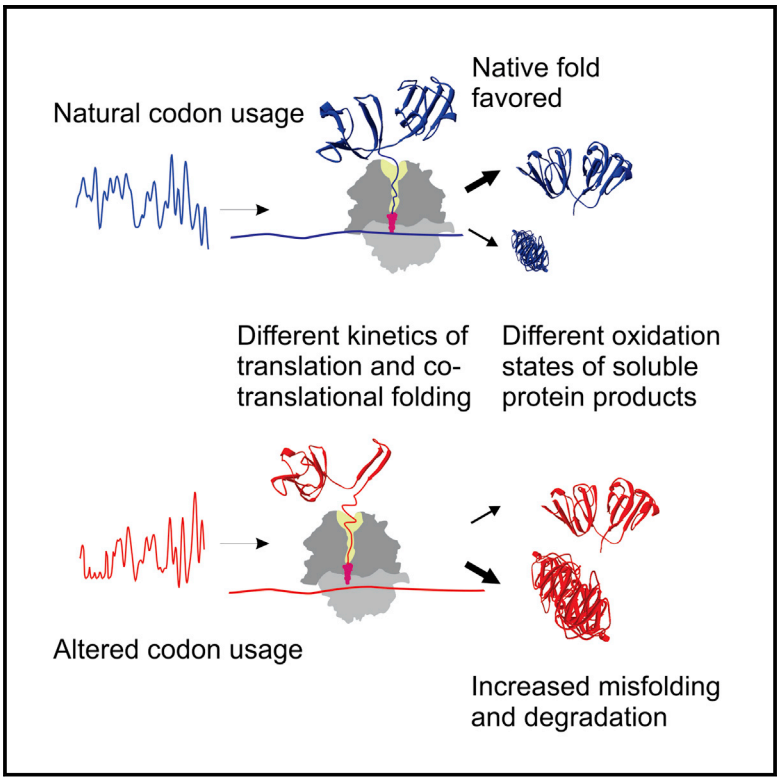
Article

Molecular Cell

Synonymous Codons Direct Cotranslational Folding toward Different Protein Conformations

Graphical Abstract



Authors

Florian Buhr, Sujata Jha, Michael Thommen, ..., Harald Schwalbe, Marina V. Rodnina, Anton A. Komar

Correspondence

schwalbe@em.uni-frankfurt.de (H.S.), rodnina@mpibpc.mpg.de (M.V.R.), a.komar@csuohio.edu (A.A.K.)

In Brief

Buhr et al. demonstrate how synonymous codon usage guides the real-time kinetics of translation and cotranslational folding of a gamma-B crystallin. Gamma-B crystallins produced from mRNAs with altered codon usage have the same primary sequence but attain different conformations in solution and in the cell.

Highlights

- Quality of protein folding in cells is guided by synonymous codon usage
- NMR reveals multiple conformational and oxidation states of synonymous variants
- Synonymous codon usage in mRNA alters translation kinetics
- Real-time cotranslational folding is guided by synonymous codon usage

Buhr et al., 2016, Molecular Cell *61*, 341–351
February 4, 2016
http://dx.doi.org/10.1016/j.molcel.2016.01.008

CellPress

Molecular Cell
Article

Synonymous Codons Direct Cotranslational Folding toward Different Protein Conformations

Florian Buhr,[1,6] Sujata Jha,[2,6] Michael Thommen,[3,6] Joerg Mittelstaet,[3] Felicitas Kutz,[1] Harald Schwalbe,[1,*] Marina V. Rodnina,[3,*] and Anton A. Komar[2,4,5,*]

[1]Center for Biomolecular Magnetic Resonance, Institute of Organic Chemistry and Chemical Biology, Johann Wolfgang Goethe-Universität Frankfurt am Main, 60438 Frankfurt am Main, Germany

[2]Center for Gene Regulation in Health and Disease and Department of Biological, Geological, and Environmental Sciences, Cleveland State University, Cleveland, OH 44115, USA

[3]Department of Physical Biochemistry, Max Planck Institute for Biophysical Chemistry, 37077 Goettingen, Germany

[4]Department of Biochemistry and Center for RNA Molecular Biology, Case Western Reserve University, Cleveland, OH 44106, USA

[5]Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, Cleveland, OH 44195, USA

[6]Co-first author

*Correspondence: schwalbe@em.uni-frankfurt.de (H.S.), rodnina@mpibpc.mpg.de (M.V.R.), a.komar@csuohio.edu (A.A.K.)

http://dx.doi.org/10.1016/j.molcel.2016.01.008

SUMMARY

In all genomes, most amino acids are encoded by more than one codon. Synonymous codons can modulate protein production and folding, but the mechanism connecting codon usage to protein homeostasis is not known. Here we show that synonymous codon variants in the gene encoding gamma-B crystallin, a mammalian eye-lens protein, modulate the rates of translation and cotranslational folding of protein domains monitored in real time by Förster resonance energy transfer and fluorescence-intensity changes. Gamma-B crystallins produced from mRNAs with changed codon bias have the same amino acid sequence but attain different conformations, as indicated by altered in vivo stability and in vitro protease resistance. 2D NMR spectroscopic data suggest that structural differences are associated with different cysteine oxidation states of the purified proteins, providing a link between translation, folding, and the structures of isolated proteins. Thus, synonymous codons provide a secondary code for protein folding in the cell.

INTRODUCTION

The genetic code is degenerate, with up to six synonymous codons encoding a given amino acid. The occurrence of synonymous codons in protein-coding open reading frames (ORFs) of genes is not random, thus revealing the existence of evolutionary pressure on codon choice (Hershberg and Petrov, 2008; Sharp et al., 2010; Plotkin and Kudla, 2011; Pechmann and Frydman, 2013; Chaney and Clark, 2015). The occurrence of synonymous codons and the abundance of the corresponding isoacceptor tRNAs are major causes of nonuniformity in mRNA translation (Ikemura, 1985; Buchan and Stansfield, 2007; Komar, 2009; Pop et al., 2014; Dana and Tuller 2014; Gardin et al., 2014), although several other factors, such as mRNA structure or the presence of anti-Shine-Dalgarno-like sequences in bacterial ORFs (Li et al., 2012), may also contribute (Pop et al., 2014; Ingolia, 2014). In a given organism, frequently used codons are translated more rapidly than infrequently used ones (Ikemura, 1985; Buchan and Stansfield, 2007; Ingolia et al., 2009; Komar, 2009; Ingolia, 2014; Dana and Tuller 2014; Gardin et al., 2014). Nonrandom and nonuniform distribution of codons in gene ORFs provides a unique organism-specific pattern that modulates local translation elongation rates (Clarke and Clark 2008; Komar, 2009; Pechmann and Frydman, 2013; Chaney and Clark, 2015) and contributes to mRNA stability (Pedersen et al., 2011; Presnyak et al., 2015).

In vivo, protein folding begins cotranslationally as nascent peptide chains emerge from the ribosome exit tunnel (Hartl and Hayer-Hartl, 2009; Komar, 2009; Kramer et al., 2009; Cabrita et al., 2010; Waudby et al., 2013; Pechmann et al., 2013; Gloge et al., 2014; Chaney and Clark, 2015). Variations in local translation rates may facilitate protein folding by allowing ordered, sequential structuring of the nascent polypeptide chains emerging from the ribosome (Tsai et al., 2008; Kramer et al., 2009; Komar, 2009; Zhang and Ignatova, 2011; Waudby et al., 2013; O'Brien et al., 2014; Gloge et al., 2014; Chaney and Clark, 2015). The significance of synonymous codon usage on protein folding is highlighted by findings showing that synonymous mutations and naturally occurring synonymous single-nucleotide polymorphisms (sSNPs) can affect proteins' activity (Komar et al., 1999; Kimchi-Sarfaty et al., 2007; Yu et al., 2015), interactions with drugs and inhibitors (Kimchi-Sarfaty et al., 2007), phosphorylation profiles (Zhou et al., 2013), sensitivity to limited proteolysis (Kimchi-Sarfaty et al., 2007; Zhang et al., 2009; Zhou et al., 2013), spectroscopic properties (Sander et al., 2014), and aggregation propensity (Hu et al., 2013; Sander et al., 2014; Kim et al., 2015), which ultimately can cause diseases (Sauna and Kimchi-Sarfaty, 2011; Hunt et al., 2014). Synonymous codon choice has also been suggested to affect efficient interaction of nascent polypeptides with the signal-recognition particle (Pechmann et al., 2014). Changes in codon context caused by synonymous mutations may also induce mistranslation, leading to protein misfolding (Drummond and Wilke, 2008). An effect of

synonymous codon usage on cotranslational folding properties of several proteins has been suggested based on computational work or experiments that utilized stalled ribosome-nascent chain complexes (for reviews, see Tsai et al., 2008; Komar, 2009; Zhang and Ignatova, 2011; O'Brien et al., 2014; Gloge et al., 2014). However, little attention has been paid to the correct amino acid composition of the synonymous protein variants under study and the direct link between the synonymous mutations, the kinetics of translation, and real-time cotranslational folding, and the structure of the released protein has not been shown to date.

To investigate how differential usage of synonymous codons affects translation kinetics, co- and posttranslational folding, and protein stability, we analyzed in vivo expression of the recombinant bovine eye-lens protein gamma-B crystallin in *Escherichia coli* cells and in vitro in a completely reconstituted high-performance translation system from *E. coli*. Gamma-B crystallin is a globular protein consisting of two homologous all-beta Greek-key domains connected by a six-residue flexible linker (Bloemendal et al., 2004). We have chosen gamma-B crystallin because translation of this two-domain protein in a mammalian-cell-free system is a nonuniform process (Komar and Jaenicke, 1995). We suggested that the codon usage and translation rates in gamma-B crystallin are optimized to tune the synthesis and folding of this protein in the cell (Komar and Jaenicke, 1995); however, the direct experimental evidence in support of this hypothesis was lacking. The two domains of gamma-B crystallin comprise a fold that is very common in globular proteins, and its folding pathway may be representative for many beta-rich folds in general (Bloemendal et al., 2004). Here, we monitored the kinetics of synthesis and cotranslational folding of gamma-B crystallin synonymous variants in real time by fluorescence and Förster resonance energy transfer (FRET). The final protein conformation was assessed by 2D NMR (nuclear magnetic resonance). The stability of the protein variants was also probed in vivo by protein expression analysis and in vitro by limited proteolysis with Proteinase K (PK). Taken together, our data show how synonymous codon usage regulates translation elongation, cotranslational folding, and protein quality in the cell.

RESULTS

Synonymous Codon Replacements Improve the Stability and Solubility of Gamma-B Crystallin

The codon distribution in the native gamma-B crystallin mRNA sequence is nonoptimal for translation in *E. coli* due to the presence of a number of codons that are more commonly used in the native host, *Bos taurus*, than in *E. coli* (Table S1, available online). To study the role of synonymous mutations, we sought to compare the expression of this mRNA with a variant that would have a codon choice more optimal for *E. coli.* Common approaches to optimize heterologous gene expression include substitution of the majority of infrequently used codons with synonymous frequently used ones, often combined with the elimination of extreme guanine-cytosine (GC) content that could contribute to formation of stable mRNA secondary structures, thereby reducing translation efficiency (Gustafsson et al., 2004). However, maximizing the global speed of translation through these strategies often yields biologically inactive products that form insoluble aggregates and have to be refolded in order to recover their native structure and biological activity (Gustafsson et al., 2004; Rosano and Ceccarelli, 2009; Komar, 2009). Moreover, even when proteins expressed in heterologous hosts remain soluble, they are not necessarily natively folded (de Marco et al., 2005). We chose to use an alternative optimization strategy and compare expression of the original *B. taurus* sequence (denoted U for unoptimized) with that of a so-called "harmonized" (Angov et al., 2008) (H) gene variant. Harmonization aims to optimize translation by introducing synonymous codons that have the most similar usage frequencies in the native and target host organisms (Angov et al., 2008; Komar, 2009) (Figure 1A; Table S1). Rather than simply replacing rare codons with frequent ones, this strategy involves alteration of both rare and frequent codons and is expected to establish mRNA translation kinetics in a heterologous host that mimic those observed in the native host organism (Angov et al., 2008; Komar, 2009). Harmonization may lead to increased GC content and a decreased Codon Adaptation Index (CAI; a measure for synonymous codon-usage bias) (Sharp and Li, 1987), which is also widely used as a measure for the likelihood of success of heterologous gene expression (Sharp and Li, 1987; Gustafsson et al., 2004). Nevertheless, the translation yields obtained with harmonized mRNA sequences may be higher than those with mRNA sequences optimized with frequent codons (Angov et al., 2008). This suggests that harmonization may facilitate proper protein folding and production of proteins with structures closely similar to their native analogs (Komar, 2009).

Harmonization of the gamma-B crystallin mRNA (H variant) resulted in a codon-usage profile more similar to that of the native host organism (*B. taurus*) than the U variant (Figure 1A). It also led to an overall increase in the number of codons frequently used in *E. coli*, as indicated by an increase in the CAI score for expression in *E. coli* from 0.64 for the U variant to 0.78 for the H variant, closer to the 0.85 score for the native (U) mRNA translated in *B. taurus*. In *E. coli*, expression of the H variant yielded about 1.5- to 1.6-fold more full-length protein than expression of the U variant (Figures 1B and 1C), despite identical mRNA levels (Figure S1A). Compared to the U variant, expression of the H variant yielded more soluble protein (Figure 1D, middle and right panels) and fewer truncated protein products (Figure 1C, upper panel). Western blotting with anti-His antibodies recognizing the C-terminal 6xHis-tag of the recombinant gamma-B crystallin protein and MS analysis indicated that the observed "truncated" products were, at least in part, fragments of the full-length proteins (Figures 1E and S1B; Table S2). Because both the C-terminal His-tag and C-terminal peptides were detected by these methods, these fragments most likely arise from degradation rather than premature termination of translation. The different degradation patterns of the two protein variants suggested that they adopt different conformations. Importantly, MS analysis and sequencing of the full-length recombinant protein products (Figure S1C; Table S3) confirmed that the amino acid sequences of the U and H variants were identical.

