Synonymous codon substitutions perturb cotranslational protein folding in vivo and impair cell fitness

Ian M. Walsh[a], Micayla A. Bowman[a], Iker F. Soto Santarriaga[a], Anabel Rodriguez[a], and Patricia L. Clark[a,b,1]

[a]Department of Chemistry & Biochemistry, University of Notre Dame, Notre Dame, IN 46556; and [b]Department of Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN 46556

Edited by Susan Marqusee, University of California, Berkeley, CA, and approved January 3, 2020 (received for review April 25, 2019)

In the cell, proteins are synthesized from N to C terminus and begin to fold during translation. Cotranslational folding mechanisms are therefore linked to elongation rate, which varies as a function of synonymous codon usage. However, synonymous codon substitutions can affect many distinct cellular processes, which has complicated attempts to deconvolve the extent to which synonymous codon usage can promote or frustrate proper protein folding in vivo. Although previous studies have shown that some synonymous changes can lead to different final structures, other substitutions will likely be more subtle, perturbing predominantly the protein folding pathway without radically altering the final structure. Here we show that synonymous codon substitutions encoding a single essential enzyme lead to dramatically slower cell growth. These mutations do not prevent active enzyme formation; instead, they predominantly alter the protein folding mechanism, leading to enhanced degradation in vivo. These results support a model in which synonymous codon substitutions can impair cell fitness by significantly perturbing cotranslational protein folding mechanisms, despite the chaperoning provided by the cellular protein homeostasis network.

elongation rate | translation | ribosome | cotranslational folding | protein design

Synonymous codon substitutions alter the mRNA coding sequence but preserve the encoded amino acid sequence. For this reason, these substitutions were historically considered to be phenotypically silent and often disregarded in studies of human genetic variation (1, 2). In recent years, however, it has become clear that synonymous substitutions can significantly alter protein function in vivo through a wide variety of mechanisms that can change protein level (3–5), translational accuracy (6, 7), secretion efficiency (8, 9), the final folded structure (1, 10–12), and posttranslational modifications (13). The full range of synonymous codon effects on protein production is, however, still emerging, and much remains to be learned regarding the precise mechanisms that regulate these effects.

One effect of synonymous codon substitutions long proposed but with scant evidence to support its significance in vivo is perturbations to cotranslational folding mechanisms. In general, rare synonymous codons tend to be translated more slowly than their common counterparts (14–17). Moreover, rare synonymous codons tend to appear in clusters, creating broader patterns of codon usage (18), many of which are conserved through evolution (19–21). The folding rates of many protein secondary and tertiary structural elements are similar to their rate of synthesis (22, 23), lending conceptual support to the hypothesis that even subtle changes in elongation rate could alter folding mechanisms (24). In theory, reducing the rate of translation elongation by synonymous common-to-rare codon substitutions could provide the N-terminal portion of a nascent protein with more time to adopt a stable tertiary structure before C-terminal portions are synthesized and emerge from the ribosome exit tunnel (25–27). Depending on the specific native structure of the encoded protein, such extra time could be either advantageous or detrimental to efficient folding (28). However, cells contain an extensive network of molecular chaperones to facilitate the folding of challenging protein structures, including several that associate with nascent polypeptide chains during translation (29–33). Thus, it remains unclear whether a synonymous codon-derived alteration to elongation rate and cotranslational folding mechanism could be sufficiently perturbative to rise above the buffering provided by the cellular chaperone network.

Here we show that synonymous codon changes in the coding sequence of an enzyme essential for *Escherichia coli* growth can have a dramatic effect on cell growth. We tested a variety of mechanistic origins for this growth defect, including changes to the folded protein structure, expression level, enzymatic activity, mRNA abundance, and/or activation of a cell stress response. Our results are consistent with synonymous substitutions altering the pattern of translation elongation, which alters the cotranslational folding mechanism and leads to the formation of a folded, active structure that is more susceptible to degradation. These results demonstrate that changes to synonymous codon usage can significantly affect protein folding in vivo, rising above the chaperoning capacity provided by the cellular protein homeostasis network. Synonymous codon usage may therefore have broad implications for effective protein design and the interpretation of disease-associated synonymous mutations.

Significance

Many proteins that are incapable of refolding in vitro nevertheless fold efficiently to their native state in the cell. This suggests that more information than the amino acid sequence is required to properly fold these proteins. Here we show that synonymous mRNA mutations can alter a protein folding mechanism in vivo, leading to changes in cellular fitness. This work demonstrates that synonymous codon selection can play an important role in supporting efficient protein production in vivo.

Author contributions: I.M.W., A.R., and P.L.C. designed research; I.M.W., M.A.B., I.F.S.S., and A.R. performed research; I.M.W., I.F.S.S., A.R., and P.L.C. analyzed data; and I.M.W. and P.L.C. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

Data deposition: All data discussed in the paper is available at GitHub, https://github.com/plclark1/SynonymousCodons/tree/master/Synonymous_codon_substitutions_perturb_cotranslational_protein_folding_in_vivo_and_impair_cell_fitness.

[1]To whom correspondence may be addressed. Email: pclark1@nd.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1907126117/-/DCSupplemental.

First published February 3, 2020.

