Nonsense suppression by near-cognate tRNAs employs alternative base pairing at codon positions 1 and 3

Bijoyita Roya, John D. Leszykb, David A. Mangusa, and Allan Jacobsona,1

Departments of *Microbiology and Physiological Systems and ²Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, MA 01655-0122

Edited by Michael Rosbash, Howard Hughes Medical Institute/Brandeis University, Waltham, MA, and approved February 6, 2015 (received for review December 16, 2014)

Premature termination codons (PTCs) in an mRNA ORF inactivate gene function by causing production of a truncated protein and destabilization of the mRNA. Readthrough of a PTC allows ribosomal A-site insertion of a near-cognate tRNA, leading to synthesis of a full-length protein from otherwise defective mRNA. To understand the mechanism of such nonsense suppression, we developed a yeast system that allows purification and sequence analysis of full-length readthrough products arising as a consequence of endogenous readthrough or the compromised termination fidelity attributable to the loss of Upf (up-frameshift) factors, or the presence of the aminoglycoside gentamicin. Unlike classical "wobble" models, our analyses showed that three of four possible near-cognate tRNAs could mispair at position 1 or 3 of nonsense codons and that, irrespective of whether readthrough is endogenous or induced, the same sets of amino acids are inserted. We identified the insertion of Gin, Tyr, and Lys at UAA and UAG, whereas Trp, Arg, and His were inserted at UGA, and the frequency of insertion of individual amino acids was distinct for specific nonsense codons and readthrough-inducing agents. Our analysis suggests that the use of genetic or chemical means to increase readthrough does not promote novel or alternative mispairing events; rather, readthrough effectors cause quantitative enhancement of endogenous mistranslation events. Knowledge of the amino acids incorporated during readthrough not only elucidates the decoding process but also may allow predictions of the functionality of readthrough protein products.

Significance

Readthrough-promoting drugs cause amino acid insertion at premature termination codons (PTCs), and thus have broad potential as a therapeutic approach to inherited disorders attributable to nonsense mutations. Because the mechanism involved in the insertion of near-cognate tRNAs at nonsense codons is unknown, we have identified the yeast translation errors ensuing from nonsense suppression occurring either inherently or enhanced by drugs or mutations that compromise termination fidelity. Our analyses of the products of nonsense suppression provide insights into the rules that govern readthrough at PTCs and delineate specific nonstandard Watson–Crick codon/anticodon base pairings critical to this process. These results should enable predictions of the likelihood of obtaining functional full-length readthrough products, and thus better clinical outcomes, with therapeutic nonsense suppression.


The authors declare no conflict of interest.

This article was a PNAS Direct Submission.

1To whom correspondence should be addressed. Email: allan.jacobson@umassmed.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1424127112/-/DCSupplemental.
bonds that they form (12). Furthermore, “geometrical mimicry” by nonstandard Watson–Crick pairing has been proposed as a mechanism that favors incorporation of near-cognate tRNAs (12, 13).

To understand how translational fidelity is maintained by the ribosome during termination, it is important to know which mismatches are allowed, even if they occur at a low frequency. Hence, the decoding events occurring during readthrough of PTCs are important phenomena that need to be elucidated. Further, from a therapeutic perspective, the identity of the amino acids that get incorporated at PTCs may be critical in predicting the potential functionality of the readthrough product when readthrough is induced pharmacologically. Several previous studies have sought to identify the amino acids incorporated at a PTC (14–23). In vitro studies in rabbit reticulocyte lysates identified Trp, Arg, and Cys incorporation at a UGA codon (20). Studies addressing the insertion of amino acids at UAA and UAG codons in viral mRNAs revealed only the presence of Gln (20, 24). However, given the nature of these studies, it has been difficult to judge the relative frequencies of insertion or to conclude that the same tRNAs would be involved during readthrough in vivo. Furthermore, the low amount of readthrough product, together with less sensitive detection techniques, could have obscured the identification of low-abundance amino acid insertions, as in the case of the study addressing insertion in place of UAA. A recent study in Saccharomyces cerevisiae tried to address these issues in vivo (25). However, to ensure maximal capture of the readthrough products, this study combined two different readthrough-inducing conditions (up/Δ and [PSI+]). Although this approach did help to identify the insertion of three different amino acids at all three PTCs, the interpretation of the results was not clear-cut because the frequencies of insertions reflected the combined effects of three independent readthrough events, namely, endogenous readthrough, UPF1 deletion, and a defect in eRF3.

Here, we have investigated the mechanisms of near-cognate tRNA insertion during PTC readthrough in yeast by identifying the amino acids that are inserted upon readthrough of each of the three nonsense codons. First, we describe an in vivo reporter system that allows for detection, purification, and mass spectrometric analyses of the endogenous unprogrammed readthrough product from WT cells without any treatment. Having identified the amino acids that get inserted at a PTC in WT cells, we compared and characterized the readthrough products that are synthesized when termination fidelity is compromised (in the absence of Upf factors, loss of functional translation termination factors, or conformational change in the decoding center by gentamicin binding). Our study provides a comprehensive analysis of termination readthrough from a PTC under different readthrough-inducing conditions and elucidates the nature of the nonstandard base pairings that are favored under physiological conditions in the cell.

