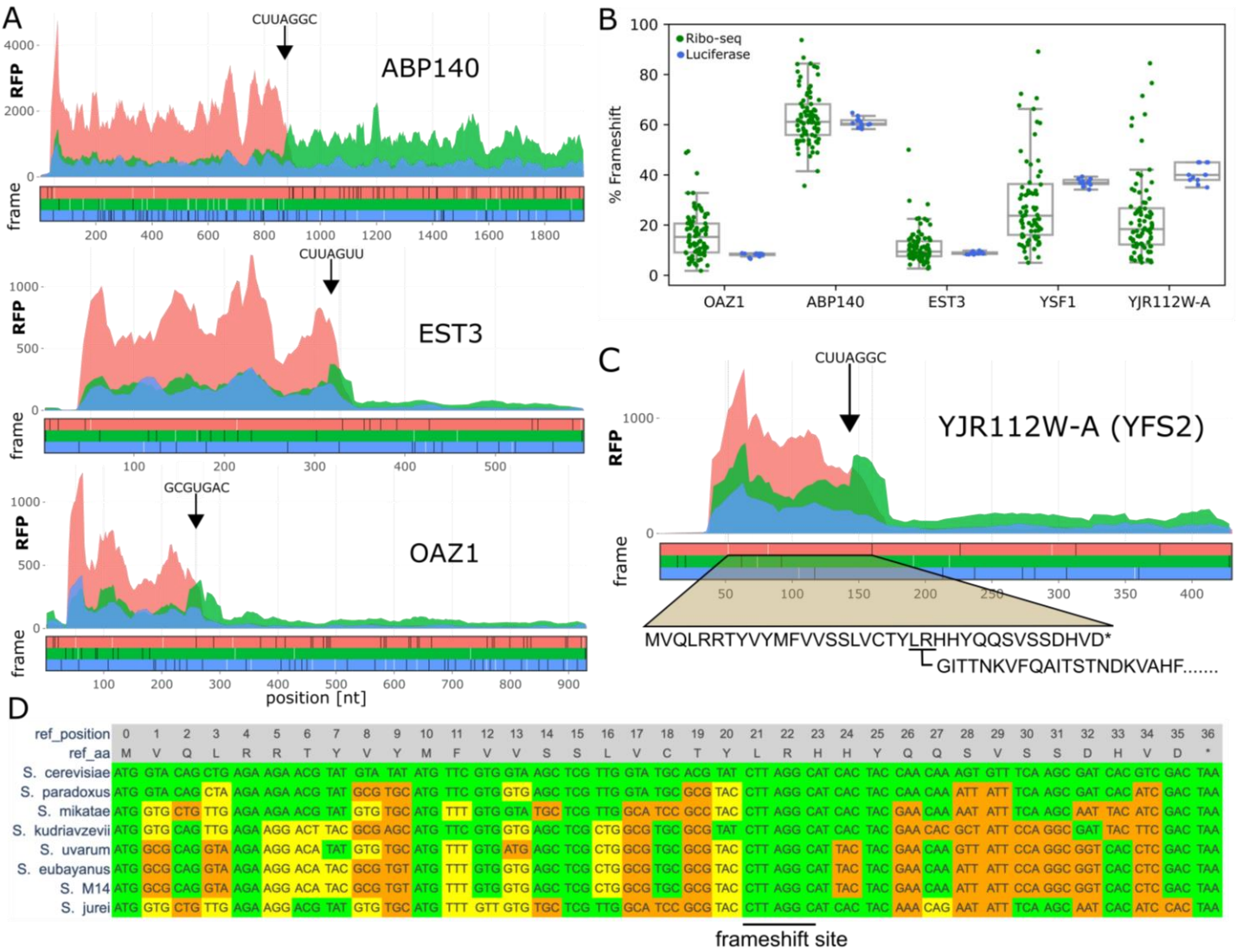

Figure 3. Analysis of publicly available ribosome profiling data. **A.** Subcodon ribosomal profiles for *ABP140*, *EST3* and *OAZ1* genes. For each gene top plots show aggregated ribosome profiling data differentially coloured based on the supported reading frame. The colours are matched to the reading frames in the ORF plot below where black and white dashes represent stop and ATG codons in each reading frame respectively. The position of the frameshifting sites are denoted with black arrows, stop codons of the zero frame ORFs are indicated with grey vertical lines. **B.** Ribosomal frameshifting inferred from 85 individual ribosome profiling datasets based on the relative ratio of footprint densities in the +1 ORF downstream of the zero frame ORF are shown as green dots. Blue dots indicate frameshifting efficiencies obtained with dual luciferase assays for each replicate (n=8-12). The parameters of distribution for both types of data are show with boxplots where centre line represents the median, box limits indicate the 25th and 75th percentiles and whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles. **C.** Same as in A, but for novel instance of frameshifting discovered in this study during translation of *YJR112W-A/YFS2* mRNA. Note the full coding regions are not displayed. **D.** Codon alignment of the 0-frame ORF of *YJR112W-A/YFS2* orthologs in *Saccharomyces* genus with the universally conserved CUU_A.GG_C indicated. Yellow codons represent synonymous changes and orange codons represent non-synonymous changes

