PNAS RESEARCH ARTICLE | BIOCHEMISTRY

Systematic analysis of nonprogrammed frameshift suppression in *E. coli* via translational tiling proteomics

Benjamin L. Springstein[a,1], Joao A. Paulo[b], Hankum Park[b,2], Kemardo Henry[a], Eleanor Fleming[a], Zoë Feder[a], J. Wade Harper[b,3], and Ann Hochschild[a,3]

Edited by Hani S. Zaher, Washington University in St. Louis, St. Louis, Missouri; received October 13, 2023; accepted December 20, 2023 by Editorial Board Member Thomas J. Silhavy

The synthesis of proteins as encoded in the genome depends critically on translational fidelity. Nevertheless, errors inevitably occur, and those that result in reading frame shifts are particularly consequential because the resulting polypeptides are typically nonfunctional. Despite the generally maladaptive impact of such errors, the proper decoding of certain mRNAs, including many viral mRNAs, depends on a process known as programmed ribosomal frameshifting. The fact that these programmed events, commonly involving a shift to the −1 frame, occur at specific evolutionarily optimized "slippery" sites has facilitated mechanistic investigation. By contrast, less is known about the scope and nature of error (i.e., nonprogrammed) frameshifting. Here, we examine error frameshifting by monitoring spontaneous frameshift events that suppress the effects of single base pair deletions affecting two unrelated test proteins. To map the precise sites of frameshifting, we developed a targeted mass spectrometry–based method called "translational tiling proteomics" for interrogating the full set of possible −1 slippage events that could produce the observed frameshift suppression. Surprisingly, such events occur at many sites along the transcripts, involving up to one half of the available codons. Only a subset of these resembled canonical "slippery" sites, implicating alternative mechanisms potentially involving noncognate mispairing events. Additionally, the aggregate frequency of these events (ranging from 1 to 10% in our test cases) was higher than we might have anticipated. Our findings point to an unexpected degree of mechanistic diversity among ribosomal frameshifting events and suggest that frameshifted products may contribute more significantly to the proteome than generally assumed.

ribosome | frameshifting | translation | proteomics | *E. coli*

Significance

Translational fidelity underlies cellular fitness by ensuring that proteins are synthesized as encoded. Although maintenance of the correct reading frame is critical to protein synthesis, programmed ribosomal frameshifting is a well-studied process enabling the production of distinct polypeptides from a single mRNA. The occurrence of nonprogrammed ribosomal frameshifting producing aberrant polypeptides is less well understood. Using genetics and an unbiased proteomic approach, we investigate the spectrum of slippage products for two unrelated mRNAs in bacteria. We observe significant frameshifting implicating numerous sites along the mRNAs, many of which do not correspond to typical "slippery" sites, indicative of an unexpected degree of mechanistic diversity. Our findings suggest that the prevalence of frameshifted products in the proteome is greater than generally assumed.

Translational fidelity is fundamental to cellular fitness, ensuring the production of a properly functioning proteome. Nonetheless, errors inevitably occur and the frequency at which different types of errors occur presumably reflects an evolutionary trade-off between the costs of achieving greater accuracy and the costs of mitigation (e.g., by protein quality control systems) (1). Among the types of errors that can occur, ribosomal frameshifting is particularly consequential because it typically results in the production of a nonfunctional polypeptide. Although the occurrence of such errors was inferred early on by establishing the leakiness of a collection of *lacZ* frameshift mutants (2), estimates in the literature suggest that the frequency of spontaneous frameshifting is low, at $<10^{-5}$ per codon translated (3, 4). However, these estimates are based on limited data, and systematic attempts to measure the frequency of these events are lacking.

Moreover, despite the detrimental effects of ribosomal frameshifting, the proper decoding of certain mRNAs depends on a process known as programmed ribosomal frameshifting. For these processes to have evolved, a less than perfectly accurate translation apparatus would seem to have been a prerequisite (1). This decoding strategy, whereby a targeted shift in reading frame augments the coding capacity of a single mRNA transcript, is particularly prevalent in the context of viruses and most often involves a shift to the −1 frame (4–9). However, both prokaryotic and eukaryotic organisms also employ programmed ribosomal frameshifting in decoding endogenous genes (9–11). Although many documented examples involve mobile genetic elements, including bacterial insertion sequence (IS) elements (12), some involve chromosomal genes that are not associated with mobile elements. In *Escherichia coli*, the *dnaX* gene (13–15), encoding the γ and τ subunits of DNA polymerase III, the *prfB* gene (16), encoding release factor 2 (RF2), and the *copA* gene (17), encoding a copper ion transporter, provide well-characterized examples involving a shift to the −1 frame in two cases (*dnaX* and *copA*) and a shift to the +1 frame in the third (*prfB*).

This article is a PNAS Direct Submission. H.S.Z. is a guest editor invited by the Editorial Board.

[1]Present address: Institute of Science and Technology, Klosterneuburg 3400, Austria.

[2]Present address: Department of Dental Science, School of Dentistry and Dental Research Institute, Seoul National University, Seoul 03080, Republic of Korea.

[3]To whom correspondence may be addressed. Email: wade_harper@hms.harvard.edu or ann_hochschild@hms.harvard.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2317453121/-/DCSupplemental.

Published January 30, 2024.

The analysis of programmed ribosomal frameshifting—specifically, –1 frameshifting—supports the existence of two general mechanisms whereby these events can occur (4). Both depend on the presence of a so-called "slippery" nucleotide sequence that is minimally required to enable efficient frameshifting (18). In one scenario, requiring a slippery heptamer with the pattern X_XXY_YYZ, where the same nucleotide (X or Y) serves as the third position of one codon and the first two positions of the next, a pair of tRNAs in the P- and A-sites of the ribosome slips in tandem during translocation such that the codon entering the now vacant A-site is in the –1 frame (Fig. 1*A*) (19). In the second scenario, requiring minimally a slippery tetramer with the pattern X_XXN followed by a codon for which the cognate aminoacyl-tRNA is in scarce supply (dubbed a "hungry" codon), a single tRNA in the P-site slips while the A-site remains vacant such that an alternative (–1) codon becomes accessible in the A-site (19, 20). Relevant to this mechanism is the demonstration that when translocation is hindered, frameshifting efficiency on a slippery sequence depends on the free-energy difference between the codon–anticodon base pairing interactions in the two frames (21). A variant single-tRNA slippage scenario that also requires a slippery tetramer (specifically, A_AAG) but lacks an adjacent hungry codon has been described as well; in such cases, it is the tRNA in the A-site that slips and the slippage is facilitated when the preceding codon is read by a tRNA that is wobble-prone (i.e., a tRNA with a greater tendency to unpair from the base at the third position of the codon) (22).