Downloaded at Lawerence Livermore National on March 5, 2020

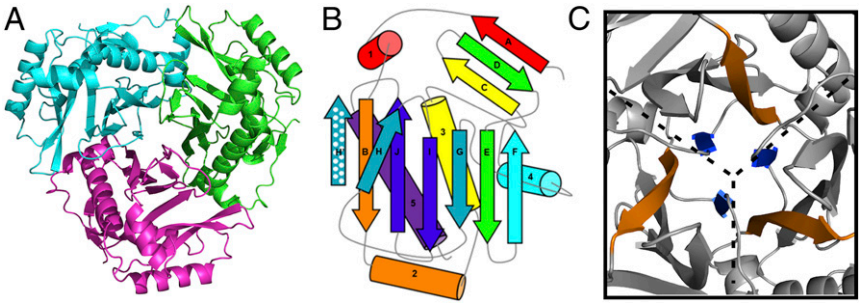

Fig. 1. Chloramphenicol acetyltransferase (CAT) has a complex tertiary and quaternary structure. (*A*) Ribbon diagram depicting the native homotrimeric structure (Protein Data Bank ID: 3CLA) (34). (*B*) Schematic representation of the complex topology of the CAT monomer structure. Secondary structure elements are shown in rainbow order. Polka dots indicate the H β-strand in the central β-sheet contributed from an adjacent monomer. (*C*) Close up of the trimer interface, with the B and H β-strands in the central β-sheets colored as in *B*. Dashed lines indicate approximate monomer boundaries.

Results

Synonymous Codon Substitutions Impair *E. coli* Growth Rate. To develop a system to test connections among synonymous codon usage, cotranslational folding, and cell fitness, we used chloramphenicol acetyltransferase (CAT), a water-soluble, homotrimeric *E. coli* enzyme with a complex tertiary structure (34) (Fig. 1). A landmark early study showed that synonymous codon substitutions near the middle of the coding sequence (Fig. 2*A* and *SI Appendix*, Fig. S1) led to lower specific activity for CAT synthesized by in vitro translation (11). CAT is essential for *E. coli* growth in the presence of chloramphenicol (cam) (35), which enabled us to use growth rate with cam as a convenient fitness assay. Furthermore, because CAT is not part of an operon or regulatory network, we hypothesized that it would be unlikely for feedback regulation of other genes to mask the effects of CAT synonymous codon changes on enzyme function (36). Crucially, although CAT cannot be refolded to its native structure after dilution from chemical denaturants, the native structure is resistant to unfolding up to 80 °C (*SI Appendix*, Fig. S2), suggesting that folding intermediates populated during and after protein synthesis are crucial for efficient folding, as once the native structure has been attained, it is not likely to populate the unfolded state over a typical bacterial lifespan.

We transformed *E. coli* with a plasmid encoding the previously described synonymous CAT coding sequence variant (11) under a titratable promoter, but detected no discernable difference in growth versus *E. coli* producing CAT from the wild-type (WT) coding sequence (Fig. 2*B* and *SI Appendix*, Fig. S3*A*). However, compared with WT-CAT, this synonymous construct contains a larger number of common codons (Fig. 2*A*), which leads to increased protein accumulation due to an overall faster translation elongation rate (11, 16, 25). Consistent with this, we detected

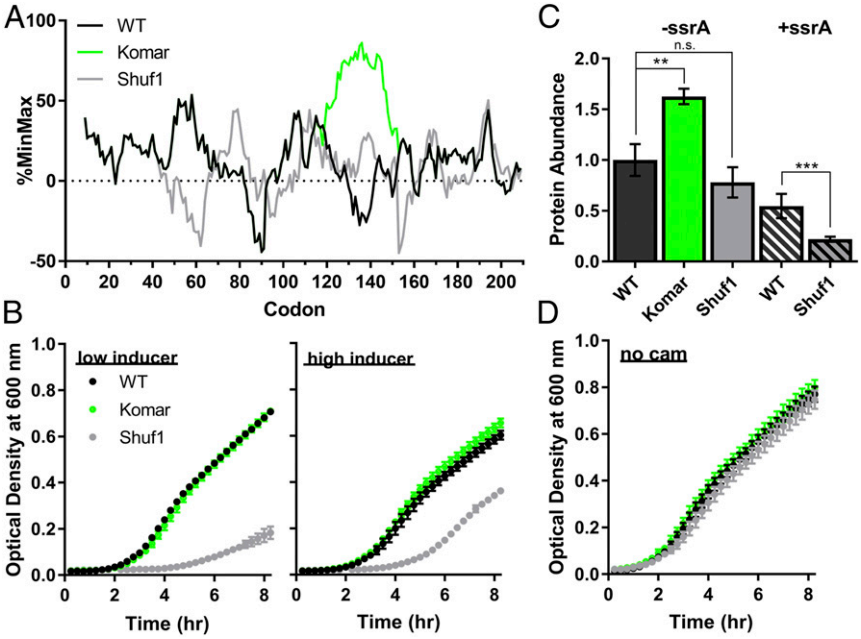

Fig. 2. CAT encoded by the synonymous Shuf1 sequence leads to impaired *E. coli* growth in the presence of cam. (*A*) Relative codon usage in WT (black), Komar (11) (green), and Shuf1 (gray) CAT coding sequences. Positive values correspond to clusters of common codons, and negative values represent clusters of rare codons, calculated over a sliding window of 17 codons (37). (*B*) Growth curves of *E. coli* expressing ssrA-tagged CAT variants challenged with cam under low (200 ng/mL) or high (1,600 ng/mL) concentrations of inducer. (*C*) Relative abundance of untagged (solid bars) or ssrA-tagged (hatched bars) CAT accumulated in cells determined by quantitative Western blotting of cell lysates. (*D*) Growth curves in the absence of cam. In all figures, data points represent the mean ± SD of at least three independent experiments; $^{**}P < 0.01$; $^{***}P < 0.001$; Welch's *t* test.