A novel case of cellular +1 frameshifting at the YJR112W-A/YFS2 locus

An independent manual analysis of ribosome profiling data revealed an unusual distribution of ribosome protected fragments (RPFs) in the *YJR112W-A* locus in *S. cerevisiae* (Fig. 3C). *YJR112W-A* is annotated as an intron containing gene (see Supplementary Fig. S1). However, publicly available RNA-seq data in GWIPS-viz (36) suggest that the annotated intron is a part of mRNA as RNA-seq density is uniform across intronic and exonic regions of this gene (Supplementary Fig. S1). The highest RPF density matches an ORF initiated

with the most 5' end AUG codon and spans the region annotated as intron and a part of the second annotated exon (Supplementary Fig. S1). The examination of this ORF sequence revealed the presence of the CUU_A.GG_C heptamer suggesting that ribosomes translating this ORF undergo +1 ribosomal frameshifting. Indeed, the ribosome footprint density is present downstream of this ORF matching the +1 reading frame, albeit at a lower density than at the zero frame ORF. We conclude that the current *S. cerevisiae* reference annotation of *YJR112W-A* CDS is incorrect, and that *YJR112W-A* does not contain introns, instead its CDS is comprised of two overlapping ORFs translated *via* ribosomal frameshifting. We named this gene *YFS2* (Yeast Frameshifting 2), following naming convention from a previous manuscript that reported a similar discovery of new +1 frameshifting case (14). The sequence alignments of YFS2/*YJR112W-A* orthologs from other *Saccharomyces* demonstrated the universal conservation of CUU_A.GG_C pattern in these species (Fig 3D) indicating that it evolves under purifying selection suggesting functional importance of this frameshifting site for these species' fitness. When tested in its native context, YFS2 frameshifting is approx. 40% efficient, like the CUU_A.GG_C heptamer efficiency. Therefore, it is unlikely this novel case contains proximal stimulating or attenuating sequence elements as tested in our luciferase system. We have named this gene *YFS2* (Yeast Frame Shift 2), following the more recent discovery of *YFS1* frameshifting (14). Both *YFS1* and *YFS2* contain the CUU_A.GG_C frameshifting heptamer close to the 5' end of the mRNA transcripts.

The ABP140 mRNA contains a frameshifting stimulator 5' of its shift site

Surprisingly, *ABP140* +1 frameshifting is 60% efficient despite it containing the same frameshifting heptamer, CUU_A.GG_C, whose frameshifting efficiency is ~40% in other contexts, suggesting a stimulatory frameshifting element within the flanking 200 nt native sequence. The unusually high efficiency of frameshifting in endogenous *ABP140* context is further supported by ribosome profiling data (Fig. 3A,B). We used RNAfold (37) to predict potential RNA secondary structures within flanking regions and identified a potential stable RNA stem loop 6 nt upstream of the frameshifting heptamer (Fig. 4A,B). An alignment of *ABP140* sequences from multiple *Saccharomyces* species shows conservation of sequence specifying the predicted stem loop at this position with several compensatory substitutions that maintain potential base pairing (Fig. 3E). To analyse whether there is an increased purifying selection acting on synonymous positions within *ABP140* which would be expected in the presence of RNA secondary structure under purifying evolutionary selection, we used Synplot2 (38). It reveals a statistically significant decrease in the synonymous substitution rates in comparison with regions outside of the potential RNA secondary structure (Fig. 3C,D).

Introduction of synonymous changes that would disrupt the predicted RNA secondary structure in the pYSGDLuc bearing the *ABP140* frameshifting cassette reduced frameshifting levels to ~40%. This level is similar to what is obtained with the frameshifting heptamer alone (Fig. 5C), suggesting that the sequence encoding the structure functions as a stimulator of frameshifting. Note that another possibility could be that *ABP140* frameshifting is stimulated by the nascent peptide within the ribosome peptide tunnel, as discovered in fungal antizyme sequences (39). To test this possibility, we generated mutations that did not change the

sequence of the encoded protein but altered the mRNA structure (Fig. 5B). This also removed the stimulatory effect suggesting that the stimulation occurs by the RNA structure and not the encoded peptide sequence. Replacing the *ABP140* stem loop structure with a shorter variant found in *Kluyveromyces marxianus* ortholog (Supplementary Fig. 2) reduced frameshifting to below 50%, but it remains well above 40% in the absence of a stem loop structure at this position (Fig. 5C). Progressive truncations from the 5' end of the *S. cerevisiae ABP140* test sequence also showed that stem loop disruption reduced frameshifting efficiency, and the entire structure is needed for the stimulation. Shortening the structure by 7 nts removed its stimulatory effect (Fig. 5D). Collectively these data provide strong support for the proposition of the upstream RNA structure as a frameshifting stimulator.