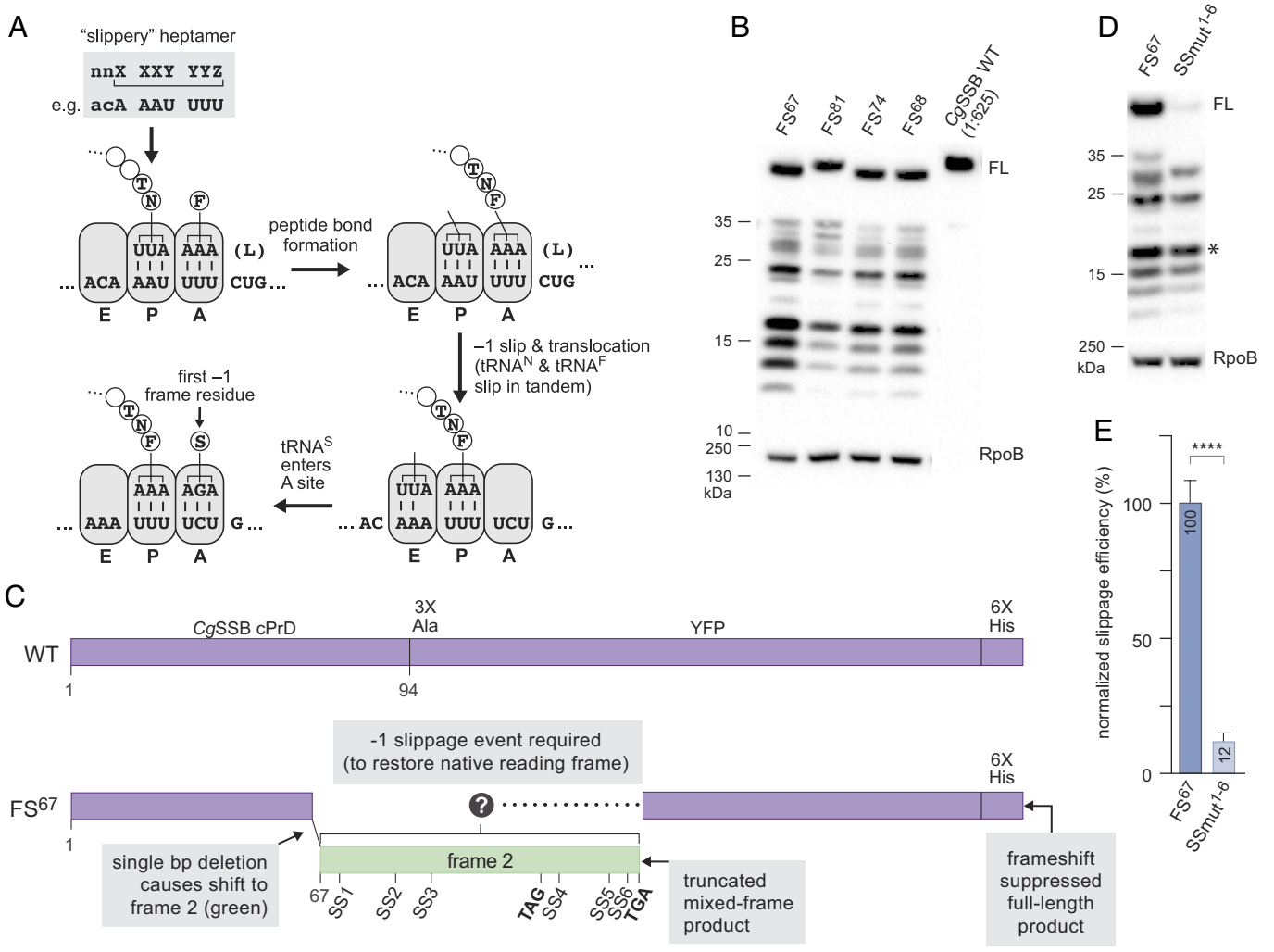

Fig. 1. Frameshift suppression of single base pair deletion mutants and effect of disrupting six candidate slippery heptamers. (*A*) Model depicting hypothetical sequence of events enabling a –1 slippage event given a "slippery" heptamer motif. Ribosomal A (aminoacyl), P (peptidyl), and E (exit) sites are shown with tRNAs in A and P sites, initially paired with cognate codons. Following peptide bond formation, a tandem –1 slip occurs during translocation, such that the A site codon now specifies the first –1 frame residue. (*B*) Western blot showing that cells producing the indicated frameshift mutants contain full-length protein (FL) as detected with an anti-His antibody. Lower molecular weight bands correspond to internal start products initiating at wild-type frame ATG or GTG codons (*SI Appendix*, Fig. S2*B*). The blot was also probed with anti-RpoB as a loading control (*Bottom* strip). Extract prepared from cells producing the wild-type fusion protein (*Cg*SSB WT) was diluted 1:625 and probed on the same blot (cut vertically to remove intervening lanes) as the undiluted frameshift mutant extracts. We note that this comparison provides an underestimate of the efficiency of frameshift suppression due to a significant difference in protein stabilities (*SI Appendix*, Fig. S4). (*C*) Schematic of pathway for frameshift suppression. The continuous purple bar (*Top*) depicts native *Cg*SSB cPrD-YFP fusion protein with C-terminal hexahistidine tag (6X His). Interrupted purple bar (*Bottom*) depicts FS^{67} fusion protein. Single bp deletion at codon 67 results in shift to +1 frame (frame 2, green). Absent a compensatory slippage event, translation will terminate prematurely at the indicated TGA stop codon; because the strain harbors an amber suppressor, translation can proceed past the upstream TAG stop codon. Restoration of the native frame can occur by virtue of a –1 slippage event at one of the six candidate slippery heptamers (SS1–SS6) or elsewhere within the green segment, resulting in a full-length mixed-frame protein detectable with an anti-His antibody. (*D*) Representative western blot showing comparison between FS^{67} and a mutant derivative (SSmut$^{1\text{-}6}$) bearing disruptive mutations in each of the six candidate slippery heptamers (*SI Appendix*, Fig. S2*E*). Full-length and native-frame internal start products were detected with anti-His. The blot was also probed with anti-RpoB as a loading control (*Bottom* strip). (*E*) Quantification based on nine biological replicates of the samples shown in *B*. Slippage efficiency of FS^{67} was set to 100% and internal start product GTG-2 (asterisk; see *SI Appendix*, Fig. S2*B* legend) was used for normalization to control for any effect of the SSmut$^{1\text{-}6}$ mutations on transcript stability. Error bars represent SEM; statistical significance was determined by the *t*-test: $P < 0.0001$ (****).

In addition to the slippage-mediated mechanisms for frameshifting, a possible third mechanism has been proposed, which involves neither a canonical slippery site nor a hungry codon. In this scenario, originally described in the context of the translation of the bacteriophage MS2 synthetase and coat protein genes (23, 24), the shift occurs when a noncognate tRNA with an out-of-register match to the target codon enters the A-site and engages only the first two nucleotides of the resident codon, thus effecting the shift to the –1 frame following translocation (25) (*SI Appendix*, Fig. S1). The efficiency of frameshifting in these cases depends on the relative concentrations of the relevant charged tRNAs, such that an excess of the noncognate aminoacyl-tRNA drives frameshifting. A recent molecular dynamics study suggests that in one of the cases studied, this noncognate tRNA-mediated frameshifting may occur via a network of canonical and noncanonical interactions involving all three anticodon bases, with a canonical base pair linking the base at the 5′ end of the anticodon and the middle base of the codon (26). Although more speculative and perhaps a special case, this third scenario is mechanistically distinct because, unlike the first two, it does not involve transient anticodon-codon dissociation.

Much has been learned about the mechanistic underpinnings of programmed ribosomal frameshifting, which typically include a stimulatory role for certain structural features of the mRNA (9, 11, 27, 28). However, considerably less is known about the scope of error (nonprogrammed) frameshifting events, which have been detected in diverse settings (see, for example, ref. 29). An analysis of the distribution of two particularly frameshift-prone sequences (A_AAA_AAG, which facilitates –1 frameshifting, and CCC_TGA, which facilitates +1 frameshifting) within the coding regions of *E. coli* suggests that neither is under strong negative selection despite supporting frameshifting at frequencies of >1%, at least for genes that are not among the most highly expressed in the genome (30). However, the overall distribution of sites at which error frameshifting events occur has not been investigated, and thus, it is not known to what extent such errors are confined to particular sequence motifs.

Here, we begin to explore this question by examining the full set of spontaneous –1 frameshift events that suppress the effect of a single base pair deletion in an overexpressed fusion gene in *E. coli*. We thus monitored the production of the full-length fusion protein, which was dependent on the occurrence of a frameshift event to enable circumvention of multiple premature stop codons in the nonnative (+1) frame that the ribosome enters by virtue of the single base pair deletion. To map the precise sites of frameshifting, we developed a targeted mass spectrometry–based method referred to as "translational tiling proteomics" for interrogating all possible –1 slippage events that could lead to the observed frameshift suppression. The method depends on the identification of diagnostic mixed-frame peptides that are not detected by conventional mass spectrometry–based protein database searches, which use canonical translational reading frames. Our findings indicate that such events occur at numerous sites along the transcript, implicating one half of the available codons, of which only about one third could be identified as plausible "slippery" sites. Moreover, the aggregate frequency of these events was higher than we might have anticipated, at about 1% (corresponding to an error frequency of ~10^{-4} per codon). Extension of the approach to an unrelated test protein emphasizes the generality of our findings with respect to both the diversity of slippage sites and also the relatively high aggregate frequency of –1 slippage events (~10% in this case, corresponding to an error frequency of >10^{-3} per codon).

Our findings i) have implications for the interpretation of overexpression studies, ii) suggest that frameshifted products may contribute more significantly to the proteome than generally appreciated, iii) suggest that certain pseudogenes may specify functional proteins, and iv) point to greater mechanistic diversity among instances of ribosomal frameshifting than has previously been documented.

Results

Genetic Screen Leads to Identification of Frameshift Mutations. The genesis of our study was an incidental finding made in the context of a genetic screen. Specifically, we set out to evaluate the prion-forming potential of a candidate bacterial prion-forming domain (cPrD) – that of the *Corynebacterium glutamicum* (*Cg*) single-stranded DNA-binding protein (SSB). To perform this analysis, we were working with a plasmid-encoded fusion protein consisting of the *Cg* SSB cPrD fused to YFP bearing a C-terminal hexahistidine (His) tag produced under the control of a strong inducible promoter. Although this fusion protein was able to access an aggregated prion-like conformation under certain conditions in the *E. coli* cytoplasm, these aggregates were not heritably transmitted over multiple generations, as inferred based on the failure to detect blue colonies using a previously described genetic assay (31).

In an effort to identify amino acid substitutions that might confer heritability on the aggregates, we introduced random mutations into the cPrD moiety of the fusion gene and screened for blue colonies (indicative of heritably transmitted prion-like aggregates) after transforming our *E. coli* reporter strain with a plasmid library encoding the mutant fusion proteins. Sequencing of plasmid DNA from four clones exhibiting the blue colony-color phenotype in each case revealed a single base-pair deletion within the portion of the fusion gene specifying the C-terminal region of the PrD moiety, such that codons 67, 68, 74, and 81 are the first out-of-frame codons (*SI Appendix*, Fig. S2*A*). We refer to these as frameshift (FS) mutants FS^{67}, FS^{68}, FS^{74}, and FS^{81}, respectively. The basis for the blue colony-color phenotype will be addressed in a separate study. In brief, we found that the frameshift mutants specify a prematurely terminated product that forms heritable aggregates and hence gives rise to the blue colony-color phenotype.

Evidence for Frameshift Suppression: –1 Slippage Events. While characterizing the proteins encoded by these frameshift mutants, we were surprised to observe a significant amount of full-length protein, as assessed by western blot analysis with an anti-His antibody, in each case (Fig. 1*B*). We note that most of the faster migrating bands detectable on the blot correspond to internal start products initiating at ATG or GTG codons present in the wild-type frame within the YFP moiety (*SI Appendix*, Fig. S2 *B–D*). From this apparent nonprogrammed frameshift suppression resulting in the generation of full-length polypeptide products, we inferred that some compensatory slippage event(s) must be occurring during the translation of the mRNAs encoding the frameshift mutants, resulting in restoration of the wild-type frame (defined as frame 1) before the first stop codon is encountered in the shifted frame (hereafter frame 2). Specifically, a "–1" slippage event would be required to compensate for a single base pair deletion.

Given both the efficiency of production of a full-length product resulting from an apparent –1 slippage event and an incomplete understanding of nonprogrammed frameshifting, we sought to determine the basis for the frameshift suppression through detailed analysis of one of the frameshift mutants (FS^{67}), which would provide the largest sequence window for a –1 slippage event to

occur prior to the first stop codon in-frame 2 (Fig. 1*C*). Because previous analyses of programmed ribosomal frameshifting (–1 frameshifting, in particular) indicate that a slippery heptanucleotide is often present and required to enable efficient frameshifting (Fig. 1*A*), we examined the relevant portion of FS^{67} sequence for potential slippery heptanucleotide motifs resembling the X_XXY_YYZ consensus (where X is any nucleotide, Y is typically A or U, and Z is typically A, U, or C) (9, 11, 18). As there are two frame-2 stop codons immediately upstream of the site of the FS^{67} single base pair deletion (precluding productive slippage events upstream of codon 67), we looked for potential slippage sites between codon 67 and the first premature stop codon in-frame 2. Moreover, because we were working with a strain that contains an amber suppressor (resulting in the insertion of a glutamine at the site of TAG stop codons) and the first frame 2 stop codon downstream of codon 67 is a TAG at codon 115 (TAG^{115}), we extended the region of interest to the next stop codon, a TGA at codon 158 (TGA^{158}) (Fig. 1*C*).