To explore the effects of local synonymous codon substitutions at certain gene regions, to further reduce intracellular protein

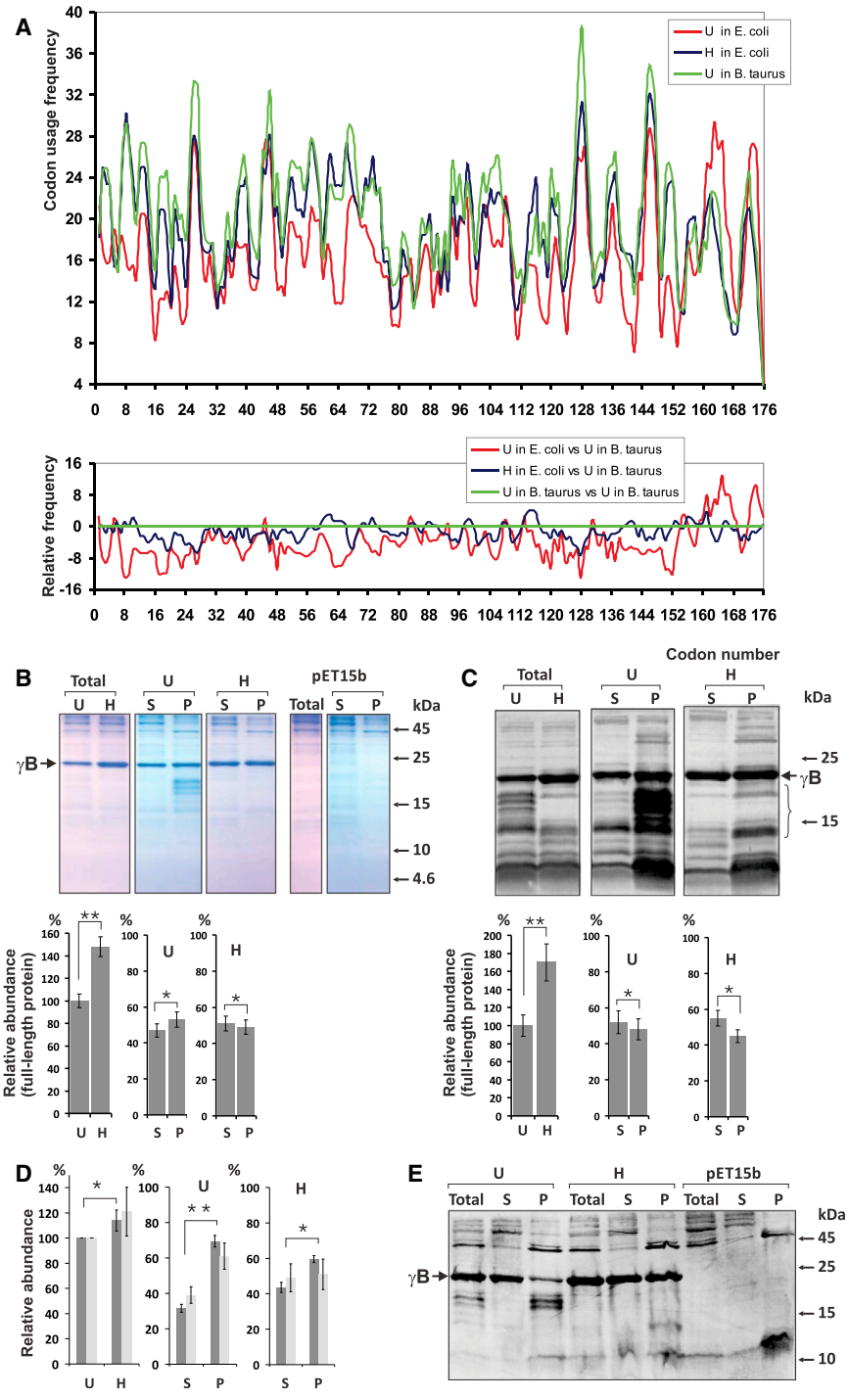


Figure 1. Effect of Synonymous Codon Choice on the Expression and Stability of Gamma-B Crystallin in *E. coli*

(A) Top: codon-frequency profiles of gamma-B crystallin variants in *B. taurus* (green) and *E. coli* U (red) and H (blue). Bottom: relative differences in usage frequencies.

(B) Expression of U and H variants of gamma-B crystallin (γB) in *E. coli*. Top panel: SDS-PAGE: total, soluble (S), and pellet (P) fractions, and pET15b empty vector control. Proteins were visualized by Coomassie Brilliant Blue (CBB) staining. Bottom panel: Quantification of full-length gamma-B crystallin and its distribution between S and P fractions. The total protein expression level (set to 100%) is the sum of S and P fractions.

(C) Western blotting using polyclonal anti-γB antibodies. The parenthesis indicates truncated products of the U variant expression that are not present or less abundant with H variant.

(D) Expression of U and H variants of gamma-B crystallin in *E. coli* based on quantitation of all bands detected by western blotting in Figure 1C (dark gray bars) or by ELISA (light gray bars). The amounts of protein in soluble and pellet fractions are represented as a fraction of total protein.

Error bars (B–D) represent SEM; *p < 0.05, **p < 0.01 by Student's t test.

(E) Detection of expression products of U and H variants that contain a C-terminal 6xHis-tag in total, soluble (S), and pellet (P) fractions of *E. coli* extracts by western blotting using monoclonal anti-polyhistidine antibody.

See also Figure S1 and Tables S1, S2, and S3.

degradation, and to increase protein solubility, we created a number of additional synonymous variants of gamma-B crystallin (Figure S2A). While protein solubility was improved up to ~70% (Figure S2B), the expression of some of these variants led to substantial levels of miscoding, in particular resulting in frequent replacement of Arg92 with Lys (Figures S2C and S2D). Thus, it could not be excluded that reduced degradation and increased solubility in some cases resulted from an altered protein sequence. We thus focused our further analysis on the U and H variants only, as the expression of these mRNAs did not give rise to any miscoding events. Together, these results indicate that synonymous codon content can lead to differences in the intracellular protein conformation and stability and, in certain cases, to increased levels of miscoding.

Purified Gamma-B Crystallin Synonymous Variants Attain Different Conformations, as Measured by 2D NMR

To investigate whether differences in protein conformations observed during intracellular U and H protein expression were preserved in the final structure of the isolated soluble proteins, we conducted NMR analyses of the U and H variants expressed in *E. coli*. Preparative ion-exchange chromatography (IEX) revealed one major peak for the U variant and two peaks for the H variant (H-P1 and H-P2, respectively; Figure 2A). Because all of the protein variants were of identical mass, the shift in the elution position between H-P1 and H-P2 suggested a lower isoelectric point (i.e., a

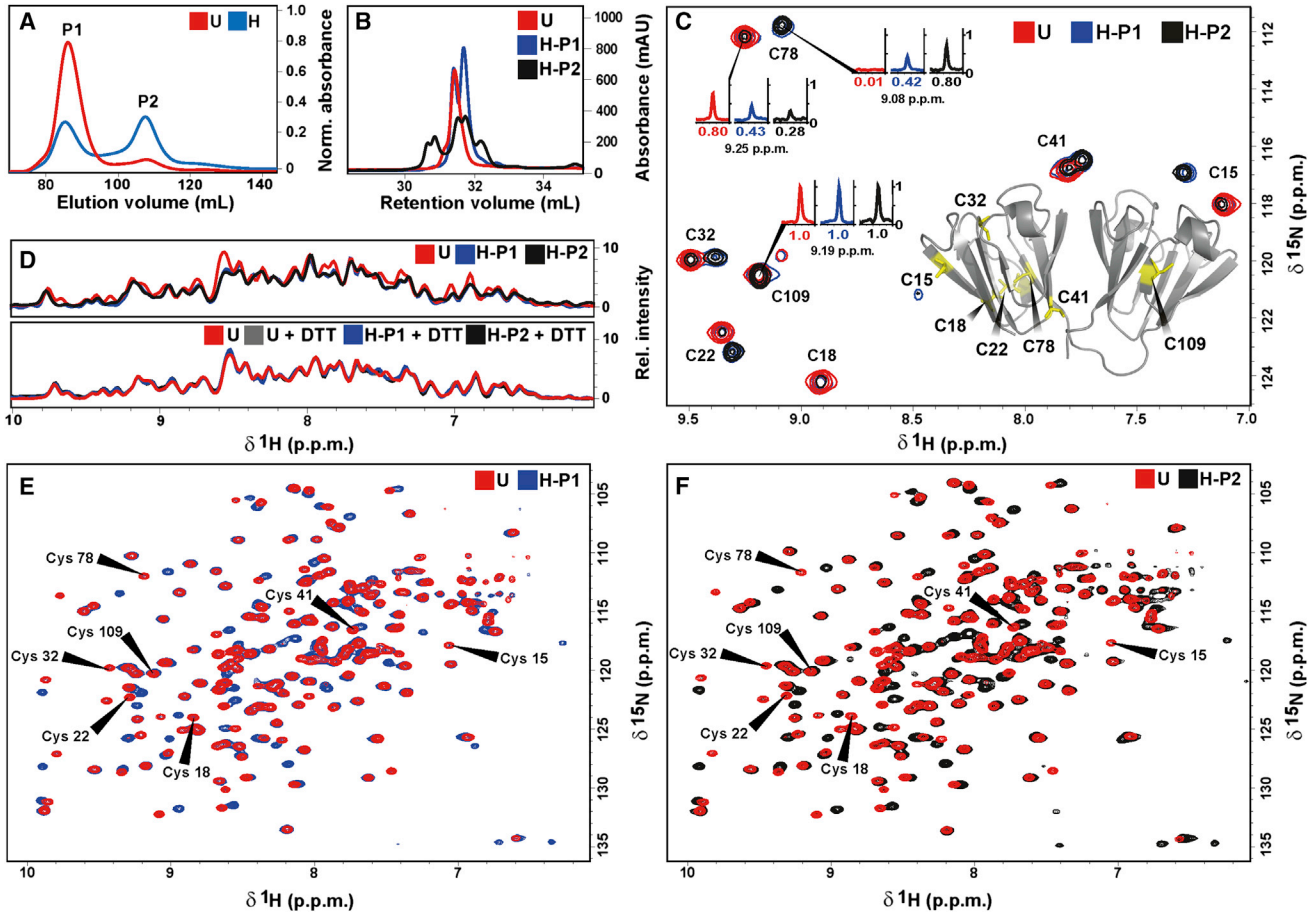


Figure 2. Physico-Chemical and Structural Properties of Gamma-B Crystallin U and H Variants Expressed in *E. coli*, as Determined by RP-HPLC and 2D-^{1}H-^{15}N-Correlation NMR

(A) Preparative pH-gradient ion-exchange chromatography (U, red; H, blue).

(B) Analytical RP-HPLC of ion-exchange fractions (U, red; H-P1, blue; H-P2, black).

(C) Overlay of 2D-^{1}H-^{15}N-correlated NMR backbone spectra of ^{15}N-cysteine-labeled U, H-P1, and H-P2. Insets show 1D rows to visualize differential cysteine-peak intensities, which are normalized against C109. Full list of peak integrals is presented in Table 1. Lower right quadrant: crystal structure of bovine gamma-B crystallin (PDB-ID 4GCR); cysteine residues highlighted.

(D) 1D summation of rows extracted from 2D-^{1}H-^{15}N backbone datasets recorded for U, H-P1, and H-P2 expressed in ^{15}N-labeled rich medium. Upper panel: spectra comparison. Lower panel: spectra after treatment of samples with 10 mM DTT.

(E and F) Overlay of 2D-^{1}H-^{15}N-correlated NMR spectra for U and H-P1 (E) and H-P2 (F). Addition of DTT resulted in full convergence to U-like spectra (Figure S3A). See also Table 1 and Figures S3B and S4.

different surface charge distribution) in H-P2. Further analysis by reversed-phase high-performance liquid chromatography (RP-HPLC) showed predominantly one peak for U, two peaks for H-P1, and a higher degree of heterogeneity for H-P2 (Figure 2B). Gamma-B crystallin contains seven cysteine residues, six of which are localized close to one another in the N-terminal domain (NTD) (Figure 2C). It cannot be excluded that these cysteines may become oxidized in the *E. coli* cytosol. Although disulfide bonds are rarely formed in *E. coli* cytoplasmic proteins, exceptions to this general rule do exist (Aslund et al., 1999; Jakob et al., 1999; Seras-Franzoso et al., 2012). Our NMR analysis of U, H-P1, and H-P2 revealed chemical shift differences in particular for the six cysteine residues at positions 15, 18, 22, 32, 41, and 78 (in NTD) (Figure 2C, inset; Table 1), both in 1D spectra (Figure 2D) and in 2D-^{1}H-^{15}N-correlated experiments (Figures 2E and 2F). Upon dithioerythritol (DTT) treatment, or refolding under reducing conditions (Figure S3A), all protein species converged to a single state identical to U, strongly suggesting that U represents the fully reduced state of the protein (Figures 2D and S3A). This notion is supported by the Cα/Cβ chemical-shift values of U-variant cysteine residues measured in 3D NMR experiments, which were characteristic for reduced cysteines. Therefore, purified U and H variants of gamma-B-crystallin expressed in *E. coli* differ structurally with respect to the oxidation state of cysteine residues within the NTD. This conclusion is further supported by the results of analysis of U, H-P1, and H-P2 protein spectra in which only cysteine residues were ^{15}N labeled (Figure 2C; Table 1). Importantly, peak doubling and chemical-shift changes

Table 1. Differential Cysteine Peak Intensities of 2D-^{1}H-^{15}N-Correlated NMR Backbone Spectra of ^{15}N-Cysteine-Labeled U, H-P1, and H-P2

Residue	U	H-P1	H-P2
Cys 109 (Reference)	1.0	1.0	1.0
Cys 15 (red)	1.05	0.53	0.30
Cys 15 (ox)	0.01	0.39	0.59
Cys 18 (red)	0.83	0.45	0.21
Cys 18 (ox)	N/A	N/A	N/A
Cys 22 (red)	0.87	0.40	0.22
Cys 22 (ox)	0.01	0.38	0.70
Cys 32 (red)	0.86	0.45	0.33
Cys 32 (ox)	0.03	0.30	0.50
Cys 41 (red)	0.97	0.55	0.41
Cys 41 (ox)	0.02	0.36	0.50
Cys 78 (red)	0.80	0.43	0.28
Cys 78 (ox)	0.01	0.42	0.80

See also Figure 2.

were also observed for other residues in 2D-^{1}H-^{15}N-correlated experiments (Figures 2E and 2F). This suggests that the oxidation of at least two cysteine residues within the NTD is a major cause of overall structural changes, which affect the conformation of the respective cysteines as well as other residues in the purified protein. These and previous data (Bloemendal et al., 2004), including a crystal structure of gamma-B crystallin (Najmudin et al., 1993), suggest the existence of two alternative conformations for Cys18 and Cys22 in the native protein purified from bovine eye lenses (Najmudin et al., 1993). Therefore, in terms of its oxidation state, the H variant produced in *E. coli* appears to be closer than the U variant produced in *E. coli* to the natural (N) protein variant purified from bovine eye lenses. Overnight incubation of U with catalytic amounts of Cu(II) under air supply (known to catalyze cysteine oxidation to cystine; Cavallini et al., 1969) resulted in a ^{15}N backbone spectrum similar to H-P1, albeit with a different peak-intensity distribution, suggesting that the respective peaks in the NMR spectra of the H protein reflect the differences in the oxidation state (Figure S3B).

Expression of U and H variants in the SHuffle T7 *E. coli* (NEB) strain (which constitutively expresses a chromosomal copy of the disulfide bond isomerase DsbC lacking its signal sequence, hence retaining the protein in the cytoplasm and thus facilitating disulfide bond formation in the *E. coli* cytoplasm; Lobstein et al., 2012) revealed overall similar differences for the two variants compared to the expression in BL21 strain (Figure S4). However, use of this strain when grown in minimal medium resulted in extensive protein misfolding (especially for the U variant; Figure S4A), possibly due to excessive oxidation and/or early formation of mixed disulfides leading to aggregation and precluding accurate NMR analysis under these conditions. Expression in SHuffle T7 *E. coli* grown in ^{15}N-rich medium resulted in yet another oxidized variant of U (with chemical shifts partially overlapping with, but still different from, those of the U variant expressed in the BL21 strain; Figure S4C). In contrast, the spectra of H-P1 were remarkably similar after expression in either strain of bacteria (Figures S4B and 4D). This suggests that the intracellular folding of H is more robust and less dependent on environmental conditions than the folding of U, possibly due to improved cotranslational folding kinetics resulting from harmonization of the codon-usage profile in the H variant. We thus believe that optimized translation kinetics can help to produce recombinant proteins more similar to their natural analogs.

Synonymous mRNA Variants Are Translated with Different Kinetics

The translational kinetics of the U and H variants was compared in a completely reconstituted high-performance, cell-free in vitro translation system from *E. coli* (Mittelstaet et al., 2013). In this system, proteins were N-terminally labeled during translation by using BodipyFL-Met-tRNA^{fMet} as the initiator tRNA (BOF-Met-tRNA^{fMet}). BOF and BOP (Bodipy 576/589), used in the experiments below, were previously utilized to study cotranslational helix formation within the ribosome (Woolhead et al., 2004) and have been validated as reporters for translation and cotranslational folding (Johnson, 2005). The ribosomes were synchronized at the translation initiation step, and translation elongation was initiated by the addition of aminoacyl-tRNAs, EF-Tu, and EF-G. After translating the mRNA to the 3′ end, the ribosomes remained bound to the mRNA, and the nascent peptide was not released from the ribosome due to the lack of a stop codon in the mRNA and absence of termination factors from the translation system. In this reconstituted system, the H mRNA was translated more rapidly than the U mRNA (Figures 3A and 3B). Full-length gamma-B crystallin protein was first detected after 43 s of H mRNA translation compared to 57 s for U mRNA (Figure 3B). The average protein-synthesis rate, calculated from the sum of the delay and synthesis times, was 2.2 aa/s for the H variant, compared to 1.8 aa/s for the U variant.