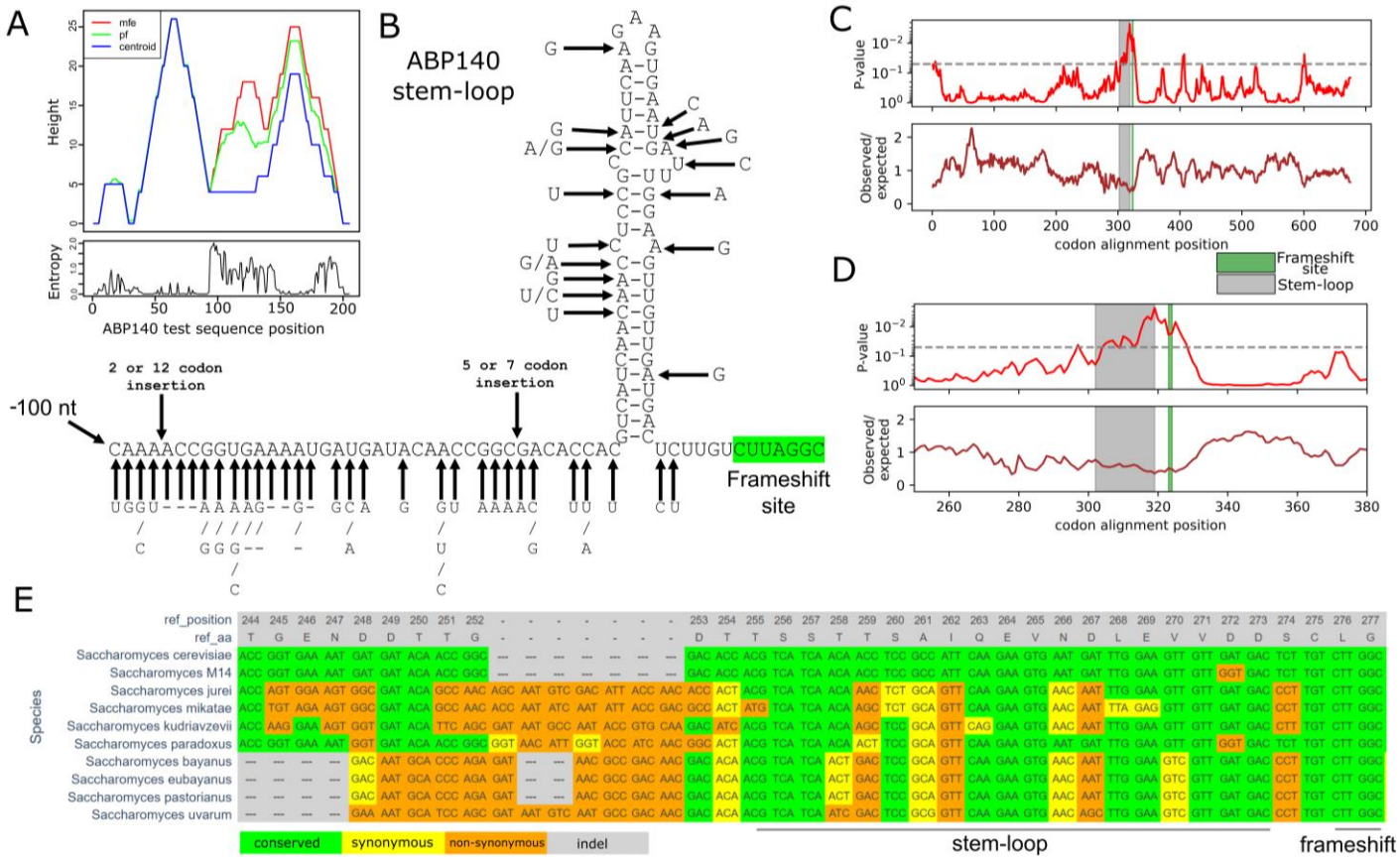

Figure 4. Conserved RNA stem-loop in *ABP140* upstream of the frameshifting site. A. RNA secondary structure (minimum free energy, partition function and centroid values) and entropy calculations across the ABP140 test sequence as calculated via RNAfold. **B.** RNA secondary diagram, base substitution in orthologous sequences Base substitutions from the *Saccharomyces* are denoted with arrows. **C.** Analysis of the rate of synonymous substitutions with Synplot2 for the ABP140 protein coding region. Observed/expected ratio of substitutions is shown at the bottom track with corresponding p-values above. Dotted line represents $p < 0.05$ threshold. Below is a zoomed region for the frameshifting cassette. Synplot2 data are visualized from a 17-codon sliding window. **E.** Codon alignment of the ABP140 coding sequence with *S. cerevisiae* as a reference.

In theory, in addition to frameshifting, the second luciferase could also be expressed via ribosome reinitiation downstream of the first ORF by non-frameshifted ribosomes that terminated at the in-frame stop codon. To exclude this possibility, we removed the second StopGo (E2A) site in the vector (between test sequence and Firefly) and caried out Western Blotting Analysis of yeast cells expressing this reporter. It showed only a

single Firefly luciferase product and no shorter products that would be expected in case of reinitiation (Supplementary Fig. 3). The frameshifting efficiency measured with luciferase assays remained the same (Fig. 5C).