Downloaded at Lawrence Livermore National on March 5, 2020

more CAT in *E. coli* transformed with this coding sequence enriched in common codons (Fig. 2*C*). We hypothesized that this higher intracellular CAT concentration could mask a defect in specific activity. To test this, we used a Monte Carlo simulation method (18, 37) (*SI Appendix, Supplemental Methods*) to design and select an alternative synonymous CAT coding sequence, Shuf1. In Shuf1, the local synonymous codon usage patterns are very different from the WT coding sequence, but the global codon usage frequencies are very similar (Fig. 2*A* and *SI Appendix*, Fig. S1), which we predicted would lead to the synthesis of a WT-like amount of CAT. To avoid known effects of 5′ synonymous codon substitutions on translation initiation (5, 38–41), the first 46 codons of Shuf1 are identical to the WT coding sequence. Consistent with our prediction, *E. coli* produced CAT from the Shuf1 coding sequence at levels indistinguishable from WT-CAT (Fig. 2*C*). However, cells expressing Shuf1-CAT grew more slowly than cells expressing WT-CAT (*SI Appendix*, Fig. S3*A*).

We hypothesized that we could further exacerbate the observed Shuf1-CAT growth defect by adapting a strategy developed by Hilvert and coworkers to couple subtle changes in enzyme function to *E. coli* growth rate (42). This strategy involves encoding a ClpXP recognition tag (ssrA) at the C terminus of the protein of interest, selectively enhancing its degradation by the *E. coli* AAA^+ protease ClpXP and leading to correspondingly lower intracellular protein concentrations. Addition of the ssrA tag did not affect CAT structure, stability, or specific activity (*SI Appendix*, Fig. S3 *B–D*), but did lead to a dramatic growth defect for *E. coli* expressing Shuf1-CAT versus ssrA-tagged WT-CAT in the presence of cam (Fig. 2*B*). This defect also led to a lower minimum inhibitory concentration for *E. coli* expressing Shuf1-CAT versus WT-CAT (*SI Appendix*, Fig. S3*E*).

Neither the Shuf1-CAT mRNA nor Protein Is Inherently Toxic. A major challenge of all in vivo experiments is discerning the precise origin of an observed effect. For example, a recent study indicated that synonymous codon substitutions can lead to toxicity at the mRNA level even in the absence of protein production (43). To test whether production of the Shuf1-CATssrA mRNA and/or protein is inherently toxic, we compared the growth of *E. coli* expressing WT or Shuf1-CATssrA in the absence of cam. These growth curves were indistinguishable (Fig. 2*D*), indicating that the Shuf1 defect is specifically related to impaired CAT enzyme function. Moreover, in the presence of cam, the growth defect was partially suppressed at higher inducer concentrations (Fig. 2*B*), contrary to the larger growth defect expected if the Shuf1-CATssrA mRNA and/or protein were inherently toxic.

To test whether Shuf1-CAT expression induces a general cell stress response, we used mass spectrometry to compare the abundances of 1,277 proteins in *E. coli* expressing ssrA-tagged CAT from either the WT or Shuf1 coding sequence. There was no significant difference detected in the level of most proteins, including known stress-associated molecular chaperones and proteases (Fig. 3). Taken together, these results support a model in which the Shuf1-CAT growth defect is due to a direct defect in active CAT protein production, rather than an indirect effect on other cell functions.

Shuf1 Coding Sequence Does Not Adversely Affect mRNA Concentration. We noticed that addition of the ssrA tag led to a larger reduction in intracellular accumulation for CAT produced from the Shuf1 versus WT coding sequence (Fig. 2*C*, hatched bars). To determine whether this decrease in Shuf1-CAT was due to a defect arising from Shuf1 transcription and/or mRNA half-life, versus a translation-related defect, we quantified the levels of WT and Shuf1 mRNA. These levels were indistinguishable (*SI Appendix*, Fig. S4*A*). Together with the indistinguishable levels of WT-CAT and Shuf1-CAT protein accumulation in the absence of the ssrA tag (Fig. 2*C*, filled bars), these results suggest a model in which the Shuf1 synonymous codon changes affect intracellular CAT concentration at the translational level, likely due to a greater susceptibility of the Shuf1-CAT protein to degradation (Fig. 4).

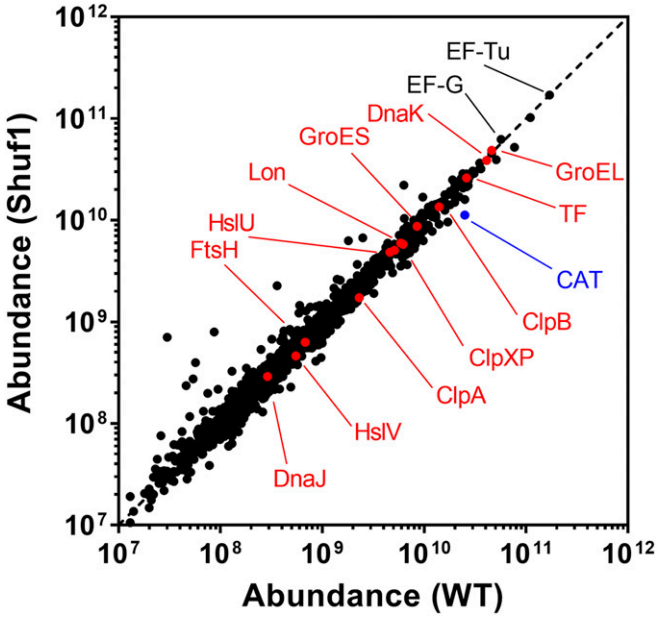

Fig. 3. Translation of CAT using Shuf1 coding sequence does not significantly perturb the *E. coli* proteome. Relative abundance of *E. coli* proteins upon expression of WT or Shuf1 CAT. Twelve *E. coli* molecular chaperones and AAA^+ ATPases are shown in red; 1,264 other *E. coli* proteins are shown in black. No significant upregulation of chaperones or ATPases was observed for *E. coli* expressing Shuf1.

