Programmed ribosomal frameshifting occurs in the *E. coli* copper transporter gene *copA*

The copper chaperone resulting from the frameshift contributes to copper tolerance

A slippery sequence, an mRNA structure, and the CopA nascent chain stimulate the event

*copA* frameshifting elements are found in several bacteria and the human *ATP7B* gene
Programmed Ribosomal Frameshifting Generates a Copper Transporter and a Copper Chaperone from the Same Gene

Sezen Meydan,1 Dorota Klepacki,1 Subbulakshmi Karthikeyan,1 Tõnu Margus,1 Paul Thomas,2 John E. Jones,3 Yousuf Khan,4 Joseph Briggs,5 Jonathan D. Dinman,3 Nora Vázquez-Laslop,1,* and Alexander S. Mankin1,4,*

1Center for Biomolecular Sciences-m/c 870, University of Illinois at Chicago, 900 S. Ashland Avenue, Chicago, IL 60607, USA
2Proteomics Center of Excellence, Northwestern University, 633 Clark Street, Chicago, IL 60208, USA
3Department of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD 20742, USA
4*Correspondence: nvazquez@uic.edu (N.V.-L.), shura@uic.edu (A.S.M.)
http://dx.doi.org/10.1016/j.molcel.2016.12.008

SUMMARY

Metal efflux pumps maintain ion homeostasis in the cell. The functions of the transporters are often supported by chaperone proteins, which scavenge the metal ions from the cytoplasm. Although the copper ion transporter CopA has been known in Escherichia coli, no gene for its chaperone had been identified. We show that the CopA chaperone is expressed in E. coli from the same gene that encodes the transporter. Some ribosomes translating copA undergo programmed frameshifting, terminate translation in the −1 frame, and generate the 70 aa-long polypeptide CopA(Z), which helps cells survive toxic copper concentrations. The high efficiency of frameshifting is achieved by the combined stimulatory action of a “slippery” sequence, an mRNA pseudoknot, and the CopA nascent chain. Similar mRNA elements are not only found in the copA genes of other bacteria but are also present in ATP7B, the human homolog of copA, and direct ribosomal frameshifting in vivo.

INTRODUCTION

Copper homeostasis is critical for organisms from all domains of life. Due to the ability of copper ions to switch between two oxidation states, Cu(I) and Cu(II), copper serves as an essential cofactor for enzymes that participate in key processes including electron transport and the oxidative stress response (Arredondo and Núñez, 2005). In excess, however, copper ions are extremely toxic for the cell, likely because of their role in generating reactive oxygen species (Rensing and Grass, 2003). Disruption of copper homeostasis has been linked to human maladies, e.g., Menkes, Wilson, and Parkinson diseases, as well as cystic fibrosis (Bull et al., 1993; Percival et al., 1999; Tórs-dóttir et al., 1999; Vulpe et al., 1993). Sophisticated systems have evolved to maintain copper homeostasis in bacterial and eukaryotic cells (Fan and Rosen, 2002; Outten and O’Halloran, 2001), in which the central role belongs to copper-translocating ATPases responsible for export of excess copper ions from the cell (Fan and Rosen, 2002; Migocka, 2015).

The operation of copper transporters in bacteria often relies on the assistance of metal chaperones (e.g., CopZ in Bacillus subtilis and Enterococcus hirae). These small soluble proteins facilitate trafficking of copper ions to the transporters and their regulators (Banci et al., 2001; Cobine et al., 1999; Palumaa, 2013). Curiously, while the CufI efflux transporter CopA operates in Escherichia coli, a gene for the CopZ-like diffusible chaperone was not identified in its genome (Fan et al., 2001; Rensing et al., 2000). A recent study demonstrated that the copper hypersensitivity caused by the artificial truncation of the N-terminal metal binding domain 1 (MBD1) of the E. coli CopA transporter could be compensated in trans by expression of the B. subtilis CopZ chaperone (Drees et al., 2015). Intriguingly, the same beneficial effect was achieved by ectopic expression of the E. coli CopA MBD1 itself, indicating that the N-terminal segment of CopA could potentially provide the copper chaperone function in E. coli. However, in contrast to the diffusible metal chaperones found in many organisms (Jordan et al., 2001; Palumaa, 2013), the putative copper chaperone of E. coli, as a part of the CopA protein, has to operate as an integral part of the membrane transporter. Such an arrangement should inevitably restrict the ability of the chaperone to scavenge copper ions from cytosolic locations and deliver them to the membrane-embedded efflux pump.

Several serendipitous observations hinted that expression of the E. coli copA gene might deviate from the conventional pathway. Gel analysis of small proteins expressed in E. coli revealed the presence of a short polypeptide with an estimated molecular weight (MW) of ~6.5 kDa, whose tryptic peptides matched those of the CopA MBD1 (Wasinger and Humphrey-Smith, 1998). Although such a protein could hypothetically be generated via proteolytic degradation of the full-size transporter, no evidence for specific cleavage of CopA by cellular proteases has been reported. Even more puzzling, recent ribosome profiling examination of protein synthesis in the E. coli strain MG1655 indicated that the number of translating ribosomes abruptly drops in the vicinity of the 70th codon of the copA gene (Li et al., 2014). A similar unexplained decrease in ribosome...
density within copA could also be seen in other ribosome profiling datasets collected from various *E. coli* strains and under different growth conditions (Balakrishnan et al., 2014; Elgamal et al., 2014; Guo et al., 2014; Kannan et al., 2014; Li et al., 2014; Mohammad et al., 2016; Oh et al., 2011). This unique pattern of distribution of ribosome progression along the gene suggested that expression of copA might be a subject of idiosyncratic regulation.

One of the mechanisms exploited by cells for expanding the spectrum of proteins expressed from a limited number of genomic open reading frames (ORFs) is translational recoding (Baranov et al., 2002). This term refers to a variety of scenarios in which interpretation of the genetic information deviates from the straightforward single-frame codon-by-codon translation of mRNA by the ribosomes. Among other options, recoding may involve programmed ribosomal frameshifting (PRF), a forward or backward slippage of the ribosome to an alternative reading frame within the ORF. PRF conceptually resembles spontaneous frameshift errors, although its frequency is usually significantly higher and could be subjected to specific regulation (Farabaugh and Björk, 1999; Kurland, 1992). In order to achieve high efficiency of recoding, particular structural features are embedded in mRNA, including PRF-prone “slippery sequences” (SSs), upstream or downstream stimulatory RNA secondary structures, or the presence of internal Shine-Dalgarno-like sequences (Caliskan et al., 2015; Farabaugh, 1996). Two core *E. coli* genes are known to be regulated by PRF. Expression of release factor 2 (RF2) from the prfB gene is controlled by +1 PRF. The PRF frequency depends on the efficiency of translation termination at the premature in-frame stop codon, which in turn requires RF2 activity. Only the full-size release factor, generated as the result of the premature in-frame stop codon, which in turn requires RF2 activity, is known to be controlled by PRF. Carries physiologically meaningful cellular function, whereas the truncated, prematurely terminated peptide seems to play no functional role (Craigin and Caskey, 1986). In contrast to prfB, the −1 PRF in the dnaX gene generates two functional polypeptides corresponding to individual subunits of the same enzyme, DNA polymerase III (Blinkova and Walker, 1999; Flower and McHenry, 1990; Tsuchihashi and Kornberg, 1990). More recently, it has been shown that −1 PRF in the gene csoS2 of *CO2*-fixating bacteria leads to production of two isoforms of the CsoS2 protein involved in the biogenesis of α-carboxisome. However, the functionality of one of the isoforms still remains unconfirmed (Chaijarasphong et al., 2016).

Here, we show that the copA gene in *E. coli* encodes two proteins, likely with related but distinct functions. Translation of the entire gene generates a membrane copper transporter CopA, while −1 PRF leading to premature termination results in synthesis of the 70 aa-long copper chaperone CopA(Z). A highly efficient −1 PRF, stimulated by an SS and specific elements encoded within the mRNA and nascent polypeptide, controls expression of the two polypeptides from a single ORF. The same SS and a similar downstream mRNA structure are present in the human *ATP7B* gene, which codes for the copper transporter homolog of bacterial CopA, and whose functional defects are implicated in Wilson disease (Bull et al., 1993; Gupta and Lutsenko, 2012). The utilization of PRF by the copper transporter genes illuminates the role of recoding in maintaining the homeostasis of an essential metal in the cell.