We identified six potential slippery heptanucleotide sites (SS1-SS6) with a reasonable match to the X_XXY_YYZ consensus (though we allowed for Y/Z = G) within the relevant 91-codon window (Fig. 1*C* and *SI Appendix*, Fig. S2*E*). To test whether or not the production of the full-length fusion protein was dependent on the functional integrity of one or more of these sites, we first constructed a FS^{67} variant (called $SSmut^{1\text{-}6}$) bearing mutations predicted to disrupt all of these sites. In general, we did this by mutating positions #1 and #4 of each heptamer (*X*_XX*Y*_YYZ, mutated nucleotides in italics), thus minimally disrupting the FS^{67} coding sequence (*SI Appendix*, Fig. S2*E*). Consistent with the occurrence of frameshift-suppressing slippage events at one or more of these sites, we detected far less full-length protein in cells containing $SSmut^{1\text{-}6}$ than in cells containing the FS^{67} parent construct (Fig. 1*D*). By using one of the faster migrating bands corresponding to an internal start product (which we termed "GTG-2"; *SI Appendix*, Fig. S2 *A* and *B*) for normalization, we were able to control for any possible effects of the mutations present in $SSmut^{1\text{-}6}$ on overall transcript stability (see *SI Appendix*, Fig. S2*B* legend). Quantification with normalization indicated that the mutations resulted in an ~10-fold decrease in the amount of full-length protein (Fig. 1*E*), suggesting that slippage at one or more of these sites accounts for ~90% of the full-length frameshift-suppressed product detected. To investigate the possibility that the unusual amino acid composition of the *Cg* SSB cPrD moiety might in some way facilitate the observed frameshift suppression, we replaced the cPrD sequence upstream of codon 67 (the site of the frameshift mutation) with an unrelated sequence (encoding portions of the SUMO protein). The extent of frameshift suppression was unaffected by this replacement (*SI Appendix*, Fig. S2 *F* and *G*). We also tested the effect of growth phase on frameshift suppression efficiency. We found that the extent of frameshift suppression was similar whether the cells were harvested during log phase or after overnight growth (our standard procedure) (*SI Appendix*, Fig. S2 *H* and *I*).

Localization of Slippage Sites. To begin to localize the site(s) of significant slippage, we introduced a series of wild-type frame stop codons across the 91-codon region of interest in the context of FS^{67} (Fig. 2*A*). With this "STOP-SCAN" series, we could determine for any given stop-codon mutant whether the majority of the slippage events occurred upstream or downstream of that site. That is, slippage events that occur upstream of a given wild-type frame stop codon will no longer result in the production of full-length protein, whereas slippage events that occur downstream of a wild-type frame stop codon will still result in the production of full-length protein. As illustrated in Fig. 2*A* for the STOP-SS3 construct (wild-type frame stop codon introduced just upstream of SS3), a truncated mixed-frame product results when a slippage event occurs upstream of the introduced stop codon (e.g., at SS2, as shown in the middle diagram). However, a full-length mixed-frame product results when a slippage event occurs downstream of the introduced stop codon (e.g., at SS3, as shown in the bottom diagram). The results of this analysis indicated that the majority of the slippage events occur at or in the vicinity of SS3 (Fig. 2 *B* and *C*; compare STOP-SS4 with STOP-SS3).

To further evaluate the individual contributions of the six candidate slippery sites to the observed frameshift suppression, we made two additional series of mutants. On the one hand, we tested the effect of mutating each site individually in the context of FS^{67} and, on the other hand, we tested the effect of restoring each site individually in the context of $SSmut^{1\text{-}6}$. The results of this reciprocal analysis supported our inference that SS3 is largely (though not exclusively) responsible for the observed frameshift suppression in FS^{67} (*SI Appendix*, Fig. S3 and Fig. 2 *D* and *E*).

In order to estimate the efficiency of the frameshift suppression in aggregate, we initially compared the amount of full-length protein in cells containing the wild-type fusion gene with the amount of full-length protein in cells containing the FS^{67} fusion gene. However, this comparison does not take into account any differences in the stabilities of the respective proteins (or the corresponding mRNAs), one of which contains only wild-type (frame 1) amino acid sequence whereas the other contains a stretch of nonnative (frame 2) residues. We controlled for this by constructing a mutant series with compensatory single base pair insertions mimicking the effects of slippage at each of the six candidate slippery sites (termed the "SS1-SS6 frameshift repair" series) (*SI Appendix*, Fig. S4*A*). In particular, we found that the mutant with the same amino acid sequence as that resulting from a SS3 slippage event (accounting for the bulk of the slippage) was present in an amount that was approximately eightfold lower than the fully wild-type frame fusion protein (*SI Appendix*, Fig. S4*B*). We therefore estimated the efficiency of frameshift suppression by comparing the amount of full-length, frameshift-suppressed protein in cells containing FS^{67} with the amount of full-length protein in cells containing the SS3 frameshift repair mutant. Based on this comparison, we estimate that the aggregate frameshift efficiency in the context of FS^{67} is ~1% (*SI Appendix*, Fig. S4*C*).

Frameshift Suppression In Vitro. To investigate whether or not the observed frameshift suppression reflects an inherent property of the bacterial ribosome, we took advantage of a commercially available reconstituted protein synthesis system (PURExpress; New England Biolabs, USA) in which all the components needed for translation have been purified from *E. coli*. Transcription in this cell-free system is carried out by the T7 RNA polymerase. Accordingly, we generated constructs directing the synthesis of the wild-type PrD-YFP fusion protein, FS^{67}, and three mutant derivatives ($SSmut^{1\text{-}6}$, $SSmut^{3}$, and $SSmut^{1\text{-}6}$ with SS3 restored) under the control of the T7 promoter. Because multiple components of the PURExpress system bear a His-tag, we replaced the C-terminal His-tag on our constructs with a C-terminal FLAG tag (also in the wild-type frame).

Western blot analysis of the products of the in vitro transcription/translation reactions revealed that frameshift suppression was at least as efficient in vitro as in cells (i.e., a bulk efficiency of approximately 4%; Fig. 3*A*). We note that for the in vitro reactions, we do not need to correct for potential differences in protein stability, as the reactions contain no detectable nuclease or protease activities. Similar to what we observed in vivo, mutation of the

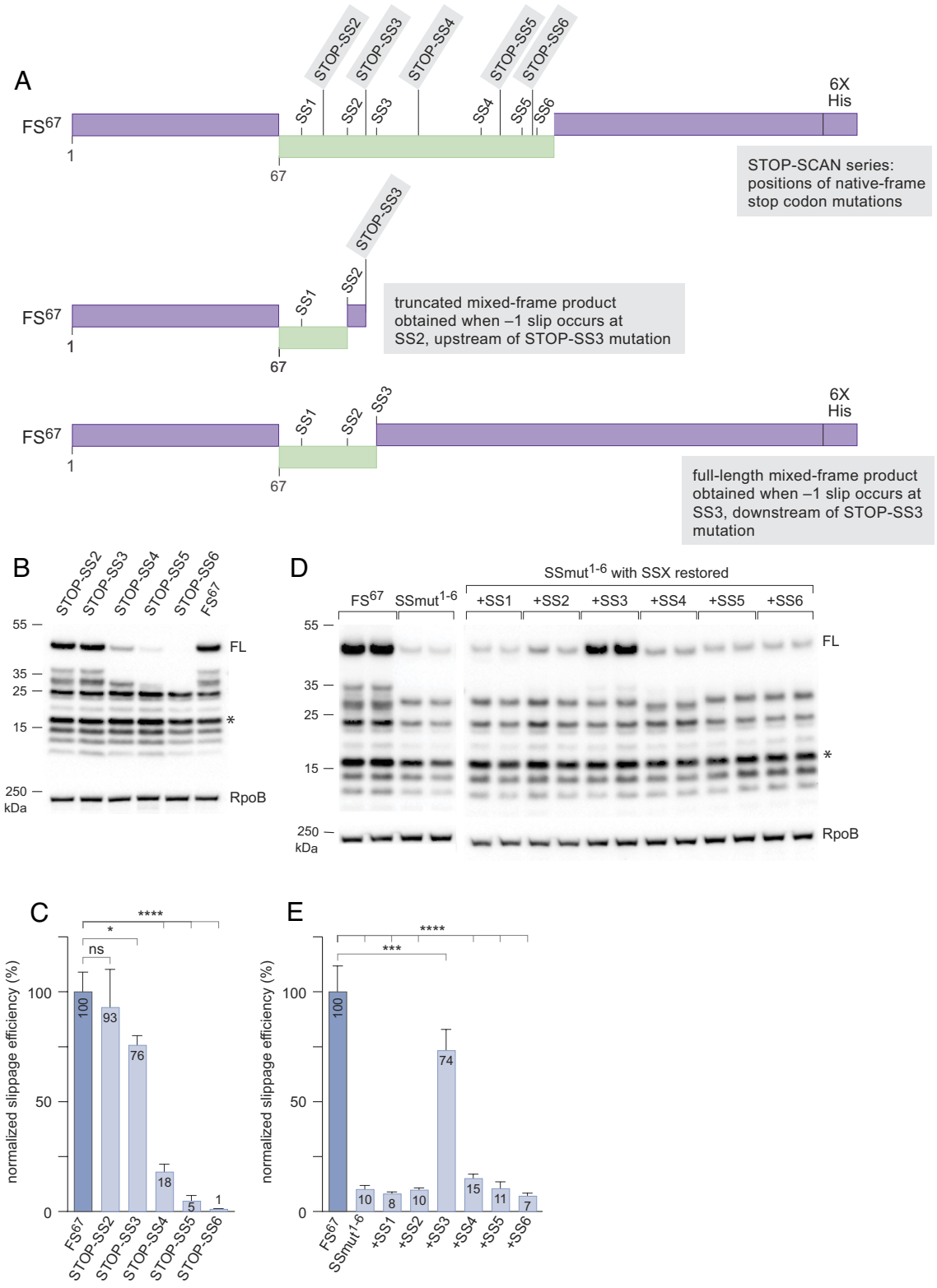

Fig. 2. Identification of SS3 as site of significant slippage. (*A*) Schematic illustrating logic of STOP-SCAN series. Top indicates the position of each of five native-frame stop codons that were individually introduced into FS^{67}. *Middle* and *Bottom* depict examples using the STOP-SS3 mutant. A –1 slippage event occurring at SS2 (upstream of the STOP-SS3 mutation) results in the truncated mixed-frame product shown in the middle, whereas a –1 slippage event occurring at SS3 (downstream of the STOP-SS3 mutation) results in the full-length mixed-frame product shown at the bottom. (*B*) Representative western blot showing STOP-SCAN series together with FS^{67}. Full-length protein (FL) and native-frame internal start products were detected with an anti-His antibody. The blot was also probed with anti-RpoB as a loading control (*Bottom* strip). (*C*) Quantification based on three biological replicates of the samples shown in *B*. Slippage efficiency of FS^{67} was set to 100% and internal start product GTG-2 (asterisk; see *SI Appendix*, Fig. S2*B*) was used for normalization. Error bars represent SEM. Statistical significance was determined by one-way ANOVA followed by Dunnett's multiple comparisons test; *P*-values: <0.0001 (****); 0.0128 (*). (*D*) Western blot showing the effect of restoring each candidate slippery heptamer individually in the context of $SSmut^{1-6}$ (duplicate samples represent biological replicates). FL and native-frame internal start products were detected with an anti-His antibody. The reference samples (FS^{67} and $SSmut^{1-6}$) were prepared at the same time as the set of single-site restoration mutants and run in parallel on a separate gel. The blots were also probed with anti-RpoB as a loading control (*Bottom* strip). (*E*) Quantification based on three biological replicates of the sample set shown in *D*, which was not used for the quantification. Quantification performed as in *C*. Error bars represent SEM. Statistical significance was determined as in C; *P*-values: <0.0001 (****); 0.0001 (***).