ClpX Deletion Indicates ClpXP Is Major Source of Shuf1-CATssrA Growth Defect. If the Shuf1 codon-dependent growth defect is due to more efficient degradation of ssrA-tagged Shuf1-CAT by cellular proteases, specifically ClpXP, deleting ClpX would be expected to ameliorate the growth defect in vivo. ClpXP is the major *E. coli* protease responsible for degrading ssrA-tagged polypeptides under log-phase growth (44, 45). In general, less stably folded proteins are more susceptible to degradation by ClpXP than more stable substrates (46–48), presumably because less energy is required for ClpX to unfold unstable protein structures and expose the polypeptide chains to the ClpP protease active sites (49). To test whether ClpXP degradation is the key mechanism impairing growth when *E. coli* expresses CAT from the Shuf1 coding sequence, we induced expression of WT-CAT and Shuf1-CAT in an *E. coli* W3110 derivative that lacks ClpX (46, 50) and compared growth in this background to the parent strain W3110 in the presence of cam. ClpX deletion enhanced growth only of cells expressing ssrA-tagged CAT from the Shuf1 coding sequence (Fig. 5*A*). Likewise, omission of the ssrA tag enhanced growth only for *E. coli* expressing ClpX; there was no effect on cells lacking ClpX (Fig. 5*B*). These results confirm that the major effect of the Shuf1 synonymous codon substitutions is enhanced degradation of ssrA-tagged CAT by ClpXP.

Native WT- and Shuf1-CAT Proteins Are Subtly Different. Synonymous codon substitutions can lead to a wide range of effects on the encoded protein, including changes to translational fidelity (decoding accuracy) (6) and the native structure (1, 10, 12, 17). As a next test of the mechanism by which Shuf1 codon changes alter cell growth rate, we compared the CAT proteins produced

Downloaded at Lawrence Livermore National on March 5, 2020

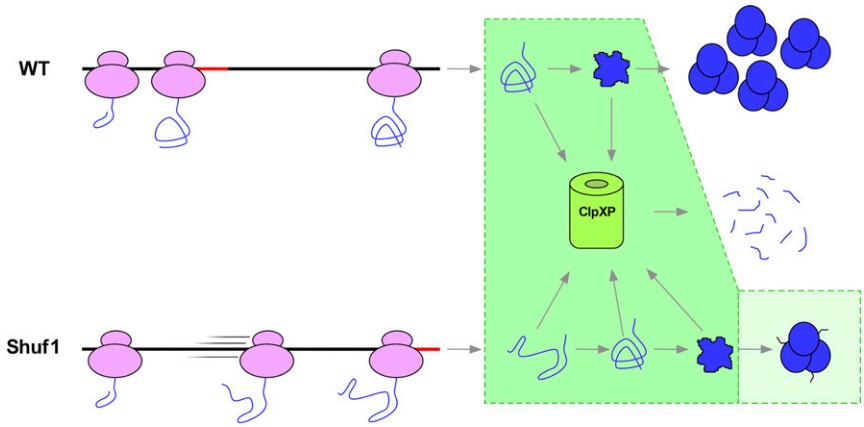

Fig. 4. Proposed model for the effects of synonymous CAT codon substitutions on ssrA-tagged CAT folding and cell fitness. Synonymous changes in the Shuf1 coding sequence alter the local rate of translation, affecting the conformation of CAT cotranslationally and persisting after release of the nascent protein from the ribosome. These altered Shuf1 folding intermediates are more susceptible to degradation by ClpXP than intermediates populated during and after translation of the WT coding sequence. Some Shuf1-CATssrA proteins evade degradation and eventually fold to an active conformation that is also more susceptible to degradation than WT-CATssrA.

from the WT and Shuf1 coding sequences. In both cases, CAT was detected only in the soluble fraction of the cell lysate (*SI Appendix*, Fig. S4*B*), indicating the Shuf1 growth defect is not due to CAT aggregation. Likewise, the secondary and tertiary structure (*SI Appendix*, Fig. S4*C*), resistance to chemical and thermal denaturation (Fig. 5*C* and *SI Appendix*, Fig. S4*D*), and specific activity (Fig. 5*D*) of purified CAT produced from the Shuf1 mRNA sequence were indistinguishable from CAT translated from the WT coding sequence. We also used mass spectrometry to compare the molecular weights of CAT translated from these coding sequences. These masses were indistinguishable to within one mass unit and matched the expected molecular weight of 25,953 Da. Taken together, these results demonstrate that CAT production from the Shuf1 coding sequence does not prohibit formation of a stable, active CAT protein structure.

Despite the native CAT structural similarities reported here, it is important to note that digestion by ClpXP requires force-mediated unfolding of a substrate protein from its C terminus, driven by ATP hydrolysis (51–53). Resistance to mechanical