At any given time during translation, individual ribosomes in the reaction carry nascent peptides of different lengths (Figure 3A). At the beginning of translation, short peptides are prevalent, but at later time points, the full-size protein becomes predominant. Analysis of the transient accumulation of nascent peptides indicated that the U and H variants of gamma-B crystallin differ not only in global translation rate but also in local translation kinetics (Figure 3A). Specifically, more translational pausing (and hence greater prevalence of shorter nascent chains <70 aa) was observed during synthesis of the NTD of the U protein, compared to the H protein (Figures 3A and 3C). The cumulative lifetime (t) of nascent NTD chains before they were converted into longer peptides was 48 s for U, compared to 25 s for H. In addition, a smaller proportion of ribosomes continued translation past the NTD on the U mRNA than on the H mRNA (Figure 3C). Slow synthesis of the NTD is consistent with the abundance of rare codons in the corresponding part of the U mRNA (Figure 1A; codons 16–24) and explains the observed delayed synthesis of the full-length U protein. In contrast, accumulation of nascent chains with lengths of ~130–160 aa was detected during translation of the H variant, but not the U variant, suggesting ribosome pausing during synthesis of the C-terminal domain (CTD) of the H variant.

To improve the time resolution of translation, we performed an experiment similar to that shown in Figure 3A in which we

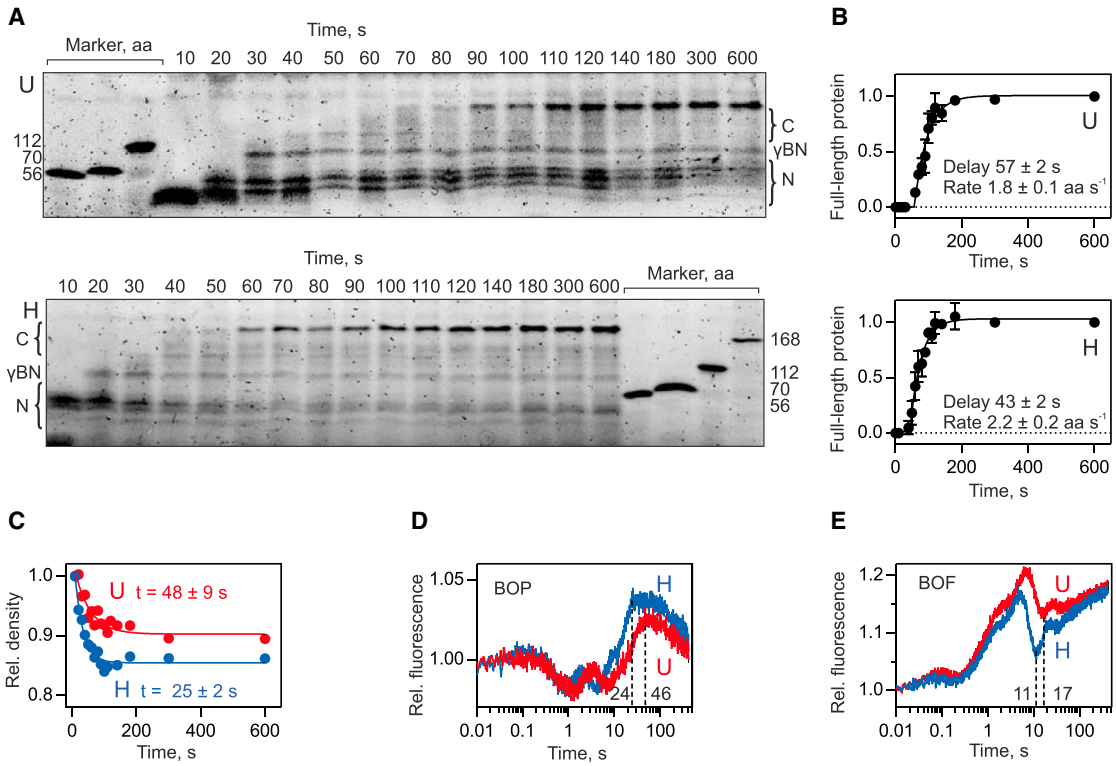


Figure 3. Different Translation Kinetics of U and H Variants in a Fully Reconstituted *E. coli* Cell-free Translation System

(A) Accumulation and size distribution of U (upper panel) and H (lower panel) in vitro translation products. Peptides were separated by SDS-PAGE and visualized by the fluorescence of the BOF label attached to the N terminus of the peptides. N and C indicate peptides arising during translation of the NTD and CTD, respectively; γBN is the fragment corresponding to the NTD.

(B) Kinetic analysis of accumulation of full-length gamma-B crystallin. "Delay" is the time before appearance of the full-length product. "Rate" is the average translation rate (aa/s). Error bars show SD calculated from n = 3 replicates. We further tested whether the difference in delay times for H and U is statistically equivalent to a single shared parameter for the delay time in synthesis of γB crystallin from the U and H mRNA (null hypothesis) or, alternatively, if two independent parameters for the delay for U or H are justified. According to the extra sum-of-squares F test, the null hypothesis of an identical delay time for U/H is rejected with a significance of ****$p < 0.0001$.

(C) Lifetime of translation intermediates corresponding to the NTD upon translation of U (red) and H (blue) mRNAs. The total intensity of all bands indicated as N in Figure 3A (U and H, respectively) after 10 s of translation was set to 1.

(D) Stopped-flow kinetics of synthesis and movement through the ribosome exit tunnel of U (red) and H (blue) nascent chains monitored by a fluorescence reporter (BOP) at the N terminus of the nascent peptides. The maximum delay in translation of the U sequence relative to H is indicated. Apparent variations in the height of fluorescence changes between U and H are due to differences in the accumulation of the respective nascent peptides resulting from the altered translation rates.

(E) Same as (D), except using BOF fluorescence as the reporter.

See also Figure S5.

monitored fluorescence changes of the BOP reporter attached to the N terminus of the nascent peptide moving through the ribosomal exit tunnel in real time (Figure 3D). This approach also demonstrated that the U variant was translated more slowly than the H variant due to an approximately 20 s delay between the H and U variants (Figure 3D). The difference in kinetics appeared early (detected after 1 s of translation) and became more pronounced through the duration of NTD synthesis (compare Figures 3C and 3D), consistent with translational pausing due to rare codons in the NTD of the U variant.

Finally, when BOF at the N terminus of the growing nascent peptide was monitored, the fluorescence changes for the U and H variants were identical initially but started to deviate after about 1–2 s of translation, again coinciding with the presence of rare codons at the beginning of the U mRNA (Figure 3E). The 6 s difference in synthesis between the U and H accounts for a delay in the production of the initial parts of the NTD; differences in translation of the C-terminal portions of the protein are not reported by BOF fluorescence. Taken together, these data indicate that the synonymous codon substitutions that distinguish the U and H mRNA variants affected the global and local kinetics of gamma-B crystallin translation. It cannot be excluded that the observed alterations in translation kinetics are not due to differences in decoding rates of tRNA selection, but rather are related to possible differences in mRNA secondary structure or other factors unknown so far. However, regardless of the underlying molecular mechanism(s), it is clear that the choice between synonymous codons can affect translation kinetics.

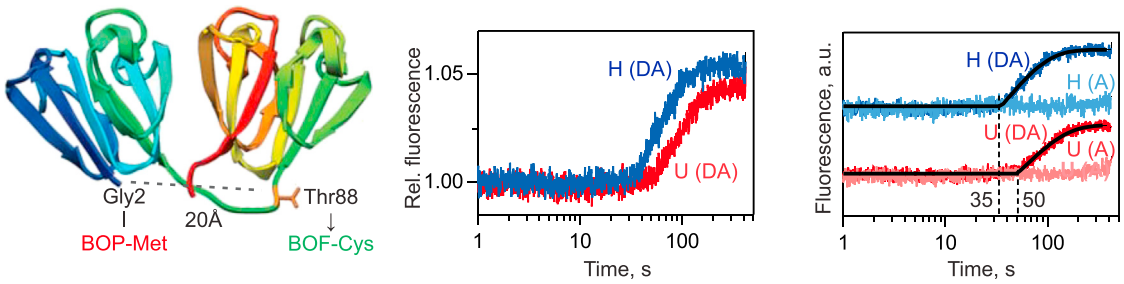


Figure 4. Cotranslational Folding of the NTD Monitored in Real Time by FRET

Left panel: positions of the donor (BOF) and acceptor (BOP) dyes in the structure of gamma-B crystallin. Middle and right panels: time-resolved folding of U (red) and H (blue) peptides monitored by FRET between BOP-Met at position 1 and BOF-Cys at position 88 in the stopped-flow apparatus. DA, both donor and acceptor dyes were present; A, control in the absence of the donor. Middle panel: direct comparison of FRET due to folding for the U and H variants. Right panel: FRET signal versus acceptor direct excitation at donor wavelength; for better comparison, the traces with and without the donor for the U or H variants, respectively, were adjusted to the same starting level; the H traces are arbitrarily shifted from the U traces for visual clarity. Time courses were evaluated using a two-step model comprising a delay phase, which shows no change in fluorescence, and an exponential phase corresponding to a monomolecular folding reaction using GraphPad Prism. Delay times are 35 ± 0.3 s for H and 50 ± 0.5 s for U. The folding times determined by exponential fitting after delay are 39 ± 1 s for H and 59 ± 1 s for U. See also Figure S5.

Synonymous Codon Usage Modulates the Kinetics of Cotranslational Folding

Next we examined whether the observed differences in translation kinetics led to differences in cotranslational folding of the U and H nascent polypeptides by performing real-time FRET measurements (Figure 4). The FRET acceptor BOP was attached at the N terminus of nascent peptides by priming translation with BOP-Met-tRNA^{fMet}. To incorporate the fluorescence donor, BOF, we introduced a UAG stop codon at position 88 of the U and H mRNAs, which was decoded by a modified amber suppressor BOF-Cys-tRNA. Thus, the fluorescence donor was introduced cotranslationally. Importantly, decoding of the UAG stop codon by BOF-Cys-tRNA did not affect the translation kinetics of either mRNA (Figure S5). In the unfolded protein, donor and acceptor fluorophores are too far apart for FRET to occur. Upon protein compaction during folding, BOP and BOF come into proximity, resulting in FRET (Figure 4). The onset of folding was significantly delayed for the U variant, compared to the H variant (50 s versus 35 s, respectively), consistent with the observed differences in H and U translation kinetics (Figures 3A–3C). Given that the synthesis of the NTD was completed after 20 s (H) or 30 s (U) (Figures 3A and S5), the delay between NTD synthesis and folding was 15 s and 20 s for H and U, respectively. At the respective translation rates of 2.2 aa/s and 1.8 aa/s (Figure 3B), this delay should allow the ribosome to synthesize 33–36 aa past the NTD, which would fill the polypeptide exit tunnel of the ribosome, such that the NTD itself is extruded from the tunnel. This suggests that global folding of the NTD (H or U), as reported by FRET, occurred cotranslationally, shortly after the corresponding portion of the polypeptide emerged from the exit tunnel. Formation of locally restricted structural elements within the exit tunnel cannot be excluded, however. Thereafter, the average time of NTD folding was 59 s for U and 39 s for H. Notably, the 20 s difference represents the divergence in the H and U folding rates after the corresponding portion of the protein has emerged from the ribosome. This reveals the specific effect of synonymous codon replacement on the kinetics of cotranslational protein folding beyond the differences in production times for nascent chains of the same length. The overall effect, taking into account also the additional delay time for the NTD and linker synthesis, gives a total folding time of 109 s for U and 74 s for H.

Gamma-B Crystallin Synonymous Variants Display Differential Protease Sensitivity In Vitro

To further distinguish between folded and unfolded states of the protein, we probed the conformation of ribosome-bound nascent chains using pulse proteolysis with PK (Figures 5A–5C). Isolated recombinant or natural gamma-B crystallin is completely resistant to PK digestion (Figures S6A–S6C), while partial denaturation with 3M guanidinium chloride (GuHCl) results in complete gamma-B crystallin degradation by PK (Figures S6A–S6C). PK treatment of ribosome-bound nascent chains produced upon in vitro translation of U or H mRNAs resulted in two major products: (1) full-length protein and (2) a fragment of about 80–90 residues in length, corresponding in size to the NTD (Figures 5A and S6D). While NTD folding was too rapid to evaluate potential differences in the folding rates of U and H variants by proteolysis, the approach did reveal that formation of a PK-resistant full-length nascent polypeptide occurred on a significantly shorter timescale with H than with U (Figure 5B). Because the synthesis of the nascent chains is essentially complete after 2 min, protection against PK digestion must represent variations in folding dynamics of the emerging nascent U and H proteins. Notably, because the ribosome tunnel occludes about 30–40 aa of the protein from the C terminus (Wilson and Beckmann, 2011), folding of the CTD is not completed until the entire nascent chain is released from the ribosome. Therefore, the appearance of the full-length protein after PK digestion is an indication of the initial stages of compaction of the emerging CTD. Overall, these results show that synonymous codon replacements in gamma-B crystallin resulted in altered conformational dynamics of ribosome-bound nascent chains.

To examine whether the folding states of the U and H variants are also different after their release from the ribosome, we tested the PK resistance of gamma-B crystallin chains released from

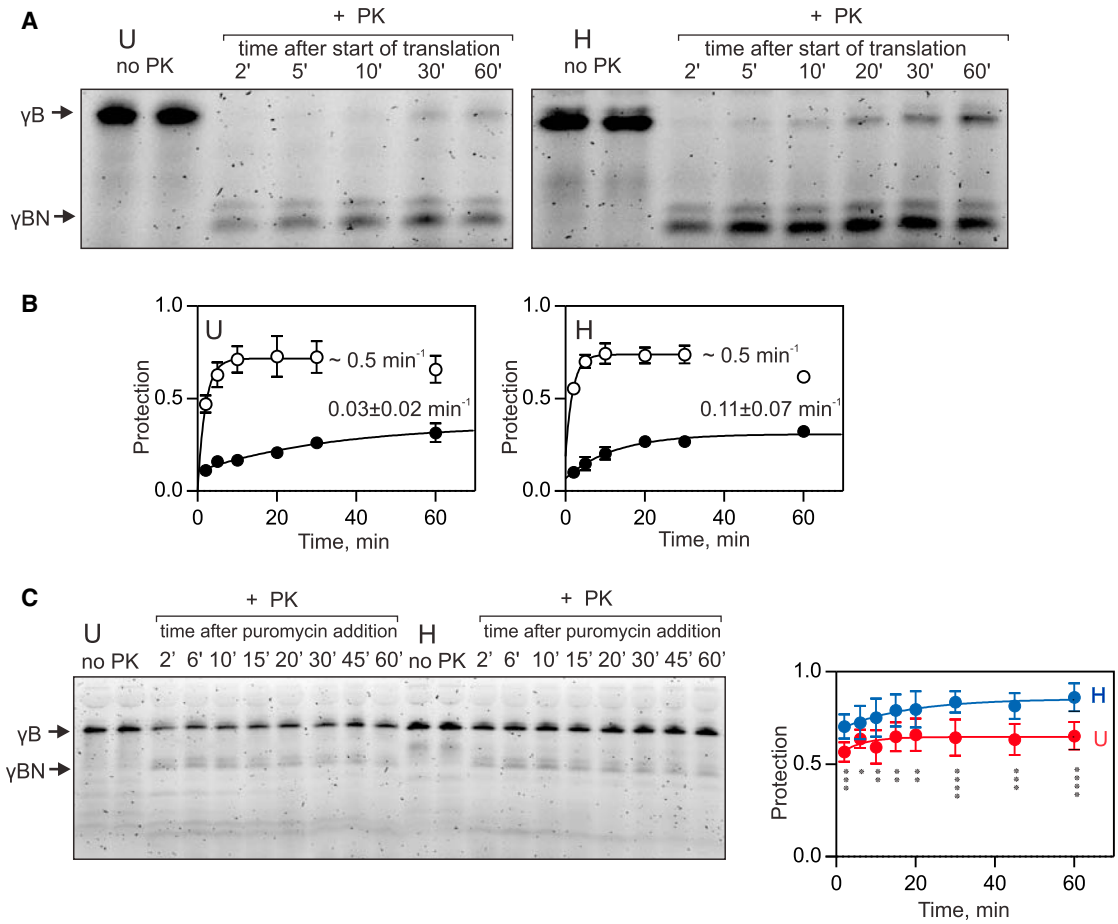


Figure 5. Sensitivity of U and H Polypeptide Chains to Pulse Proteolysis

(A) Proteolysis of ribosome-bound gamma-B crystallin chains assessed by SDS-PAGE. At different time points of in vitro translation, chains were digested with 5.4 pmol PK for 2 min at 37°C. γB (full-length protein) and γBN (N-terminal domain) protease-resistant products are indicated by arrows.