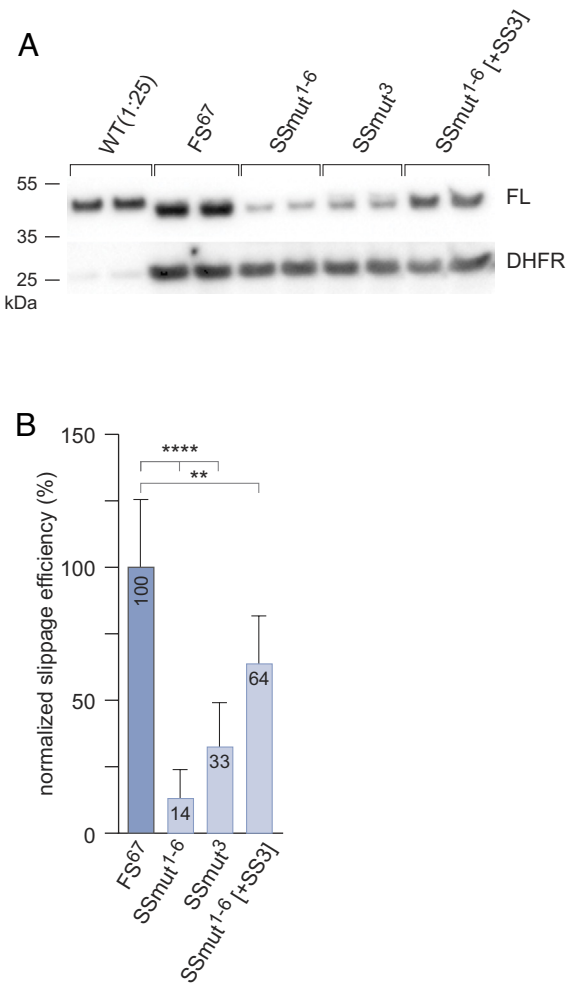

Fig. 3. Detection of frameshift suppression in vitro. (*A*) Representative western blot showing translation products obtained using an *E. coli*-based reconstituted protein synthesis system (PURExpress; New England Biolabs) (*SI Appendix*, Fig. S2*J*). Constructs were generated directing the synthesis of the indicated proteins (provided with a C-terminal FLAG tag) under the control of a T7 promoter because transcription in the PURExpress system is carried out by the T7 RNA polymerase. The duplicate samples represent technical replicates. The reaction directing the synthesis of the wild-type fusion protein (WT) was diluted 1:25. Plasmid-encoded *E. coli* dihydrofolate reductase (DHFR), one of the components provided with the PURE system, serves as an internal control for the functionality of the PURE system. DHFR (His-tagged) was probed as a separate strip with anti-His. (*B*) Quantification based on three experimental replicates, including the experiment shown in *A*. As in *A*, technical replicates of each sample were loaded on the gel and the values for the technical replicates were averaged. Slippage efficiency of FS^{67} was set to 100% and error bars represent SEM. Statistical significance was determined by one-way ANOVA followed by Dunnett's multiple comparisons test; *P*-values: <0.0001 (****); 0.0082 (**).

six candidate slippery sites resulted in a major reduction in the amount of full-length (FLAG-reactive) frameshift-suppressed fusion protein, and a substantial fraction of this decrease was attributable to mutation of SS3 (Fig. 3 *A* and *B*). Nonetheless, as was observed in vivo, mutation of the six candidate slippery sites did not fully eliminate frameshift suppression.

Effects of Transcription–Translation Coupling and Ribosomal Protein L9 on Frameshift Suppression. Our in vitro findings indicate that frameshift suppression does not depend on transcription by the *E. coli* multisubunit RNA polymerase (RNAP), as the results are qualitatively similar to those obtained in vivo despite the use of the single subunit T7 RNAP for the in vitro reactions. Consistently, we found that the extent of frameshift suppression in vivo was similar regardless of whether FS^{67} was produced under the control of our standard promoter (P*tac*, which is recognized by *E. coli* RNAP) or a T7 promoter (recognized by T7 RNAP), an experiment we performed using an *E. coli* strain that encodes T7 RNAP polymerase under the control of an inducible promoter (NiCo21 [DE3] from NEB, USA) (*SI Appendix*, Fig. S5 *A* and *B*). In fact, the quantification (with normalization to the GTG-2 internal start product; *SI Appendix*, Fig. S2*B*) suggested that the extent of frameshift suppression was modestly higher when FS^{67} was produced under the control of the T7 promoter (*SI Appendix*, Fig. S5*B*). We considered the possibility that the relevant difference is that transcription and translation are coupled in *E. coli* (via connection between the multisubunit RNAP and the lead ribosome mediated by the elongation factor NusG) (32–34), whereas no such coupling occurs with the T7 RNAP. Experiments with a strain encoding a coupling deficient NusG mutant (32, 34) revealed a modest increase in frameshift suppression when coupling was disrupted (~1.7-fold; *SI Appendix*, Fig. S5 *C* and *D*), suggestive of a small hindrance of coupling on –1 frameshifting.

We also performed an experiment to test the effect on frameshift suppression of deleting the gene encoding ribosomal protein L9 (*rplI*), mutations in which have been shown previously to affect both frameshifting and another slippage-related phenomenon known as ribosomal hopping that is required during translation of bacteriophage T4 gene 60 in *E. coli* (35–37). Consistent with the previous studies, which have documented an increase in –1 frameshifting in the absence of L9, we observed a modest increase in frameshift suppression (~1.5-fold) with our construct when the strain carried an *rplI* deletion (*SI Appendix*, Fig. S5 *E* and *F*).

Translational Tiling Proteomics Identifies Diverse Mixed Frame Translation Products. We next turned to mass spectrometry to determine the precise identities of the mixed-frame polypeptides that were present in cells containing the FS^{67} mutant and in our in vitro transcription–translation reactions. We anticipated detecting products corresponding to a SS3 slippage event; however, because even $SSmut^{1-6}$ supported the production of some residual full-length protein, we attempted to take an unbiased approach for identifying any mixed-frame products that might be present. To do this, we developed a method we refer to as translational tiling proteomics wherein all possible mixed-frame products arising from –1 slippage events occurring at sequential codons in the region between the site of the frameshift mutation at codon 67 and the first TGA stop at codon 158 were incorporated into the search process for peptides identified by mass spectrometry (Fig. 4 and Dataset S1). We note that such diagnostic mixed-frame products would not be detected using conventional mass spectrometry–based searches that rely on canonical reading frames. Having generated a mixed-frame database, we used two complementary experimental methods to purify and subsequently identify mixed-frame peptides that corresponded to our list of hypothetical mixed-frame slippage products.

First, we prepared extracts from cells producing FS^{67} and isolated His-tagged full-length frameshift-suppressed product by means of Ni-NTA affinity chromatography, followed by separation from shorter fragments (*SI Appendix*, Fig. S2*B*) and other contaminating species via excision from SDS-PAGE. Excised full-length species were subjected to in-gel digestion with either chymotrypsin or Asp-N to liberate peptides for analysis by translational tiling proteomics. To increase the number of potentially diagnostic peptides generated by Asp-N, we replaced the Gln residue at position 66 (immediately preceding the site of the frameshift mutation) with Asp and we also replaced either Val 117 or Gly 121 with Asp. Thus, we carried out a parallel analysis with

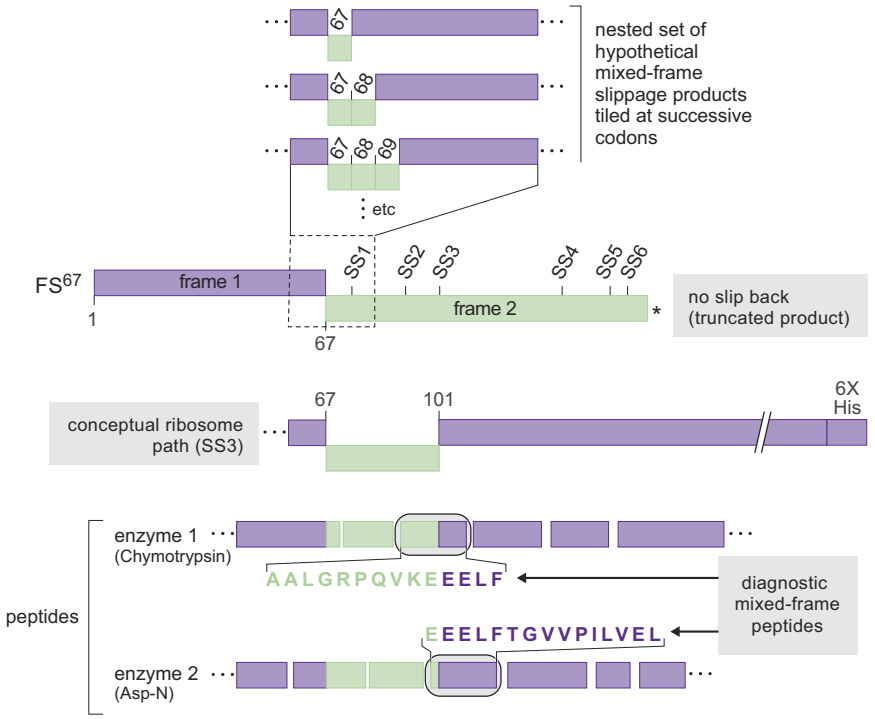

Fig. 4. Translational tiling proteomics scheme for identifying mixed-frame translation products. (*Upper* portion) The first three of a nested set of hypothetical mixed-frame products arising from −1 slippage events occurring at sequential codons in the target region between codon 67 and the frame-2 TGA stop at codon 158; this set comprised the database that was used to identify peptides detected by mass spectrometry. Depicted just below these is the truncated mixed-frame product that results in the absence of any frameshift-suppressing slippage event. (*Lower* portion) Depiction of the expected full-length mixed-frame product arising if a −1 slippage event occurs at SS3. Shown below are sample cleavage patterns generated by two different proteases. Circled are two mixed-frame peptides (with amino acid sequence) identified by mass spectrometry that are diagnostic of the SS3 slippage event.

the two resulting variants bearing substitutions Q66D/V117D or Q66D/G121D (Dataset S2), both of which displayed comparable amounts of full-length product as the parental FS^{67} mutant (*SI Appendix*, Fig. S6*A*).