Results

Reporter System for Studying Readthrough. To identify the specific amino acids incorporated when PTCs are subject to readthrough, we constructed a luciferase gene (LUC) reporter system with several key attributes, including the ability to yield highly purified readthrough products in sufficient quantities for MS even when those products are expressed at low levels and the ability to derive unambiguous sequencing information from well-resolved peptides. Starting with a codon optimized firefly LUC gene, we inserted an in-frame 3x HA tag at the N terminus and adjacent StrepII and FLAG (SF) tags at the C terminus of the luciferase ORF and directed transcription of the construct with the potent TP11 promoter (Fig. L4). Three separate alleles were constructed with each of the possible PTCs inserted at codon 20 of the LUC ORF (Table S1), rendering each respective mRNA an NMD substrate (Fig. 1B). Readthrough of each of the LUC PTCs in WT [PSI+] cells resulted in the translation of full-length luciferase that also retains activity (Fig. 2A), indicating that these mRNAs are susceptible to basal levels of readthrough under normal growth conditions. The efficiency of readthrough from each PTC under different conditions known to compromise termination fidelity (the absence of Upf factors, the presence of defective translation termination factors, or the presence of the aminoglycoside gentamicin) was then assessed as a measure of luciferase activity (normalized to total protein content of the sample as well as to the level of the LUC mRNA) (Fig. 2B). As expected (26), readthrough was observed from all three nonsense codons, with UGA being the most permissive (Fig. 2B). Maximall readthrough was observed when strains were either defective in translation termination factor eRF1 (sup45-2; for UAA) or treated with gentamicin (for UAG and UGA) (Fig. 2B). Taken together, these results demonstrate that the HA-LUC(PTC20)-SF constructs act as bona fide NMD reporters and validate their use as a quantitative readthrough reporter system.

Characterization of the Endogenous Readthrough Product from PTCs. To gain insight into the misreading events that occur when a termination codon is present at the A-site of the ribosome, it is important, first, to characterize these events when the ribosome is not predisposed to any other readthrough stimulus. Accordingly,
to ensure that the readthrough observed was solely a readout of basal readthrough events, we expressed the HA-LUC\textsubscript{PTC20}-SF reporters in WT cells with a \textsubscript{PSF} background [i.e., cells devoid of the prion form of Sup35 (eRF3)]. Hence, any detectable readthrough in these cells is unlikely to be attributable to Sup35 aggregation (27). Strep-Tactin (IBA) purification of WT luciferase (Fig. 4A) and the readthrough products expressed from all three PTC alleles was followed by gel electrophoretic analysis to resolve the full-length products. Following silver staining of the gel, the band corresponding to full-length luciferase was excised and subjected to endo-LysC digestion and liquid chromatography-tandem MS analysis. Analyses of the readthrough products purified from WT \textsubscript{PSF} cells showed that when UAA was the PTC, three amino acids were incorporated at that position, namely, Gln, Lys, and Tyr (Fig. 4A). Three independent experiments revealed that the frequencies of insertion of Tyr and Gln were similar (45.5 ± 5% and 54 ± 7%, respectively) but that Lys was inserted at a much lower frequency (0.5 ± 0.1%) (Fig. 4A). The same sets of amino acids were inserted when readthrough from a UAG occurred in WT \textsubscript{PSF} cells, but with different frequencies. Gln was identified as the predominant amino acid, with an insertion frequency of 80 ± 6% (Fig. 4B). Tyr (11 ± 4%) and Lys (9 ± 2%), on the other hand, had similar insertion frequencies (Fig. 4B). With UGA as the PTC, we identified the incorporation of Trp, Arg, and Cys (Fig. 4C and Fig. S1). Trp was inserted at a higher frequency (86 ± 4%) than both Arg (7 ± 3%) and Cys (7 ± 2%) (Fig. 4C). The identification of Trp, Arg, and Cys as a consequence of UGA readthrough confirms observations from earlier studies (20). For UAG readthrough products, identification of Gln as the most abundant amino acid suggests that an earlier study with viral mRNAs that identified only Gln, and not Lys or Tyr, might reflect the lower frequencies of insertion for the latter two amino acids. This observation suggests that our approach for identifying the amino acids inserted during nonsense suppression is suitable for the detection of low-frequency events.

**Readthrough in the Absence of Upf Factors.** NMD requires the regulatory factors, Upf1, Upf2, and Upf3 (7). In yeast, the loss of any of these factors results in the stabilization and increased accumulation of PTC-containing mRNAs, with little or no effect on the abundance and stability of most WT transcripts (7). In addition to their role in NMD, the yeast Upf factors control the fidelity of translation termination, as manifested by increased nonsense suppression when these factors are mutated or deleted (10, 28). This effect on nonsense suppression is largely attributable to NMD regulation of the uORF (upstream open reading frame)-containing \textit{ALR1} mRNA and the altered intracellular Mg\textsuperscript{2+} concentration that occurs when this mRNA has been stabilized (29). Recent work in yeast suggests that the nonsense suppression phenotypes of Upf-deficient cells may also include direct effects of the Upf factors on the regulation of translation termination (30). Thus, as expected, our analyses showed that the loss of any of the three Upf factors led to increased readthrough from the HA-LUC\textsubscript{PTC20}-SF constructs (Fig. 2B). Because we observed a distinct pattern of amino acid insertion for endogenous readthrough in WT \textsubscript{PSF} cells, we assayed the readthrough products arising in the absence of the Upf factors for their amino acid incorporation patterns at PTCs. Readthrough products were purified from all three PTC-containing reporters expressed in \textit{upf\textsubscript{1}} (Fig. 3B and Fig. S2).