(B) Time courses of accumulation of PK-resistant N-terminal (open circles) and full-length (closed circles) products relative to the undigested protein. The last data point (60 min) was excluded from the exponential fitting, as it represents a decrease in the portion of PK-resistant NTD due to accumulation of the full-length protein. Error bars show the SEM for n = 7 replicates.

(C) Left panel: PK proteolysis of puromycin-released chains after 3 min of translation. Right panel: quantification of PK resistance of released U (red symbols) and H (blue symbols) products from the data shown in the left panel. Error bars show the SD for n = 9 replicates. *p < 0.05, **p < 0.01, ***p < 0.005, ****p < 0.0001 by Student's two-tailed unpaired t test. See also Figure S6.

the ribosome by puromycin (Figure 5C). Upon release, both U and H proteins rapidly fold into a PK-resistant conformation similar to that of the purified recombinant protein (Figure S6). The yield of folded U protein did not change significantly with time; however, with H, another 16% of the protein became more PK resistant within 30 min after release from the ribosome. This suggests that, after release from the ribosome, the conformational ensemble of H is sufficiently preorganized to allow more efficient posttranslational folding. The final yield of PK-resistant protein was greater for the H variant (86%) than for the U variant (65%). The peptides that did not fold properly were PK sensitive and likely represent part of the protease-sensitive protein found in the insoluble fraction in vivo (Figure 1). Thus, differences in velocities of cotranslational folding resulted in different conformational states of U and H variants on the ribosome and after release from the ribosome.

DISCUSSION

Our analysis of the effects of synonymous codon choice on the translation of gamma-B crystallin mRNAs shows that codon bias alters local and global translation rates and results in the formation of alternative conformations of the nascent protein on the ribosome and in solution after the release of the completed protein (Figure 6). Rare codons at the beginning of the U mRNA slow down the synthesis of the NTD and lead to significantly slower global NTD folding of the U variant compared to the H variant. It is possible that slow translation allows for partitioning of elemental folding events to alternative pathways (O'Brien et al., 2014). The spectrum of conformations attained by the nascent chains on the ribosome may determine different trajectories of their folding in the vestibule region (Tu et al., 2014; Lu and Deutsch, 2014; Holtkamp et al., 2015) of the polypeptide exit

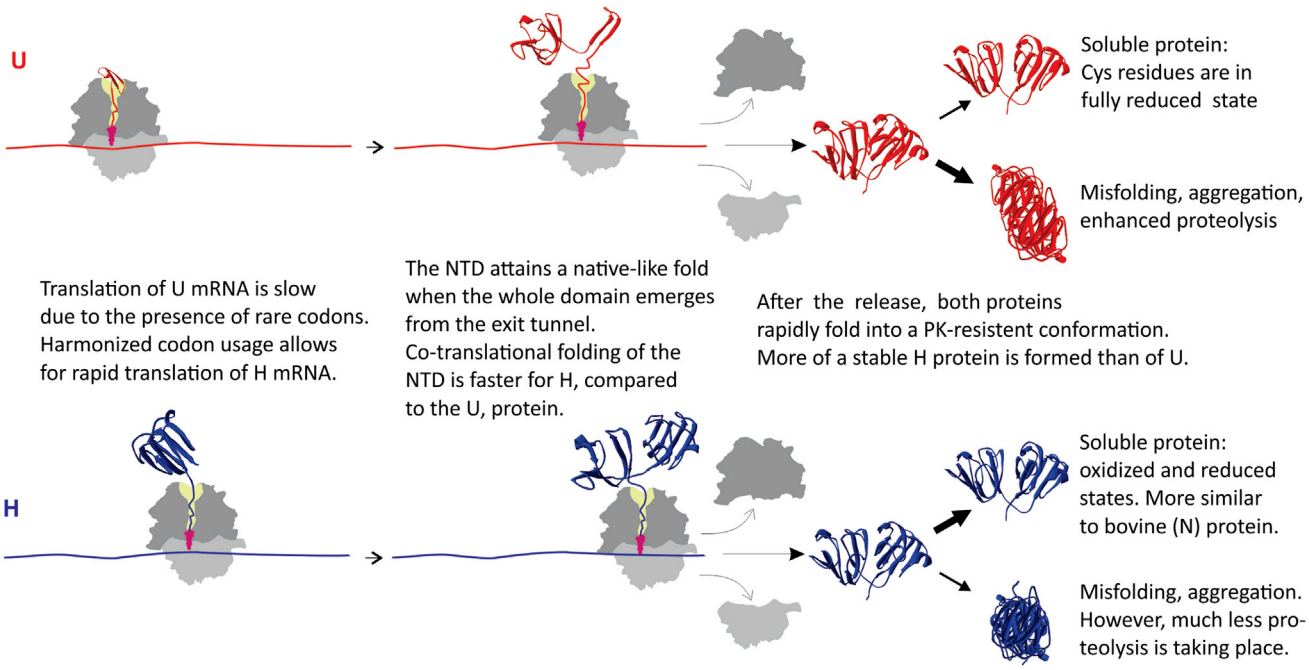


Figure 6. Synonymous Codon Usage Directs Cotranslational Folding toward Different Protein Conformations

tunnel (Wilson and Beckmann, 2011) and/or after the chain emerges from the tunnel and is released from the ribosome. This, in turn, may change the local environment and orientation of cysteine residues leading to distinct patterns of disulfide bridges and affecting the resulting protein's stability and propensity to form aggregates.

Alternatively, the slower folding of U compared to H may be due to the decreased rate of CTD synthesis, which is manifested, e.g., by the different rates of translation past the NTD for U and H (Figures 3A and 3C). In this case, folding of the NTD may be affected by interaction(s) with the ribosome (Kaiser et al., 2011; Holtkamp et al., 2015) and the rate with which the NTD moves away from the ribosome surface and/or altered interactions with the emerging CTD and the dynamics of the sampling between NTD and CTD. While such mechanistic details remain to be clarified, our results clearly demonstrate that different synonymous codon usage influences the kinetics of translation, resulting in different cotranslational folding trajectories and ultimately different final conformations of the protein. Thus, synonymous codon usage serves as a secondary code that guides in vivo protein folding and constitutes an additional source of conformational variability of proteins.

EXPERIMENTAL PROCEDURES

Plasmids, *E. coli* Cells, and Protein Expression

Bovine gamma-B crystallin cDNA (Hay et al., 1987) was a gift of Dr. J. Mark Petrash (University of Colorado). For expression in *E. coli*, the unoptimized (U) gamma-B crystallin cDNA was PCR amplified, fused to a C-terminal 6xHis-tag, and cloned into the pET15b vector (Novagen) via NcoI and XhoI sites. Variants H of the gamma-B crystallin gene were chemically synthesized by GeneArt AG (Life Technologies) and cloned into pET15b to obtain a construct with 5′ and 3′ untranslated regions identical to those of the U variant. Analytical-scale protein expression was performed in *E. coli* cells as described in Supplemental Experimental Procedures.

Analysis of Protein Expression

For analysis of the relative expression levels of the U and H variants and comparison of soluble (supernatant) and insoluble (pellet) protein levels, Coomassie Brilliant Blue (CBB) staining of proteins separated by Tris-Tricine SDS-PAGE (Schägger and von Jagow, 1987), western blotting (WB), and enzyme-linked immunosorbent assay (ELISA) were used. See Supplemental Experimental Procedures for detailed descriptions.

In Vitro Translation

Single-turnover in vitro translation was carried out in a reconstituted minimum-translation system (Doerfel et al., 2013; Mittelstaet et al., 2013). Components of the reconstituted in vitro translation system were prepared as described (Mittelstaet et al., 2013); see Supplemental Experimental Procedures for details.

Pulse Proteolysis of In Vitro Translation Products

Pulse proteolysis of translation products was carried out using either ribosome-bound nascent chains or the released chains obtained after puromycin (1.3 mM final concentration) treatment. Routinely, 25 μl translation reaction mix was subjected to digestion with PK as indicated in Figures 4 and S6 for 2 min at 37°C. Following digestion, PMSF was added to a final concentration of 10 mM to inactivate the protease, and samples were immediately flash frozen in liquid nitrogen. Samples were thawed and subjected to RNaseA treatment (5 μL, 2 mg/mL RNase A, 30 min at 37°C). After RNase digestion, 2× Tris-tricine loading buffer was added, and peptides were resolved by Tris-tricine SDS-PAGE as described above.

Protein Expression and Purification for NMR Spectroscopy

Protein expression was performed in BL21(DE3) and Shuffle T7 *E. coli* strains. For uniform ^{15}N labeling, a commercial rich medium was used. $^{15}N/^{13}C$-labeled protein for backbone assignment was expressed in M9 minimal medium containing [^{15}N]NH_4Cl and [U-^{13}C]glucose. For selective ^{15}N-cysteine labeling, a modified M9 medium containing [^{15}N]cysteine was used, as described (Muchmore et al., 1989). Cells were harvested by centrifugation,

and lysates were purified by nickel-affinity chromatography (HisTrap HP Ni-NTA), size-exclusion chromatography (Superdex 75 pg), and ion-exchange chromatography (HiTrap Q XL). NMR buffer was as follows: 50 mM Tris (pH 8.0), 200 mM NaCl, 10% D_2O, 0.1% DSS. Samples for C-4 reverse-phase HPLC were taken before and after ion-exchange chromatography. See Supplemental Experimental Proceduresfor details.

NMR Spectroscopy

NMR experiments were conducted on Bruker spectrometers (AV600–AV900). Processing and analysis of the data were performed using the software programs TopSpin 2.1–3.2 (Bruker BioSpin, Rheinstetten) and Cara 1.9.0 (Rochus L.J. Keller). 2D ^{1}H-^{15}N backbone spectra were recorded using the BEST-TROSY experiment. Backbone-resonance assignment was performed using standard 3D pulse sequences. Sample concentrations were 250–350 μM after expression in ^{15}N-labeled rich media and 0.8–1.2 mM after expression in minimal media. For further details, see Supplemental Experimental Procedures (Tables S4–S7).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2016.01.008.

AUTHOR CONTRIBUTIONS

H.S., M.V.R., and A.A.K. supervised the research. F.B., S.J., M.T., J.M., and F.K. performed the research. F.B., S.J., M.T., J.M, F.K., H.S., M.V.R., and A.A.K. analyzed data. H.S., M.V.R., and A.A.K. wrote the manuscript, and all authors contributed to the final version of the manuscript.

ACKNOWLEDGMENTS

We thank Drs. Jun Qin and Xianqing Song for many helpful discussions; Patricia Stanhope Baker for help with manuscript preparation; Shrutee Jakhanwal for help in preparing functionally active $tRNA^{Cys}$; and Olaf Geintzer, Christina Kothe, Sandra Kappler, Anna Pfeifer, Theresia Uhlendorf, Tanja Wiles, and Franziska Hummel for expert technical assistance. This work was supported by grants of the Human Frontier Science Program (grant #RGP0024/2010 to A.A.K., M.V.R., and H.S.), 13GRNT17070025 (American Heart Association), 1R15HL121779-01A1 (NIH; to A.A.K.), and the Deutsche Forschungsgemeinschaft (FOR1805 to M.V.R.).

Received: September 10, 2015
Revised: November 12, 2015
Accepted: December 24, 2015
Published: February 4, 2016

REFERENCES

Angov, E., Hillier, C.J., Kincaid, R.L., and Lyon, J.A. (2008). Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. PLoS ONE *3*, e2189.

Aslund, F., Zheng, M., Beckwith, J., and Storz, G. (1999). Regulation of the OxyR transcription factor by hydrogen peroxide and the cellular thiol-disulfide status. Proc. Natl. Acad. Sci. USA *96*, 6161–6165.

Bloemendal, H., de Jong, W., Jaenicke, R., Lubsen, N.H., Slingsby, C., and Tardieu, A. (2004). Ageing and vision: structure, stability and function of lens crystallins. Prog. Biophys. Mol. Biol. *86*, 407–485.

Buchan, J.R., and Stansfield, I. (2007). Halting a cellular production line: responses to ribosomal pausing during translation. Biol. Cell *99*, 475–487.

Cabrita, L.D., Dobson, C.M., and Christodoulou, J. (2010). Protein folding on the ribosome. Curr. Opin. Struct. Biol. *20*, 33–45.

Cavallini, D., De Marco, C., Duprè, S., and Rotilio, G. (1969). The copper catalyzed oxidation of cysteine to cystine. Arch. Biochem. Biophys. *130*, 354–361.

Chaney, J.L., and Clark, P.L. (2015). Roles for synonymous codon usage in protein biogenesis. Annu. Rev. Biophys. *44*, 143–166.

Clarke, T.F., 4th, and Clark, P.L. (2008). Rare codons cluster. PLoS ONE *3*, e3412.

Dana, A., and Tuller, T. (2014). The effect of tRNA levels on decoding times of mRNA codons. Nucleic Acids Res. *42*, 9171–9181.

de Marco, A., Vigh, L., Diamant, S., and Goloubinoff, P. (2005). Native folding of aggregation-prone recombinant proteins in Escherichia coli by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones. Cell Stress Chaperones *10*, 329–339.

Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M.V. (2013). EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. Science *339*, 85–88.

Drummond, D.A., and Wilke, C.O. (2008). Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. Cell *134*, 341–352.

Gardin, J., Yeasmin, R., Yurovsky, A., Cai, Y., Skiena, S., and Futcher, B. (2014). Measurement of average decoding rates of the 61 sense codons in vivo. eLife *3*, 03735.

Gloge, F., Becker, A.H., Kramer, G., and Bukau, B. (2014). Co-translational mechanisms of protein maturation. Curr. Opin. Struct. Biol. *24*, 24–33.

Gustafsson, C., Govindarajan, S., and Minshull, J. (2004). Codon bias and heterologous protein expression. Trends Biotechnol. *22*, 346–353.

Hartl, F.U., and Hayer-Hartl, M. (2009). Converging concepts of protein folding in vitro and in vivo. Nat. Struct. Mol. Biol. *16*, 574–581.

Hay, R.E., Woods, W.D., Church, R.L., and Petrash, J.M. (1987). cDNA clones encoding bovine gamma-crystallins. Biochem. Biophys. Res. Commun. *146*, 332–338.

Hershberg, R., and Petrov, D.A. (2008). Selection on codon bias. Annu. Rev. Genet. *42*, 287–299.

Holtkamp, W., Kokic, G., Jäger, M., Mittelstaet, J., Komar, A.A., and Rodnina, M.V. (2015). Cotranslational protein folding on the ribosome monitored in real time. Science *350*, 1104–1107.

Hu, S., Wang, M., Cai, G., and He, M. (2013). Genetic code-guided protein synthesis and folding in *Escherichia coli*. J. Biol. Chem. *288*, 30855–30861.

Hunt, R.C., Simhadri, V.L., Iandoli, M., Sauna, Z.E., and Kimchi-Sarfaty, C. (2014). Exposing synonymous mutations. Trends Genet. *30*, 308–321.

Ikemura, T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. Mol. Biol. Evol. *2*, 13–34.

Ingolia, N.T. (2014). Ribosome profiling: new views of translation, from single codons to genome scale. Nat. Rev. Genet. *15*, 205–213.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science *324*, 218–223.

Jakob, U., Muse, W., Eser, M., and Bardwell, J.C. (1999). Chaperone activity with a redox switch. Cell *96*, 341–352.

Johnson, A.E. (2005). The co-translational folding and interactions of nascent protein chains: a new approach using fluorescence resonance energy transfer. FEBS Lett. *579*, 916–920.

Kaiser, C.M., Goldman, D.H., Chodera, J.D., Tinoco, I., Jr., and Bustamante, C. (2011). The ribosome modulates nascent protein folding. Science *334*, 1723–1727.