With this method, we detected mixed-frame peptides diagnostic of slippage at five of the six candidate slippery sites we had originally identified based on their resemblance to the X_XXY_YYZ motif (i.e. SS1, SS2, SS3, SS4, and SS6) (*SI Appendix*, Table S1 [columns 1 & 2]; Dataset S3). The majority of mixed-frame peptides we detected were derived from the chymotrypsin digestions. However, in the case of SS4, a single diagnostic peptide was detected following Asp-N digestion of $FS^{67/Q66D/V117D}$ (Dataset S3). Consistent with our genetic analyses, peptides diagnostic of slippage at SS3 and an adjacent position were represented by higher numbers of peptide spectral matches (PSMs) (7–12) than the peptides diagnostic of slippage at SS1, SS2, SS4, or SS6 (3 PSMs or less) (Dataset S3). Strikingly, we also identified peptides diagnostic of slippage at multiple additional sites (for a total of 27 sites out of the 91 possible sites that were sampled based on the translational tiling search space; *SI Appendix*, Table S1 and Dataset S3). Peptides diagnostic of several of these additional slippage sites were represented by 4 or 5 PSMs. Moreover, for 8 of the 27 detected slippage sites (including slippage at SS3 and the adjacent position), two or more unique mixed-frame diagnostic peptides were identified (Dataset S3). Although only semiquantitative, PSM counting suggests that numerous slippage sites can be identified using this method, but with SS3 yielding somewhat higher recovery of diagnostic peptides.

The power of the gel excision approach is the ability to uniquely isolate full-length products, which have necessarily undergone frameshift suppression. However, in order to ensure that in-gel digestion did not bias detection of potential frame-shifted products, we sought to circumvent the gel electrophoresis step. Additionally, having identified multiple mixed-frame products that were not associated with a motif resembling the X_XXY_YYZ consensus, we sought to increase the possibility of detecting low-abundance mixed-frame peptides that might be inefficiently liberated from the gel matrix. Thus, we created a set of truncated constructs with the wild-type frame His-tag positioned immediately after the first TGA stop in the FS^{67} frame (i.e., for both the Q66D/V117D and the Q66D/G121D variants). The removal of the remaining YFP sequences eliminated the production of contaminating wild-type frame internal start products, allowing us to purify the His-tagged frameshift-suppressed products by Ni-NTA affinity chromatography for direct in-solution digestion with chymotrypsin or Asp-N. For this set of experiments, we extended our analysis to $SSmut^{1\text{-}6}$ and to $SSmut^{1\text{-}6}$ with SS3 restored ($SSmut^{1\text{-}6}$ [+SS3]) in the context of the Q66D/G121D variant. This set of 4 samples was digested and subjected to mass spectrometry in two independent experiments (Dataset S2). Through this approach, we again detected mixed-frame peptides diagnostic of slippage at SS1, SS2, SS3, SS4, and SS6 (*SI Appendix*, Table S1 [columns 1, 3, 4, & 5]; Dataset S3). The majority of peptides were detected as chymotryptic products, with the exception of a single Asp-N-derived peptide diagnostic of slippage at SS4 with $FS^{67/Q66D/V117D}$ (as before), and a single Asp-N-derived peptide diagnostic of slippage at SS6 with both $FS^{67/Q66D/V117D}$ and $FS^{67/Q66D/G121D}$ in the two replicate experiments (Dataset S3). Also, as before, the peptides diagnostic of slippage at SS3 and the adjacent position were represented by higher numbers of PSMs than the peptides diagnostic of slippage at SS1, SS2, SS4, or SS6, though in one instance, a similar number of PSMs was found for a peptide diagnostic of slippage at SS1 (Dataset S3).

Like the in-gel digestion protocol, the in-solution digestion protocol resulted in the identification of slippage at multiple additional sites—an even greater number than with the in-gel digestion

protocol. Considering only $FS^{67/Q66D/V117D}$ and $FS^{67/Q66D/G121D}$, we detected mixed-frame peptides diagnostic of slippage at a total of 39 sites out of the 91 possible sites that were sampled (*SI Appendix*, Table S1 [columns 1 & 3]). Moreover, this number increased to 46 when we included mixed-frame peptides detected with the $SSmut^{1-6}$ sample and the $SSmut^{1-6}$ [+SS3] sample (*SI Appendix*, Table S1 [columns A, C, D, & E]).

As expected, when we subjected the $SSmut^{1-6}$ sample to in-solution digestion and mass spectrometry, we no longer detected any mixed-frame peptides diagnostic of slippage at SS3 or the adjacent position, whereas these peptides were again detected with the $SSmut^{1-6}$ [+SS3] sample (*SI Appendix*, Table S1 [columns 1, 4 & 5]; Dataset S3). Moreover, neither of these samples yielded mixed-frame peptides diagnostic of slippage at the other mutated sites with one exception. In the case of SS2, a unique diagnostic peptide (represented by one PSM) was detected with both the $SSmut^{1-6}$ and the $SSmut^{1-6}$ [+SS3] samples (Dataset S3). We note that the SS2 heptamer is followed by a rare arginine codon, a potential "hungry" codon (*Introduction*), with codon usage frequency serving as a proxy for cognate tRNA abundance (38). Thus, slippage at SS2 may involve a single P-site tRNA (19, 20), and the presence of the rare arginine codon may permit some residual slippage even when a disruptive base pair substitution is introduced into the X_XXN tetramer that precedes the rare codon.

Translational Tiling Proteomics Identifies In Vitro Frameshift Products. Having determined that frameshift suppression also occurs when translation is performed in vitro using the bacterial PURExpress transcription/translation system, we analyzed the products of these reactions by mass spectrometry, as well. $FS^{67/Q66D/G121D}$ harboring a C-terminal FLAG tag in the wild-type frame was employed, thereby allowing purification of full-length frameshift-suppressed product without copurification of His-tagged components in the PURExpress system. α-FLAG antibody-purified full-length products were isolated after gel electrophoresis and subjected to in-gel digestion and translational tiling proteomics.

In total, we detected mixed-frame peptides diagnostic of slippage at 12 different codon positions (the majority of which were detected following chymotrypsin digestion), including SS2 and SS3 (*SI Appendix*, Table S1 [columns 1 & 6]; Dataset S3). As in other experiments, the peptides diagnostic of slippage at SS3 and the adjacent position were represented by higher numbers of PSMs than other diagnostic peptides (in this case, 9 and 5 PSMs, respectively, compared with 1-3 PSMs for the other diagnostic peptides) (Dataset S3). Among the 12 codon positions we identified in this experiment, 11 were positions that were also identified by mixed-frame peptides detected in one or more of the experiments involving samples prepared from cell extracts (*SI Appendix*, Table S1).

Overall, these data indicate the presence of multiple slippery sites that conform at least loosely to previously identified –1 frameshift sequences, but also by tiling across all possible codons, we identified many additional peptides that are diagnostic of frameshifts at unexpected sites (*Discussion*).

Direct Validation of Mixed Frame Peptides Identified by Translational Tiling. We next sought to verify the identities of selected mixed-frame peptides by mass spectrometry. Specifically, we selected seven mixed-frame peptides for validation that were identified frequently within the data described above, but whose presence could not be explained by obvious slippery sequences (*SI Appendix*, Fig. S7*A*). Synthetic peptides were subjected to proteomic analysis and fragment ions compared with those obtained from *E. coli* or in vitro translation (*SI Appendix*, Fig. S7 *B–H*). In each case, there was a strong correspondence between the major b and y fragment ions identified in the synthetic peptide and ions present in frameshift-suppressed samples. These data strongly support the occurrence of extensive nonprogrammed frameshift suppression of FS^{67} both in *E. coli* and in vitro.

Overview of Translational Tiling Results. To provide an overview of our findings with FS^{67}, we tabulated the diagnostic mixed-frame peptides we detected across all experiments according to the number of samples that yielded each particular peptide "envelope." We define a given peptide envelope as those (overlapping) mixed-frame peptides that are diagnostic of slippage at a specific codon position, but reflect distinct cleavage events produced by chymotrypsin or Asp-N. With $FS^{67/Q66D/V117D}$ and $FS^{67/Q66D/G121D}$, we analyzed a total of seven samples (six obtained from cell extracts and one from the in vitro reaction), each of which was digested with either chymotrypsin or Asp-N. Thus, the peptide sets were classified as Tier 1-4 according to the following scheme: Peptide envelopes that were detected in six or seven of the samples (whether following chymotrypsin digestion or Asp-N digestion or both) were designated Tier 1, peptide envelopes that were detected in four or five samples were designated Tier 2, peptide envelopes that were detected in two or three samples were designated Tier 3, and peptide envelopes that were detected in 1 sample only were designated Tier 4. With this classification, there were 7, 10, 20, and 8 peptide envelopes designated Tier 1, 2, 3, and 4, respectively, corresponding to a total of 45 distinct codon positions (Table 1).