RESULTS

In Vivo Expression of the copA Gene Results in the Formation of the 70 aa-Long CopA(Z)

Ribosome profiling revealed an abrupt decrease of ribosome density in the vicinity of the 70th codon of the *E. coli* copA gene (Balakrishnan et al., 2014; Elgamal et al., 2014; Guo et al., 2014; Kannan et al., 2014; Li et al., 2014; Mohammad et al., 2016; Oh et al., 2011) (Figure 1A). We hypothesized that this unique pattern of copA translation could reflect the specific expression of an approximately 70 aa residues-long truncated CopA protein via an as yet undefined mechanism. This putative polypeptide would encompass the entire MB1 of the CopA transporter (Figure 1A) and thus would closely correspond to the N-terminal segment of CopA proposed to serve as a transporter-linked copper chaperone in *E. coli* (Drees et al., 2015). In order to explore whether the N-terminal segment of CopA is indeed synthesized in the cell as an individual polypeptide, we used the copA-containing ASKA library plasmid (Kitagawa et al., 2005) to express the N-terminally His6-tagged CopA (we will refer to this plasmid as pCopA). Introduction of pCopA in the ΔcopA *E. coli* strain BW25113 (Baba et al., 2006) led to the appearance of not only the full-size CopA transporter, but also of a polypeptide that migrated in an SDS gel as a ~10–12 kDa protein (Figures 1B and 2B). The size of this shorter product, as estimated from its electrophoretic mobility, closely matched the predicted MW for the His6-tagged MB1 of CopA (~9 kDa). To determine the precise size of the expressed truncated CopA polypeptide, which we named CopA(Z), we purified it using an Ni2+ affinity column (Figure 1B) and determined its exact MW by mass spectrometry (Figure 1C). The experimentally determined MW of the expressed His6-tagged CopA(Z) was 9,323 Da. This value was in reasonable agreement with the profiling results, where the ribosomal density dropped immediately after the 70th codon; the predicted MW of the polypeptide encoded by the first 70 codons of the copA gene would be 9,337 Da. Thus, the results of the mass spectrometry analysis suggested that a diffusible 70 aa-long CopA(Z) protein encompassing MB1 is expressed from the copA gene alongside the full-size 834 aa-long CopA, a transmembrane metal transporter. Such a scenario would resolve the mystery of the missing copper chaperone in *E. coli*. However, the mechanism of generation of CopA(Z), as well as the origin of the difference of 14 Da between its predicted and experimentally determined MWs (Figure 1C), remained puzzling.

PRF Resulting in Premature Termination of Translation Is Responsible for the Production of CopA(Z)

We explored the reason for the discrepancy between the predicted and experimentally determined MW of CopA(Z), hoping that it held the clue to the mechanism of its generation. The molecular mass difference between the estimated and experimental CopA(Z) molecular mass could be accounted for by the replacement of the CopA(Z) C-terminal alanine, encoded in the copA 70th codon, with glycine (Figure 1C). Such a change could be brought about by a single nucleotide substitution in the 70th codon, converting it from GCT to GGT (Gly). However, neither the chromosomal gene of the parental BW25113 strain nor the
pressed in the IPTG-induced cells and purified by Ni
(C) Top-down mass spectrometric analysis of the purified His
indicated by arrowheads.
of His multiply charged species of the intact protein; for reference, the peak corre-
coded in the 70th codon of the predicted mass of CopA(Z) (9,337.60 Da), where the C-terminal Ala, en-
pA(Z). The deconvoluted monoisotopic mass of 9,323.61 Da is consistent with
from the control cells in
in the
binding domains, MBD1 and MBD2, of the CopA transporter protein encoded
70th codon of the
(A) Ribosome profiling reveals an abrupt drop in ribosome density around the
70th codon of the copA gene (black arrow). The boundaries of the two metal
binding domains, MBD1 and MBD2, of the CopA transporter protein encoded
in the copA gene are indicated; MBD1 is colored gray. The profiling data are
from the control cells in Kannan et al. (2014).
(B) Two left panels: Coomassie-stained 15% SDS gel showing low-MW pro-
teins in the control and IPTG-induced ∆copA E. coli cells carrying the pCopA
plasmid, which encodes N-terminally His6-tagged CopA. Right: silver-stained
4%-20% SDS gel showing the N-terminally His6-tagged short protein ex-
pressed in the IPTG-induced cells and purified by Ni²⁺-affinity chromatog-
raphy. The short protein with the electrophoretic mobility approximately corre-
sponding to the His6-tagged MBD2 of CopA, which we named CopA(Z), is
indicated by arrowheads.
(C) Top-down mass spectrometric analysis of the purified His6-tagged Co-
pA(Z). The deconvoluted monoisotopic mass of 9,323.61 Da is consistent with
the predicted mass of CopA(Z) (9,337.60 Da), where the C-terminal Ala, en-
coded in the 70th codon of copA, is replaced with Gly. The spectrum shows
multiply charged species of the intact protein; for reference, the peak corre-
sponding to the charge +10 is indicated by an arrow. The amino acid sequence
of His6-tagged CopA(Z) is shown below and the C-terminal Gly is boxed.

Figure 1. The Short Polypeptide CopA(Z) Is Translated in the Cell
from the E. coli copA Gene
(A) Ribosome profiling reveals an abrupt drop in ribosome density around the
70th codon of the copA gene (black arrow). The boundaries of the two metal
binding domains, MBD1 and MBD2, of the CopA transporter protein encoded
in the copA gene are indicated; MBD1 is colored gray. The profiling data are
from the control cells in Kannan et al. (2014).
(B) Two left panels: Coomassie-stained 15% SDS gel showing low-MW pro-
teins in the control and IPTG-induced ∆copA E. coli cells carrying the pCopA
plasmid, which encodes N-terminally His6-tagged CopA. Right: silver-stained
4%-20% SDS gel showing the N-terminally His6-tagged short protein ex-
pressed in the IPTG-induced cells and purified by Ni²⁺-affinity chromatog-
raphy. The short protein with the electrophoretic mobility approximately corre-
sponding to the His6-tagged MBD2 of CopA, which we named CopA(Z), is
indicated by arrowheads.
(C) Top-down mass spectrometric analysis of the purified His6-tagged Co-
pA(Z). The deconvoluted monoisotopic mass of 9,323.61 Da is consistent with
the predicted mass of CopA(Z) (9,337.60 Da), where the C-terminal Ala, en-
coded in the 70th codon of copA, is replaced with Gly. The spectrum shows
multiply charged species of the intact protein; for reference, the peak corre-
sponding to the charge +10 is indicated by an arrow. The amino acid sequence
of His6-tagged CopA(Z) is shown below and the C-terminal Gly is boxed.

Secondary Structure of the copA mRNA
Stimulates –1 PRF
The precipitous drop in ribosome density after the 70th codon of
the copA gene observed in profiling experiments (Figure 1A) sug-
gests that a large fraction of the ribosomes that initiate translation
at the start codon of the gene could shift to the –1 frame and
prematurely terminate after translating 70 codons. Because the mere
presence of a 7 nt-long SS is insufficient to account for
PRF (Giedroc et al., 2000), we hypothesized that additional
proteins were detected using anti-His containing the WT (B) Immunoblot of the lysates of following stop codon, generating the CopA(Z) protein.

coding segment. Slippage of the ribosome to the codons of (C) In vitro transcription-translation of a DNA template containing the first 94

tionated in a 4%–20% SDS gel. The SS-disrupting mutations (which do not

Figure 2. Production of CopA(Z) Depends on the Integrity of the SS Present in the copA Gene

(A) The SS 5’-CCCAGG-3’ present in the copA gene at the end of the MBD1-coding segment. Slippage of the ribosome to the –1 frame would cause the ribosome to incorporate Gly-70 after Lys-69 and terminate translation at the following stop codon, generating the CopA(ΔZ) protein.
(B) Immunoblot of the lysates of E. coli cells carrying either the pCopA plasmid containing the WT copA gene or pCopA-mSS in which the SS was mutated to prevent –1 PRF. The CopA (gray arrowhead) and CopA(Z) (black arrowhead) proteins were detected using anti-His- tag antibodies. Proteins were fractionated in a 4%–20% SDS gel. The SS-disrupting mutations (which do not change the sequence of the encoded protein) are underlined.
(C) In vitro transcription-translation of a DNA template containing the first 94 codons of copA, followed by an engineered stop codon. The 16.5% Tris-Tricine SDS gel shows the [35S]-labeled products corresponding to the complete 94 aa-long protein encoded in the 0 frame (gray arrowhead) or the 70 aa-long CopA(Z) produced via PRF (black arrowhead). Disruption of SS in the template mSS by two synonymous mutations (shown in B) prevents production of CopA(Z) in vitro.