![Fig. 3. Strep-Tactin purification of HA-LUC\textsubscript{SF} and HA-LUC\textsubscript{PTC20}-SF products. (A) Silver-stained SDS-PAGE gel showing Strep-Tactin purification of the luciferase reporter protein from WT \textsubscript{PSF} cells. (B) Silver-stained gel showing the readthrough products purified from \textit{upf\textsubscript{1}} \textsubscript{PSF} cells expressing HA-LUC\textsubscript{PTC20}-SF reporters. The asterisks denote full-length WT (A) and readthrough (B) luciferase products.](https://www.pnas.org/content/112/9/3040)

![Fig. 4. Amino acids inserted at PTCs during termination readthrough. HA-LUC\textsubscript{PTC20}-SF protein products purified from WT \textsubscript{PSF} cells, \textit{upf}\textsubscript{1} \textsubscript{PSF} mutants, termination factor mutants, or WT \textsubscript{PSF} cells treated with gentamicin (WT + gent) were subjected to MS analyses. The numbers denote the frequency (±SD of the mean) of insertion of the amino acids at UAA (A), UAG (B), or UGA (C) (n = 3). The type and position of mispairing in the codon is represented for each amino acid inserted.](https://www.pnas.org/content/112/9/3040)
upf2Δ, or upf3Δ cells that were also [PSI+] to ensure that the insertions observed were due solely to loss of the Upf factors. Readthrough from UAA to UAG in the absence of each of the Upf factors showed the insertion of Tyr, Gln, and Lys in the position of the PTC (Fig. 4A). The frequencies of insertion of the amino acids were similar in all three deletion strains, with equal distribution of Tyr [upf1Δ (53 ± 5%), upf2Δ (47 ± 4%), and upf3Δ (54 ± 3%)], and Gln [upf1Δ (43 ± 3%), upf2Δ (52 ± 6%), and upf3Δ (45.4 ± 2%)] and a very low frequency of insertion for Lys [upf1Δ (4 ± 1%), upf2Δ (0.8 ± 0.5%), and upf3Δ (0.6 ± 0.4%)].

With UAG as a PTC, the same set of amino acids was inserted (i.e., Tyr, Gln, Lys) (Fig. 4B). Insertion of Gln was found to be predominant in all three strains [upf1Δ (69 ± 5%), upf2Δ (70 ± 5%), and upf3Δ (81 ± 2%)], with Lys being the least favored insertion [upf1Δ (5 ± 1%), upf2Δ (2.0 ± 0.5%), and upf3Δ (9 ± 1%)]. However, the frequency of insertion of Tyr was lower in the upf3Δ strain (9 ± 1%) than observed for the upf1Δ and upf2Δ strains (26 ± 14% and 28 ± 3%, respectively). Analysis of the readthrough products from UGA as the PTC showed the insertion of Arg, Trp, and Cys, similar to amino acid insertions observed from WT cells (Fig. 4C). Trp insertion was favored [upf1Δ (71 ± 5%), upf2Δ (72 ± 3%), and upf3Δ (89 ± 1%)] over insertion of either Arg [upf1Δ (21.2 ± 2%), upf2Δ (25 ± 2%), and upf3Δ (6.2 ± 2%) or Cys [upf1Δ (7.9 ± 2%), upf2Δ (3 ± 0.5%), and upf3Δ (5 ± 0.5%)].

The results of Fig. 4 indicate that upf1Δ and upf2Δ cells exhibited similar amino acid insertion patterns for UAG and UGA, which were distinct from those insertion patterns manifested by upf3Δ cells (the frequencies of insertion of Tyr and Gln at UAG and Trp and Arg at UGA differed significantly, with P < 0.001, as determined by one-way ANOVA followed by post hoc Tukey analyses). It is not clear why deletion of Upf3 would affect readthrough differently than the other Upf factors, but the observation supports the notion that enhanced readthrough in the absence of Upf factors is not likely to be exclusively due to increased magnesium levels in the cell. Rather, the Upf factors might be affecting readthrough in more than one way, independently (Upf1 or Upf2).

**Readthrough in the Presence of Defective Translation Termination Factors.** Translation termination is mediated by eRF1 and eRF3 (Sup45 and Sup35 in yeast), and defects in these factors are known to increase readthrough at PTCs (8, 31). The sup45-2 strain is a temperature-sensitive eRF1 mutant that fails to bind the ribosome upon a temperature shift to 37°C, leading to increased readthrough activity (8). Readthrough from the HA-LUC[PTC20]-SF constructs in sup45-2 [PSI+] cells increased approximately fivefold after a 30-min shift to 37°C (Fig. S3). Readthrough in sup45-2 [PSI+] cells that were not temperature-shifted was similar to readthrough of WT cells (Fig. S3). Characterization of the readthrough products purified from sup45-2 [PSI+] cells after 30 min at 37°C showed that when UAA was the premature stop codon, Tyr, Gln, and Lys were inserted with frequencies of 46.9 ± 5%, 53 ± 8%, and 0.03 ± 0.05%, respectively (Fig. 4A). Gln (87.7 ± 3%) was the predominant insertion at a UAG codon, whereas Tyr was inserted with a frequency of 12 ± 2% and Lys was the least favored insertion (0.3 ± 0.1%) (Fig. 4B). Readthrough at a UGA resulted in the insertion of Trp (82 ± 5%). The results of Fig. 4 indicate that upf1Δ and upf2Δ cells showed over twofold more insertion of Gln (72 ± 3%) than Tyr (27.3 ± 5%) (Fig. 4A). Similar to sup45-2 [PSI+] cells, Lys was the least favored insertion (0.7 ± 1%). At a UAG, [PSI+] cells showed Gln insertion at a frequency of 77 ± 6% and an equal distribution of Tyr (11 ± 2%) and Lys (12 ± 5%) (Fig. 4B). Trp was inserted at a frequency of 89 ± 4% in UGA readthrough products, whereas Cys insertion was observed at 3 ± 0.5%. Interestingly, Arg insertion (8 ± 3%) was an order of magnitude lower than Arg insertion found in sup45-2 [PSI+] cells (Fig. 4C).