We also categorized these 45 slippage sites according to whether they were associated with a plausible match to a slippery heptamer or tetramer motif (16/45) and whether they were in principle compatible with one or more previously described mechanisms for a –1 frameshift event (*SI Appendix*, Table S2). Among the 29 sites that were not associated with a plausible match to a slippery motif, three were associated with a rare codon appropriately positioned to support a P-site slippage event (though in one of those cases, slippage would seem to be maximally unfavorable, leading to mismatches at each codon–anticodon position). Thus, well over half of the –1 frameshift events we detected were not readily explicable based on the understanding of programmed ribosomal frameshifting. We considered one alternative mechanism that involves a near-cognate mispairing event followed by slippage (*Discussion*). Such events could in principle occur if the tRNA that would normally read the first shifted-frame codon (which is the wild-type frame in the case of FS^{67}) is a near-cognate match to the 3′ overlapping original-frame codon (i.e., the FS^{67} frame). Inspection of the relevant overlapping codon sequences associated with the 29 slippage sites that lacked any recognizable slippery motif revealed that 12 (~40%) satisfied this criterion (*SI Appendix*, Table S2).

High Efficiency Frameshift Suppression with a Bacteriophage Lambda Protein. To begin to explore the generality of our observations with the *Cg* SSB PrD-YFP fusion protein, we sought to investigate the occurrence of frameshift suppression (–1 slippage events) in the context of one of the most well-studied proteins from the bacterial domain of life—the CI protein of bacteriophage lambda (λ). To do this, we first generated two λCI frameshift mutants, each of which bore a single base pair deletion (one within codon 8 and the other within codon 58) and a C-terminal His-tag in the wild-type frame (Fig. 5*A*). As we had seen with the *Cg* SSB PrD-YFP frameshift mutants, production of the λCI frameshift mutants (λCI FS^{8} and λCI FS^{58}) resulted in a significant amount

Table 1. Overview of FS[67]-derived mixed-frame products ranked by tier[*]

Transl. Tile #	Tier	Transl. Tile #	Tier	Transl. Tile #	Tier
5	4	40	3	63	3
6/7 (SS1)	3	41	4	65	4
8	4	42	3	68	3
9	3	43	3	73	
15	3	44	2	77/78	3
20	2	45	2	79	3
25	1	46		81	
27	2	47	4	82	3
28	3	50	2	83	1
30	3	51	3	84	1
31		52	2	85	3
31/32 (SS2)	1	53	2	86 (SS6)	2
33	4	55	3	87	4
34		57/58	3	88	4
36/37	1	59	1	89	3
38 (SS3)	1	60/61 (SS4)	3	93	3
39	2	62	2		

[*]See *SI Appendix*, Table S1 for a list that includes the compiled results for the three separate experiments that were tabulated to assign each tier number. Tier 1 entries are highlighted in blue font. Entries lacking a tier number (31, 34, 46, 73, 81) designate mixed-frame products that were detected only with a mutant variant of FS[67] (*SI Appendix*, Table S1).

of full-length frameshift-suppressed protein, as assessed by western blot analysis. In fact, the efficiency of frameshift suppression was ~10% for each of the mutants (even without correcting for any possible decrease in the stabilities of the frameshift-suppressed proteins relative to the wild-type parent protein), considerably more efficient than for the *Cg* SSB PrD-YFP FS[67] mutant (Fig. 5 *B* and *C*).

We again turned to mass spectrometry to identify the sites at which the –1 slippage events were occurring. As before, we define the wild-type frame as frame 1 and the shifted frame as frame 2. In the case of λCI FS[8], the generation of a full-length frameshift-suppressed product would require a slippage event within a 56-codon window before the first frame-2 stop codon is encountered, whereas in the case of λCI FS[58], a slippage event within a 77-codon window would be required (Dataset S4). (Note that in each case, the window extends several codons upstream of the site of the frameshift mutation, where a slippage event could preemptively restore the correct reading frame without stop codon interference.)

For λCI FS[8], we detected mixed-frame peptides diagnostic of slippage at ten different sites within the target window (*SI Appendix*, Table S3 and Datasets S5 and S6). Peptides diagnostic of slippage at one of these sites were represented by a peak of PSMs (26 to 34 across three technical replicates) and peptides diagnostic of slippage at three of the remaining nine sites were represented by 4-8 PSMs (*SI Appendix*, Table S3 and Dataset S6). Nucleotide sequence inspection revealed the presence of a particularly favorable slippery heptamer motif (A_AAA_AAG) (18, 39) that could explain slippage at the peak site, as well as an overlapping heptamer (A_AAG_AAA) that would explain slippage at the adjacent codon (*SI Appendix*, Table S4). To test the overall contribution of these overlapping heptamers to the observed frameshift suppression, we introduced mutations predicted to disrupt slippage at the first heptamer (c_AAc_AAG, termed mut[2,3]) or at both heptamers (c_AAc_AAc, termed mut[2,3,4]). Western blot analysis revealed that these mutations decreased the production of the full-length frameshift-suppressed product substantially (by a factor of ~7 with both heptamers disrupted) (Fig. 5 *D* and *E*).

As for FS[67], we categorized all the identified λCI FS[8] slippage sites according to whether they were associated with a plausible match to a slippery heptamer or tetramer motif (5/10) and whether they were in principle compatible with one or more previously described mechanisms for a –1 frameshift event (*SI Appendix*, Table S4). Among the five sites that were associated with a plausible match to a slippery motif, two were also associated with a rare arginine codon appropriately positioned to support a P-site slippage event (*SI Appendix*, Table S4). Among the remaining five sites, none of which were associated with a rare codon, one was in principle capable of supporting a near-cognate mispairing event that could lead to the observed –1 frameshift event (*Discussion*).

For λCI FS[58], we detected mixed-frame peptides diagnostic of slippage at 39 different sites within the target window (*SI Appendix*, Table S5 and Datasets S5 and S6). Peptides diagnostic of slippage at two of these sites were represented by similar PSM peaks (11 to 21 and 17 to 19 across three technical replicates, respectively); other peptides were represented by up to 7 PSMs (*SI Appendix*, Table S5 and Dataset S6). Neither of these preferred sites appeared to be defined by a candidate slippery heptamer motif, though one of them was associated with a slippery tetramer motif (T_TTT) (*SI Appendix*, Table S6). To test the overall contribution of each of these two sites to the observed frameshift suppression, we introduced single base pair substitutions predicted to create a severe mismatch at position 1 of the tetramer involved in the slippage (G_AGA to c_AGA, termed mut[86] and T_TTT to g_TTT, termed mut[106]). The latter mutation decreased the production of the full-length frameshift-suppressed product modestly, whereas the former mutation had no detectable effect (Fig. 5 *F* and *G*). These results raise the possibility that, in the case of λCI FS[58], slippage at multiple sites may contribute incrementally to the overall efficiency of frameshift suppression.

Among the 39 identified slippage sites, 11 were associated with a plausible match to a slippery heptamer or tetramer motif and one of these was also associated with a rare codon appropriately positioned to support a P-site slippage event (*SI Appendix*, Table S6). Among the remaining 28 sites, two were associated with an analogously positioned rare codon (though in one of these cases slippage would seem to be maximally unfavorable), seven were in principle capable of supporting a near-cognate mispairing event that could lead to the observed –1 frameshift event, and one had a sequence resembling that described in the context of bacteriophage MS2 that supports a noncognate tRNA-mediated –1 frameshift not involving slippage (23, 24) (see Introduction). Thus, 18 of the 39 frameshift events we detected with λCI FS[58] occurred in the absence of any readily identifiable sequence features suggestive of one or another frameshift mechanism.

Frameshift Suppression in Two Other Bacteria. Having performed our experiments in *E. coli* (or in vitro with *E. coli* ribosomes), we wondered whether other bacteria would support similar levels of apparently promiscuous (i.e., nonprogrammed) –1 frameshifting. To begin to investigate this, we selected another gram-negative bacterium, *Pseudomonas aeruginosa* (*Pa*) and one gram-positive bacterium, *Bacillus subtilis* (*Bsu*). We tested the two λCI frameshift mutants, λCI FS[8] and λCI FS[58], in these bacteria. When produced in *Pa*, λCI FS[8] supported efficient frameshift suppression (at an efficiency of ~5%), whereas we detected no frameshift suppression with λCI FS[58] (*SI Appendix*, Fig. S8*A*). However, when produced in *Bsu*, λCI FS[8] supported only very inefficient (barely detectable) frameshift suppression, and we again detected