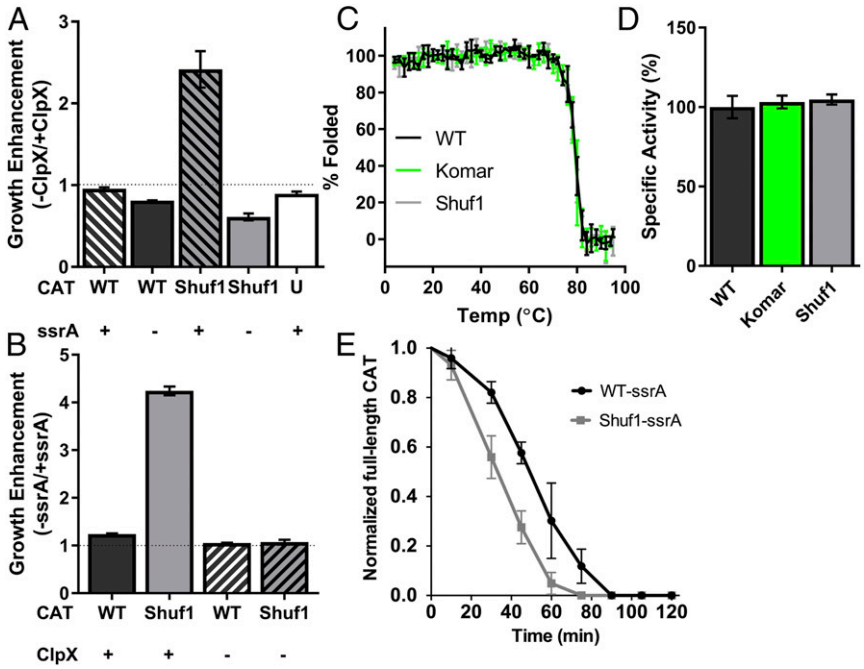

Fig. 5. Shuf1 CAT is more susceptible to ClpXP degradation than WT CAT, despite several other indistinguishable characteristics. (*A* and *B*) Selective effects of ssrA-tagging and ClpX deletion on the Shuf1 growth defect. (*A*) In the ClpX deletion strain (W3110 Δ*clpX*), a large increase in growth rate relative to the parent strain is observed only for ssrA-tagged Shuf1. Other constructs grow slightly slower in the absence of ClpX. U, uninduced cell culture. (*B*) Cell growth data from *A* plotted to highlight the effect on growth rate of removing the ssrA tag. Omitting the ssrA tag has no effect on growth in the ClpX knockout (hatched bars). In the presence of ClpX (filled bars), there is a much larger increase in growth upon ssrA tag deletion for Shuf1 than WT, indicating Shuf1 is more susceptible to ClpXP degradation than WT. (*C*) Thermal denaturation of CAT monitored by far-UV CD spectroscopy at 205 nm. (*D*) Acetyltransferase activity of purified, native CAT, normalized to WT. (*E*) In vitro ClpXP degradation of native, purified, ssrA-tagged CAT trimers (43). In all panels, data points represent the mean ± SD; $n = 3$ biological replicates.

Downloaded at Lawerence Livermore National on March 5, 2020

force reports on a distinct aspect of protein stability from resistance to chemical or thermal denaturation (54–57). To directly test whether the Shuf1 synonymous codon substitutions lead to a native CAT structure that is more susceptible to force-mediated unfolding and degradation, we subjected native, purified ssrA-tagged CAT produced in vivo from the WT or Shuf1 coding sequences to an in vitro ClpXP degradation assay (44, 58). Although both proteins exhibit resistance to ClpXP degradation, CAT synthesized from the Shuf1 coding sequence was degraded more rapidly than WT-CAT (Fig. 5*E* and *SI Appendix*, Fig. S5), in direct contrast to the indistinguishable behavior observed in our other analyses (e.g., Fig. 5 *C* and *D*). Shuf1-CATssrA was more susceptible to ClpXP degradation even when a second, control substrate was added to the reaction and degraded at the same rate (*SI Appendix*, Fig. S6), demonstrating the robustness of this surprising result. This result demonstrates that CAT produced from the Shuf1 mRNA sequence is more susceptible to degradation by ClpXP, perhaps both before as well as after acquiring its native structure. Crucially, the differential susceptibility to ClpXP degradation provides direct evidence of the impact of the Shuf1 codon substitutions on CAT folding, as proteins with identical amino acid sequences would arrive at different native structures only via distinct folding mechanisms. Because the ssrA tag is located at the CAT C terminus, we expect that degradation by ClpXP is predominantly posttranslational, occurring after release of the nascent chain from the ribosome.

mRNA Secondary Structural Stability Does Not Explain Shuf1 Growth Defect. The results above suggest the Shuf1 synonymous codon substitutions impair CAT cotranslational folding by altering local patterns of translation elongation. In vitro, synonymous codons have been shown to alter elongation rate either by altering the rate of decoding (59) or by altering downstream mRNA stability, which can impede ribosome translocation (60). In vivo, there is some evidence that stable mRNA stem-loop structures can alter the elongation rate of the ribosome (61–63), although other studies have detected no difference (38, 64, 65), likely due to destabilization of mRNA structure by polysomes and/or the helicase activity of the ribosome. Although the overall predicted mRNA stability of the WT and Shuf1 genes are similar, a predicted stable 3′ stem-loop structure in Shuf1 is not present in the WT coding sequence (*SI Appendix*, Fig. S7*A*). To test whether this structure is responsible for the Shuf1 growth defect, we created chimeric mRNA sequences with only the 5′, middle, or 3′ portion of the WT sequence substituted with the Shuf1 sequence (*SI Appendix*, Fig. S7*B*), but observed no growth defect, for the chimera bearing the 3′ portion of Shuf1 had no impact on growth rate (*SI Appendix*, Fig. S7*C*). Moreover, growth rates for these chimeras correlated more closely with the difference in relative codon usage frequencies than measures of mRNA stability (*SI Appendix*, Fig. S7*D*). Taken together, these results indicate that translation elongation rate differences arising from changes in codon usage frequencies is a more likely origin of the Shuf1 growth defect than changes in mRNA secondary structure.

Discussion

Most of our current understanding of protein folding mechanisms is derived from studies of small proteins that refold reversibly when diluted from chemical denaturants. However, only a small number of proteins can refold robustly in vitro, even though many more can be maintained in a stable state once extracted from the cell (24, 66, 67). This suggests both that the conformations adopted early during the folding process are crucial to successful folding and that the cellular environment supports the formation of early folding intermediates that are distinct from the conformations populated upon dilution from denaturant. Indeed, there is substantial evidence that molecular chaperones are crucial to the successful folding of many complex proteins in vivo (29–33). Although it has been hypothesized that synonymous codon changes could alter elongation rate and modify folding mechanisms in vivo, it has thus far been challenging to find evidence to support this hypothesis from experiments performed in vivo, possibly due to buffering provided by molecular chaperones.