In light of the joint roles played by eRF1 and eRF3 in termination (2) it is of interest to determine whether defects in either factor yield comparable readthrough results. For readthrough from UAA, defects in eRF1/Sup45 resembled those defects seen in WT [PSI+] cells, whereas [PSI+] strains exhibited prevalent Gln insertion rather than the approximately equal ratio of Tyr/Gln insertions observed under the other readthrough-inducing conditions. Amino acid insertion frequencies at UAG and UGA in sup45-2 strains were also similar to WT [PSI+], whereas [PSI+] resembled the Upf factor deletion category for UAG insertions and WT [PSI+] for UGA insertions. The resemblance in amino acid insertion frequencies between sup45-2, [PSI+], and WT [PSI+] suggests that the basal level of readthrough that occurs in the WT [PSI+] cells may have its basis in the reduced eRF binding at PTCs thought to occur because of their distance from the normal 3′-UTR of the mRNA (7). The [PSI+] strains resemble the WT [PSI+] readthrough profile for all but UAA. It is not clear at this point what leads to this difference, but it should be noted that our understanding of the [PSI+] prion form of Sup35 is just emerging (32).

**Readthrough in the Presence of an Aminoglycoside.** Having characterized the distribution of amino acids incorporated at PTCs in different genetic conditions that predispose the ribosome to read through a PTC, we then compared the corresponding events when a small-molecule drug is used to induce readthrough. Gentamicin, a commonly used aminoglycoside, can suppress PTCs and restore protein function in vivo by promoting readthrough (2, 3). Binding of gentamicin to the ribosome decoding center can promote termination by induction of A-site structural changes (33). As in previous experiments, we used WT [PSI+] cells to ensure that the readthrough observed is solely due to gentamicin treatment. WT [PSI+] cells treated with varying concentrations of gentamicin showed an increase in readthrough from the HA-LUC[PTC20]-SF reporters, as measured by luciferase activity (34) (Fig. S4). Readthrough products were purified from WT [PSI+] cells treated with gentamicin at a final concentration of 100 μg/mL and the frequencies of amino acid insertion determined as above. Amino acids inserted at the UAA codon included Gln (64 ± 4%), Tyr (34 ± 3%), and Lys (0.5 ± 0.1%) (Fig. 4A). Gln was inserted with a frequency of 67 ± 2%, Tyr with a frequency of 13 ± 4%, and Lys with a frequency of 20 ± 5% when UAG was the PTC (Fig. 4B). Readthrough products from UGA showed predominant insertion of Trp (98 ± 1%), followed by Arg (1.2 ± 0.5%) and then Cys (0.4 ± 0.1%) (Fig. 4C). Even though gentamicin treatment resulted in the insertion of the same amino acids observed in other conditions, the relative frequencies of the respective insertions were significantly different from any other condition (P < 0.001, as determined by one-way ANOVA with post hoc Tukey analysis). The most dramatic of these differences were observed for Tyr and Gln insertion at UAA, Lys and Gln insertion at UAG, and Trp insertion at UGA. These observations suggest that the conformational changes at the A-site due to gentamicin binding induce readthrough in a way that prefers the selection of some near-cognate trRNAs over others, even though cellular RNA abundance should be unaffected.

**Discussion.**

Protein synthesis is a highly efficient and accurate process that allows an organism to translate genomic information into functional
protein products. Despite its apparent accuracy, translation has a misincorporation rate of $10^{-2} - 10^{-3}$ (11). Like other steps in gene expression where the maintenance of fidelity is crucial (e.g., DNA replication, DNA transcription), translation relies on the complementarity of nucleotides to choose the right substrate. The selection of cognate aminoacyl-ribozymes is based on codon/anticodon base pairing and is central to maintaining translation fidelity. Although failures in translational fidelity generally lead to nonfunctional proteins, not all mistranslation events have deleterious effects (35). For example, the [PSI+] state in yeast appears to increase adaptability to environmental cues in parallel with increased readthrough of NTCs (36).

Stop codon readthrough can also be beneficial to a cell when the readthrough event takes place at a PTC. Ordinarily, the occurrence of a PTC in an ORF will lead to the formation of a truncated, nonfunctional polypeptide and a vast reduction in the affected mRNA as a consequence of NMD, and will essentially yield a null allele (37, 38). These consequences are at least partially overridden when the PTC is translated (i.e., read through by the ribosome), leading to in-frame elongation and production of the full-length protein product. Stop codon readthrough at both PTCs and NTCs takes place when recognition of the termination codon by a class I release factor is superseded by readthrough of PTCs in yeast re-

and the phenomenon has been exploited for therapeutic pur-

poses. Here, we have used readthrough of PTCs in yeast to deter-

mine (i) which amino acids are inserted in place of a termination codon when the ribosome reads through a premature stop codon, (ii) whether readthrough induced under different conditions has varying effects on amino acid insertions, and (iii) whether the pattern of near-cognate tRNA insertion in response to novel readthrough effectors can be predicted.