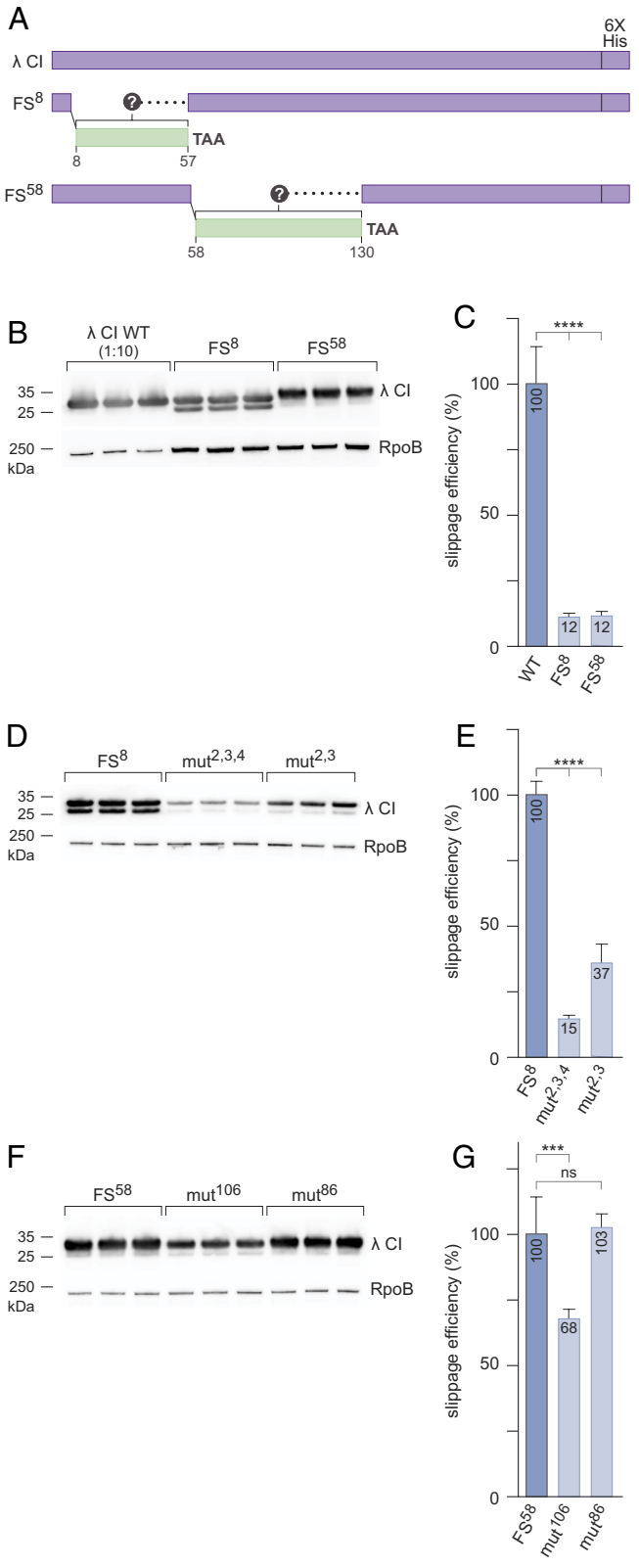

Fig. 5. Frameshift suppression of single base pair deletion mutants of λCI. (*A*) Continuous purple bar (*Top*) depicts wild-type λCI with C-terminal hexahistidine tag (6X His). Interrupted purple bars depict frameshift mutants FS^8 and FS^{58}. Single bp deletion at codon 8 or codon 58 results in shift to +1 frame (frame 2, green). Absent a compensatory slippage event, translation will terminate prematurely at the indicated TAA stop codon. Restoration of the native frame by virtue of a −1 slippage event within the green segment results in a full-length mixed-frame protein detectable with an anti-His antibody. (*B*) Western blot showing that cells producing frameshift mutants FS^8 and FS^{58} contain full-length protein as detected with anti-His antibody. Biological triplicates of extracts prepared from cells producing wild-type λCI were diluted 1:10. The blot was also probed with anti-RpoB as a loading control (*Bottom* strip). FS^8 runs as a doublet, likely due to an N-terminal proteolytic cleavage event that is facilitated by the frameshift mutation, as mutations that reduce the amount of the full-length protein cause a proportional reduction in the amount of the faster migrating species (see below). (*C*) Quantification of the biological triplicates shown in *B*. The amounts of frameshift-suppressed FS^8 (both bands) and FS^{58} were compared to the amount of wild-type λCI, which was set to 100% (taking into account the 10-fold dilution). Error bars represent SEM. Statistical significance was determined by one-way ANOVA followed by Dunnett's multiple comparisons test; *P*-values: <0.0001 (****). (*D*) Western blot (biological triplicates) showing comparison between FS^8 and mutant derivatives ($mut^{2,3,4}$ and $mut^{2,3}$) bearing disruptive mutations in overlapping slippery heptamers. Detection of λCI and RpoB was as described for *B*. (*E*) Quantification of the biological triplicates shown in *D*. Slippage efficiency of FS^8 was set to 100% and error bars represent SEM. Statistical significance was determined as in *C*; *P*-values: <0.0001 (****). (*F*) Western blot (biological triplicates) showing comparison between FS^{58} and mutant derivatives (mut^{106} and mut^{86}) bearing mutations in two preferred slippage sites as identified by mass spectrometry. Detection of λCI and RpoB was as described for *B*. (*G*) Quantification of the biological triplicates shown in *F*. Slippage efficiency of FS^{58} was set to 100% and error bars represent SEM. Statistical significance was determined as in *C*; *P*-values: 0.0001 (***).

no frameshift suppression with λCI FS^{58} (*SI Appendix*, Fig. S8*B*). From these findings, we infer that the biochemical determinants for frameshifting are distinct in different bacteria. We note that another form of translational infidelity—mistranslation—is known to function as an adaptive stress response in certain situations, with organism-specific differences (40).

Discussion

Here, we investigate the occurrence of nonprogrammed ribosomal frameshifting—specifically, −1 slippage events—during translation in *E. coli*. Specifically, we studied slippage events that restored the native frame to frameshift mutants. In the examples we studied, these restoration events occurred at frequencies ranging from 1 to 10%, corresponding to error frequencies of ~10^{-4} per codon and >10^{-3} per codon, respectively. To enable an unbiased approach to the detection of restorative slippage events, we developed a targeted mass-spectrometry method that we call translational tiling proteomics. With this method, we were able to interrogate all possible −1 slippage events that could lead to the observed frameshift suppression.

Diverse −1 Frameshift Events Detected by Mass Spectrometry. Our findings indicated that slippage could occur at a remarkably high fraction of the available sites. In fact, we detected slippage events at approximately one half of the target codons in two of the three examples we studied (at 45 out of 91 codons surveyed in one case and at 39 out of 77 codons surveyed in the other). The identified slippage events included those occurring at sites that we had identified and validated genetically based on their resemblance to canonical "slippery" heptanucleotide motifs defined through the study of programmed ribosomal frameshifting. We also detected slippage events coincident with the presence of slippery tetranucleotide motifs (X_XXY), which have also been found to support programmed ribosomal frameshifting (18). However, approximately two thirds of the slippage events we detected occurred at sites that bore little or no resemblance to slippery heptamers or tetramers (*SI Appendix*, Tables S2, S4, and S6), indicative of greater mechanistic diversity than might have been inferred based on the study of programmed ribosomal frameshifting. Importantly, we tested seven synthetic peptides corresponding to identified mixed-frame peptides whose presence was not explained by obvious slippery sequences. In every case, diagnostic fragment ions matched the spectra obtained from *E. coli* samples, thereby experimentally

validating the occurrence of –1 frameshifting in the absence of recognizable slippery sites.

We propose one alternative mechanism to explain frameshift events that occurred in the absence of obvious slippery site motifs. Specifically, we hypothesized that the cognate tRNA corresponding to the first shifted-frame codon (i.e., frame-1 codon in our examples) might initially pair with the overlapping original-frame codon (i.e., frame-2 codon) and then undergo an A-site slippage event (*SI Appendix*, Fig. S9). This mechanism would presumably be favored in cases where the tRNA in question is a near-cognate match (41) to the original-frame codon it initially engages. Examination of the relevant sequences for the 62 cases (including FS^{67} and both λCI FS^{8} and λCI FS^{58}) lacking obvious slippery site motifs revealed that 20 (~1/3) could in principle occur via such a mechanism (*SI Appendix*, Tables S2, S4, and S6). We note that such events might be particularly favorable in cases where the near-cognate aminoacyl-tRNA is in significant excess over the competing cognate aminoacyl-tRNA.

Among all the –1 frameshift events we detected (a total of 94 for FS^{67}, λCI FS^{8}, and λCI FS^{58}), we observed the presence of an appropriately positioned rare codon to support a P-site slippage event in 10 cases, of which 5 were associated with a plausible match to a slippery tetramer motif (*SI Appendix*, Tables S2, S4, and S6). We also noted 4 cases that could in principle be explained by a noncognate tRNA-mediated event analogous to those described in the context of bacteriophage MS2, which do not involve slippage but rather an out-of-register A-site pairing resulting in the shift to the –1 frame upon translocation (23, 24) (*Introduction*). In the documented cases of this mechanism, the incoming aminoacyl-tRNA is either $tRNA^{Ser3}$, which mispairs with an alanine codon, or $tRNA^{Thr3}$, which mispairs with a proline codon. Among the 4 cases in our data that could be explained by this mechanism, two would involve an incoming $tRNA^{Ser3}$ and the other two an incoming $tRNA^{Thr3}$ (*SI Appendix*, Tables S2, S4, and S6). Whether or not any other tRNA-codon combinations might mediate nonslippage dependent frameshifting by a similar mechanism is unclear. But if so, then this class of event might explain at least some of the events we detected at sites that were not associated with any obvious slippery motif.

In addition to the presence of a slippery heptanucleotide or tetranucleotide, which appears to play an essential role in diverse examples of programmed –1 frameshifting across organisms, mRNA structural features (e.g., stem-loop structures and pseudoknots) can significantly augment the effect of the slippery sequence (9, 11, 27, 28). Further studies are necessary to understand the contribution of mRNA sequence or structure to the occurrence of basal ribosomal frameshifting.

Potential Impact of Translation Errors that Result in Frameshift Events. Although we are not aware of previous attempts to look in a systematic manner for ribosomal slippage events resulting in mixed frame polypeptides, translation errors resulting in frameshift events have been described previously. A particularly dramatic example involves the isolation of a viable strain of *E. coli* containing a single base pair insertion in an essential gene that encodes the β subunit of RNA polymerase (RpoB) (42). In this case, frameshift suppression was shown to occur via a specific +1 slippage event facilitated by the presence of a particularly favorable sequence (consisting of a rare AGG codon preceded by a slippery CCC codon) unveiled by the frameshift mutation, resulting in a full-length polypeptide with just three nonnative amino acids that confer resistance to rifampicin. Although the efficiency of the frameshift suppression was 5%, the protein was produced at near wild-type levels due to a feedback loop causing translational upregulation in response to insufficient RpoB. In our examples, involving –1 slippage events, the proteins were produced in a manner that was devoid of autoregulation, but the example of RpoB highlights a mechanism whereby the effects of frameshift suppression can be dramatically amplified regardless of whether the events in question occur at a unique site or at an ensemble of sites.

Organism-Specific Differences in Translational Fidelity. Whereas our detailed analyses were carried out with proteins produced in *E. coli* or in *E. coli*-based in vitro reactions, we also selected two other bacteria, *P. aeruginosa* (a gram-negative organism) and *B. subtilis* (a gram-positive organism) for comparison. We found that the overall efficiency of ribosomal frameshifting differed among the three bacteria. With one of the two frameshift mutants that we tested, we observed efficient frameshift suppression in both *E. coli* and *P. aeruginosa* (though the efficiency was somewhat higher in *E. coli*), suggestive of some mechanistic conservation between the two organisms. In contrast, *B. subtilis* supported only barely detectable frameshift suppression in the case of the same mutant. However, with the other frameshift mutant, which was efficiently suppressed in *E. coli*, we observed no detectable frameshift suppression in the other two bacteria. It will be interesting to investigate the basis for the apparently greater translational fidelity in *P. aeruginosa* and especially *B. subtilis*. We speculate that there may be evolutionary trade-offs that dictate these differences, depending on the physiological requirements imposed by the environments experienced by the different organisms.