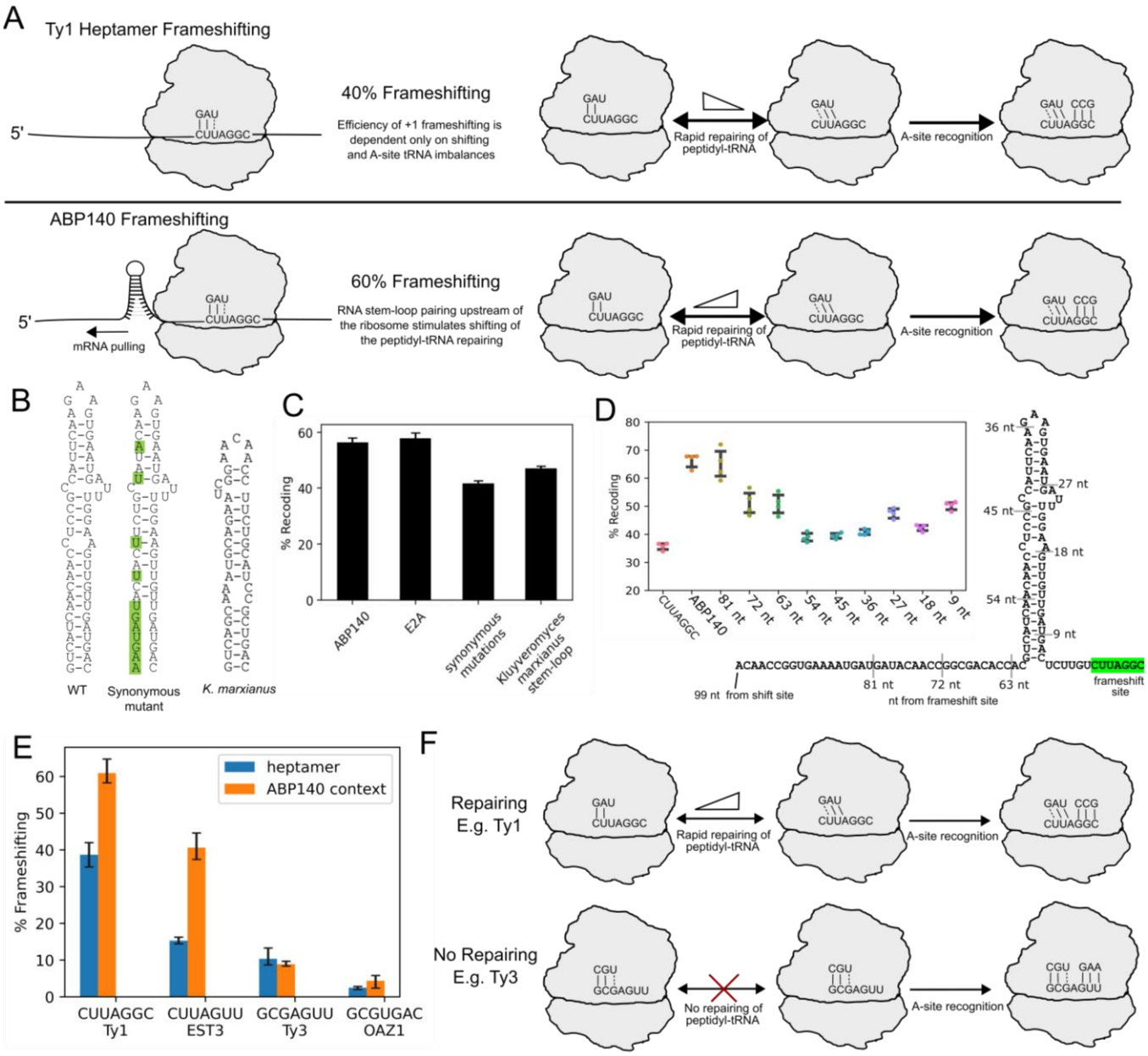

Figure 5. The stimulatory mode of upstream RNA structure affirms two distinct mechanisms of +1 ribosomal frameshifting. A. The proposed model of +1 ribosomal frameshifting by the upstream RNA structure stimulator. **B.** Tested RNA structures, WT, mutants (via synonymous codon substitutions) and the stem loop of *Kluyveromyces marxianus*. **C.** Luciferase results showing frameshifting on the *ABP140* test sequence for wildtype and mutants. E2A represents deletion of the second E2A sequence positioned between the test sequence and firefly luciferase coding sequence. **D.** Luciferase assay results showing truncations of the 5' mRNA context of the *ABP140* test sequence. As a control, the Ty1 heptamer and the original *ABP140* test sequence was tested alongside the truncations. **E.** Luciferase assay results showing wildtype frameshifting efficiencies of the heptamers (blue bars) and when these frameshifting heptamers are placed in the context of *ABP140* (orange bars).

The stimulatory RNA structure promotes P-site tRNA slippage and affirms existence of two distinct mechanisms of +1 ribosomal frameshifting.

Most RNA structures stimulating ribosomal frameshifting have been found downstream of frameshifting sites promoting -1 ribosomal frameshifting (4, 40, 41). In SARS-CoV-2 and likely other instances of ribosomal frameshifting, the stimulation is achieved as a result of the tensions in mRNA between the decoding centre of the ribosome (A and P-sites) and mRNA tunnel entrance where the stimulatory structure is located. This tension prevents triplet movement of mRNA together with tRNAs during the translocation (42). However, in bacteria frameshifting is often stimulated by Shine-Dalgarno (SD) and anti-SD interactions. This is likely based on a similar principle of tensions within mRNA caused by these interactions. A short distance between the SD and decoding centre pulls mRNA out of the decoding centre promoting +1 frameshifting while a long distance pushes mRNA towards decoding centre promoting -1 frameshifting (43, 44). Furthermore, a somewhat similar phenomenon occurs during transcription in bacteria. Transcription terminators consist of RNA secondary structures upstream of polyU in RNA:DNA hybrid within the RNA polymerase bubble and termination involves pulling polyU towards the structure during its formation (45). Transcriptional slippage occurring due to the realignment of RNA relative to its DNA template is also known to be stimulated by RNA secondary structures upstream (46). Inspired by these examples we proposed a model of how the RNA stem-loop stimulator discovered in this study may stimulate +1 frameshifting during decoding of the *ABP140* mRNA (Fig. 5A). When mRNA is exiting the ribosome, it starts forming the RNA secondary structure eventually pulling mRNA out of the ribosome, thus promoting forward movement of the P-site tRNA and its realignment from the zero frame CUU codon with the +1 UUA codon.

If this model is correct the RNA secondary structure should only promote +1 frameshifting that involves P-site tRNA movement relative to mRNA (top two mechanisms in Figure 1, but not the bottom one). Thus, we explored the *ABP140* frameshifting stimulator within the context of other known frameshifting heptamers. To achieve this, we substituted the native CUU_A.GG_C frameshifting site with the frameshifting heptamers from *EST3*, *Ty3*, and *OAZ1*. If the model is correct and we observe a significant increase in frameshifting efficiency, it would imply that all frameshifting sites involve the repairing of peptidyl-tRNA to the overlapping +1 frame codon. We observed a substantial increase in +1 frameshifting from 15% to 40% for the EST3 heptamer CUU_A.GU_U, but no significant increase for the Ty3 (GCG_A.GU_U) and OAZ1 (GCG_U.GA_C) heptamers (Fig. 5E). Therefore, these results strongly suggest that ribosomal frameshifting on heptamers with GCG in the P-site do not involve tRNA repairing with the +1 codon as originally suggested by (22) and instead, a distinct +1 frameshifting mechanism is present.

DISCUSSION

Prior work on ribosomal frameshifting in *S. cerevisiae* provided widely varying estimates of frameshifting efficiencies, likely because of high variability in assays and the frameshifting cassettes used (47–50).