Several important general conclusions can be drawn from our analysis of the readthrough events in response to readthrough effectors. First, our comparisons of the readthrough products from WT [PSI+] cells with cells subjected to various modes of readthrough enhancement indicate that although readthrough efficiency is enhanced by genetic and pharmacological effectors, the phenomenon has been exploited for therapeutic pur-

poses (2, 3), but the mechanism of near-cognate tRNA selection and its enhancement by readthrough-inducing agents is unknown. Here, we have used readthrough of PTCs in yeast reporter genes to determine (i) which amino acids are inserted in place of a termination codon when the ribosome reads through a premature stop codon, (ii) whether readthrough induced under different conditions has varying effects on amino acid insertions, and (iii) whether the pattern of near-cognate tRNA insertion in response to novel readthrough effectors can be predicted.

Several important general conclusions can be drawn from our analysis of the readthrough events in response to readthrough effectors. First, our comparisons of the readthrough products from WT [PSI+] cells with cells subjected to various modes of readthrough enhancement indicate that although readthrough efficiency is enhanced by genetic and pharmacological effectors, the phenomenon has been exploited for therapeutic pur-

poses (2, 3), but the mechanism of near-cognate tRNA selection and its enhancement by readthrough-inducing agents is unknown. Here, we have used readthrough of PTCs in yeast to deter-

mine (i) which amino acids are inserted in place of a termination codon when the ribosome reads through a premature stop codon, (ii) whether readthrough induced under different conditions has varying effects on amino acid insertions, and (iii) whether the pattern of near-cognate tRNA insertion in response to novel readthrough effectors can be predicted.

Several important general conclusions can be drawn from our analysis of the readthrough events in response to readthrough effectors. First, our comparisons of the readthrough products from WT [PSI+] cells with cells subjected to various modes of readthrough enhancement indicate that although readthrough efficiency is enhanced by genetic and pharmacological effectors, the phenomenon has been exploited for therapeutic pur-

poses (2, 3), but the mechanism of near-cognate tRNA selection and its enhancement by readthrough-inducing agents is unknown. Here, we have used readthrough of PTCs in yeast to deter-

mine (i) which amino acids are inserted in place of a termination codon when the ribosome reads through a premature stop codon, (ii) whether readthrough induced under different conditions has varying effects on amino acid insertions, and (iii) whether the pattern of near-cognate tRNA insertion in response to novel readthrough effectors can be predicted.

Several important general conclusions can be drawn from our analysis of the readthrough events in response to readthrough effectors. First, our comparisons of the readthrough products from WT [PSI+] cells with cells subjected to various modes of readthrough enhancement indicate that although readthrough efficiency is enhanced by genetic and pharmacological effectors, the phenomenon has been exploited for therapeutic pur-

poses (2, 3), but the mechanism of near-cognate tRNA selection and its enhancement by readthrough-inducing agents is unknown. Here, we have used readthrough of PTCs in yeast to deter-

mine (i) which amino acids are inserted in place of a termination codon when the ribosome reads through a premature stop codon, (ii) whether readthrough induced under different conditions has varying effects on amino acid insertions, and (iii) whether the pattern of near-cognate tRNA insertion in response to novel readthrough effectors can be predicted.

Several important general conclusions can be drawn from our analysis of the readthrough events in response to readthrough effectors. First, our comparisons of the readthrough products from WT [PSI+] cells with cells subjected to various modes of readthrough enhancement indicate that although readthrough efficiency is enhanced by genetic and pharmacological effectors, the phenomenon has been exploited for therapeutic pur-

poses (2, 3), but the mechanism of near-cognate tRNA selection and its enhancement by readthrough-inducing agents is unknown. Here, we have used readthrough of PTCs in yeast to deter-

mine (i) which amino acids are inserted in place of a termination codon when the ribosome reads through a premature stop codon, (ii) whether readthrough induced under different conditions has varying effects on amino acid insertions, and (iii) whether the pattern of near-cognate tRNA insertion in response to novel readthrough effectors can be predicted.

Several important general conclusions can be drawn from our analysis of the readthrough events in response to readthrough effectors. First, our comparisons of the readthrough products from WT [PSI+] cells with cells subjected to various modes of readthrough enhancement indicate that although readthrough efficiency is enhanced by genetic and pharmacological effectors, the phenomenon has been exploited for therapeutic pur-

poses (2, 3), but the mechanism of near-cognate tRNA selection and its enhancement by readthrough-inducing agents is unknown. Here, we have used readthrough of PTCs in yeast to deter-

mine (i) which amino acids are inserted in place of a termination codon when the ribosome reads through a premature stop codon, (ii) whether readthrough induced under different conditions has varying effects on amino acid insertions, and (iii) whether the pattern of near-cognate tRNA insertion in response to novel readthrough effectors can be predicted.
readthrough efficiencies from either codon position 1 or position 3 mispairing (Fig. 5B) shows higher readthrough from position 3 mispairing events. Higher readthrough from position 3 mispairing might explain why readthrough from UGA, with predominant position 3 mispairing, occurs more readily than it does with the other two terminators. Taken together, our study shows that the decoding center of the ribosome accommodates alternative noncanonical base-pairing events during termination, allowing expansion of the genetic code.