See also Figure S2.

The -1 PRF is often promoted by mRNA structural elements that hinder the forward movement of the ribosome along the transcript, thereby stimulating the backward slippage of the ribosome-tRNA complex (Plant et al., 2003). Computational modeling of the possible folding of the copA mRNA segment following the –1 PRF site predicted the formation of a stem-loop structure, which may be a part of a stable pseudoknot (PK) (ΔG = –19.7 kcal/mole) (Figure 3A). Furthermore, the location of the first stem (S1) of the predicted PK relative to the SS is compatible with its stimulatory role in promoting efficient –1 PRF (Giedroc and Cornish, 2009).

We tested the contribution of the downstream mRNA segment containing the putative PK to PRF efficiency by analyzing in vitro expression of CopA(ΔZ) from a series of 3’ truncated copA templates (Figure 3B). The longest template, copA1–312, contained the first 104 codons of copA (the region encompassing the SS and the entire PK), followed by an engineered stop codon that would direct translation termination in the 0 frame. Expression of this construct yielded both the full-length encoded protein (104 aa long) and the shorter (70 aa long) CopA(ΔZ) (Figure 3C, lane “1–312”). The –1 PRF efficiency, calculated from the relative intensities of the major products bands, was approximately 45% (Figure 3D). The efficiency of –1 PRF gradually diminished in the progressively 3’ truncated templates in which the second (S2) PK stem (constructs copA1–261, copA1–222, and copA1–222) or both PK stems (construct copA1–222) were eliminated (Figure 3B). CopA(ΔZ) production was largely abrogated when only 15 nucleotides downstream of SS were present in the construct (Figures 3C and 3D). Collectively, the results of these experiments indicate that the mRNA segment downstream of SS likely adopts a specific secondary structure fold and is essential for promoting highly efficient –1 PRF that mediates the generation of CopA(Z).

The CopA Nascent Peptide Modulates –1 PRF

Nascent peptide-ribosome interactions often play an important role in the regulation of translation (Ito and Chiba, 2013) and can influence certain recoding events (Chen et al., 2015; Gupta et al., 2013; Samatova et al., 2014; Weiss et al., 1990; Yordanova et al., 2015). We considered the possibility that the CopA nascent chain might affect the frequency of –1 PRF and hence play a role in controlling the relative expression of CopA(Z) and CopA proteins from the same gene. In order to test the effect of the CopA nascent peptide on PRF, we introduced two sets of compensatory single-nucleotide indel mutations within the template copA1–312. The first set of mutations (construct copA1–312-NP1; Figures 4A and S1, available online) altered the sequence of CopA amino acid residues 4–29 located outside the ribosomal exit tunnel when the ribosome reaches the PRF site (Figures 4A and S1). With the second set of mutations, the sequence of the CopA segment 31–67, which resides in the exit tunnel of the frameshifting ribosome, was altered (construct copA1–312-NP2). The integrity of the SS and PK regions was preserved in both mutants. Because the amino acid composition of either of the mutant proteins is different from that of the WT CopA(Z), control templates were prepared (constructs CopA(Z)NP1 and CopA(Z)NP2 in Figure S1) to generate the corresponding

elements residing either in the copA mRNA and/or the encoded CopA peptide could stimulate recoding.

The –1 PRF is often promoted by mRNA structural elements that hinder the forward movement of the ribosome along the transcript, thereby stimulating the backward slippage of the ribosome-tRNA complex (Plant et al., 2003). Computational modeling of the possible folding of the copA mRNA segment following the –1 PRF site predicted the formation of a stem-loop structure, which may be a part of a stable pseudoknot (PK) (ΔG = –19.7 kcal/mole) (Figure 3A). Furthermore, the location of the first stem (S1) of the predicted PK relative to the SS is compatible with its stimulatory role in promoting efficient –1 PRF (Giedroc and Cornish, 2009).

We tested the contribution of the downstream mRNA segment containing the putative PK to PRF efficiency by analyzing in vitro expression of CopA(Z) from a series of 3’ truncated copA templates (Figure 3B). The longest template, copA1–312, contained the first 104 codons of copA (the region encompassing the SS and the entire PK), followed by an engineered stop codon that would direct translation termination in the 0 frame. Expression of this construct yielded both the full-length encoded protein (104 aa long) and the shorter (70 aa long) CopA(Z) (Figure 3C, lane “1–312”). The –1 PRF efficiency, calculated from the relative intensities of the major products bands, was approximately 45% (Figure 3D). The efficiency of –1 PRF gradually diminished in the progressively 3’ truncated templates in which the second (S2) PK stem (constructs copA1–261, copA1–222, and copA1–222) or both PK stems (construct copA1–222) were eliminated (Figure 3B). CopA(Z) production was largely abrogated when only 15 nucleotides downstream of SS were present in the construct (Figures 3C and 3D). Collectively, the results of these experiments indicate that the mRNA segment downstream of SS likely adopts a specific secondary structure fold and is essential for promoting highly efficient –1 PRF that mediates the generation of CopA(Z).

The CopA Nascent Peptide Modulates –1 PRF

Nascent peptide-ribosome interactions often play an important role in the regulation of translation (Ito and Chiba, 2013) and can influence certain recoding events (Chen et al., 2015; Gupta et al., 2013; Samatova et al., 2014; Weiss et al., 1990; Yordanova et al., 2015). We considered the possibility that the CopA nascent chain might affect the frequency of –1 PRF and hence play a role in controlling the relative expression of CopA(Z) and CopA proteins from the same gene. In order to test the effect of the CopA nascent peptide on PRF, we introduced two sets of compensatory single-nucleotide indel mutations within the template copA1–312. The first set of mutations (construct copA1–312-NP1; Figures 4A and S1, available online) altered the sequence of CopA amino acid residues 4–29 located outside the ribosomal exit tunnel when the ribosome reaches the PRF site (Figures 4A and S1). With the second set of mutations, the sequence of the CopA segment 31–67, which resides in the exit tunnel of the frameshifting ribosome, was altered (construct copA1–312-NP2). The integrity of the SS and PK regions was preserved in both mutants. Because the amino acid composition of either of the mutant proteins is different from that of the WT CopA(Z), control templates were prepared (constructs CopA(Z)NP1 and CopA(Z)NP2 in Figure S1) to generate the corresponding
mutant CopA(Z) proteins to serve as electrophoretic mobility markers. In vitro –1 PRF efficiencies on the mutant copA$_{312}$-NP1 and copA$_{312}$-NP2 constructs were calculated from the intensities of the gel bands corresponding to the truncated and full-size products encoded in the 104 codons of the WT and mutant templates. Changing the sequence of the N-terminal segment of CopA (the copA$_{312}$-NP1 construct) had little effect on –1 PRF frequency in vitro, which remained comparable to the WT level (Figures 4B and 4C). However, when the inner-tunnel segment of the CopA nascent chain was altered, PRF efficiency dropped nearly 2-fold (Figures 4B and 4C).

We further tested whether the contribution of the nascent peptide to –1 PRF in copA is manifested in vivo. For this, we introduced the WT or mutant versions of the copA$_{312}$ sequence into the dual luciferase reporter plasmid pEK4 (Grentzmann et al., 1998; Kramer and Farabaugh, 2007) (Figure 4D). In this plasmid, the downstream firefly luciferase (Fluc) coding sequence is in the –1 frame relative to the preceding Renilla luciferase (Rluc) ORF. A –1 PRF is required to generate functional Fluc, whereas Rluc serves as an internal control for the 0-frame translation. Expression of the reporter carrying the WT copA$_{312}$ sequence inserted after the Rluc ORF resulted in –24% 32% –1 PRF efficiency (for the purpose of the subsequent comparison, we took the frequency of WT –1 PRF as 100%; Figure 4E). As expected, –1 PRF was abrogated when the SS was disrupted (mSS reporter in Figures 4B and 4E). Consistent with the results obtained in vitro, altering the copA codons 31–67 in pEK4 to generate the construct NP2 (Figures S1 and 4D) reduced the –1 PRF efficiency by ~40% (Figure 4E). We therefore concluded that the CopA nascent chain residing inside the ribosomal exit tunnel modulates the efficiency of –1 PRF during translation of the copA gene both in vitro and in vivo.