Therefore, to be able to compare different cases of frameshifting, it is important to measure them uniformly. In this work, we set out to explore all known instances of +1 ribosomal frameshifting in *S. cerevisiae* using a combination of approaches: reporter assays, ribosome profiling and comparative sequence analysis. While the aim of the study was to provide a comprehensive and uniform characterisation of frameshifting cassettes across different *S. cerevisiae* genes utilizing ribosomal frameshifting in their expression, several unexpected findings arose.

First, we found a gene, *YJR112W-A*, that is misannotated in the *S. cerevisiae* genome annotation as an intron-containing gene, while it is apparently transcribed into an intronless mRNA expressing a protein encoded in two ORFs, that are decoded as a single protein via +1 ribosomal frameshifting. This finding illustrates that even nowadays, the protein coding catalogues of eukaryotic genes are not compete even in species with comparatively low frequencies of splicing such as *S. cerevisiae*. It demonstrates that the current annotation pipelines do not adequately capture the complexity of mRNA decoding mechanisms and fail to detect rare decoding strategies. This finding also demonstrates the utility of ribosome profiling data at overcoming these limitations in identifying novel translated regions.

Second, unexpectedly we identified a novel mechanism of ribosomal frameshifting stimulation. Hitherto frameshifting stimulatory structures were identified downstream of frameshifting sites and are believed to act upon elongating ribosomes by slowing down its movement because of the requirement to unwind these structures (42). Here, we identified a stimulatory structure upstream of frameshifting site in *ABP140*, suggesting that it acts on the ribosome downstream by promoting its forward movement. Such a possibility is supported by a previous study, in which a stem-loop was shown to stimulate +1 frameshifting when artificially placed upstream of the Ty1 frameshifting heptamer in *S. cerevisiae* (51). It is likely that such a mode of frameshifting stimulation is not limited to *S. cerevisiae* and may work in other organisms.

Third, we utilized this novel stimulator to address a long-standing question in the ribosomal frameshifting field. How can +1 frameshifting work in the absence of apparent repairing of the P-site tRNA with an overlapping P-site codon? Placing the structure upstream of different frameshifting sites revealed its specificity to those frameshifting sites that allow for P-site tRNA repairing. It provided no stimulation at frameshifting sites where repairing is seemingly impossible, suggesting that in these cases, frameshifting does not involve repositioning of the P-site codon from the zero to the +1 frame. This finding argues against a unified model of ribosomal frameshifting in which the forward movement of the P-site tRNA is required irrespective of its ability to repair with the +1 codon (12). On the contrary, it suggests that a distinct mechanism exists enabling A-site tRNA acceptance at the +1 codon that does not require P-site tRNA repositioning as suggested earlier (22). The existence of two such mechanisms may not be specific to yeast, frameshifting sites with seemingly impossible P-site codon repairing have been reported in glass sponge mitochondria (52) and under severe limitations of specific amino acids in *E. coli* (53, 54).

Finally, while in this study, we focused on the *ABP140* stem-loop, other cases of frameshifting also showed intriguing results. Both the *Ty1* and *EST3* test sequences with native contexts showed a decrease in frameshifting efficiency suggesting that these contexts may have attenuator elements. Surprisingly, *EST3* supposedly contains a stimulatory element (50) surrounding the frameshifting heptamer and we do not

observe any such effect with both experimental reporters and publicly available ribosome profiling data analysis.

Methods

Cloning

SnapGene (www.snapgene.com) was used to codon optimise (based on codon usage) Renilla and Firefly luciferase sequences from pSGDLuc (24) and synthesized as a gblock from IDT (Integrated DNA Technologies, Leuven, Belgium). The reporter cassette was inserted into the pGREG-505 plasmid which was linearized with SalI (NEB, #R3138S) and size selected to remove the *HIS3* coding sequence. For insertion of short frameshifting heptamers, annealed oligos were ligated into the NdeI (NEB, #R0111S) and NheI (NEB, #R3131S) digested vector with T4 DNA ligase (NEB, #M0202S) and incubated overnight at room temperature in a 5 µL reaction. For insertion of longer frameshifting cassettes, 60 ng of NdeI and NheI digested plasmid and ~7 ng of ~200 bp inserts were added to 15 µL of homemade Gibson assembly master mix (55) and incubated at 50ºC for 1 hour in a thermocycler. 5 µL of the Gibson assembly or T4 ligation reaction was transformed into chemically competent *E. coli* DH5a cells.

Yeast culturing and transformation

5 mL of YPD broth was inoculated with a single colony of *S. cerevisiae* (strain BY4741) and incubated overnight at 30ºC and 200 RPM. The next day, 200-400 ng of plasmid was transformed into cells *via* the Lithium acetate/Salmon sperm carrier method (56). Cultures were plated onto SC 2% glucose media with an amino acid drop out -Leucine mix (Formedium, #DSCK052) agar plates and incubated at 30ºC for 2-3 days until colonies formed.

Luciferase Assays

Two independent yeast colonies were inoculated in 5 mL 2% glucose -Leucine medium in a 50 mL Falcon tube and incubated overnight at 30ºC and 200 rpm. The next day, the A600 (O.D. 600) values were measured and the culture was diluted to an A600 of ~0.1. For the 96 well plate assay, 200 µL of culture was plated into each well (a total of six wells per biological replicate). The plate was sealed using gas-permeable tissue culture seals (4titude, #4ti-0516/384), placed on a plate shaker inside a 30ºC incubator and incubated for ~6 hours. 50 µL of the culture was then transferred to a white full area 96 well plate and 50 µL of 2X Passive Lysis Buffer (Promega) was added to each well and incubated for 50-60 minutes with shaking at room temperature to allow lysis. 50 µL of homemade LAR (substrate for Firefly luciferase) and 50 µL homemade StopGlow (substrate for Renilla luciferase) were injected into each well (57).

Western Blotting

Yeast transformants were inoculated in 5 mL -Leucine media in a 50 mL falcon tube and incubated overnight at 30ºC with 200 RPM. A600 values were measured and ~10 A600 units were transferred to a 1.5 mL tube.