An understanding of the decoding events taking place during premature termination may provide useful insights into the utilization of this process for therapeutic purposes. For example, knowledge of the amino acids inserted at a specific PTC should help to predict the likelihood of success with a readthrough-inducing small molecule, particularly when the targeted PTC would lead to insertion of a novel amino acid in an essential functional site of a protein. Further, because premature termination is less efficient than normal termination (2, 3, 7) it is plausible that some miscoding events that are allowed in the context of a PTC may not happen during normal termination (when there appears to be more efficient eRF sampling and the ribosome does not pause as long) (3). It will thus be of interest to compare the amino acid insertions at NTCs to determine if the rules of near-cognate tRNA insertion are different.

Materials and Methods

Plasmid Construction

The wild-type (HA-LUC-5F) and the PTC-containing luciferase (HA-LUC-PTC20-5F) constructs comprise, from 5′ to 3′: the T7 promoter, followed by an N-terminal 3x: Hemagglutinin tag fused in-frame with the first LUC gene (without an in-frame stop codon at LUC position 20), in-frame C-terminal StrepII and Flag tags, and the T7 3′-UTR. The ORF spans from the 3′HA tag to the StrepIIFlag tags. Plasmids harboring either HA-LUC-5F or HA-LUC-PTC20-5F were generated by PCR and standard molecular cloning techniques. In each case, oligonucleotides (listed in Table S1) containing restriction sites were used for PCR amplification, and the resulting fragments were inserted into pYplac181 after digestion of the respective restriction sites. Premature termination codons (UAU, UAG, or UGA) were inserted at codon 20 of the LUC ORF using site-directed mutagenesis and oligonucleotides described in Table S1.

RNA and Protein Analyses

Procedures for RNA analysis, protein analysis, and luciferase assays were as described previously (29) or in SI Materials and Methods. Readthrough products were purified according to instructions provided by the manufacturer of Strept-Tactin resin (IBA). Details are provided in SI Materials and Methods. Procedures used for mass spectrometric analysis are described in SI Materials and Methods.

ACKNOWLEDGMENTS

We thank Drs. Andrei A. Korostelev and Christopher R. Trotta for helpful discussions. This work was supported by NIH Grant R37 GM27757-35 (to A.J.).
Supporting Information

Roy et al. 10.1073/pnas.1424127112

SI Materials and Methods

Luciferase Assays. Luciferase assays were performed on clarified cells, as previously described (1). Briefly, cell pellets were resuspended in buffer W [100 mM Tris (pH 8.0), 150 mM NaCl, and 1 mM EDTA; supplemented with 0.1 mM DTT, 1 mM PMSF, and 1× protease inhibitor mixture], and enzyme activity was measured after the addition of 400 μL of luciferase assay buffer [25 mM Hepes (pH 7.8), 15 mM potassium phosphate (pH 7.8), 15 mM magnesium sulfate, 4 mM EGTA, 2 mM ATP, and 1 mM DTT]. Assays were performed in a GloMax 20/20 luminometer by injecting 100 μL of 200 mM D-luciferin (Analytical Luminescence Labs). Luciferase activity was normalized to total protein concentration of the lysates measured with the Bradford assay (BioRad).

RNA Analyses. Cells expressing either HA-LUC-SF or HA-LUC_{PTC20}-SF were grown in synthetic complete (SC) medium lacking Leu. Total RNA isolation and Northern blot analyses were performed as described previously (1). Random-primer DNA probes made from the 1.6-Kb Ncol-XhoI luciferase fragment were used to detect LUC mRNA. Randomly primed DNA probes made from the 0.6-Kb EcoRI-HindIII CYH2 fragment or the 0.5-Kb EcoRI SCR1 fragment were used to detect the CYH2 pre-mRNA, CYH2 mRNA, and SCR1 RNA as controls for NMD and normalization, respectively. Transcript-specific signals were determined with a FUJI BAS-2500 analyzer (Fuji Electric).

Western Analyses. Cells representing 0.4 OD_600 units were harvested and resuspended in 200 μL of sample buffer as described previously (2). Samples were resolved by 8% SDS/PAGE, transduced, and resuspended in 200 μL of a 100-mM iodoacetamide solution and allowed to react for 30 min. The gel slices were washed twice with 1-mL aliquots of distilled, deionized water. The water was removed, and the gel slices were further dried in a Speed Vac (Savant Instruments). Gel slices were rehydrated in 100 μL of 4 ng/μL endoproteaseinase Lys-C (Roche Diagnostics) in 0.01% Protease-MAX Surfactant (Promega) and 50 mM ammonium bicarbonate. Additional bicarbonate buffer was added to ensure complete submersion of the gel slices. Samples were incubated at 37 °C for 18 h. The supernatant of each sample was then removed and placed in individual 0.5-μL Eppendorf tubes. Gel slices were further extracted with 200 μL of 80:20 (acetonitrile/1% formic acid). The extracts were combined with the supernatants of each sample. The samples were then completely dried down in a Speed Vac.

Liquid Chromatography/Tandem MS Analysis. Endoproteaseinase Lys-C peptide digests were reconstituted in 20 μL of 5% acetonitrile containing 0.1% (vol/vol) trifluoroacetic acid and separated on a NanoAcquity (Waters) UPLC (ultraperformance liquid chromatography). In brief, a 4-μL injection was loaded in 5% acetonitrile containing 0.1% formic acid at 4 μL·min⁻¹ for 4 min onto a 100-μm i.d. fused-silica precolumn packed with 2 μm of 5-μm (200 Å) Magic C18AQ (Bruker–Microm) and eluted using a gradient at 300 nL·min⁻¹ onto a 75-μm i.d. analytical column packed with 25 cm of 3-μm (100 Å) Magic C18AQ particles to a gravity-pulled tip. The solvents were 0.1% formic acid (solvent A) and acetonitrile in 0.1% formic acid (solvent B). A linear gradient was developed from 5% (vol/vol) solvent A to 35% (vol/vol) solvent B in 35 min. Ions were introduced by positive electrospray ionization via liquid junction into a Q Exactive hybrid mass spectrometer (Thermo). Mass spectra were acquired over m/z 300–1,750 at a resolution of 70,000 (m/z 200), and data-dependent acquisition selected the top 12 most abundant precursor ions for tandem MS by high-energy collisional dissociation fragmentation using an isolation width of 1.2 Da, collision energy of 27, and resolution of 35,000. The run conditions follow the “sensitive” settings recommended by Kelstrup et al. (3) for optimizing the Q Exactive mass spectrometer for low-abundance proteins.