**Diffusible CopA(Z) Generated via –1 PRF Facilitates Cell Survival at Elevated Concentrations of Copper**

Our experimental data argue that the long-missing enigmatic copper chaperone of E. coli is generated via –1 PRF during translation of the copA gene. If high-frequency PRF within the copA gene is the result of an evolutionary selection rather than a genetic aberrance, production of the diffusible CopA(Z) protein should facilitate the maintenance of copper ion homeostasis. The recent work of Lübben and coworkers has shown that the
expression of CopA with an N-terminal truncation, i.e., a version of the transporter protein lacking the MBD1 domain, increased the sensitivity of *E. coli* to copper (Drees et al., 2015). This defect could be rescued by the ectopic expression of the CopA N-terminal protein fragment encompassing MBD1 (Drees et al., 2015), whose identity closely matched the naturally produced CopA(Z) that we observed in our experiments (Figures 1 and 2). Although the result described by Drees et al. hinted that the MBD1 of CopA could possibly function as a metal chaperone, the experimental set-up did not precisely match the true cellular scenario, where CopA(Z) is co-expressed from the *copA* gene via −1 PRF alongside the full-size CopA transporter. To test...
whether CopA(Z) contributes to copper tolerance when expressed from the intact copA gene, we compared survival of *E. coli* during exposure to copper, when production of the full-size CopA was unaffected but generation of CopA(Z) was either allowed (WT) or prevented by the SS mutations. Two synonymous codon mutations, which disrupted the SS sequence by changing it from 5’-CAC201-CCA204-AAG-GC-3’ (WT) to 5’-CAT201-CCT204-AAG-GC-3’ (mSS mutant) but did not change the encoded amino acid sequence, were introduced in the copA gene of the BW25113 chromosome. We then mixed equal numbers of WT and mSS cells and passaged the culture for several generations at a mildly toxic concentration of CuSO4 (4 mM). Changes in the mSS/WT cell ratio were monitored by PCR amplifying the copA gene from the genomic DNA of the mixed culture and analyzing the Sanger sequencing chromatogram peaks corresponding to the gene’s positions 201 and 204, those that differ in the WT and mSS cells (Figures 5A–5C). After 50 generations of growth in the presence of 4 mM CuSO4, most of the cells in the co-culture carried the WT copA, whereas the mSS cells had practically disappeared (Figures 5C and 5D). Thus, the production of diffusible CopA(Z) via −1 PRF enables cell survival in the presence of toxic concentrations of copper. This result argues that the presence of PRF signal in the copA gene is an evolutionarily selected trait. This conclusion is further supported by the observation that the mRNA structural elements promoting −1 PRF (SS and a predicted PK) are present at the edge of the MBD1-encoding copA segment in a wide range of bacterial species (Figures S2 and S3), indicating that co-expression of the copper ion transporter and its chaperone from the same gene could be advantageous for many bacterial species.

**A −1 PRF Signature Could Be Identified in the Human Homolog of the Bacterial copA Gene**

The transporter protein ATP7B in human cells is homologous to the bacterial CopA and, similar to CopA, catalyzes the efflux of copper ions (Gupta and Lutsenko, 2012). ATP7B carries six MBDs that are involved in copper ion trafficking (Barry et al., 2010; Cater et al., 2004). Interestingly, the heptanucleotide 5’-CCCAAAAG-3’, identical to the SS of *E. coli* copA, is present between the MBD2- and MBD3-coding segments of ATP7B (Figure 6A). Furthermore, the SS in the human gene is immediately followed by an mRNA sequence predicted to form a stable I-type PK (Theis et al., 2008) (ΔG = −27 kcal/mole) (Figure 6B). The arrangement of these elements, which are known to stimulate recoding, suggests that −1 PRF may take place during translation of the ATP7B mRNA. Because a stop codon is present in the −1 frame at a short distance after the −1 PRF signal (Figures 6A and 6B), the putative recoding event could result in production of a truncated protein composed of MBD1 and MBD2 of ATP7B.

---

**Figure 5. Production of the Diffusible Metal Chaperone CopA(Z) Helps Survival during Copper Stress**

Co-growth competition of the WT and mSS mutant *E. coli* cells. Production of CopA(Z) in the mSS mutants was prevented by introduction of two synonymous mutations that disrupted the SS in the chromosomal copA gene. Equal numbers of WT and mutant cells were mixed and passaged in liquid culture in the absence or presence of toxic (4 mM) concentration of CuSO4. Sequencing chromatograms of the copA segment PCR amplified from the genomic DNA isolated from the co-cultures were used to assess the ratio of WT and mSS cells (A) at the onset of the experiment, (B) after 30 generations, or (C) after 50 generations of growth in the presence of CuSO4. The WT copA sequence contains C at position 201 and A at position 204 within the 5’-CCCAAAAG-3’ SS; these residues were mutated to T in the mSS mutant. The arrows indicate the sequencing chromatogram peaks corresponding to these nucleotides.

(D) The ratios between WT and mSS cells in the co-culture were computed from the height of the chromatogram peaks. Error bars show SD from the mean in three independent experiments.
To assess whether the combination of SS and predicted PK of ATP7B could promote −1 PRF, the sequence encompassing these two elements was introduced into a eukaryotic dual luciferase reporter construct (Grentzmann et al., 1998) (Figure 6C) and tested in HEK293T cells. Measuring the relative activities of Rluc and Fluc revealed that the putative recoding elements from the ATP7B gene promoted −1 PRF with 12% efficiency (Figure 6D), a level comparable to that mediated by the well-characterized HIV −1 PRF signal (Jacks et al., 1988b). When a termination codon was introduced prior to the ATP7B recoding elements (PTC in Figure 6C) or when a stop codon was inserted at the beginning of the −1 frame Fluc ORF (OOF in Figure 6C), the Fluc expression was abrogated, ensuring that expression of Fluc is indeed driven by PRF that takes place within the ATP7B-derived segment of the reporter. A similar arrangement of the SS followed by a putative downstream PK and a stop codon in −1 frame is found in ATP7B homologs of higher primates and some other mammals (Figure S4), an observation that opens the possibility that in these organisms, similar to bacteria, a −1 PRF-based recoding may be involved in generating two ATP7B-encoded proteins with related but distinct functions in copper management.

**DISCUSSION**

While several cases of PRF are known in bacteria (Atkins et al., 2016), there is essentially only one well-substantiated example where −1 PRF leads to the generation of two functional proteins from one gene (dnaX) (Blinkowa and Walker, 1990). Our finding that −1 PRF in the E. coli gene copA directs synthesis of two functional proteins illuminates a possible broader penetrance of this distinctive mechanism of protein coding and gene regulation. Several lines of evidence support the view that production of the short CopA(Z) polypeptide along with the full-size CopA from the same gene is a result of evolutionary selection for improved cell fitness, rather than a spontaneous non-consequential genetic aberrance. First, the location of the −1 PRF site is ideal for generating a protein nearly precisely corresponding to the functional MBD1 domain, which, as it has been shown, retains its metal ion binding properties (Drees et al., 2015). Second, if spontaneous appearance of a generic heptameric SS (XXXXYYZ) could be a relatively frequent scenario, its co-occurrence with the mRNA structure and the specific sequence of the CopA nascent chain contribute to the highly efficient PRF and hence stimulate production of CopA(Z). Third, the conservation of the SS in the copA genes of a range of bacterial species, together with the presence of a characteristic downstream mRNA structure at a proper distance from the SS to efficiently promote −1 PRF, argues that −1 PRF in copA, both the mRNA structure and the specific sequence of the CopA nascent chain contribute to the highly efficient PRF and hence stimulate production of CopA(Z).
Although exploring the direct function of CopA(Z) in the maintenance of copper homeostasis was beyond the scope of our study, previous reports strongly argue that this protein plays the role of the “missing” copper chaperone in *E. coli*. A recent study has demonstrated that ectopically expressed MBDB1 of CopA can bind copper ions with high affinity and transfer them to the transporter—both characteristics of a metal ion chaperone (Drees et al., 2015). However, because there was no clear evidence that the MBDB1 could be naturally produced as an independent protein in *E. coli*, it remained unknown how MBDB1 could efficiently scavenge copper ions from the cytoplasm while remaining an integral part of the membrane-embedded transporter. Our finding that CopA MBDB1 in fact exists in two forms, independent functioning in the same biochemical pathway. This scenario is distinct from –1 PRF in dnaA (Tsuschilahishi and Kornberg, 1990) and possibly cssS2 (Chajjarasphong et al., 2016), where PRF produces two polypeptides of a multi-subunit complex that function as a single enzyme.