Cells were washed once in sterile water and boiled in 280 µL of 1X SDS-sample buffer for 3 minutes. 5-10 µL was transferred onto NuPage protein gels (Invitrogen) and run for 200V for ~24 minutes with 1X MES running buffer (Invitrogen). Proteins were transferred to a Nitrocellulose membrane using a BioRad Transblot. Membranes were blocked with 5% low fat milk in 1X PBS-T for 45 minutes at room temperature with gentle agitation. The membranes were treated with a primary antibody solution consisting of mouse anti-Renilla (Millipore) and goat anti-Firefly (Promega) in 1% BSA and PBS-T and was incubated overnight at 4ºC. The membrane was washed three times in PBS-T for 5 minutes each before and after secondary antibodies (anti-mouse red and anti-goat green). Membranes were visualized using an Odyssey imager. Densitometry analyses were carried out using imageJ.

Ribosome Profiling Analysis

Data from each study was obtained via the Ribocrypt browser (https://ribocrypt.org/). To determine the +1-frameshifting efficiency, the number of RPFs on each frame was normalized by the number of codons. The normalized number of RPFs in the +1 frame was divided by the number in the 0-frame and multiplied by 100 to obtains the % frameshifting.

Bioinformatic and Computational Analyses

ABP140 homologs were found using TBLASTN (58) against the genomes of budding yeast species. To produce in-frame alignments for homologs that do not contain frameshifting, the +1-frame was fused to the 0-frame by replacing CTTAGGC with CTTGGC. RNAfold was used to determine RNA secondary structures (37). Orthologous sequences were aligned using MUSCLE (59) and converted to a codon alignment using PAL2NAL (60). A codon alignment viewer was generated with python. The rate of synonymous across multiple sequence alignment was analysed with Synplot2 (38), using a window size of 17 codons.

Author Contributions

DAF and MB carried out experiments. DAF and MS carried out bioinformatic analyses. DAF, JFA and PVB conceived the study. DAF, GL and MMY designed experiments and analysed results. DAF, JFA and PVB drafted the manuscript and all authors participating in manuscript editing. PVB, JK and JFA secured funding and provided supervision.

Acknowledgements

We thank Yousuf Khan (Stanford University) and Sinéad O'Loughlin for technical advice in the early stages of this project. We thank Javier Valera (University College Cork) for supplying the pGREG-505 plasmid. This research was funded in whole or in part by the Wellcome Trust [210692/Z/18/] and Taighde Éireann – Research Ireland [20/FFP-A/8929] to P.V.B. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.

This work was also supported by grants from Poland National Science Centre [UMO-2021/41/B/NZ2/03036] to J.K. and Irish Research Council [IRCLA/2019/74] to J.F.A.

Conflict of Interests

P.V.B. and G.L. are cofounders and shareholders of EIRNABio Ltd.

References

1. Gesteland,R.F., Weiss,R.B. and Atkins,J.F. (1992) Recoding: reprogrammed genetic decoding. *Science (New York, N.Y.)*, **257**, 1640–1641.

2. Baranov,P. V, Atkins,J.F. and Yordanova,M.M. (2015) Augmented genetic decoding: global, local and temporal alterations of decoding processes and codon meaning. *Nature reviews. Genetics*, **16**, 517–529.

3. Rodnina,M. V (2023) Decoding and Recoding of mRNA Sequences by the Ribosome. *Annual review of biophysics*, **52**, 161–182.

4. Atkins,J.F., Loughran,G., Bhatt,P.R., Firth,A.E. and Baranov,P. V (2016) Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. *Nucleic Acids Research*, **44**, 7007–7078.

5. Gaydukova,S.A., Moldovan,M.A., Vallesi,A., Heaphy,S.M., Atkins,J.F., Gelfand,M.S. and Baranov,P. V (2023) Nontriplet feature of genetic code in Euplotes ciliates is a result of neutral evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **120**, e2221683120.

6. Wilson,W., Malim,M.H., Mellor,J., Kingsman,A.J. and Kingsman,S.M. (1986) Expression strategies of the yeast retrotransposon Ty: a short sequence directs ribosomal frameshifting. *Nucleic acids research*, **14**, 7001–7016.

7. Clare,J.J., Belcourt,M. and Farabaugh,P.J. (1988) Efficient translational frameshifting occurs within a conserved sequence of the overlap between the two genes of a yeast Ty1 transposon. *Proceedings of the National Academy of Sciences of the United States of America*, **85**, 6816–6820.

8. Mellor,J., Fulton,S.M., Dobson,M.J., Wilson,W., Kingsman,S.M. and Kingsman,A.J. (1985) A retrovirus-like strategy for expression of a fusion protein encoded by yeast transposon Ty1. *Nature*, **313**, 243–246.

9. Clare,J. and Farabaugh,P. (1985) Nucleotide sequence of a yeast Ty element: evidence for an unusual mechanism of gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, **82**, 2829–2833.

10. Belcourt,M.F. and Farabaugh,P.J. (1990) Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. *Cell*, **62**, 339–352.

11. Chan,P.P. and Lowe,T.M. (2019) tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. *Methods in molecular biology (Clifton, N.J.)*, **1962**, 1–14.

12. Baranov,P. V, Gesteland,R.F. and Atkins,J.F. (2004) P-site tRNA is a crucial initiator of ribosomal frameshifting. *RNA (New York, N.Y.)*, **10**, 221–230.

13. Asakura,T., Sasaki,T., Nagano,F., Satoh,A., Obaishi,H., Nishioka,H., Imamura,H., Hotta,K., Tanaka,K., Nakanishi,H., *et al.* (1998) Isolation and characterization of a novel actin filament-binding protein from Saccharomyces cerevisiae. *Oncogene*, **16**, 121–130.