Data Analysis. Raw data files were processed with Proteome Discoverer (version 1.4; Thermo) or Mascot Distiller (version 2.4; Matrix Science, Inc.) before database searching with Mascot Server (version 2.4; Matrix Sciences, Ltd.) against the UniProt index of Saccharomyces cerevisiae containing construct sequences of all 20 potential codon 20 mutations of the LUC gene. Search parameters included endoproteaseinase Lys-C specificity with two missed cleavages and the variable modifications of oxidized Met, pyroglutamic acid for N-terminal Gln, N-terminal acetylation of the protein, and a fixed modification for carbamidomethyl Cys. The mass tolerances were 10 ppm for the precursor and 0.05 Da for the fragments. Precursor intensity data of individual peptides...
were extracted using Mascot Distiller or Proteome Discoverer before loading into the Scaffold viewer (Proteome Software, Inc.). To calculate the relative abundance of each mutant in the sample, the corresponding precursor intensity of each endo Lys-C peptide containing codon 20 was added to produce a total peptide abundance that was then used to calculate the percentage of contribution of each specific mutant. All charge states observed were taken into account in calculating relative abundances, and the assumption was made that single mutations within a larger peptide did not significantly alter the peptide’s ionization potential.

**MALDI-TOF/TOF Analysis.** Samples digested as described in liquid chromatography-tandem MS (MS/MS) analysis were further purified via a micro–zip-tip cleanup. Briefly, 10-μL volumes were acidified with 1.5 μL of 5% (vol/vol) TFA. Samples were loaded on a μC18 zip-tip (Millipore Corp.) after pre-equilibration in 0.1% TFA. After washing with 2 × 10-μL aliquots of 0.1% TFA, samples were deposited directly onto the MALDI sample target in 1 μL of acetonitrile in 0.1% TFA, followed by the addition of 0.5 μL of matrix solution, which consisted of 5 mg/mL alpha-cyano-4-hydroxycinnamic acid in acetonitrile in 0.1% TFA. Samples were allowed to air-dry before insertion into the mass spectrometer. Analyses were performed on a Shimadzu Biotech Axima TOF² (Shimadzu Instruments) MALDI-TOF mass spectrometer. Peptides were analyzed in positive ion reflectron mode. The instrument was externally calibrated using a local spot to the sample of interest with angiotensin II (1,046.54 Da), P14R (1,533.86 Da), and ACTH (18–39) (adrenocorticotropic hormone residues 18–39) (2,465.20 Da). Collisionally induced dissociation (CID) analysis was performed on the same instrument using a dual-timed ion gate for high-resolution precursor selection with a laser power about 20% higher than for MS acquisition. CID fragments were separated in a curved field reflectron, which allowed for a seamless full mass range acquisition of the MS/MS spectrum. All spectra were processed with Mascot Distiller before database searching. Database searches were performed in-house with Mascot Server. For MS searches, the peptide mass fingerprint program was used with a peptide mass tolerance of 100 ppm. For MS/MS searching (CID spectra), the MS/MS ion search program was used with a precursor tolerance of 100 ppm and a fragment tolerance of 1.5 Da.

Roy et al.

Fig. S1. (A) Representative MS/MS spectra generated via high-energy collisional dissociation (HCD) of the endo-LysC peptides showing WT luciferase containing Gly and luciferase readthrough products containing Trp, Arg, and Cys. Fragment ions are designated as either "y," where charge retention is on the C termini, or "b," where charge retention is on the N termini. The mass difference between subsequent fragment ions of the same type is equal to the mass of the amino acid in the sequence. The mutated amino acid is clearly defined by the corresponding mass difference between the y8 and y9 fragment ions in each spectrum. The corresponding precursor mass and charge state with mass error in parts per million are shown in the upper right corner of each panel. (B) MALDI-TOF MS spectra showing the relative abundances of the singly charged mutant endo-LysC peptides containing Trp, Arg, and Cys (Upper) in place of Gly (Lower) in HA-Luc-UGA_{SF} readthrough products.
**Fig. S2.** Strep-Tactin purification of HA-LUC\textsubscript{PTC20}-SF products. Western blot analysis was performed using anti-FLAG antibody of the full-length readthrough products purified from upf1\textsuperscript{Δ} [PSI\textsuperscript{−}] cells expressing HA-LUC\textsubscript{UGA20}-SF reporter. Ec pooled (fractions E3–E7) and concentrated elution fractions.