One of the intriguing questions about copA expression is whether the production of CopA(Z) via –1 PRF occurs with invariable frequency in every growth condition or if, alternatively, the efficiency of PRF and the ratio of CopA(Z) to CopA are subject to regulation. It is conceivable that the formation and/or stability of the protein complex that function as a single enzyme.

**Experimental Procedures**

**Strains and Plasmids**

The ASKA collection plasmid that we named pCopA (Table S2) carries the *E. coli* copA gene encoding the N-terminally His6-tagged CopA protein (Kita-gawa et al., 2009). The pCopA plasmid was introduced into the Keio collection *E. coli* strain JW0473-3 lacking chromosomal copA (copA::kan) (Baba et al., 2006). The point mutations that generated the PRF-deficient copA variant in the pCopA-mSS plasmid were engineered using the QuikChange Lightning Multi-Site-Directed Mutagenesis kit (Agilent Technologies) and primer #1 (all primer sequences are listed in Table S1). The PRF-deficient variant of the chromosomal copA (mSS), in which the SS sequence 5'-CACGCT-3' was mutated to 5'-CACTGG-3', was engineered in the BW25113 strain by homologous recombination using the pKOV plasmid (Link et al., 1997). Two PCR products were generated: primers #11 and #12 were used to generate the first one, using plasmid pCopA-mSS as the template, and primers #13 and #14 were used to prepare the second product using genomic DNA of *E. coli* BW25113 as the template. Both PCR products, which contain overlapping sequences, were introduced by Gibson assembly into the pKOV plasmid cut with NotI and BamHI restriction enzymes. The resulting pKOV-mSS plasmid, carrying the sequence starting 800 nucleotides upstream and encoding 1,207 nucleotides downstream of the stop codon of the wild-type copA gene, was transformed into BW25113 cells by electroporation. The transformants were plated onto Luria-Bertani (LB)/agar supplemented with 30 μg/mL chloramphenicol, and plates were incubated overnight at 42°C to induce integration of the plasmid into the chromosome. Several colonies were resuspended in 1 mL LB and dilutions were plated on LB agar plates supplemented with 5% (w/v) sucrose to induce resolution of the vector. After overnight incubation of the plates at 37°C, the loss of the vector plasmids was confirmed by replica plating, which showed the sensitivity of the cells to chloramphenicol. The resulting mSS was used in the competition experiments. The bacterial dual luciferase reporter plasmids were prepared on the basis of the pEK4 plasmid (Greentzmann et al., 1998; Kramer and Farabaugh, 2007). To engineer the reporter plasmids with the copA –1 PRF element, three PCR fragments (DLR1, DLR2, and DLR3) were used to generate using plasmid pCopA as the template. The PCR products DLR1 (primers #15 and #16), DLR2 (primers #17 and #18), and DLR3 (primers #19 and #20), which contain overlapping sequences, were introduced by Gibson assembly into pEK4 cut with SacI/SalI to generate pEK4-copA3–294-WT. The resulting plasmid contained codons 2–98 of the WT copA inserted in frame after the first (Fluc) gene in pEK4. Similarly, pEK4-copA3–294-mSS was assembled from the same DLR1, DLR3 fragments, and the DLR2-mSS PCR product, which was amplified from the pCopA-mSS plasmid using primers #21 and #22. Plasmid pEK4-copA3–294-NP2 was assembled from the DLR1-NP2 fragment combined with DLR2-NP2 amplified from the synthetic copA::kan fragment with gBlock (Table S1) using primers #15 and #16 or primers #23 and #24, respectively. In all three reporter plasmids, the copA sequence carried the deletion of nucleotide T210, located downstream of the entire –1 PRF signal, introduced to eliminate an in-frame stop codon, and allowed for translation of the Fluc ORF upon –1 PRF. Control plasmid pEK4-copA3–294-C, which was used for normalization of the relative levels of Fluc and Fluc expression in the absence of –1 PRF, contained Fluc and Fluc in the 0 frame, whereas the copA –1 PRF was disabled by the mSS mutations. This plasmid was obtained by introducing the DLR-C fragment, amplified from pCopA-mSS using primers #25 and #26, into the SalI/SacI cut pEK4 plasmid. The 0-frame version of pEK4-copA3–294-NP2-C control plasmid was constructed using fragments DLR1-NP2, DLR2-NP2-C (primers #27 and #28), and DLR3-NP2-C (primers #19 and #20), which were introduced into SalI/SalI cut pEK4 plasmid by Gibson assembly.

The Gibson assembly reactions were transformed into *E. coli* JM109 cells. Plasmids with the desired sequences were then introduced into BW25113 host cells in order to carry out the dual luciferase reporter assays.

**Overexpression of copA and Purification of CopA(Z)**

The *E. coli* ΔcopA cells (JW0473-3) carrying the pCopA plasmid were grown at 37°C in LB medium supplemented with kanamycin (50 μg/mL) and chloramphenicol (30 μg/mL). Upon reaching an *A*~660~ of 0.6, cultures were induced with 0.1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) and incubation
continued for 3 hr. Cells were collected by centrifugation at 4 °C, resuspended in buffer (20 mM Tris [pH 8.0], 300 mM NaCl, 30 mM imidazole) supplemented with 1× Halot Protease Inhibitor (Life Technologies), and lysed using a French Press at 16,000 psi. The lysate was centrifuged at 15,000 rpm (rotor JA-25.50) for 1 hr at 4 °C and then filtered through a 0.2 μm cellulose acetate filter. The lysate was passed through a HisTrap HP column (GE) using an AKTA/Unicorn FPLD system (GE). Bound protein was eluted using a linear 30–300 mM imidazole gradient. Fractions in which the CopA2 protein was detected, as assessed by SDS gel electrophoresis (4%–20% TGV [Bio-Rad]), were pooled together and subjected to an additional round of purification on the HisTrap HP column, this time using a stepwise 30–300 mM imidazole elution. The purified protein was dialyzed using Spectra/Por membrane (MWCO 3,500) against 10 mM ammonium acetate (pH 8.0) for 3 hr at 4 °C. The recovered sample was concentrated in Amicon concentrator (MWCO 3,000). Concentration of the isolated protein was measured using the bicinchoninic acid assay (BCA) kit (Thermo Scientific) and its purity was confirmed by gel electrophoresis and silver staining (Chevallet et al., 2006).

Mass Spectrometry Analysis
Purified protein (~100 μg) was precipitated with cold acetone to remove salts and resuspended in 4% SDS prior to the addition of GELFREE loading buffer. Separation was performed as described (Tran and Doucette, 2009) using a 10% gel on a commercial GELFREE 8100 system (Expedeon). The fraction containing CopA2 was isolated and SDS was removed using the methanol-chloroform-water method (Wessel and Flugge, 1984). After SDS removal, proteins were resuspended in 40 μL sodium acetate (94.9% H2O, 5% acetic acid, and 0.1% formic acid) and 5 μL were injected onto a trap column (150 μM ID × 3 cm) coupled with a nanobore analytical column (75 μM ID × 15 cm). The trap and analytical column were packed with polymeric reverse-phase media (5 μm, 1,000 Å pore size) (PLRP-S, Phenomenex). Samples were separated using a linear gradient of solvent A and solvent B (4.9% water, 95% acetonitrile, and 0.1% formic acid) and 5 μL/min. Mass spectrometry data were obtained on an Orbitrap Elite mass spectrometer (Thermo Scientific) fitted with a custom nanospray ionization source. Intact mass spectrometry data were obtained at a resolving power of 120,000 (m/z 400). The top 2 m/z species were isolated within the Velos ion trap and fragmented using higher-energy collisional dissociation (HCD). Data were analyzed with ProSightPC against a custom CopA database.