14. Ivanov,I.P., Gaikwad,S., Hinnebusch,A.G. and Dever,T.E. (2020) Conserved +1 translational frameshifting in the S. cerevisiae gene encoding YPL034W. *bioRxiv*, 10.1101/2020.04.29.069534.

15. Morris,D.K. and Lundblad,V. (1997) Programmed translational frameshifting in a gene required for yeast telomere replication. *Current biology : CB*, **7**, 969–976.

16. Farabaugh,P.J., Zhao,H. and Vimaladithan,A. (1993) A novel programed frameshift expresses the POL3 gene of retrotransposon Ty3 of yeast: frameshifting without tRNA slippage. *Cell*, **74**, 93–103.

17. Sundararajan,A., Michaud,W.A., Qian,Q., Stahl,G. and Farabaugh,P.J. (1999) Near-Cognate Peptidyl-tRNAs Promote +1 Programmed Translational Frameshifting in Yeast. *Molecular Cell*, **4**, 1005–1015.

18. Palanimurugan,R., Scheel,H., Hofmann,K. and Dohmen,R.J. (2004) Polyamines regulate their synthesis by inducing expression and blocking degradation of ODC antizyme. *The EMBO journal*, **23**, 4857–4867.

19. Ivanov,I.P., Gesteland,R.F. and Atkins,J.F. (2006) Evolutionary specialization of recoding: frameshifting in the expression of S. cerevisiae antizyme mRNA is via an atypical antizyme shift site but is still +1. *RNA (New York, N.Y.)*, **12**, 332–337.

20. Kurian,L., Palanimurugan,R., Gödderz,D. and Dohmen,R.J. (2011) Polyamine sensing by nascent ornithine decarboxylase antizyme stimulates decoding of its mRNA. *Nature*, **477**, 490–494.

21. Vimaladithan,A. and Farabaugh,P.J. (1994) Special peptidyl-tRNA molecules can promote translational frameshifting without slippage. *Molecular and Cellular Biology*, **14**, 8107–8116.

22. Pande,S., Vimaladithan,A., Zhao,H. and Farabaugh,P.J. (1995) Pulling the ribosome out of frame by +1 at a programmed frameshift site by cognate binding of aminoacyl-tRNA. *Molecular and cellular biology*, **15**, 298–304.

23. Loughran,G., Andreev,D.E., Terenin,I.M., Namy,O., Mikl,M., Yordanova,M.M., McManus,C.J., Firth,A.E., Atkins,J.F., Fraser,C.S., *et al.* (2025) Guidelines for minimal reporting requirements, design and interpretation of experiments involving the use of eukaryotic dual gene expression reporters (MINDR). *Nat Struct Mol Biol*, 10.1038/s41594-025-01492-x.

24. Loughran,G., Howard,M.T., Firth,A.E. and Atkins,J.F. (2017) Avoidance of reporter assay distortions from fused dual reporters. *RNA (New York, N.Y.)*, **23**, 1285–1289.

25. Stahl,G., Bidou,L., Rousset,J.P. and Cassan,M. (1995) Versatile vectors to study recoding: conservation of rules between yeast and mammalian cells. *Nucleic acids research*, **23**, 1557–1560.

26. Shah,A.A., Giddings,M.C., Parvaz,J.B., Gesteland,R.F., Atkins,J.F. and Ivanov,I.P. (2002) Computational identification of putative programmed translational frameshift sites. *Bioinformatics (Oxford, England)*, **18**, 1046–1053.

27. Grentzmann,G., Ingram,J.A., Kelly,P.J., Gesteland,R.F. and Atkins,J.F. (1998) A dual-luciferase reporter system for studying recoding signals. *RNA (New York, N.Y.)*, **4**, 479–486.

28. Fixsen,S.M. and Howard,M.T. (2010) Processive selenocysteine incorporation during synthesis of eukaryotic selenoproteins. *Journal of molecular biology*, **399**, 385–396.

29. Luke,G. and Ryan,M. (2024) The 2A Story: The End of the Beginning. In Seifi,M. (ed). IntechOpen, Rijeka.

30. Donnelly,M.L.L., Luke,G., Mehrotra,A., Li,X., Hughes,L.E., Gani,D. and Ryan,M.D. (2001) Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *The Journal of general virology*, **82**, 1013–1025.

31. Loughran,G., Fedorova,A.D., Khan,Y.A., Atkins,J.F. and Baranov,P. V (2022) Lack of evidence for ribosomal frameshifting in ATP7B mRNA decoding. *Molecular cell*, **82**, 3745-3749.e2.

32. Souza-Moreira,T.M., Navarrete,C., Chen,X., Zanelli,C.F., Valentini,S.R., Furlan,M., Nielsen,J. and Krivoruchko,A. (2018) Screening of 2A peptides for polycistronic gene expression in yeast. *FEMS Yeast Research*, **18**.

33. Barry,J.K. and Miller,W.A. (2002) A -1 ribosomal frameshift element that requires base pairing across four kilobases suggests a mechanism of regulating ribosome and replicase traffic on a viral RNA. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 11133–11138.

34. Smith,A.M., Costello,M.S., Kettring,A.H., Wingo,R.J. and Moore,S.D. (2019) Ribosome collisions alter frameshifting at translational reprogramming motifs in bacterial mRNAs. *Proceedings of the National Academy of Sciences of the United States of America*, **116**, 21769–21779.

35. Cook,G.M., Brown,K., Shang,P., Li,Y., Soday,L., Dinan,A.M., Tumescheit,C., Mockett,A.P.A., Fang,Y., Firth,A.E., *et al.* (2022) Ribosome profiling of porcine reproductive and respiratory syndrome virus reveals novel features of viral gene expression. *eLife*, **11**.