Results presented here indicate that during synthesis, the folding of nascent CAT polypeptide chains is sensitive to synonymous codon-induced changes to translation elongation rate. Although the nascent chains produced using different synonymous codon patterns remain broadly capable of achieving a stable, active CAT trimer structure, translation using the synonymous Shuf1 mRNA sequence leads to CAT proteins that are more susceptible to degradation by the cellular protease ClpXP than WT-CAT, leading to a dramatic cell growth defect. Given that the ClpXP ssrA degradation tag is attached to the very C terminus of CAT, it is likely that the majority of this digestion occurs only posttranslationally, after the CAT nascent chain is released from the ribosome (Fig. 4). Remarkably, even native Shuf1-CATssrA protein is more susceptible to degradation than native CATssrA translated using the WT coding sequence, demonstrating that the impact of the codon-induced perturbations persists long after translation and folding is complete. Buffering by the cellular protein homeostasis network is therefore not sufficient to mask the impact of the Shuf1-CAT folding defect on cell growth.

These results are consistent with a small but growing number of studies indicating that synonymous codon substitutions can perturb protein folding mechanisms (1, 10, 12, 68). The ssrA tagging approach developed here provides a general strategy to uncover such perturbations in other coding sequences, even when they do not lead to dramatic remodeling of the final protein structure. In contrast to the translation rate-sensitive effects we observed for CAT folding, recent in vitro single-molecule force-unfolding experiments have shown that some small, ribosome-bound natively folded domains can fold via similar mechanisms on and off the ribosome (69, 70). However, as these studies noted, forced unfolding measured by molecular tweezers cannot capture the transient folding of a nascent chain during its synthesis (33), and hence what is measured in these experiments is the effect of close proximity of the ribosome surface, rather than cotranslational folding. The very robust folding behavior of these well-characterized, reversible folding models may indeed lead to indistinguishable folding behavior during translation, a model supported by recent force-feedback folding measurements (71). However, the model proteins selected for these studies are smaller than >75% of proteins in the *E. coli* proteome (24), whereas all known examples of synonymous codon-derived alterations to cotranslational folding are much larger (e.g., refs. 1, 9, 10, and 72). We are not aware of an in vitro folding mechanism for a protein >175 aa long that is preserved during cotranslational folding. Synonymous codon-derived modulation of elongation rate may therefore play a broad role in the efficient folding of larger, more complex proteins.

Our CAT results demonstrate that synonymous changes to mRNA coding sequences can significantly perturb folding of the WT protein sequence even in the presence of the cellular repertoire of molecular chaperones. This result suggests that mRNA sequences have likely evolved alongside molecular chaperones to most efficiently support folding of the broad repertoire of protein structures produced in vivo. Although our understanding of cotranslational folding mechanisms is still in its infancy, these results imply that it should ultimately be possible to rationally design mRNA coding sequences to enhance in vivo folding yield and to identify disease-associated synonymous codon substitutions most likely to adversely affect protein cotranslational folding, particularly for large or otherwise complex proteins.

Downloaded at Lawerence Livermore National on March 5, 2020

Methods

Cell Growth Assays. A single colony of *E. coli* KA12 (73) or W3110 (50) transformed with a pKT-CAT plasmid from a freshly streaked LB-amp plate was used to inoculate 20 mL of LB plus 100 μg/mL ampicillin (LB-amp) and grown overnight with shaking at 37 °C. Unless otherwise specified, all cultures contained 100 μg/mL ampicillin and no tetracycline. Overnight cultures were used to inoculate fresh LB-amp to an optical density at 600 nm (OD_{600}) of 0.05, to which was added 35 μg/mL chloramphenicol (unless otherwise specified) and the indicated concentration of tetracycline inducer (0 to 1,600 ng/mL), transferred to one well of a 12-well plate and incubated at 37 °C with continuous shaking in a Synergy H1 microplate reader (BioTek). Growth was measured as the increase in OD_{600}. The linear portion of the growth curve was fit to a straight line, and the slope was taken as the growth rate.

Data Availability Statement. All data discussed in the paper is available at GitHub, https://github.com/plclark1/SynonymousCodons/tree/master/Synonymous_codon_substitutions_perturb_cotranslational_protein_folding_in_vivo_and_impair_cell_fitness.

ACKNOWLEDGMENTS. We thank Matt Champion and the Notre Dame Mass Spectrometry & Proteomics Facility for performing the mass spectrometry experiments, Don Hilvert for the kind gift of the pKT and pKTS plasmids, and Peter Chein, Don Hilvert, Jeff Nivala and Mark Akeson for sharing *E. coli* strains with us. We are grateful to Gabriel Wright and Scott Emrich for helpful discussions. This project was supported by grants GM120733 and GM105816 from the National Institutes of Health and a research grant from the W. M. Keck Foundation.