Western Blotting
Five micrograms of total protein from lysates of uninduced or IPTG-induced BW25113 cells transformed with pCOpA or pCOpA-mSS were loaded on a 4%–20% polyacrylamide-SDS gel (TGX, Bio-Rad). Resolved proteins were transferred to a PVDF membrane (Bio-Rad) by electroblotting (Bio-Rad Trans-Blot Cell, 100 mA, 0.5 V, Dry Transfer Cell, 100 V, 0.5 mA at 20 V). The membrane was blocked with 3% BSA (Sigma-Aldrich) in TBST buffer (10 mM Tris-Cl [pH 7.4], 150 mM NaCl, and 0.1% Tween-20) and probed with 6x-His Epitope Antibody HRP conjugate (Thermo Scientific) at 1:2,000 dilution. The membranes were developed using Clarity Western ECL substrate (BioRad) and visualized using FluorChem R System (Protein Simple).

Growth Competition Experiments
Overnight cultures of WT (BW25113) and mSS cells were diluted 1:100 in fresh LB and grown to mid-log phase (A600~0.5). The cell densities in both cultures were adjusted to the same value, and equal volumes of the WT and mSS cultures were mixed. The 1:1 cell mixture (starting culture) was diluted to A600~0.5 and grown overnight in LB without or with the addition of 4 mM CuSO4. Five microliters of the overnight culture were diluted into 5 mL LB with or without 4 mM CuSO4. Cultures were diluted 1:100-fold upon reaching saturation for a total of five passages. Aliquots from every passage were used to isolate genomic DNA, and the SS region of copA was PCR amplified using primers #6 and #11 (Table S1) and the following PCR conditions: 94 °C, 2 min; followed by 34 cycles, 94 °C, 2 min; 52 °C, 30 s; 68 °C, 15 s, and final extension at 68 °C for 2 min. The PCR fragments were subjected to Sanger sequencing using primer #6. The ratio of the mSS to WT cells was estimated by quantifying and then averaging the heights of the sequencing chromatogram peaks corresponding to copA residues 201 (C for WT, T for mSS) and 20 (C for WT, T for mSS).

In Vitro Measurement of –1 PRF Efficiency
Coupled in vitro transcription-translation reactions in the E. coli lysate were performed using an S30 transcription-translation system for linear DNA (Promega). DNA templates (0.6 pmol) were PCR amplified from either E. coli BW25113 genomic DNA or the plasmid pCOpA-mSS (using primers #2 to #10 for both), or from synthetic gBlocks (Table S1). The resulting templates carrying the gene segments of interest controlled by the Tpr promoter were translated in 5 μL reactions containing 2 μCi [35S]-L-methionine (specific activity 1,175 Ci/mmol) (IMP Biomedicals). Reactions were incubated at 37 °C for 30 min and translation products were precipitated with 8 vol cold acetone. After the recovery of the pellet by centrifugation, proteins were resolved on 16.5% Tricine-SDS polyacrylamide gels (Schägger and von Jagow, 1987). The gels were dried, exposed to a phosphorimager screen, and visualized with a Typhoon scanner (GE). Protein bands corresponding to CopA2 (the ~1 PRF product) and the relevant truncated CopA reference (originated from 0-frame translation) were quantified using the ImageJ software (https://imagej.nih.gov/ij/). PRF efficiency was estimated as a ratio of the density of the CopA2 band to the combined density of CopA2 and full-protein gel bands.

Identification of –1 PRF Elements in Different Species
Bacterial orthologs of copA (E. coli) were retrieved from OrthologDB (Waterhouse et al., 2013). Coordinates of MBs of CopA were determined by HMMSearch and Pfam model PF00403 (Finn et al., 2014; Wistrand and Sonnhammer, 2005). For prediction of –1 PRF elements analogous to –1 PRF signal in the E. coli copA, the bacterial copA genes encoding two sequential MBs were selected. The gene segments separating two MBs were scanned for the presence of –1 PRF signals (~1 frame SS, nearby ~1 frame stop codon, and a downstream mRNA structure) using KnotInFrame (Theis et al., 2008). For the genes with the interdomain-coding segment shorter than 50 nt, the last 50 nt of the MBD1 coding sequence and the first 150 nt of the MBD2 coding sequence were also included in the KnotInFrame search. After removing redundancy for the strains of the same species, the predicted –1 PRF is found in 35 copA genes belonging primarily to the proteobacteria branch of the bacterial phylogenetic tree. Maximum likelihood (ML) tree for 35 copA sequences was computed (Huerta-Cepas et al., 2016), and the resultant tree (Figure S2) was decorated with phyla/class-level taxonomy information. The structure of copA mRNA was initially modeled by using iPon (Sato et al., 2011) and KnotInFrame (Theis et al., 2008). The Simulprot program (Meyer and Miklós, 2007) was used to generate the final consensus structure and the covariational mutations (Figures 3 and S3). Aligned cDNA sequences of human ATP7B and its orthologs (Figure S4) were retrieved from Ensembl (release 85) (Aken et al., 2016). The prediction of ATP7B mRNA downstream structure was based on KnotInFrame (Theis et al., 2008) by submitting downstream the 100 nt-long sequence as input.

Bacterial Dual Luciferase Reporter Assay
The E. coli BW25113 cells carrying the derivatives of the peK4 plasmid were grown at 37 °C in 5 mL LB medium supplemented with 50 μg/mL ampicillin. Upon reaching A600~0.5, cells were collected by centrifugation at 4 °C and then resuspended in 200 μL lysis buffer (1 mg/mL lysozyme, 10 mM Tris-HCl [pH 8.0], and 1 mM EDTA). The lysates were prepared by freezing-thawing as previously described (de Wet et al., 1985). Five microliters of the extracts were used for Fluc and Rluc activities using the Dual-Luciferase Reporter Assay System (Promega). Luminescence was measured in 96-well plates in Top Count NXT (Perkin Elmer). The PRF efficiency (pPRF) was calculated using the equation pPRF = Fluc/control (Rluc/control) (Grentzmann et al., 1998), where test plasmids were peK4-copycopA−292, peK4-copycopA−241−mSS, and peK4-copycopA−159−N2P2, and the control plasmids were peK4-copycopA−292−C and peK4-copycopA−159−N2P2−C (Table S2). Experimental replicates were performed using lysates prepared from bacterial cultures grown from three independent colonies.

Dual Luciferase Assay for –1 PRF in Human Cells
The ATP7B-derived sequence indicated in Figure 6B, acquired as a gBlock (Integrated DNA Technologies) (Table S1, #2–34), was cloned into Sall/SacI-cut pLuci plasmid so that a –1 PRF event would direct ribosomes elongating through the upstream RUC ORF into the downstream Fluc ORF. Two additional

ACKNOWLEDGMENTS

We thank Pavel Baranov and John Atkins for helpful advice; Kurt Fredrick and Hyunwoo Lee for providing the pEK4 and pKOV plasmids, respectively; Andrew Jin for help with some experiments; and Dimple Modi, Nikolay Alekshin, and Hao Lei for advice with some experimental procedures. This work was supported by the grants MCB 1244455 and MCB 1615851 (both to A.S.M. and N.V.-L.) from the National Science Foundation and R01 GM117177 and R01 HL119439-01A1 (to J.D.D.) from the NIH. Proteomics analysis was performed at the Northwestern Proteomics Core Facility, supported by the grants PCI CCSG P30 CA060553 and P41 GM108569 from the NIH.

Received: October 31, 2016
Revised: November 23, 2016
Accepted: December 13, 2016
Published: January 19, 2017

REFERENCES

Aken, B.L., Ayling, S., Barrett, D., Clarke, L., Curwen, V., Fairley, S., Fernandez, J.E.J., Y.K., and J.B. performed the experiments. J.D.D., J.E.J., and Y.K. designed the experiments, and wrote the paper; S.M., N.V.-L., and A.S.M. conceived the project, designed the experiments, analyzed the data, and wrote the paper; J.D.D., J.E.J., and Y.K. designed some experiments and analyzed the data; and S.M., D.K., S.K., T.M., P.T., J.E.J., Y.K., and J.B. performed the experiments.