**Fig. S3.** sup45-2 [PSI\textsuperscript{−}] cells show increased readthrough after temperature shift. Luciferase activity representing the extent of readthrough from sup45-2 [PSI\textsuperscript{−}] cells expressing HA-LUC\textsubscript{PTC20}-SF reporters before (gray bar) and after (black bar) temperature shift to 37 °C for 30 min is shown. Luciferase activity is expressed as relative luciferase units (RLU) per microgram of protein/RNA units (n = 3; error bars represent SD from the mean).
Fig. S4. Readthrough efficiency after gentamicin treatment. Luciferase activity is expressed as RLU per microgram of protein after treatment with varying concentrations of gentamicin ($n = 3$; error bars represent SD from the mean).

Fig. S5. Readthrough from position 1 and position 3 mispairing. Comparison of position 1 (white bar) or position 3 (black bar) mispairing for each termination codon leading to amino acid insertion at PTCs under multiple conditions.
**Table S1. Oligonucleotides used in this study**

<table>
<thead>
<tr>
<th>Oligonucleotide name</th>
<th>Nucleotide sequence (5′–3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TPI1 promoter-PstI (For)</td>
<td>AACTGCAGGAGTTATAATAATCCTACGTTAGTGAGCGGG</td>
</tr>
<tr>
<td>TPI1 promoter-SalI (Rev)</td>
<td>AGCCGTGCACTTTTAAGTTATAGTTGTTTGGTTTTGTAGTTATAGTTATTAAGC</td>
</tr>
<tr>
<td>3HA-SalI (For)</td>
<td>CGCGTCGACAAAATGGGATATCCATTCAATATG</td>
</tr>
<tr>
<td>3HA-NcoI (Rev)</td>
<td>GGCGCTAGGCGGTAACCTGGCACGTC</td>
</tr>
<tr>
<td>Luc-Ncol (For)</td>
<td>CCAGCCATGAGAGGAGGCCAAACATATTAGAGAGAGGAGGCGGCGGCC</td>
</tr>
<tr>
<td>Luc-Xhol (Rev)</td>
<td>CTGGGCTGAGGCGGCATCTGGGCGCCCTCTCTGGSC</td>
</tr>
<tr>
<td>SF-Xhol (For)</td>
<td>GGCGCTAGGCGGCTGGTTCCCACCCCC</td>
</tr>
<tr>
<td>SF-XbaI (Rev)</td>
<td>GGGCTCTAGATTATTGGTCATACG</td>
</tr>
<tr>
<td>TPI1 3′ UTR-XbaI (For)</td>
<td>GGTCTAGATTAATATATTAATAATATATAAAAAATATATATATATATATATATAGG</td>
</tr>
<tr>
<td>TPI1 3′ UTR-EcoRI (Rev)</td>
<td>GCGAATTCGAGTACCCCATATAGATGACATCACAATTTACAAGG</td>
</tr>
<tr>
<td>UAA20 (For)</td>
<td>CAACGAGAGCTAAACCGCCGGC</td>
</tr>
<tr>
<td>UAA20 (Rev)</td>
<td>CGCCGGCGGTTAGTGTCCTGAGTC</td>
</tr>
<tr>
<td>UAG20 (For)</td>
<td>CAACGAGAGCTAAACCGCCGGC</td>
</tr>
<tr>
<td>UAG20 (Rev)</td>
<td>CGCCGGCGGTTAGTGTCCTGAGTC</td>
</tr>
<tr>
<td>UGA20 (For)</td>
<td>CAACGAGAGCTAAACCGCCGGC</td>
</tr>
<tr>
<td>UGA20 (Rev)</td>
<td>CGCCGGCGGTTAGTGTCCTGAGTC</td>
</tr>
</tbody>
</table>

For, forward; Rev, reverse.

**Table S2. Possible mispairings at PTCs**

<table>
<thead>
<tr>
<th>Position of mispairing in codon</th>
<th>Potential decoding tRNA sequence (5′–3′)</th>
<th>AA inserted</th>
<th>BP created</th>
<th>Potential decoding tRNA sequence (5′–3′)</th>
<th>AA inserted</th>
<th>BP created</th>
<th>Potential decoding tRNA sequence (5′–3′)</th>
<th>AA inserted</th>
<th>BP created</th>
</tr>
</thead>
<tbody>
<tr>
<td>UAA (position 1)</td>
<td>UUU Lys</td>
<td>U-U</td>
<td></td>
<td>UAG (position 1)</td>
<td>CUG Lys</td>
<td>U-U</td>
<td>UGA (position 1)</td>
<td>UCU Arg</td>
<td>U-U</td>
</tr>
<tr>
<td></td>
<td>UUG Gin</td>
<td>U-G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UCC Gly</td>
<td>U-C</td>
</tr>
<tr>
<td></td>
<td>UUC Glu</td>
<td>U-C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UCA* U-A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UUA*</td>
<td>U-A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UUA* U-A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UGA Ser</td>
<td>A-G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UGA* U-A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UCA*</td>
<td>A-C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UCA* U-A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UAA Leu</td>
<td>A-A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UAA* Leu</td>
<td>G-A</td>
</tr>
<tr>
<td></td>
<td>UUA*</td>
<td>A-U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UUA* G-U</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GUAA Tyr</td>
<td>A-G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>GUA Cys</td>
<td>A-G</td>
</tr>
<tr>
<td></td>
<td>CUA*</td>
<td>A-C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CUA* G-C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AUA Tyr†</td>
<td>A-A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AUA Tyr†</td>
<td>G-A</td>
</tr>
</tbody>
</table>

Underlining denotes codon positions that are mispaired.
* tRNA unknown in available database.
† tRNA detectable in some organisms but absent in yeast.

Roy et al. www.pnas.org/cgi/content/short/1424127112