Kim, S.J., Yoon, J.S., Shishido, H., Yang, Z., Rooney, L.A., Barral, J.M., and Skach, W.R. (2015). Protein folding. Translational tuning optimizes nascent protein folding in cells. Science *348*, 444–448.

Kimchi-Sarfaty, C., Oh, J.M., Kim, I.W., Sauna, Z.E., Calcagno, A.M., Ambudkar, S.V., and Gottesman, M.M. (2007). A "silent" polymorphism in the MDR1 gene changes substrate specificity. Science *315*, 525–528.

Komar, A.A. (2009). A pause for thought along the co-translational folding pathway. Trends Biochem. Sci. *34*, 16–24.

Komar, A.A., and Jaenicke, R. (1995). Kinetics of translation of gamma B crystallin and its circularly permutated variant in an in vitro cell-free system: possible relations to codon distribution and protein folding. FEBS Lett. *376*, 195–198.

Komar, A.A., Lesnik, T., and Reiss, C. (1999). Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. FEBS Lett. *462*, 387–391.

Kramer, G., Boehringer, D., Ban, N., and Bukau, B. (2009). The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. Nat. Struct. Mol. Biol. *16*, 589–597.

Li, G.W., Oh, E., and Weissman, J.S. (2012). The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. Nature *484*, 538–541.

Lobstein, J., Emrich, C.A., Jeans, C., Faulkner, M., Riggs, P., and Berkmen, M. (2012). SHuffle, a novel Escherichia coli protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microb. Cell Fact. *11*, 56.

Lu, J., and Deutsch, C. (2014). Regional discrimination and propagation of local rearrangements along the ribosomal exit tunnel. J. Mol. Biol. *426*, 4061–4073.

Mittelstaet, J., Konevega, A.L., and Rodnina, M.V. (2013). A kinetic safety gate controlling the delivery of unnatural amino acids to the ribosome. J. Am. Chem. Soc. *135*, 17031–17038.

Muchmore, D.C., McIntosh, L.P., Russell, C.B., Anderson, D.E., and Dahlquist, F.W. (1989). Expression and nitrogen-15 labeling of proteins for proton and nitrogen-15 nuclear magnetic resonance. Methods Enzymol. *177*, 44–73.

Najmudin, S., Nalini, V., Driessen, H.P., Slingsby, C., Blundell, T.L., Moss, D.S., and Lindley, P.F. (1993). Structure of the bovine eye lens protein gammaB(gammaII)-crystallin at 1.47 A. Acta Crystallogr. D Biol. Crystallogr. *49*, 223–233.

O'Brien, E.P., Ciryam, P., Vendruscolo, M., and Dobson, C.M. (2014). Understanding the influence of codon translation rates on cotranslational protein folding. Acc. Chem. Res. *47*, 1536–1544.

Pechmann, S., and Frydman, J. (2013). Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. Nat. Struct. Mol. Biol. *20*, 237–243.

Pechmann, S., Willmund, F., and Frydman, J. (2013). The ribosome as a hub for protein quality control. Mol. Cell *49*, 411–421.

Pechmann, S., Chartron, J.W., and Frydman, J. (2014). Local slowdown of translation by nonoptimal codons promotes nascent-chain recognition by SRP *in vivo*. Nat. Struct. Mol. Biol. *21*, 1100–1105.

Pedersen, M., Nissen, S., Mitarai, N., Lo Svenningsen, S., Sneppen, K., and Pedersen, S. (2011). The functional half-life of an mRNA depends on the ribosome spacing in an early coding region. J. Mol. Biol. *407*, 35–44.

Plotkin, J.B., and Kudla, G. (2011). Synonymous but not the same: the causes and consequences of codon bias. Nat. Rev. Genet. *12*, 32–42.

Pop, C., Rouskin, S., Ingolia, N.T., Han, L., Phizicky, E.M., Weissman, J.S., and Koller, D. (2014). Causal signals between codon bias, mRNA structure, and the efficiency of translation and elongation. Mol. Syst. Biol. *10*, 770.

Presnyak, V., Alhusaini, N., Chen, Y.H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R., and Coller, J. (2015). Codon optimality is a major determinant of mRNA stability. Cell *160*, 1111–1124.

Rosano, G.L., and Ceccarelli, E.A. (2009). Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted Escherichia coli strain. Microb. Cell Fact. *8*, 41–50.

Sander, I.M., Chaney, J.L., and Clark, P.L. (2014). Expanding Anfinsen's principle: contributions of synonymous codon selection to rational protein design. J. Am. Chem. Soc. *136*, 858–861.

Sauna, Z.E., and Kimchi-Sarfaty, C. (2011). Understanding the contribution of synonymous mutations to human disease. Nat. Rev. Genet. *12*, 683–691.

Schägger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. *166*, 368–379.

Seras-Franzoso, J., Affentranger, R., Ferrer-Navarro, M., Daura, X., Villaverde, A., and García-Fruitós, E. (2012). Disulfide bond formation and activation of Escherichia coli β-galactosidase under oxidizing conditions. Appl. Environ. Microbiol. *78*, 2376–2385.

Sharp, P.M., and Li, W.H. (1987). The codon Adaptation Index–a measure of directional synonymous codon usage bias, and its potential applications. Nucleic Acids Res. *15*, 1281–1295.

Sharp, P.M., Emery, L.R., and Zeng, K. (2010). Forces that influence the evolution of codon bias. Philos. Trans. R. Soc. Lond. B Biol. Sci. *365*, 1203–1212.

Tsai, C.J., Sauna, Z.E., Kimchi-Sarfaty, C., Ambudkar, S.V., Gottesman, M.M., and Nussinov, R. (2008). Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. J. Mol. Biol. *383*, 281–291.

Tu, L., Khanna, P., and Deutsch, C. (2014). Transmembrane segments form tertiary hairpins in the folding vestibule of the ribosome. J. Mol. Biol. *426*, 185–198.

Waudby, C.A., Launay, H., Cabrita, L.D., and Christodoulou, J. (2013). Protein folding on the ribosome studied using NMR spectroscopy. Prog. Nucl. Magn. Reson. Spectrosc. *74*, 57–75.

Wilson, D.N., and Beckmann, R. (2011). The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. Curr. Opin. Struct. Biol. *21*, 274–282.

Woolhead, C.A., McCormick, P.J., and Johnson, A.E. (2004). Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. Cell *116*, 725–736.

Yu, C.H., Dang, Y., Zhou, Z., Wu, C., Zhao, F., Sachs, M.S., and Liu, Y. (2015). Codon usage influences the local rate of translation elongation to regulate co-translational protein folding. Mol. Cell *59*, 744–754.

Zhang, G., and Ignatova, Z. (2011). Folding at the birth of the nascent chain: coordinating translation with co-translational folding. Curr. Opin. Struct. Biol. *21*, 25–31.

Zhang, G., Hubalewska, M., and Ignatova, Z. (2009). Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. Nat. Struct. Mol. Biol. *16*, 274–280.

Zhou, M., Guo, J., Cha, J., Chae, M., Chen, S., Barral, J.M., Sachs, M.S., and Liu, Y. (2013). Non-optimal codon usage affects expression, structure and function of clock protein FRQ. Nature *495*, 111–115.

Molecular Cell

Supplemental Information

Synonymous Codons Direct Cotranslational Folding toward Different Protein Conformations

Florian Buhr, Sujata Jha, Michael Thommen, Joerg Mittelstaet, Felicitas Kutz, Harald Schwalbe, Marina V. Rodnina, and Anton A. Komar

Figure S1

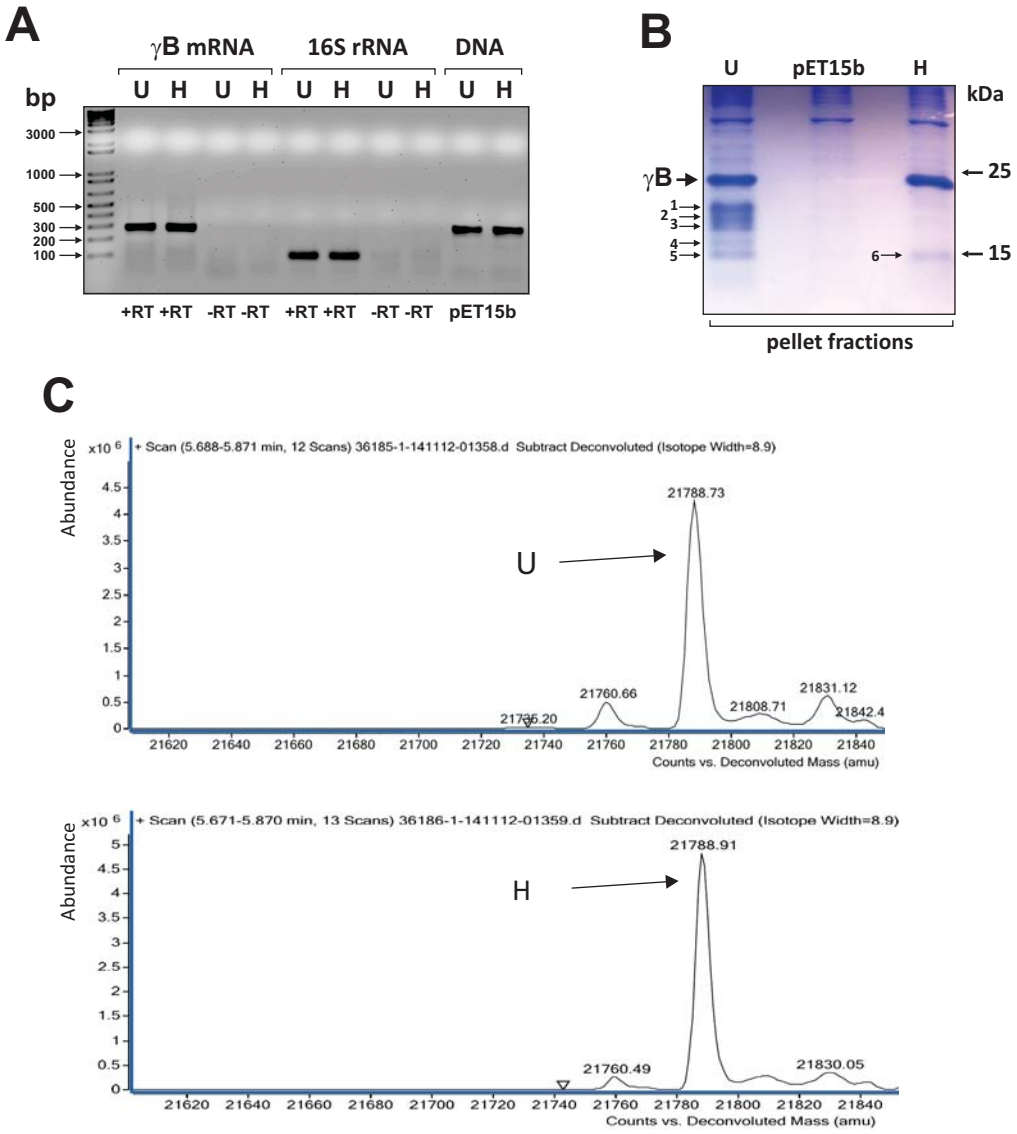


Figure S1, related to Figure 1. Analysis of gamma-B crystallin expression in *E. coli*

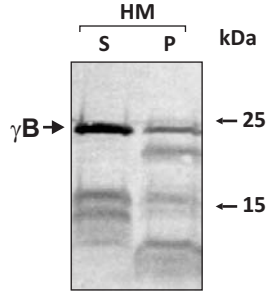
Figure S2

A

90.7% identity

```
                10        20        30        40        50        60        70        80
H       ATGGGAAAAATCACTTTCTATGAAGATCGCGGTTTTCAGGGTCATTGTTATGAATGTAGCAGCGATTGTCCGAACCTGCA
        :::::::::::::::::::: ::::: :: ::::: ::::: :: ::::: :::::::      :: ::::::::::::::
HM      ATGGGAAAAATCACTTTCTACGAAGACCGGGGTTTCCAGGGCCACTGTTACGAATGTTCTTCTGACTGTCCGAACCTGCA
                10        20        30        40        50        60        70        80

                90       100       110       120       130       140       150       160
H       GCCGTACTTTAGCCGCTGTAACAGCATCCGCGTTGATAGCGGTTGTTGGATGCTGTACGAACGCCCGAACTATCAGGGTC
        ::::::::::::::: ::::::   ::::::::::::    :: :::::::::::::::::::::::::::::::: ::::: :
HM      GCCGTACTTTAGCCGGTGTAACTCTATCCGCGTTGACTCTGGCTGTTGGATGCTGTACGAACGCCCGAACTACCAGGGCC
                90       100       110       120       130       140       150       160

                170       180       190       200       210       220       230       240
H       ATCAGTATTTTCTGCGCCGCGGTGATTATCCGGATTATCAGCAGTGGATGGGTTTTAACGATAGCATCCGCAGCTGTCGC
        : ::::: :: :::::::::::::::::: :::::::: :::::::::::::::::::::::::    ::::: ::::::::
HM      ACCAGTACTTCCTGCGCCGCGGTGATTACCCGGATTACCAGCAGTGGATGGGTTTTAACGATTCTATCCGGAGCTGTCGG
                170       180       190       200       210       220       230       240

                250       260       270       280       290       300       310       320
H       TTAATCCCACAACATACCGGTACTTTTCGGATGCGGATCTACGAACGGGACGATTTTCGGGGACAGATGTCCGAAATCAC
        :::::::: ::::: :: :: :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
HM      TTAATCCCCCAACACACTGGCACTTTTCGGATGCGGATCTACGAACGGGACGATTTTCGGGGACAGATGTCCGAAATCAC
                250       260       270       280       290       300       310       320

                330       340       350       360       370       380       390       400
H       TGATGACTGTCCGTCATTGCAAGATCGCTTTCATTTAACTGAAGTACATAGCTTAAACGTTCTGGAAGGGAGCTGGGTCT
        ::::::::::::::: :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
HM      TGATGACTGTCCGTCACTGCAAGATCGCTTTCATTTAACTGAAGTACATAGCTTAAACGTTCTGGAAGGGAGCTGGGTCT
                330       340       350       360       370       380       390       400

                410       420       430       440       450       460       470       480
H       TATACGAAATGCCGAGCTATCGCGGACGCCAGTATCTGCTGCGCCCGGGAGAATATCGCCGGTACTTGGATTGGGGAGCT
        :::::::::::::: ::::: :: ::::: ::::: :::::::: :: :::::::: :: ::::::::::::::::::::
HM      TATACGAAATGCCAAGCTACCGGGGACGGCAGTACCTGCTGCGACCAGGAGAATACCGGCGGTACTTGGATTGGGGAGCT
                410       420       430       440       450       460       470       480

                490       500       510       520
H       ATGAACGCAAAGGTAGGGTCACTACGGCGCGTTATGGACTTCTAC
        :::::::::::::::::::::::::::::: ::::::::::::::
HM      ATGAACGCAAAGGTAGGGTCACTACGGCGGGTTATGGACTTCTAC
                490       500       510       520
```

B



C

Composite: trypsin only, chymotrypsin only,
trypsin and chymotrypsin **93% Sequence Coverage**

```
1    MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR
81   LIPQHTGTFR MKIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA
161  MNAKVGSLRR VMDFY
```

D

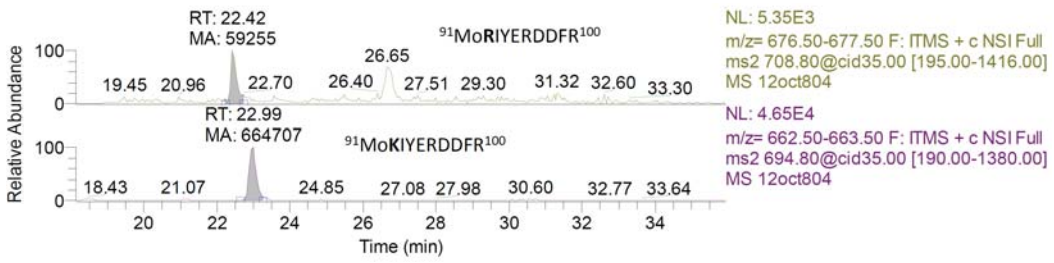


Figure S2, related to Figure 1. Analysis of gamma-B crystallin expression and sequence integrity in *E. coli*