Supplemental Information

Programmed Ribosomal Frameshifting Generates a Copper Transporter and a Copper Chaperone from the Same Gene

Sezen Meydan, Dorota Klepacki, Subbulakshmi Karthikeyan, Tõnu Margus, Paul Thomas, John E. Jones, Yousuf Khan, Joseph Briggs, Jonathan D. Dinman, Nora Vázquez-Laslop, and Alexander S. Mankin
Figure S1. Generation of mutant nascent peptide templates, Related to Figure 4.

The nucleotide sequences of the DNA templates used for in vitro translation and the amino acid sequences of the proteins encoded in 0 frame. The $\text{copA}_{1-312}$ template carries the wt $\text{copA}$ sequence except for insertion of two stop codons (UAA) after codon 104. The wt amino acid sequence (residues in orange) was changed in the $\text{copA}_{1-312}$-NP1 (amino acids shown in dark blue) and $\text{copA}_{1-312}$-NP2 (residues in light blue) templates by introducing compensatory indel mutations (indicated by the arrows). Additional mutations introduced to avoid the appearance of premature stop codons in the mutant constructs are indicated with bold and underlined characters. Notice that none of the changes in the templates disrupted the integrity of the CCCAAAG slippery sequence (boxed). The cartoons, which represent ribosomes positioned at the $\text{copA}$ slippery sequence, illustrate the portion of the CopA nascent chain that has been altered. Also shown are the sequences of the marker templates $\text{copA}(Z)_{m}$, $\text{copA}(Z)_{m}$-NP1, and $\text{copA}(Z)_{m}$-NP2, in which the relevant CopA(Z) polypeptides, used as markers for the gel electrophoresis, are encoded in the 0 frame.
Figure S2. The presence of the slippery sequence at the end of the MBD1-coding segment in the copA gene of a variety of bacterial species, Related to Figure 2

Phylogenetic tree of bacterial species whose copA genes contain an SS heptamer followed by a nearby -1 frame stop codon. The colored vertical bars indicate the phylum of each bacterial species. The criteria applied for the selection of the species included in the tree is described in STAR methods.
Figure S3. Conservation of the predicted mRNA pseudoknot that facilitates -1 PRF in the copA gene among different bacterial species, Related to Figure 3

The stems S1 and S2 of the predicted pseudoknot downstream from the slippery sequence (boxed) in the copA gene of different bacteria are indicated by green and blue arches, respectively. Note that the covariation of residues within the S1 and S2 sequences is consistent with the folding of mRNA into the pseudoknot structure.
Figure S4. Slippery sequence is found in the ATP7B gene homologs of several eukaryotic species, Related to Figure 6

Sequence alignment of the ATP7B gene sequences in the vicinity of the slippery sequence (shaded red). The downstream stop codons in the -1 frame are boxed.
Table S1: Oligonucleotides used in this study, Related to the STAR Methods

<table>
<thead>
<tr>
<th>Sequence (5' to 3')</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ACGCATCTGTAAGCCTACCTAAGGCTAAACCCTCGG</td>
<td>Site directed mutagenesis for generation of pcopA-mSS</td>
</tr>
<tr>
<td>2 TTATTAACCTCTGGATGGCTTACAGATGCGTCAATAC</td>
<td>Reverse primer for CopA(Z)M-NP1 and CopA(Z)M-NP2</td>
</tr>
<tr>
<td>3 TTATTAACCTCTGGATGGCATACAGATGCGTCAATAC</td>
<td>Reverse primer for CopA(Z)M-NP2</td>
</tr>
<tr>
<td>4 CATGATTGTGACAAATTAATCATCGGCTGTATAATGTGTGGACTTAAACCGCTGGGA</td>
<td>Forward primer for CopA(Z)M-NP1</td>
</tr>
<tr>
<td>5 CATGATTGTGACAAATTAATCATCGGCTGTATAATGTGTGGACTTAAACCGCTGGGA</td>
<td>Forward primer for all templates used in vitro</td>
</tr>
<tr>
<td>6 TTA TTA CAG CAA CTG CTG GCT G</td>
<td>Reverse primer for copA1-312</td>
</tr>
<tr>
<td>7 TTA TTA GGT CGC TGC CGG</td>
<td>Reverse primer for copA1-282</td>
</tr>
<tr>
<td>8 TTA TTA AGA AAC CGC TGT CAG TG</td>
<td>Reverse primer for copA1-261</td>
</tr>
<tr>
<td>9 TTA TTA TTC CGA CGG GAT TGA TGA</td>
<td>Reverse primer for copA1-243</td>
</tr>
<tr>
<td>10 TTA TTA CGC CAG CGG TTT AGC</td>
<td>Reverse primer for copA1-222</td>
</tr>
<tr>
<td>11 TCACAAACTATCGAccCTGACC</td>
<td>Forward primer for Fragment-1</td>
</tr>
<tr>
<td>12 GGGCGGACACGCCCATTCCCGCCCTCGACGACTCTAGAGGCACCCTCACC</td>
<td>Reverse primer for Fragment-1</td>
</tr>
<tr>
<td>13 ATGCTATTGCTGTGTTATTCGCTGACCAGGGATCCTCGGCCTATCCGGCCTGACCACAG</td>
<td>Forward primer for Fragment-2</td>
</tr>
<tr>
<td>14 TTACAGATCTGACCTAACCCTG</td>
<td>Reverse primer for Fragment-2</td>
</tr>
<tr>
<td>15 TTAACGGAGTTCTCAAAATGAACAAATGTGACGTACAAAATCGACCTGACC</td>
<td>Forward primer for DLR1 and DLR1-NP2</td>
</tr>
<tr>
<td>16 GATGGCGTACATTCAGGCTGACC</td>
<td>Reverse primer for DLR1 and DLR1-NP2</td>
</tr>
<tr>
<td>17 ACCATCAAACAAGCGGGTATTGACGACATCTGTAAGCCACTCCAAACCCTAAAAGCCA</td>
<td>Forward primer for DLR2</td>
</tr>
<tr>
<td>18 GCTTCCGACGGGATGATGACCTCCTCGGACGCTATCCGCTTTGGGTGGCTTACAGATGCG</td>
<td>Reverse primer for DLR2</td>
</tr>
<tr>
<td>19 GCTGCCGAGCTCATCAATCC</td>
<td>Forward primer for DLR3</td>
</tr>
<tr>
<td>20 CCTTTCTTATGTTTTTGGCGCTTTCCATGAGCTCATCGTCATCGTCGCGGCTGTC</td>
<td>Reverse primer for DLR3</td>
</tr>
<tr>
<td>21 ACCATCAAACAAGCGGGTATTGACGACATCTGTAAGCCACTCCAAACCCTAAAAGCCA</td>
<td>Forward primer for DLR2-mSS</td>
</tr>
<tr>
<td>22 GCTTCCGACGGGATGATGACCTCCTCGGACGCTATCCGCTTTGGGTGGCTTACAGATGCG</td>
<td>Reverse primer for DLR2-mSS</td>
</tr>
<tr>
<td>23 CCATCAAACAAGCGGGTATTGACGACATCTGTAAGCCACTCCAAACCCTAAAAGCCA</td>
<td>Forward primer for DLR2-NP2</td>
</tr>
<tr>
<td>24 GCTTCCGACGGGATGATGACCTCCTCGGACGCTATCCGCTTTGGGTGGCTTACAGATGCG</td>
<td>Reverse primer for DLR2-NP2</td>
</tr>
<tr>
<td>25 AATGTCGAGCTCAAAACTATCGACCTGACC</td>
<td>Forward primer for DLR-C</td>
</tr>
<tr>
<td>26 CATGACCTCATCGTCATCGCGGTC</td>
<td>Reverse primer for DLR-C</td>
</tr>
<tr>
<td>27 CATCAAACAAGCGGGTATTGACGACATCTGTAAGCCACTCCAAACCCTAAAAGCCA</td>
<td>Forward primer for DLR2-NP2-C</td>
</tr>
<tr>
<td>28 CCTCCGACGGGATGATGACCTCCTCGGACGCTATCCGCTTTGGGTGGCTTACAGATGCG</td>
<td>Reverse primer for DLR2-NP2-C</td>
</tr>
<tr>
<td>29 CCTTTCTTATGTTTTTGGCGCTTTCCATGAGCTCATCGTCATCGCGGCTGTC</td>
<td>Reverse primer for DLR3-NP2-C</td>
</tr>
<tr>
<td>30</td>
<td>CATGGATTCTTGACAAATATCATGGCTCGTATAATGTGTGGACTTAA GTATAAGGAGGAAAAATTGTCACAACATATCGACCAGACCCTGGACG GCCTGTCTTGCGCTACCTGCTTTAAAGCCGAGAAAGGATCTGGGACAGCAGGGTTATGACGCAATATCGACCTGACCCTGGACGGCCTGTCCTGCGGTCACTGCGTTAAACGCGTGAAAGAAAGTCTTGAA AGCGTCCGGATGGTTGAGCAGGCGGATGTGTCTATCACTGAAGCGCA GTTACCCTGACGCAGCAGCAGCAGCAGTTGCTGTAATAA</td>
</tr>
<tr>
<td>31</td>
<td>CATGGATTCTTGACAAATATCATGGCTCGTATAATGTGTGGACTTAA GTATAAGGAGGAAAAATTGTCACAACATATCGACCAGACCCTGGACG GCCTGTCTTGCGCTACCTGCTTTAAAGCCGAGAAAGGATCTGGGACAGCAGGGTTATGACGCAATATCGACCTGACCCTGGACGGCCTGTCCTGCGGTCACTGCGTTAAACGCGTGAAAGAAAGTCTTGAA AGCGTCCGGATGGTTGAGCAGGCGGATGTGTCTATCACTGAAGCGCA GTTACCCTGACGCAGCAGCAGCAGCAGTTGCTGTAATAA</td>
</tr>
<tr>
<td>32</td>
<td>TTCGTTGAGCGAGTTCTTCTAAAATGAACAAATGTCGACCCCAAGGAC CTGGATTTATTGCTTCTGGCAACGAGTTCTTCATCTTGGGACCAAGGACGCATGTGGTCACCCTCCAACTGAGAATAGATGGAATGCACGGATCCCCCGGGGAGCTCATGGAAGACTAAGCCAAAAACATAAAGAAAAGCCCGGCGCCATTCTATCCTCTAGAGGATGGAACCGCTGG</td>
</tr>
<tr>
<td>33</td>
<td>TTCGTTGAGCGAGTTCTTCTAAAATGAACAAATGTCGACCCCAAGGAC CTGGATTTATTGCTTCTGGCAACGAGTTCTTCATCTTGGGACCAAGGACGCATGTGGTCACCCTCCAACTGAGAATAGATGGAATGCACGGATCCCCCGGGGAGCTCATGGAAGACTAAGCCAAAAACATAAAGAAAAGCCCGGCGCCATTCTATCCTCTAGAGGATGGAACCGCTGG</td>
</tr>
<tr>
<td>34</td>
<td>GTTGAGGAGTTCTTCTAAAATGAACAAATGTCGACCCCAAGGACCTTTATCTGCTTGCAACGAGGTTCTTCTGAGACCTTGGGACCAAGGACGCATGTGGTCACCCTCCAACTGAGAATAGATGGAATGCACGGATCCCCCGGGGAGCTCATGGAAGACTAAGCCAAAAACATAAAGAAAAGCCCGGCGCCATTCTATCCTCTAGAGGATGGAACCGCTGG</td>
</tr>
</tbody>
</table>
### Table S2. Reagents and Resources