1. C. Kimchi-Sarfaty *et al*., A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science* **315**, 525–528 (2007).
2. A. A. Komar, The yin and yang of codon usage. *Hum. Mol. Genet.* **25**, R77–R85 (2016).
3. A. R. Subramaniam *et al*., A serine sensor for multicellularity in a bacterium. *eLife* **2**, e01501 (2013).
4. A. Radhakrishnan *et al*., The DEAD-box protein Dhh1p couples mRNA decay and translation by monitoring codon optimality. *Cell* **167**, 122–132. (2016).
5. S. Bhattacharyya *et al*., Accessibility of the shine-dalgarno sequence dictates N-terminal codon bias in E. coli. *Mol. Cell* **70**, 894–905. (2018).
6. D. A. Drummond, C. O. Wilke, Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell* **134**, 341–352 (2008).
7. V. Daidone *et al*., An apparently silent nucleotide substitution (c.7056C>T) in the von Willebrand factor gene is responsible for type 1 von Willebrand disease. *Haematologica* **96**, 881–887 (2011).
8. S. Pechmann, J. W. Chartron, J. Frydman, Local slowdown of translation by non-optimal codons promotes nascent-chain recognition by SRP in vivo. *Nat. Struct. Mol. Biol.* **21**, 1100–1105 (2014).
9. G. Zhang, M. Hubalewska, Z. Ignatova, Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat. Struct. Mol. Biol.* **16**, 274–280 (2009).
10. I. M. Sander, J. L. Chaney, P. L. Clark, Expanding Anfinsen's principle: Contributions of synonymous codon selection to rational protein design. *J. Am. Chem. Soc.* **136**, 858–861 (2014).
11. A. A. Komar, T. Lesnik, C. Reiss, Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. *FEBS Lett.* **462**, 387–391 (1999).
12. F. Buhr *et al*., Synonymous codons direct cotranslational folding toward different protein conformations. *Mol. Cell* **61**, 341–351 (2016).
13. F. Zhang, S. Saha, S. A. Shabalina, A. Kashina, Differential arginylation of actin isoforms is regulated by coding sequence-dependent degradation. *Science* **329**, 1534–1537 (2010).
14. M. A. Sørensen, S. Pedersen, Absolute in vivo translation rates of individual codons in Escherichia coli. The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate. *J. Mol. Biol.* **222**, 265–280 (1991).
15. T. Tuller *et al*., An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* **141**, 344–354 (2010).
16. P. Shah, M. A. Gilchrist, Explaining complex codon usage patterns with selection for translational efficiency, mutation bias, and genetic drift. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10231–10236 (2011).
17. G. Boël *et al*., Codon influence on protein expression in E. coli correlates with mRNA levels. *Nature* **529**, 358–363 (2016).
18. T. F. Clarke, 4th, P. L. Clark, Rare codons cluster. *PLoS One* **3**, e3412 (2008).
19. J. L. Chaney *et al*., Widespread position-specific conservation of synonymous rare codons within coding sequences. *PLOS Comput. Biol.* **13**, e1005531 (2017).
20. W. M. Jacobs, E. I. Shakhnovich, Evidence of evolutionary selection for cotranslational folding. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 11434–11439 (2017).
21. S. Pechmann, J. Frydman, Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat. Struct. Mol. Biol.* **20**, 237–243 (2013).
22. K. A. Dill, J. L. MacCallum, The protein-folding problem, 50 years on. *Science* **338**, 1042–1046 (2012).
23. K. W. Plaxco, K. T. Simons, D. Baker, Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* **277**, 985–994 (1998).
24. E. Braselmann, J. L. Chaney, P. L. Clark, Folding the proteome. *Trends Biochem. Sci.* **38**, 337–344 (2013).
25. J. L. Chaney, P. L. Clark, Roles for synonymous codon usage in protein biogenesis. *Annu. Rev. Biophys.* **44**, 143–166 (2015).
26. G. Hanson, J. Coller, Codon optimality, bias and usage in translation and mRNA decay. *Nat. Rev. Mol. Cell Biol.* **19**, 20–30 (2018).
27. C. E. Brule, E. J. Grayhack, Synonymous codons: Choose wisely for expression. *Trends Genet.* **33**, 283–297 (2017).
28. G. N. Jacobson, P. L. Clark, Quality over quantity: Optimizing co-translational protein folding with non-'optimal' synonymous codons. *Curr. Opin. Struct. Biol.* **38**, 102–110 (2016).
29. D. Balchin, M. Hayer-Hartl, F. U. Hartl, In vivo aspects of protein folding and quality control. *Science* **353**, aac4354 (2016).
30. G. Kramer, A. Shiber, B. Bukau, Mechanisms of cotranslational maturation of newly synthesized proteins. *Annu. Rev. Biochem.* **88**, 337–364 (2019).
31. S. Pechmann, F. Willmund, J. Frydman, The ribosome as a hub for protein quality control. *Mol. Cell* **49**, 411–421 (2013).
32. S. Preissler, E. Deuerling, Ribosome-associated chaperones as key players in proteostasis. *Trends Biochem. Sci.* **37**, 274–283 (2012).
33. K. Liu, K. Maciuba, C. M. Kaiser, The ribosome cooperates with a chaperone to guide multi-domain protein folding. *Mol. Cell* **74**, 310–319. (2019).
34. A. G. Leslie, Refined crystal structure of type III chloramphenicol acetyltransferase at 1.75 A resolution. *J. Mol. Biol.* **213**, 167–186 (1990).
35. W. V. Shaw, The enzymatic acetylation of chloramphenicol by extracts of R factor-resistant Escherichia coli. *J. Biol. Chem.* **242**, 687–693 (1967).
36. M. Eames, T. Kortemme, Cost-benefit tradeoffs in engineered lac operons. *Science* **336**, 911–915 (2012).
37. A. Rodriguez, G. Wright, S. Emrich, P. L. Clark, %MinMax: A versatile tool for calculating and comparing synonymous codon usage and its impact on protein folding. *Protein Sci.* **27**, 356–362 (2018).
38. A.M. Mustoe *et al*., Pervasive regulatory functions of mRNA structure revealed by high-resolution SHAPE probing. *Cell* **173**, 181–195. (2018).
39. D. B. Goodman, G. M. Church, S. Kosuri, Causes and effects of N-terminal codon bias in bacterial genes. *Science* **342**, 475–479 (2013).
40. G. Kudla, A. W. Murray, D. Tollervey, J. B. Plotkin, Coding-sequence determinants of gene expression in Escherichia coli. *Science* **324**, 255–258 (2009).
41. M. Welch *et al*., Design parameters to control synthetic gene expression in Escherichia coli. *PLoS One* **4**, e7002 (2009).
42. M. Neuenschwander, M. Butz, C. Heintz, P. Kast, D. Hilvert, A simple selection strategy for evolving highly efficient enzymes. *Nat. Biotechnol.* **25**, 1145–1147 (2007).
43. P. Mittal, J. Brindle, J. Stephen, J. B. Plotkin, G. Kudla, Codon usage influences fitness through RNA toxicity. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 8639–8644 (2018).
44. S. Gottesman, E. Roche, Y. Zhou, R. T. Sauer, The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev.* **12**, 1338–1347 (1998).
45. K. C. Keiler, P. R. Waller, R. T. Sauer, Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**, 990–993 (1996).
46. J. M. Flynn, S. B. Neher, Y. I. Kim, R. T. Sauer, T. A. Baker, Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol. Cell* **11**, 671–683 (2003).
47. C. Lee, M. P. Schwartz, S. Prakash, M. Iwakura, A. Matouschek, ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. *Mol. Cell* **7**, 627–637 (2001).
48. R. T. Sauer *et al*., Sculpting the proteome with AAA(+) proteases and disassembly machines. *Cell* **119**, 9–18 (2004).
49. J. Ortega, S. K. Singh, T. Ishikawa, M. R. Maurizi, A. C. Steven, Visualization of substrate binding and translocation by the ATP-dependent protease, ClpXP. *Mol. Cell* **6**, 1515–1521 (2000).
50. S. Gottesman, W. P. Clark, V. de Crecy-Lagard, M. R. Maurizi, X. Clp, ClpX, an alternative subunit for the ATP-dependent Clp protease of Escherichia coli. Sequence and in vivo activities. *J. Biol. Chem.* **268**, 22618–22626 (1993).
51. A. O. Olivares, T. A. Baker, R. T. Sauer, Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat. Rev. Microbiol.* **14**, 33–44 (2016).
52. J. M. Miller, J. Lin, T. Li, A. L. Lucius, E. coli ClpA catalyzed polypeptide translocation is allosterically controlled by the protease ClpP. *J. Mol. Biol.* **425**, 2795–2812 (2013).
53. S. A. Mahmoud, P. Chien, Regulated proteolysis in bacteria. *Annu. Rev. Biochem.* **87**, 677–696 (2018).
54. A. J. Wilcox, J. Choy, C. Bustamante, A. Matouschek, Effect of protein structure on mitochondrial import. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15435–15440 (2005).
55. D. K. West, D. J. Brockwell, P. D. Olmsted, S. E. Radford, E. Paci, Mechanical resistance of proteins explained using simple molecular models. *Biophys. J.* **90**, 287–297 (2006).
56. M. Carrion-Vazquez *et al*., Mechanical and chemical unfolding of a single protein: A comparison. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 3694–3699 (1999).
57. E. J. Guinn, B. Jagannathan, S. Marqusee, Single-molecule chemo-mechanical unfolding reveals multiple transition state barriers in a small single-domain protein. *Nat. Commun.* **6**, 6861 (2015).
58. J. Nivala, D. B. Marks, M. Akeson, Unfoldase-mediated protein translocation through an α-hemolysin nanopore. *Nat. Biotechnol.* **31**, 247–250 (2013).
59. N. Haase, W. Holtkamp, R. Lipowsky, M. Rodnina, S. Rudorf, Corrigendum: Decomposition of time-dependent fluorescence signals reveals codon-specific kinetics of protein synthesis. *Nucleic Acids Res.* **46**, 12186–12187 (2018).
60. J. Choi *et al*., How messenger RNA and nascent chain sequences regulate translation elongation. *Annu. Rev. Biochem.* **87**, 421–449 (2018).