Figure S3

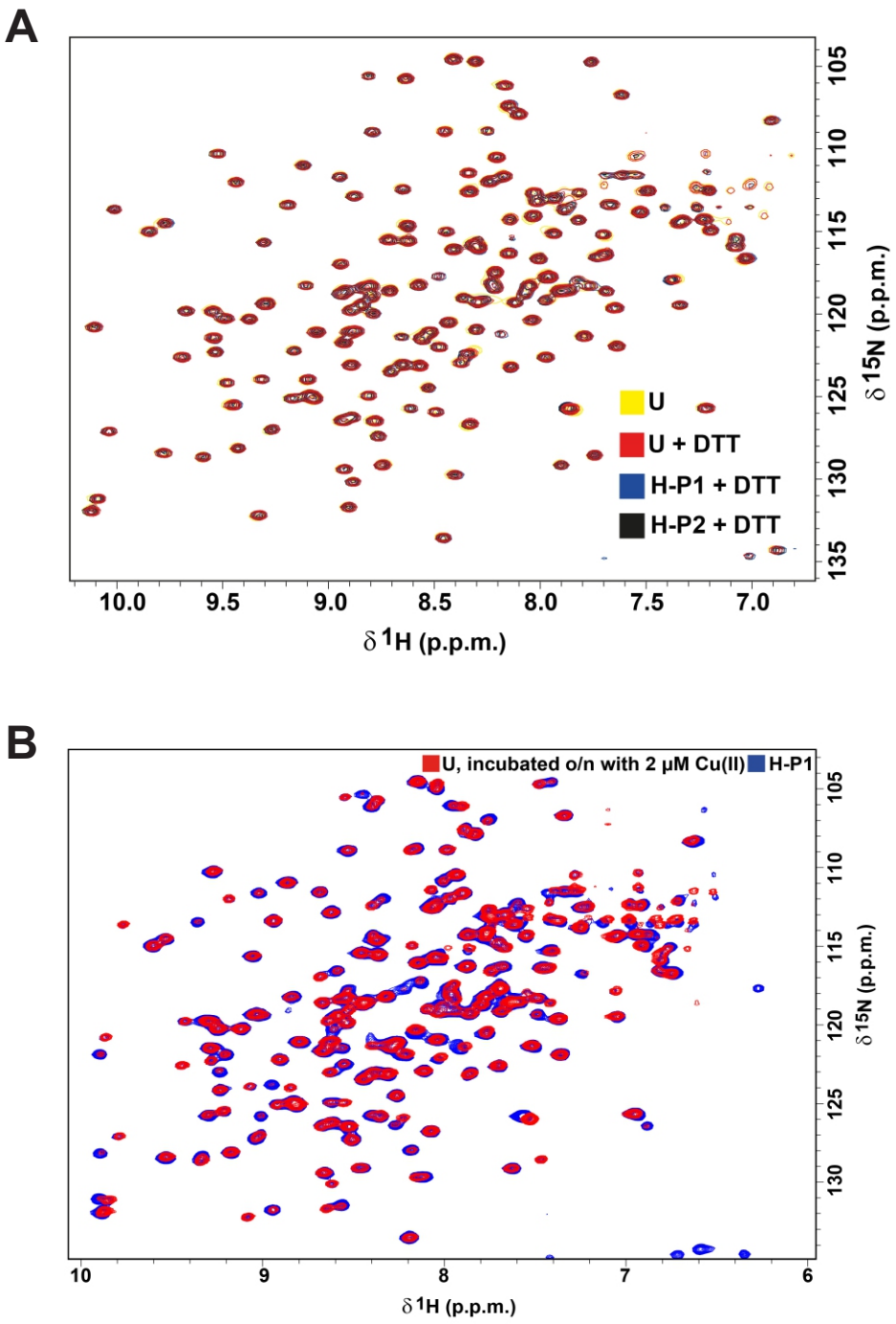


Figure S3, related to Figure 2. Overlay of 2D-1H-15N correlated NMR backbone spectra of gamma-B crystallin variants

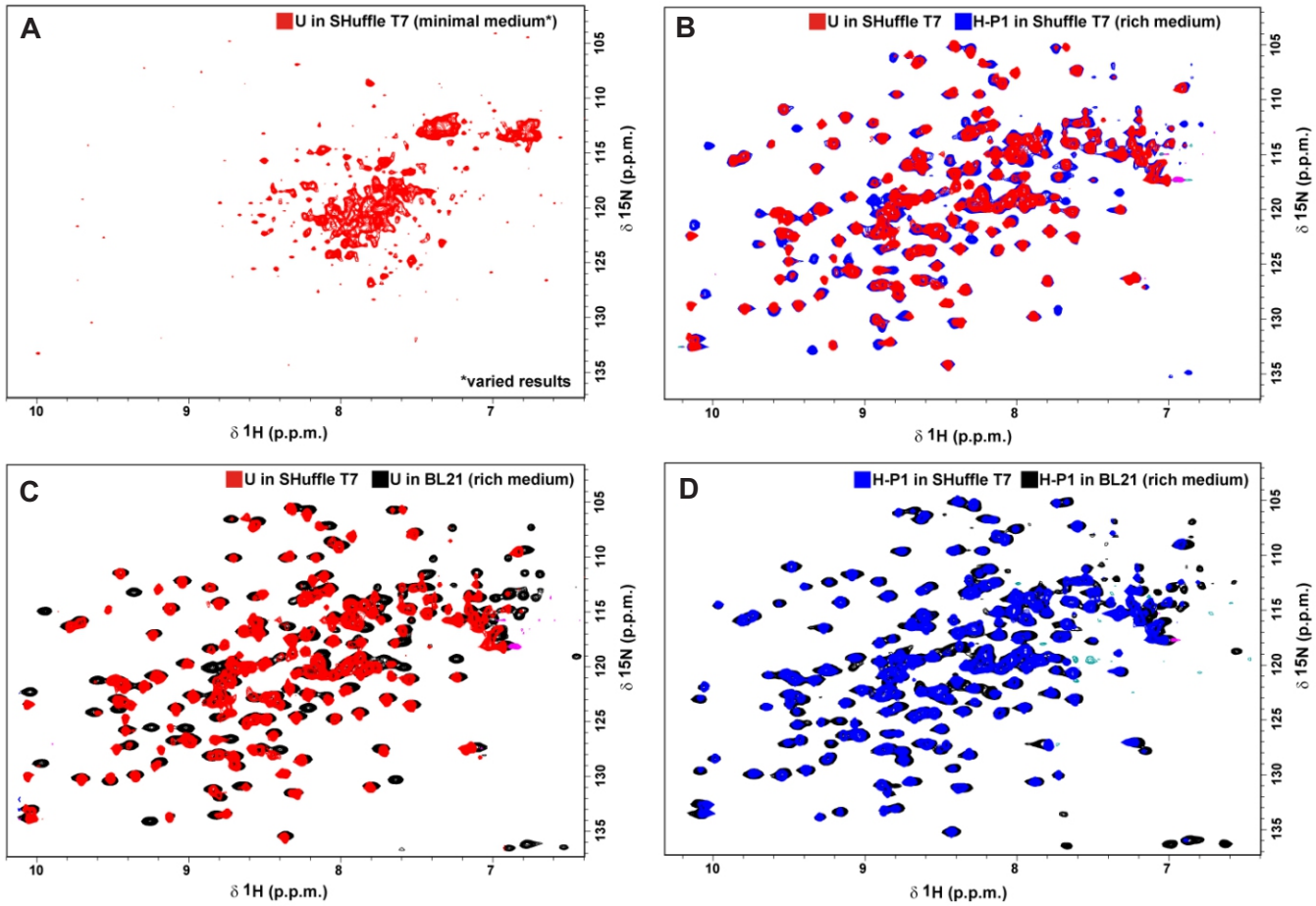


Figure S4, related to Figure 2. 2D-^{1}H-^{15}N correlated NMR backbone spectra of gamma-B crystallin variants U and H expressed in *E. coli* under normal (BL21 strain) and enhanced oxidizing conditions (SHuffle T7 strain)

Figure S5

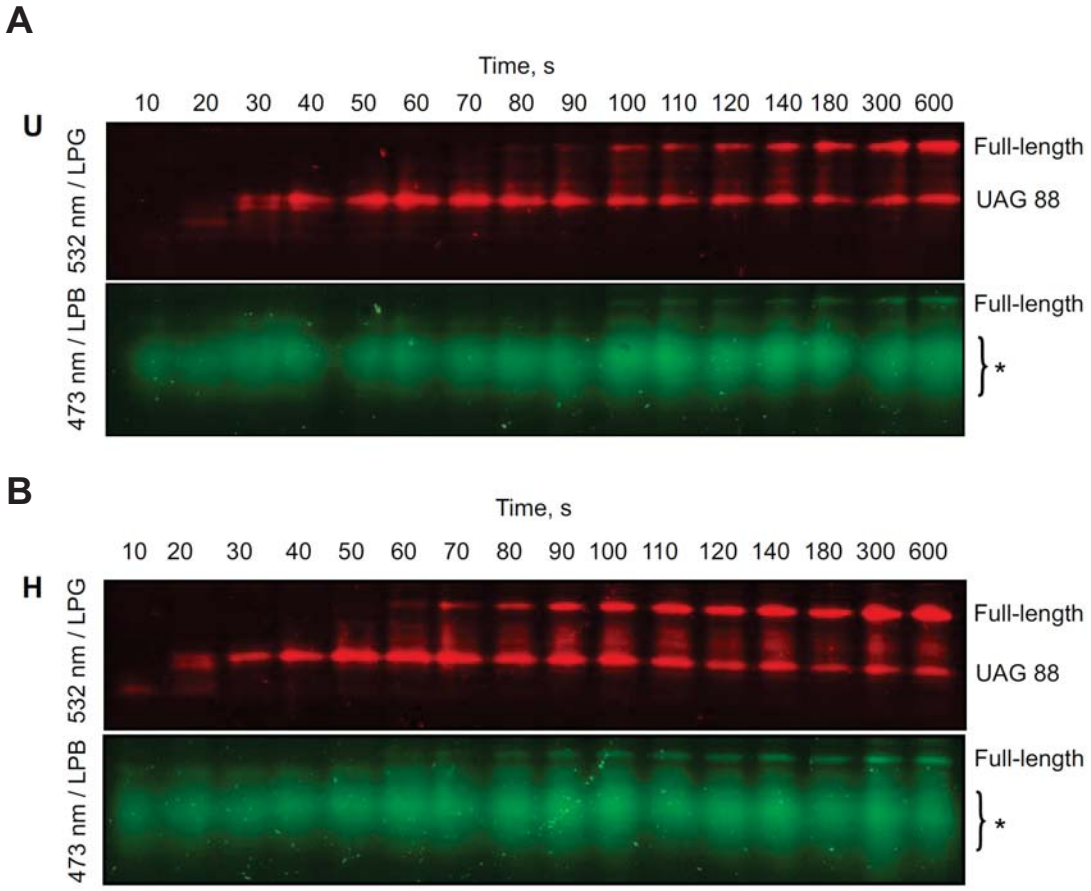


Figure S5, related to Figure 4. Kinetics of synthesis of U and H peptides with N-terminal BOP and BOF incorporated at the amber stop codon at position 88

Figure S6

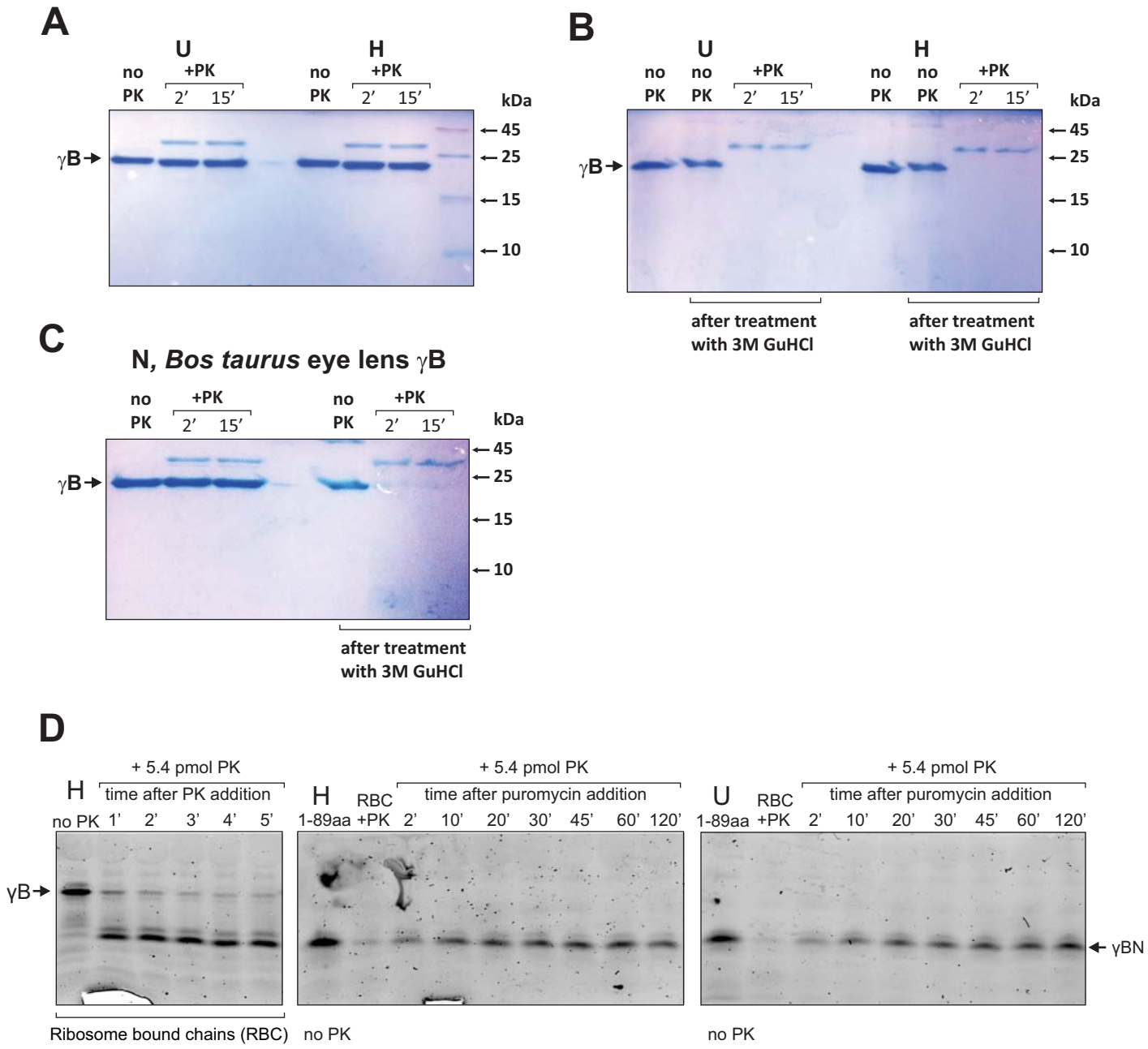


Figure S6, related to Figure 5. PK proteolysis of purified and *in vitro* translated gamma-B crystallin variants

SUPPLEMENTAL FIGURE LEGENDS

Figure S1, related to Figure 1. Analysis of gamma-B crystallin expression in *E. coli*
(A) Quantitative RT-PCR on RNA isolated from *E. coli* cells expressing either U or H variants. DNAse I pretreatment and PCR in the absence of RT served as a control (-RT). 16S rRNA PCR served as a reference. pET15b DNA was used to control for the size of the amplified fragments.
(B) SDS-PAGE of U and H pellet fractions (10 μg). CBB staining. Bands 1-6 were excised from the gel and subjected to MS analysis and microsequencing (*see* also Table S2).
(C) Mass spectra of the purified recombinant U and H proteins. The predicted mass of the 6xHis-tagged protein (U or H) is 21788.37 Da.