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antibodies</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouse monoclonal anti 6x-His tag</td>
<td>Thermo Scientific</td>
<td>Cat# MA1-21315-HRP</td>
</tr>
<tr>
<td><strong>Chemicals, Peptides, and Recombinant Proteins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Halt Protease Inhibitor Cocktail (100X)</td>
<td>Thermo Scientific</td>
<td>Cat# 78430</td>
</tr>
<tr>
<td>L-Methionine, [35S], In vitro Translation Grade</td>
<td>MP Biomedicals</td>
<td>Cat# IC51001H.5</td>
</tr>
<tr>
<td><strong>Critical Commercial Assays</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>QuickChange Lightning Multi Site-Directed Mutagenesis Kit</td>
<td>Agilent Technologies</td>
<td>Cat# 210518</td>
</tr>
<tr>
<td>Pierce BCA Protein Assay Kit</td>
<td>Thermo Scientific</td>
<td>Cat# 23225</td>
</tr>
<tr>
<td>Clarity Western ECL substrate</td>
<td>Bio-Rad</td>
<td>Cat# 170-5060</td>
</tr>
<tr>
<td>E. coli S30 Extract System for Linear Templates</td>
<td>Promega</td>
<td>Cat# L1030</td>
</tr>
<tr>
<td>Dual-Luciferase Reporter Assay System</td>
<td>Promega</td>
<td>Cat# E1910</td>
</tr>
<tr>
<td><strong>Experimental Models: Cell Lines</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human: 293T cells</td>
<td>ATCC</td>
<td>ATCC: 3216</td>
</tr>
<tr>
<td><strong>Experimental Models: Organisms/Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Escherichia coli</em> strain JW0473-3: F-, Δ(araD-araB)567, ΔlacZ4787::rrnB-3, ΔcopA767::kan, λ-, rph-1, Δ(traD-rrhB)568, hsdR514</td>
<td>(Baba et al., 2006)</td>
<td>CGSC# 8625</td>
</tr>
<tr>
<td><em>Escherichia coli</em> strain BW25113: F-, Δ(araD-araB)567, ΔlacZ4787::rrnB-3, λ-, rph-1, Δ(traD-rrhB)568, hsdR514</td>
<td>(Datsenko and Wanner, 2000)</td>
<td>CGSC# 7636</td>
</tr>
<tr>
<td><em>Escherichia coli</em> strain JM109: F' (traD36, proAB+ lacIq, Δ(lacZ)M15) emdA1 recA1 hsdR17(rk-, mcrA) supE44 λ- gyrA96 relA1 Δ(lac- proAB) thi-1</td>
<td>Promega</td>
<td>Cat# P9751</td>
</tr>
<tr>
<td><em>Escherichia coli</em> strain BW25113-mSS: F-, Δ(araD-araB)567, ΔlacZ4787::rrnB-3, copA::201C&gt;T, 204C&gt;T, λ-, rph-1, Δ(traD-rrhB)568, hsdR514</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Recombinant DNA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pCopA</td>
<td>(Kitagawa et al., 2005)</td>
<td>N/A</td>
</tr>
<tr>
<td>pCopA-mSS</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>pKOV</td>
<td>(Link et al., 1997)</td>
<td>N/A</td>
</tr>
<tr>
<td>pKOV-mSS</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>pEK4</td>
<td>(Kramer and Farabaugh, 2007)</td>
<td>N/A</td>
</tr>
<tr>
<td>pEK4-copA294-wt</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>pEK4-copA294-mSS</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>pEK4-copA294-NP2</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>pEK4-copA294-C</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>pEK4-copA294-NP2-C</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>p2luci</td>
<td>(Grentzmann et al., 1998)</td>
<td>N/A</td>
</tr>
<tr>
<td>p2luci-HIV</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>p2luci-ATP7B</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>p2luci-ATP7B-PTC</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>p2luci-ATP7B-OOF</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Sequence-Based Reagents</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>See Table S1 for the primer list used in this study</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Software and Algorithms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ImageJ</td>
<td>(Schneider et al., 2012)</td>
<td><a href="http://rsbweb.nih.gov/ij/">http://rsbweb.nih.gov/ij/</a></td>
</tr>
<tr>
<td>ProSightPC Software</td>
<td>Thermo Scientific</td>
<td>Cat# PROSIGHTPC10</td>
</tr>
<tr>
<td>Software</td>
<td>Authors</td>
<td>Date</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>KnotInFrame</td>
<td>Theis et al., 2008</td>
<td></td>
</tr>
<tr>
<td>IPknot</td>
<td>Sato et al., 2011</td>
<td></td>
</tr>
</tbody>
</table>

**Supplemental References**