Conclusions

Our findings suggest that basal levels of ribosomal frameshifting are higher than might have been anticipated and that the sites at which such events occur exhibit an unexpected diversity. The mechanistic implications of this diversity remain to be explored and could also inform studies of programmed ribosomal frameshifting. This form of translational infidelity highlights the need for caution in making assumptions about the relationship between genotype and phenotype (e.g., pseudogenes containing frameshift mutations cannot necessarily be assumed to behave as null alleles). The impact of these occurrences of frameshifting is likely to be particularly significant in the context of highly expressed genes, as modeled here in the context of an inducible promoter. A major unanswered question raised by our findings is the extent to which frameshifted products shape the proteome and potentially augment the functional proteome. Finally, our findings and others point to basal ribosomal frameshifting as an underappreciated source of evolutionary plasticity, setting the stage for the evolution of programmed frameshifting at specific sites.

Materials and Methods

Detailed procedures related to bacterial strains and growth conditions, plasmid construction, random and directed plasmid mutagenesis, western blotting, in vitro protein synthesis, mass spectrometry (sample preparation, data acquisition, and data analysis) are described in *SI Appendix, Materials and Methods*. Statistical analyses were performed using GraphPad Prism (v. 9.4.1). All error bars represent SEM and statistical significance was determined by one-way ANOVA followed by Dunnett's multiple comparisons test or by the *t*-test, as specified in the corresponding figure legends.

Data, Materials, and Software Availability. Key plasmids developed in this work will be deposited with Addgene; all other plasmids developed in this work will be made available upon request with no restrictions. Proteomic

data (.RAW files) and SDRF (sample and data relationship format) files are available via ProteomeXchange with identifier PXD047675 (43).

ACKNOWLEDGMENTS. We thank S. L. Dove for valuable discussion and comments on the manuscript and R. Hellmiss for artwork. This work was supported by NIH grants GM136247 to A.H., AG011085 to J.W.H., and GM132129 to J.A.P.

Author affiliations: [a]Department of Microbiology, Harvard Medical School, Boston MA 02115; and [b]Department of Cell Biology, Harvard Medical School, Boston MA 02115

Author contributions: B.L.S., J.A.P., H.P., J.W.H., and A.H. designed research; B.L.S., J.A.P., H.P., K.H., E.F., and Z.F. performed research; B.L.S., J.A.P., H.P., J.W.H., and A.H. analyzed data; and J.W.H. and A.H. wrote the paper.

Competing interest statement: J.W.H. is a consultant and founder of Caraway Therapeutics and is a founding board member of Interline Therapeutics. Other authors declare no competing interests.

1. D. A. Drummond, C. O. Wilke, The evolutionary consequences of erroneous protein synthesis. *Nat. Rev. Genet.* **10**, 715–724 (2009).
2. J. F. Atkins, D. Elseviers, L. Gorini, Low activity of -galactosidase in frameshift mutants of Escherichia coli. *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1192–1195 (1972).
3. C. G. Kurland, Translational accuracy and the fitness of bacteria. *Annu. Rev. Genet.* **26**, 29–50 (1992).
4. M. V. Rodnina *et al.*, Translational recoding: Canonical translation mechanisms reinterpreted. *Nucleic Acids Res.* **48**, 1056–1067 (2020).
5. T. Jacks, H. E. Varmus, Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science* **230**, 1237–1242 (1985).
6. T. Jacks *et al.*, Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **331**, 280–283 (1988).
7. T. Jacks, H. D. Madhani, F. R. Masiarz, H. E. Varmus, Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. *Cell* **55**, 447–458 (1988).
8. J. D. Dinman, Mechanisms and implications of programmed translational frameshifting. *Wiley Interdiscip. Rev. RNA* **3**, 661–673 (2012).
9. J. F. Atkins, G. Loughran, P. R. Bhatt, A. E. Firth, P. V. Baranov, Ribosomal frameshifting and transcriptional slippage: From genetic steganography and cryptography to adventitious use. *Nucleic Acids Res.* **44**, 7007–7078 (2016).
10. Y. A. Khan *et al.*, Evaluating ribosomal frameshifting in CCR5 mRNA decoding. *Nature* **604**, E16–E23 (2022).
11. N. Caliskan, F. Peske, M. V. Rodnina, Changed in translation: mRNA recoding by -1 programmed ribosomal frameshifting. *Trends Biochem. Sci.* **40**, 265–274 (2015).
12. P. V. Baranov, R. F. Gesteland, J. F. Atkins, Release factor 2 frameshifting sites in different bacteria. *EMBO Rep.* **3**, 373–377 (2002).
13. Z. Tsuchihashi, A. Kornberg, Translational frameshifting generates the gamma subunit of DNA polymerase III holoenzyme. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 2516–2520 (1990).
14. A. L. Blinkowa, J. R. Walker, Programmed ribosomal frameshifting generates the Escherichia coli DNA polymerase III gamma subunit from within the tau subunit reading frame. *Nucleic Acids Res.* **18**, 1725–1729 (1990).
15. A. M. Flower, C. S. McHenry, The gamma subunit of DNA polymerase III holoenzyme of Escherichia coli is produced by ribosomal frameshifting. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 3713–3717 (1990).
16. W. J. Craigen, C. T. Caskey, Expression of peptide chain release factor 2 requires high-efficiency frameshift. *Nature* **322**, 273–275 (1986).
17. S. Meydan *et al.*, Programmed ribosomal frameshifting generates a copper transporter and a copper chaperone from the same gene. *Mol. Cell* **65**, 207–219 (2017).
18. V. Sharma *et al.*, Analysis of tetra- and hepta-nucleotides motifs promoting -1 ribosomal frameshifting in Escherichia coli. *Nucleic Acids Res.* **42**, 7210–7225 (2014).
19. N. Korniy, E. Samatova, M. M. Anokhina, F. Peske, M. V. Rodnina, Mechanisms and biomedical implications of -1 programmed ribosome frameshifting on viral and bacterial mRNAs. *FEBS Lett.* **593**, 1468–1482 (2019).
20. Z. Barak, D. Lindsley, J. Gallant, On the mechanism of leftward frameshifting at several hungry codons. *J. Mol. Biol.* **256**, 676–684 (1996).
21. L. V. Bock *et al.*, Thermodynamic control of -1 programmed ribosomal frameshifting. *Nat. Commun.* **10**, 4598 (2019).
22. P. Licznar *et al.*, Programmed translational -1 frameshifting on hexanucleotide motifs and the wobble properties of tRNAs. *EMBO J.* **22**, 4770–4778 (2003).
23. J. F. Atkins, R. F. Gesteland, B. R. Reid, C. W. Anderson, Normal tRNAs promote ribosomal frameshifting. *Cell* **18**, 1119–1131 (1979).
24. T. J. Dayhuff, J. F. Atkins, R. F. Gesteland, Characterization of ribosomal frameshift events by protein sequence analysis. *J. Biol. Chem.* **261**, 7491–7500 (1986).
25. J. F. Atkins *et al.*, "Poking a hole in the sanctity of the triplet code: Inferences for framing" in *The Ribosome. Structure, Function, Antibiotics and Cellular Interactions*, R. A. Garrett *et al.*, Eds. (ASM Press, Washington D.C., 2000), pp. 369–383.
26. T. Caulfield, M. Coban, A. Tek, S. C. Flores, Molecular dynamics simulations suggest a non-doublet decoding model of -1 frameshifting by tRNA(Ser3). *Biomolecules* **9**, 745 (2019).
27. P. J. Carmody *et al.*, Coordination of -1 programmed ribosomal frameshifting by transcript and nascent chain features revealed by deep mutational scanning. *Nucleic Acids Res.* **49**, 12943–12954 (2021).
28. T. E. Dever, J. D. Dinman, R. Green, Translation elongation and recoding in Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **10**, a032649 (2018).
29. C. Govarts, K. Somers, P. Stinissen, V. Somers, Frameshifting in the p6 cDNA phage display system. *Molecules* **15**, 9380–9390 (2010).
30. O. L. Gurvich *et al.*, Sequences that direct significant levels of frameshifting are frequent in coding regions of Escherichia coli. *EMBO J.* **22**, 5941–5950 (2003).
31. E. Fleming, A. H. Yuan, D. M. Heller, A. Hochschild, A bacteria-based genetic assay detects prion formation. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 4605–4610 (2019).
32. B. M. Burmann *et al.*, A NusE:NusG complex links transcription and translation. *Science* **328**, 501–504 (2010).
33. C. Wang *et al.*, Structural basis of transcription-translation coupling. *Science* **369**, 1359–1365 (2020).
34. R. S. Washburn *et al.*, Escherichia coli NusG links the lead ribosome with the transcription elongation complex. *iScience* **23**, 101352 (2020).
35. K. L. Herbst, L. M. Nichols, R. F. Gesteland, R. B. Weiss, A mutation in ribosomal protein L9 affects ribosomal hopping during translation of gene 60 from bacteriophage T4. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12525–12529 (1994).
36. A. J. Herr, C. C. Nelson, N. M. Wills, R. F. Gesteland, J. F. Atkins, Analysis of the roles of tRNA structure, ribosomal protein L9, and the bacteriophage T4 gene 60 bypassing signals during ribosome slippage on mRNA. *J. Mol. Biol.* **309**, 1029–1048 (2001).
37. A. M. Smith, M. S. Costello, A. H. Kettring, R. J. Wingo, S. D. Moore, Ribosome collisions alter frameshifting at translational reprogramming motifs in bacterial mRNAs. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 21769–21779 (2019).
38. Y. Liu, Q. Yang, F. Zhao, Synonymous but not silent: The codon usage code for gene expression and protein folding. *Annu. Rev. Biochem.* **90**, 375–401 (2021).
39. K. S. Koutmou *et al.*, Ribosomes slide on lysine-encoding homopolymeric A stretches. *Elife* **4**, e05534 (2015).
40. K. Mohler, M. Ibba, Translational fidelity and mistranslation in the cellular response to stress. *Nat. Microbiol.* **2**, 17117 (2017).
41. S. Rudorf, R. Lipowsky, Protein synthesis in E. coli: Dependence of codon-specific elongation on tRNA concentration and codon usage. *PLoS One* **10**, e0134994 (2015).
42. D. L. Huseby, G. Brandis, L. Praski Alzrigat, D. Hughes, Antibiotic resistance by high-level intrinsic suppression of a frameshift mutation in an essential gene. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 3185–3191 (2020).
43. J. A. Paulo, J.W. Harper, "Systematic analysis of non programmed frameshift suppression in E. coli via translational tiling proteomics" PRoteomics IDEntifications (PRIDE) Archive database. https://www.ebi.ac.uk/pride/archive/projects/PXD047675. Deposited 12 October 2023.