Figure S2, related to Figure 1. Analysis of gamma-B crystallin expression and sequence integrity in *E. coli*
(A) Sequence alignment of H and HM gamma-B crystallin variants. HM (Harmonized with Multiple additional synonymous substitutions) sequence was created by introducing additional synonymous changes to the H variant. Regions carrying synonymous substitutions are highlighted in red in H and blue in HM. HM is 90.7% identical to H.
(B) Expression of HM gamma-B crystallin (γB) variant in *E. coli*. Western blotting using polyclonal anti-γB antibodies; Soluble (S) and pellet (P) fractions are shown.
(C) Mascot Search Results showing sequence coverage and location of the MS-identified peptides in HM variant containing R92K substitution (K is enlarged).
(D) LC-MS/MS chromatogram showing that HM variant consists of a mixture of wild type and mutated peptides. Trypsin digest upper panel, chymotrypsin – bottom panel.

Figure S3, related to Figure 2. Overlay of 2D-^{1}H-^{15}N correlated NMR backbone spectra of gamma-B crystallin variants
(A) Addition of DTT resulted in full convergence of the backbone spectra of U (red), H-P1 (blue) and H-P2 (black) to a U-like spectrum (shown in yellow for comparison). These results corroborate the data presented in Figure 2D-F.
(B) Spectra of gamma-B crystallin U, incubated overnight with 2 μM Cu(II) under air supply and H-P1 variants
Both proteins were expressed in ^{15}N-rich labeling medium.

Figure S4, related to Figure 2. 2D-^{1}H-^{15}N correlated NMR backbone spectra of gamma-B crystallin variants U and H expressed in *E. coli* under normal (BL21 strain) and enhanced oxidizing conditions (SHuffle T7 strain)
(A) The spectrum of U expressed in SHuffle T7 cells in minimal medium.
(B) Overlay of U and H-P1 spectra, expressed in SHuffle T7 cells in rich medium.
(C) Overlay of spectra of U expressed in SHuffle T7 and BL21 cells (rich medium).
(D) Overlay of H-P1 spectra expressed in SHuffle T7 and BL21 cells (rich medium).

Figure S5, related to Figure 4. Kinetics of synthesis of U and H peptides with N-terminal BOP and BOF incorporated at the amber stop codon at position 88
(A) U variant. Translation products were resolved by SDS-PAGE and visualized by fluorescence scanning with different excitation wavelengths and emission filters for detection of BOP (top panel) and BOF fluorescence (bottom panel). Incomplete suppression of the amber stop codon (UAG88) and hydrolysis products of BOF-Cys-tRNA$^{Cys}_{CUA\ U32C}$ (asterisks) are indicated.
(B) Same as in **A,** except for the H variant. LPB, 510 nm long-pass filter; LPG, 575 nm long-pass filter.

Figure S6, related to Figure 5. PK proteolysis of purified and *in vitro* translated gamma-B crystallin variants
(A) PK proteolysis of U, H and N variants under native conditions (see Supplemental Experimental Procedures for details). CBB-stained Tris-tricine SDS-PAGE.
(B) After denaturation with 3M GuHCl, 24 h at 37°C.
(C) Natural purified bovine eye lens protein under native and denaturing (3M GuHCl, 24 h, 37°C) conditions.
(D) Proteolysis of *in vitro*-translated gamma-B crystallin H variant and the N-terminal domains (amino acids 1-89). Left panel: time course of PK proteolysis of ribosome-bound gamma-B crystallin chain (H variant; γB) obtained after 20 min of translation. A characteristic PK-resistant fragment has a size similar to that of isolated NTD (γBN). Middle and right panels: pulse proteolysis of ribosome-bound and puromycin-released gamma-B crystallin U and H, respectively.

SUPPLEMENTAL TABLES

Table S1. related to Figure 1. Nucleotide sequences of the U and H variants (gamma-B crystallin ORFs) and the respective codon usage frequencies in *B. taurus* and *E. coli*
***See* file Table S1.xls**

Table S2, related to Figure 1. Mascot Search Results showing sequence coverage and location of the MS-identified peptides (bands 1 to 6 from Figure S1A). All bands were digested in-gel using trypsin

Band #	Mascot Search Results Sequence Info	Trypsin digest Sequence coverage %
1	**Acc. #:** 61888870 **Species:** BOS TAURUS **Name:** gamma-crystallin B	**85%**
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
2	**Acc. #:** 61888870 **Species:** BOS TAURUS **Name:** gamma-crystallin B	**85%**
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
3	**Acc. #:** 61888870 **Species:** BOS TAURUS **Name:** gamma-crystallin B	**71%**
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
4	**Acc. #:** 61888870 **Species:** BOS TAURUS **Name:** gamma-crystallin B	**65.1%**
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
5	**Acc. #:** 61888870 **Species:** BOS TAURUS **Name:** gamma-crystallin B	**65.1%**
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
6	**Acc. #:** 61888870 **Species:** BOS TAURUS **Name:** gamma-crystallin B	**47.4%**
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	

Table S3, related to Figure 1 and S5. Mascot Search Results showing sequence coverage and location of the MS-identified peptides U, H and N (natural gamma-B crystallin isolated from bovine eye lenses) protein variants
Protein samples were digested using trypsin and/or chymotrypsin

Protein	Mascot Search Results Sequence Info	Sequence coverage %
U	**Acc. #:** 61888870 **Species:** BOS TAURUS **Name:** gamma-crystallin B trypsin	55%
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
U	**Acc. #:** 61888870 **Species:** BOS TAURUS **Name:** gamma-crystallin B chymotrypsin	87%
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
U	Composite: trypsin only, chymotrypsin only, trypsin and chymotrypsin	95%
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
H	**Acc. #:** 61888870 **Species:** BOS TAURUS **Name:** gamma-crystallin B trypsin	61%
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
H	**Acc. #:** 61888870 **Species:** BOS TAURUS **Name:** gamma-crystallin B chymotrypsin	89%
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
H	Composite: trypsin only, chymotrypsin only, trypsin and chymotrypsin	95%
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
N	**Acc. #:** 61888870 **Species:** BOS TAURUS **Name:** gamma-crystallin B trypsin	60%
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
N	**Acc. #:** 61888870 **Species:** BOS TAURUS **Name:** gamma-crystallin B chymotrypsin	91%
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	
N	Composite: trypsin only, chymotrypsin only, trypsin and chymotrypsin	95%
Sequence	1 MGKITFYEDR GFQGHCYECS SDCPNLQPYF SRCNSIRVDS GCWMLYERPN YQGHQYFLRR GDYPDYQQWM GFNDSIRSCR 81 LIPQHTGTFR MRIYERDDFR GQMSEITDDC PSLQDRFHLT EVHSLNVLEG SWVLYEMPSY RGRQYLLRPG EYRRYLDWGA 161 MNAKVGSLRR VMDFY	

Table S4, related to Figure 2, S3 and S4. Summary of NMR data collection

NMR data in Figure	**2C**	**2D, 2E, 2F, S3, S4 (H-P1)**
Expression strain	BL21(DE3)	BL21(DE3)
Expression medium	Modified M9, supplemented with amino acids	Celltone Complete (rich medium)
Labelling	^{15}N cysteine-selective	^{15}N uniform
Sample concentrations	0.8 – 1.2 mM	300 - 350 μM
Spectrometer	Bruker AV600	Bruker AV600
Console	Bruker Avance II	Bruker Avance III
Software	Bruker TopSpin 3.1	Bruker TopSpin 3.2
Probe	5 mm TCI cryogenic probe ^{1}H,^{15}N,^{13}C Z-gradients	5 mm TXI cryogenic probe ^{1}H,^{15}N,^{13}C Z-gradients
B_0 field	600 MHz	600 MHz
Pulse Sequence	BEST-[^{15}N,^{1}H]-TROSY	BEST-[^{15}N,^{1}H]-TROSY
Number of Scans	64	384
Number of complex points (TD/2)	512 (^{1}H) 128 (^{15}N)	384 (^{1}H) 128 (^{15}N)
Relaxation Delay	500 ms	500 ms
Spectral width	14.0261 p.p.m. (^{1}H) 49.0822 p.p.m. (^{15}N)	13.0234 p.p.m. (^{1}H) 49.9982 p.p.m. (^{15}N)
B_1 field for ^{15}N hard pulses	6250 Hz	6410 Hz
$1/4J_{HN}$	3 ms	3.1 ms
Offset for amide band selective proton pulses	8.5 p.p.m.	8.5 p.p.m.
Lengths of selective proton pulses	3000 μs (PC9) 2000 μs (REBURP) 1920 μs (EBURP-2)	3000 μs (PC9) 2000 μs (REBURP) 1920 μs (EBURP-2)

Table S5, related to S3A and S4A. Summary of NMR data collection

NMR data in Figure	**S3A (U, oxidized)**	**S4A**
Expression strain	BL21(DE3)	Shuffle T7
Expression medium	M9 minimal medium	M9 minimal medium
Labelling	^{15}N uniform	^{15}N uniform
Sample concentrations	1.0 mM	0.8 mM
Spectrometer	Bruker AV700	Bruker AV700
Console	Bruker Avance I	Bruker Avance I
Software	Bruker TopSpin 2.1	Bruker TopSpin 2.1
Probe	5 mm TXI cryogenic probe ^{1}H,^{15}N,^{13}C Z-gradients	5 mm TXI cryogenic probe ^{1}H,^{15}N,^{13}C Z-gradients
B_0 field	700 MHz	700 MHz
Pulse Sequence	BEST-[^{15}N,^{1}H]-TROSY	BEST-[^{15}N,^{1}H]-TROSY
Number of Scans	32	32
Number of complex points (TD/2)	896 (^{1}H) 256 (^{15}N)	512 (^{1}H) 128 (^{15}N)
Relaxation Delay	500 ms	500 ms
Spectral width	16.0250 p.p.m. (^{1}H) 34.2614 p.p.m. (^{15}N)	13.0312 p.p.m. (^{1}H) 49.9759 p.p.m. (^{15}N)
B_1 field for ^{15}N hard pulses	6850 Hz	6667 Hz
$1/4J_{HN}$	2.5 ms	2.8 ms
Offset for amide band	8.5 p.p.m.	8.6 p.p.m.

selective proton pulses		
Lengths of selective proton pulses	2200 µs (PC9) 1500 µs (REBURP) 1400 µs (EBURP-2)	2570 µs (PC9) 1710 µs (REBURP) 1640 µs (EBURP-2)

Table S6, related to S4. Summary of NMR data collection

NMR data in Figure	**S4B, S4C - D (Shuffle T7)**	**S4C - D (BL21)**
Expression strain	Shuffle T7	BL21(DE3)
Expression medium	Celltone Complete (rich medium)	Celltone Complete (rich medium)
Labelling	^{15}N uniform	^{15}N uniform
Sample concentrations	250 µM	300 - 350 µM
Spectrometer	Bruker AV700	Bruker AV600
Console	Bruker Avance I	Bruker Avance III
Software	Bruker TopSpin 2.1	Bruker TopSpin 3.2
Probe	5 mm TXI cryogenic probe ^{1}H,^{15}N,^{13}C Z-gradients	5 mm TXI cryogenic probe ^{1}H,^{15}N,^{13}C Z-gradients
B_0 field	700 MHz	600 MHz
Pulse Sequence	BEST-[^{15}N,^{1}H]-TROSY	BEST-[^{15}N,^{1}H]-TROSY
Number of Scans	512	384
Number of complex points (TD/2)	608 (^{1}H) 128 (^{15}N)	384 (^{1}H) 128 (^{15}N)
Relaxation Delay	500 ms	500 ms
Spectral width	13.0312 p.p.m. (^{1}H) 49.9759 p.p.m. (^{15}N)	13.0234 p.p.m. (^{1}H) 49.9982 p.p.m. (^{15}N)
B_1 field for ^{15}N hard pulses	6667 Hz	6410 Hz
$1/4J_{HN}$	2.8 ms	3.1 ms
Offset for amide band selective proton pulses	8.5 p.p.m.	8.5 p.p.m.
Lengths of selective proton pulses	2570 µs (PC9) 1710 µs (REBURP) 1640 µs (EBURP-2)	3000 µs (PC9) 2000 µs (REBURP) 1920 µs (EBURP-2)

Table S7, related to Figure 2, S3 and S4. Backbone resonance assignment.

Backbone resonance assignment		
Expression strain	BL21(DE3)	
Expression medium	M9 minimal medium	
Labelling	^{15}N, ^{13}C uniform	
Sample concentrations	1.2 mM	
Spectrometer	Bruker AV800	Bruker AV900
Console	Bruker Avance I	Bruker Avance I
Software	Bruker TopSpin 2.1 Cara 1.9.0	Bruker TopSpin 2.1 Cara 1.9.0
Probe	5 mm TXI cryogenic probe ^{1}H,^{15}N,^{13}C Z-gradients	5 mm TXI cryogenic probe ^{1}H,^{15}N,^{13}C Z-gradients
B_0 field	800 MHz	900 MHz
Pulse sequences	BEST-TROSY HNCO BEST-TROSY-HNCA TROSY-HN(CO)CACB	BEST-TROSY TROSY-HNCACB

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Analytical scale protein expression

Analytical-scale protein expression was performed in BL21 (DE3) *Δtig::Kan* strain (Deuerling et al., 1999) (B, F^-,*dcm, ompT, hsd*S(r_B^-,m_B^-), *gal*, (DE3) *Δtig::Kan*, a gift of Dr. Bernd Bukau, ZMBH, Heidelberg, Germany), which lacks the ribosome-associated chaperone Trigger Factor. Freshly transformed *E. coli* cells harboring either U or H gamma-B crystallin variants were grown in flasks with 50 mL of liquid Luria Bertani (LB) medium containing 100 µg/mL ampicillin at 37°C with shaking at 230 rpm until the cultures reached an OD_{600} of 0.8. Protein expression was induced by addition of 1 mM IPTG (isopropyl-β-d-thiogalactopyranoside) and incubation was continued for an additional 1.5 h at 37°C with shaking. Cells were harvested by centrifugation (3200×g, 15 min). Cell pellets were re-suspended in 1 mL of 50 mM Tris-HCl buffer, pH 8.0, containing 500 mM NaCl and 1 tablet of EDTA-free protease inhibitor cocktail (Roche), and sonicated (20 s, 3 times) using an ultrasonic processor (Fisher Scientific). Lysates were pelleted at 17,000×g for 30 min, using a TLA-55 fixed-angle rotor (Beckman Coulter) to separate insoluble/aggregated proteins (pellet) from soluble protein (supernatant). The supernatant was transferred into a clean tube. The pellet was re-suspended in an equal (to that of supernatant) volume of 50 mM Tris-HCl buffer, pH 8.0, containing 500 mM NaCl and EDTA-free protease inhibitor cocktail (Roche).

Quantitative RT-PCR

For analysis of relative mRNA expression levels, total RNA was isolated from *E. coli* BL21 (DE3) *Δtig::Kan* cells expressing U and H variants using Trizol Reagent (Life Technologies). Total RNA (2 µg) was subjected to DNAse treatment (Promega) and 1 µg of DNAse-treated RNA was used for cDNA amplification with the Invitrogen Superscript III RT kit (Life Technologies). PCR reactions were carried out for 21 cycles (to avoid saturation) using the following gamma-B crystallin U/H variant-specific primers:

U forward: 5'-AAAAACCATGGGGAAGATCACTTTTTAC-3'
U reverse: 5'-GAAAGTGCCGGTGTGTTGCGG-3
H forward: 5'-AAAAACCATGGGAAAAATCACTTTCTATG-3'
H reverse: 5'-CCGAAAAGTACCGGTATGTTGTGG-3',

Both the U and H primer pairs generated products 264 bp length. PCR without RT, using DNAse-treated RNA as the template was used to control for plasmid DNA contamination. Amplified fragments of 16S rRNA were used as reference standards (16S rRNA forward primer: 5'- GCTACAATGGCGCATACAAA-3'; 16S rRNA reverse primer: 5'-TTCATGGAGTCGAGTTGCAG-3'; 101 bp PCR product).