Downloaded at Lawerence Livermore Nationa on March 5, 2020

61. C. Chen *et al*., Dynamics of translation by single ribosomes through mRNA secondary structures. *Nat. Struct. Mol. Biol.* **20**, 582–588 (2013).
62. A. G. Nackley *et al*., Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. *Science* **314**, 1930–1933 (2006).
63. D. H. Burkhardt *et al*., Operon mRNAs are organized into ORF-centric structures that predict translation efficiency. *eLife* **6**, e22037 (2017).
64. K. Nakahigashi *et al*., Effect of codon adaptation on codon-level and gene-level translation efficiency in vivo. *BMC Genomics* **15**, 1115 (2014).
65. J. D. Beaudoin *et al*., Analyses of mRNA structure dynamics identify embryonic gene regulatory programs. *Nat. Struct. Mol. Biol.* **25**, 677–686 (2018).
66. K. Xia *et al*., Identifying the subproteome of kinetically stable proteins via diagonal 2D SDS/PAGE. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 17329–17334 (2007).
67. C. Park, S. Zhou, J. Gilmore, S. Marqusee, Energetics-based protein profiling on a proteomic scale: Identification of proteins resistant to proteolysis. *J. Mol. Biol.* **368**, 1426–1437 (2007).
68. C. H. Yu *et al*., Codon usage influences the local rate of translation elongation to regulate co-translational protein folding. *Mol. Cell* **59**, 744–754 (2015).
69. E. J. Guinn, P. Tian, M. Shin, R. B. Best, S. Marqusee, A small single-domain protein folds through the same pathway on and off the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 12206–12211 (2018).
70. C. M. Kaiser, D. H. Goldman, J. D. Chodera, I. Tinoco, Jr, C. Bustamante, The ribosome modulates nascent protein folding. *Science* **334**, 1723–1727 (2011).
71. P. Tian *et al*., Folding pathway of an Ig domain is conserved on and off the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E11284–E11293 (2018).
72. S. J. Kim *et al*., Protein folding. Translational tuning optimizes nascent protein folding in cells. *Science* **348**, 444–448 (2015).
73. P. Kast, M. Asif-Ullah, N. Jiang, D. Hilvert, Exploring the active site of chorismate mutase by combinatorial mutagenesis and selection: The importance of electrostatic catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 5043–5048 (1996).

Downloaded at Lawerence Livermore National on March 5, 2020