36. Michel,A.M., Kiniry,S.J., O'Connor,P.B.F., Mullan,J.P. and Baranov,P. V (2018) GWIPS-viz: 2018 update. *Nucleic Acids Research*, **46**, D823–D830.

37. Lorenz,R., Bernhart,S.H., Höner Zu Siederdissen,C., Tafer,H., Flamm,C., Stadler,P.F. and Hofacker,I.L. (2011) ViennaRNA Package 2.0. *Algorithms for molecular biology : AMB*, **6**, 26.

38. Firth,A.E. (2014) Mapping overlapping functional elements embedded within the protein-coding regions of RNA viruses. *Nucleic acids research*, **42**, 12425–12439.

39. Yordanova,M.M., Wu,C., Andreev,D.E., Sachs,M.S. and Atkins,J.F. (2015) A Nascent Peptide Signal Responsive to Endogenous Levels of Polyamines Acts to Stimulate Regulatory Frameshifting on Antizyme mRNA. *The Journal of biological chemistry*, **290**, 17863–17878.

40. Hill,C.H. and Brierley,I. (2023) Structural and Functional Insights into Viral Programmed Ribosomal Frameshifting. *Annual review of virology*, **10**, 217–242.

41. Korniy,N., Samatova,E., Anokhina,M.M., Peske,F. and Rodnina,M. V (2019) Mechanisms and biomedical implications of -1 programmed ribosome frameshifting on viral and bacterial mRNAs. *FEBS letters*, **593**, 1468–1482.

42. Bhatt,P.R., Scaiola,A., Loughran,G., Leibundgut,M., Kratzel,A., Meurs,R., Dreos,R., O'Connor,K.M., McMillan,A., Bode,J.W., *et al.* (2021) Structural basis of ribosomal frameshifting during translation of the SARS-CoV-2 RNA genome. *Science*, **372**, 1306–1313.

43. O'Connor,P.B.F., Li,G.-W., Weissman,J.S., Atkins,J.F. and Baranov,P. V (2013) rRNA:mRNA pairing alters the length and the symmetry of mRNA-protected fragments in ribosome profiling experiments. *Bioinformatics (Oxford, England)*, **29**, 1488–1491.

44. Atkins,J.F., Baranov,P. V, Fayet,O., Herr,A.J., Howard,M.T., Ivanov,I.P., Matsufuji,S., Miller,W.A., Moore,B., Prère,M.F., *et al.* (2001) Overriding standard decoding: implications of recoding for ribosome function and enrichment of gene expression. *Cold Spring Harbor symposia on quantitative biology*, **66**, 217–232.

45. Peters,J.M., Vangeloff,A.D. and Landick,R. (2011) Bacterial transcription terminators: the RNA 3'-end chronicles. *Journal of molecular biology*, **412**, 793–813.

46. Penno,C., Kumari,R., Baranov,P. V, van Sinderen,D. and Atkins,J.F. (2017) Stimulation of reverse transcriptase generated cDNAs with specific indels by template RNA structure: retrotransposon, dNTP balance, RT-reagent usage. *Nucleic acids research*, **45**, 10143–10155.

47. Guarraia,C., Norris,L., Raman,A. and Farabaugh,P.J. (2007) Saturation mutagenesis of a +1 programmed frameshift-inducing mRNA sequence derived from a yeast retrotransposon. *RNA (New York, N.Y.)*, **13**, 1940–1947.

48. Kawakami,K., Pande,S., Faiola,B., Moore,D.P., Boeke,J.D., Farabaugh,P.J., Strathern,J.N., Nakamura,Y. and Garfinkel,D.J. (1993) A rare tRNA-Arg(CCU) that regulates Ty1 element ribosomal frameshifting is essential for Ty1 retrotransposition in Saccharomyces cerevisiae. *Genetics*, **135**, 309–320.

49. Kirchner,J., Sandmeyer,S.B. and Forrest,D.B. (1992) Transposition of a Ty3 GAG3-POL3 fusion mutant is limited by availability of capsid protein. *Journal of virology*, **66**, 6081–6092.

50. Taliaferro,D. and Farabaugh,P.J. (2007) An mRNA sequence derived from the yeast EST3 gene stimulates programmed +1 translational frameshifting. *RNA (New York, N.Y.)*, **13**, 606–613.

51. Cho,C.-P., Lin,S.-C., Chou,M.-Y., Hsu,H.-T. and Chang,K.-Y. (2013) Regulation of Programmed Ribosomal Frameshifting by Co-Translational Refolding RNA Hairpins. *PLOS ONE*, **8**, e62283.

52. Haen,K.M., Pett,W. and Lavrov,D. V (2014) Eight new mtDNA sequences of glass sponges reveal an extensive usage of +1 frameshifting in mitochondrial translation. *Gene*, **535**, 336–344.

53. Peter,K., Lindsley,D., Peng,L. and Gallant,J.A. (1992) Context rules of rightward overlapping reading. *The New biologist*, **4**, 520–526.

54. Lindsley,D. and Gallant,J. (1993) On the directional specificity of ribosome frameshifting at a ‘hungry’ codon. *Proceedings of the National Academy of Sciences of the United States of America*, **90**, 5469–5473.

55. Rabe,B.A. and Cepko,C. (2020) A Simple Enhancement for Gibson Isothermal Assembly. *bioRxiv*, 10.1101/2020.06.14.150979.

56. Gietz,R.D. and Schiestl,R.H. (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, **2**, 31–34.

57. Dyer,B.W., Ferrer,F.A., Klinedinst,D.K. and Rodriguez,R. (2000) A noncommercial dual luciferase enzyme assay system for reporter gene analysis. *Analytical biochemistry*, **282**, 158–161.

58. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *Journal of molecular biology*, **215**, 403–410.

59. Edgar,R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, **32**, 1792–1797.

60. Suyama,M., Torrents,D. and Bork,P. (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research*, **34**, W609–W612.