Analysis of recombinant protein expression

Protein concentrations in samples were determined using a Coomassie (Bradford) Protein Assay Kit (Thermo Scientific) with bovine serum albumin (BSA) as a reference standard and equal total protein amounts were analyzed for each sample. For Western blotting, 2 µg total protein (soluble and/or insoluble fractions) was separated under reducing conditions using Tris-Tricine SDS-PAGE using a 4% stacking and a 16.5% separation gel (49.5% T, 3% C). Proteins were transferred to a 0.45 µm pore size PVDF membrane (EMD Millipore) using a wet Mini Trans-Blot blotting system (Bio-Rad Laboratories) at 110V for 1.5 h. The membrane was blocked (7.5% non-fat dry milk, TBS, 0.1% Tween) for 1 h at room temperature, washed and probed with either anti-gamma-B crystallin antibody (1:1000 dilution, P-18, Santa Cruz Biotechnology) (Figure 1C) or with mouse monoclonal anti-poly-histidine antibody (1:3000, H1029, Sigma) (Figure 1E). Membranes were then incubated with secondary donkey anti-goat IgG-HRP (1:2000, Santa Cruz Biotechnology) or goat anti-mouse IgGH&L(HRP) (1:4000, ab6789, Abcam) antibodies. Signals were detected using ECL Plus reagent (GE Healthcare) and a Typhoon 9410 imaging scanner (GE Healthcare). Band intensities were quantified using either Image Quant (v. 2005) or ImageJ software. For ELISA, 2 µg of total cellular protein in 200 mM sodium carbonate/bicarbonate buffer, pH 9.5 was coated on 96 well plates and incubated overnight, at 4°C. Protein was blocked using 300 µL blocking buffer containing 7.5% non fat dry milk (Promega) in 0.1% Tris-Buffered Saline (TBS) per well and incubated for 1 hr at room temperature. After 1 hr, blocking buffer was removed and 200 µL of polyclonal anti-gamma-B crystallin antibody (P-18, Santa Cruz Biotechnology, 1:1000) diluted in 7.5% non fat dry milk, 0.1% tween TBS was added to each well and incubation was continued on shaking platform for 1 hr at room temperature. Wells were washed with 300 µl wash buffer (0.1% tween TBS) for 5 min (×3). After washing, secondary donkey anti-goat IgG-HRP antibody (1:2000) (Santa Cruz Biotechnology) diluted in 7.5% non fat dry milk, 0.1% tween TBS was added and incubation was continued for 1 hr on shaking platform followed by washing with 300 µL of wash buffer (0.1% tween TBS) for (5 min x 3 times). For detection, 100 µL of TMB One Solution (Promega) was added to each well and the samples were incubated for 10 min with shaking in the dark at room temperature. Reactions were stopped by adding 100 µL of 1N hydrochloric acid (HCl) to each well. Protein was detected by recording the absorbance at 450 nm on Victor3™ Plate reader (PerkinElmer). Percentages of the soluble and pellet fractions were calculated as a fraction of the total protein.

Recombinant protein purification

Gamma-B crystallin protein variants expressed as described above were isolated using immobilized metal-affinity chromatography (IMAC) on Ni-Nitrilotriacetic Acid Agarose (Ni-NTA) (Qiagen) followed by size-exclusion

chromatography (SEC) on a Superdex 75 gel-filtration column (1.6 cm×60 cm) (GE Healthcare) using an ÄKTA purifier liquid chromatography system (GE Healthcare). Gamma-B crystallin fractions were collected and concentrated using an Amicon Ultra-4 centrifugal device (EMD Millipore). Aliquots were flash frozen in liquid nitrogen and stored at -80°C. The purified protein samples were subjected to MS analysis and microsequencing at either the Cleveland Clinic proteomics core facility or the Harvard Mass Spectrometry facility.

Gamma-B crystallin purification from bovine eye lenses

Gamma-B crystallin from bovine eye lenses (native protein, N) was purified generally as described (Slingsby and Miller, 1983) with the following modifications. Frozen lenses collected from young cows (less than 30 months old) were purchased from Animal Technologies (Tyler, TX, USA). Lenses were thawed, decapsulated and homogenized in 50 mM Tris-HCl buffer, pH 7.2, containing 1 tablet of EDTA-free protease inhibitor cocktail (Roche). After the homogenate was centrifuged at 27,000×g for 30 min, the pellet was discarded and the supernatant was filtered and used for further purification. Gamma crystallins were separated from other proteins (including alpha- and beta-crystallins) by SEC using a Superdex 75 gel-filtration column (1.6 cm×60 cm, GE Healthcare) on an ÄKTA purifier liquid chromatography system (50 mM Tris-HCl buffer, pH 7.2, flow-rate 0.75 mL/min). Fractions containing gamma crystallins were collected and dialyzed against 20 mM Tris-acetate, pH 6.0 (buffer A). Proteins were further separated by cation exchange chromatography on a Mono S 5/50 Column (GE Healthcare). The column was equilibrated and washed with buffer A and a gradient and stepwise elution of buffer B (20 mM Tris-Acetate, 0.4 M sodium acetate, pH 6.0) was used. Elution was done by running a 5-10% buffer B gradient for 2.5 min, 10% isocratic buffer B for 10 min, followed by a 10-40% buffer B gradient for 15 min (flow rate 0.5 mL/min). To identify fractions containing gamma-B crystallin, proteins from each isolated peak were run on Tris-tricine SDS-PAGE gels, Coomassie stained, and submitted for MS analysis and microsequencing (Table S3).

Proteolysis of purified gamma-B crystallin

For proteinase K (PK) proteolysis under native/non-denaturing conditions, 190 pmol (4 μg) of purified gamma-B crystallin (U, H and/or N) in 50 mM Tris-HCl pH 8.0, 200 mM NaCl were treated with 38 pmol (1.1 μg) PK at 37°C. Reactions were stopped by PMSF addition (10 mM final concentration) and samples were immediately flash frozen in liquid nitrogen, thawed, mixed with 2X Tris-tricine loading buffer and resolved on Tris-tricine SDS-PAGE gels (Schägger and von Jagow, 1987). For proteolysis under denaturing conditions, gamma-B crystallin (U, H and/or N) in 50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 3M GuHCl was first denatured at 37°C for 24 h and then subjected to PK treatment as described above.

Mass spectrometry

Identification of gamma-B crystallin protein products was carried out by LC-MS/MS analysis at the Cleveland Clinic Mass Spectrometry Laboratory for Protein Sequencing (Cleveland, OH) and the Harvard Mass Spectrometry facility (Cambridge, MA). Finnigan LTQ-Obitrap Elite hybrid mass spectrometer systems were used. Purified proteins or gel-separated protein products were subjected to MS analysis. For in-gel protein digestion, bands were excised, washed with water, and dehydrated in acetonitrile. The bands were then reduced with DTT and alkylated with iodoacetamide prior to in-gel digestion. Bands were digested in-gel by adding either 5 μL 10 ng/μL trypsin or 10 ng/μL chymotypsin in 50 mM ammonium bicarbonate and incubating overnight at room temperature to achieve complete digestion. Peptides were extracted with a 50% acetonitrile, 5% formic acid solution, evaporated to <10 μL using a Speedvac and resuspended in 1% acetic acid to a final volume of 30 μL for LC-MS/MS analysis.

mRNA for *in vitro* translation

mRNAs coding for full-length U and H protein variants, or the N-terminal fragment (amino acids 1-89) of U and H, were prepared by T7 transcription using pET15b-based constructs containing the corresponding sequences. Neither construct contained the C- terminal 6xHis tag or a stop codon. The T7 forward primer (5'- GATCCCGCGAAATTAATACGACTC-3') was used with the following reverse primers:

U (full-length): 5'-ATAAAAATCCATCACCCGTCTTAAAGAACC-3';
U (1-89): 5'- GAAAGTGCCGGTGTGTTGCG-3';
H (full-length) 5'- GTAGAAGTCCATAACGCGCCGTA-3'; and
H (1-89): 5'-AAAAGTACCGGTATGTTGTGGGATTAAG-3'.

For FRET experiments, the threonine codon at position 88 (counting from the initiator Met) of full-lengths U or H mRNA was replaced by an amber stop codon UAG. The position 88 was chosen because it is located (i) close to the end of the NTD (position 80), which should allow for monitoring the NTD folding as soon its synthesis is completed, and (ii) at a distance that would allow efficient FRET, and because the linker can be replaced without a change in protein stability (Mayr et al., 1994).

Fluorescence-labeled tRNAs

To obtain BOP-Met-tRNA^{fMet}, [^{3}H]Met-tRNA^{fMet} (15 μM) was labeled with Bodipy576/589-NHS (5 mM) (MolecularProbes) in 100 mM sodium acetate pH 5.0, 80% DMSO for 18 h at 4°C. Unreacted fluorophore was removed by extraction with 50:50 phenol:chloroform mixture followed by ethanol precipitation. BOP-Met-tRNA^{fMet} was purified by HPLC on a

LiChrospher WP300 RP-18 column using a gradient of 5 to 40% ethanol in 20 mM ammonium acetate, 10 mM $MgCl_2$, 400 mM NaCl pH 4.5.
The gene coding for *E. coli* $tRNA^{Cys}$ under a T7 promotor was generated by primer overlap PCR and was ligated into the SmaI site of pUC19. Genes coding for amber, opal and ochre suppressor $tRNA^{Cys}$ variants were constructed by site-directed mutagenesis of the anticodon. $tRNA^{Cys}$ variants were generated by *in vitro* transcription and purified on a HiTrapQ column (GE healthcare) using a linear gradient of 0 to 1.1 M NaCl in 50 mM sodium acetate pH 5.0 and 10 mM $MgCl_2$. Aminoacylation was carried out in 100 mM HEPES pH 7.5, 15 mM KCl, 7 mM $MgCl_2$, 5 mM DTT with 2 mM ATP, 150 µM L-cysteine, 35 A_{260} units/mL of $tRNA^{Cys}$ variant, and 6 µM CysRS-His_6 for 45 min at 37°C. The yield of aminoacylation was 50% for all variants. Opal suppressor Cys-$tRNA^{Cys}$ from *E. coli* was reported to be active in translation in rabbit reticulocyte and wheat germ extracts (Gubbens et al., 2010). In contrast, in the *E. coli* system, only wt $tRNA^{Cys}$ and the amber suppressor $tRNA^{Cys}_{CUA}$ were active. To improve the decoding properties of $tRNA^{Cys}_{CUA,}$ we introduced a further replacement into the anticodon loop, U32 to C, which restored the rate of decoding to the level of wt $tRNA^{Cys}$ (Olejniczak and Uhlenbeck, 2006). Cys-$tRNA^{Cys}_{CUA\ U32C}$ (30 µM) was labeled with BodipyFL-C_1-IA (Molecular Probes) (1 mM) in 50 mM Hepes pH 7.5, 50% DMF for 45 min at 25°C. Unreacted fluorophore was removed by extraction with 50:50 phenol:chloroform mixture followed by ethanol precipitation. BOF-Cys-$tRNA^{Cys}_{CUA\ U32C}$ was separated from deacylated and unlabeled tRNA by HPLC on a LiChrospher WP300 RP-18 column using a gradient of 5 to 40% ethanol in 20 mM ammonium acetate, 10 mM $MgCl_2$, 400 mM NaCl.

***In-vitro* translation**

Translation in the fully reconstituted single-turnover *in vitro* translation system was carried out as described (Doerfel et al., 2013; Mittelstaet et al., 2013). 70S initiation complexes (IC) were prepared by incubating 70S ribosomes (purified from *E. coli* MRE 600 strain (Rodnina, and Wintermeyer, 1995) (0.75 µM) with mRNA (1.5 µM), BOF- or BOP-labeled [^{3}H]Met-$tRNA^{fMet}$ (0.5 µM), a mixture of initiation factors IF1, IF2, IF3 (1.5 µM each), and GTP (2 mM) in $TAKM_7$ buffer (50 mM Tris-HCl, pH 7.5, 70 mM NH_4Cl, 30 mM KCl, 7 mM $MgCl_2$ 2 mM DTT) for 1 h at 37°C. Initiation efficiency was verified by nitrocellulose filtration and [^{3}H]-radioactivity counting (Milon et al., 2007); typically, all added fluorescence-labeled [^{3}H]Met-$tRNA^{fMet}$ was found in the complex with the ribosome. IC stock (0.125 µM) was diluted to HiFi translation buffer (50 mM Tris-HCl, pH 7.5, 70 mM NH_4Cl, 30 mM KCl, 3.5 mM $MgCl_2$, 0.5 mM spermidine, 8 mM putrescine, 2 mM DTT). Factor mix (FM) was formed by incubating elongation factor (EF) EF-Tu–GDP (250 µM), GTP (2 mM), phosphoenolpyruvate (6 mM), pyruvate kinase (2%), and EF-G (7.5 µM) at 37°C for 15 min, followed by addition of purified total *E. coli* aa-tRNA (100 µM) aminoacylated with the mixture of amino acids containing [^{14}C]Phe and additional incubation for 2 min at 37°C. The concentrations of the components in the *in vitro* system were optimized to achieve the speed and fidelity similar to that *in vivo* (Wohlgemuth et al., 2010; Rudorf et al., 2014). *In vitro* translation was initiated by rapid mixing of IC (0.015 µM) with FM (40 µM ternary complex) at 37°C. At indicated time points after the start of translation, aliquots (25 µl) were withdrawn and the reaction was stopped by adding 5 µl of 2 M NaOH and digested for 30 min at 37°C. The efficiency of translation was tested by precipitating the peptides by cold trichloroacetic acid (5%) and measuring [^{3}H]Met and [^{14}C]Phe incorporation into proteins; typically, 70-80% of the IC carried peptides of expected length. 2 M Hepes acid (5.5 µL) and 2X loading buffer (50 mM Tris-HCl, pH 6.8, 12% (w/v) glycerol, 2% β-mercaptoethanol, 4% SDS) was added to each sample and the translation products were separated by Tris-tricine SDS-PAGE using 4% stacking, 10 % spacer and 16.5 % separation gels (49.5% T, 3% C). The gel was washed with water and scanned on a FLA-9000 fluorescence imager (Fujifilm Lifescience). Fluorescence was excited at 473 nm and monitored after passing a LPB (510LP) filter. Because the mRNAs used for translation did not have a stop codon, translation was limited to a single round per ribosome and the nascent peptides did not dissociate from the ribosome, unless released by puromycin treatment.
For FRET experiments, translation was carried out with an IC containing BOP-Met-$tRNA^{fMet}$ and an mRNA variant containing a UAG stop codon at position 88. The stop codon was decoded by an amber suppressor $tRNA^{Cys}_{CUA\ U32C}$ (Gubbens et al., 2010) aminoacylated with Cys and thiol-modified by BOF. The fluorophores BOF or BOP were exited at 470 nm and 540 nm, respectively. The emission of BOF was monitored after passing a 500 nm cut-off filter; the emission of BOP after a 570 nm cut-off filter. As the acceptor-only control, unlabeled Cys-$tRNA^{Cys}_{CUA\ U32C}$ was used. Cys-$tRNA^{Cys}_{CUA\ U32C}$ was separated from deacylated tRNA by purifying the ternary complex EF-Tu•GTP•Cys-$tRNA^{Cys}_{CUA\ U32C}$ on a Superdex 75 10/300 size exclusion column.

Protein expression and purification for NMR spectroscopy

Aliquots of Invitrogen OneShot BL21(DE3) *E. coli* cells (Life Technologies) or SHuffle T7 *E. coli* cells (NEB) were freshly transformed with the pET15b plasmid (U or H) and incubated overnight in 50 mL LB medium containing 100 mg/L ampicillin (37°C, 220 rpm shaking). Pre-cultures were pelleted (15 min, 5000×g, 4°C) and resuspended in a small volume of the final expression medium, which was then used to inoculate main cultures to a starting OD_{600} of 0.2 – 0.3 for BL21(DE3) and 0.5 for SHuffle T7 cells.
All expression media contained 100 mg/L ampicillin. A rich medium for uniform ^{15}N-labeling, Celltone Complete (98% ^{15}N, Cambridge Isotope Labs) was used (incubation conditions: 250 mL culture volume in 1 L baffled flasks, 37°C, 160-180 rpm). ^{15}N/^{13}C–labeled gamma-B crystallin for backbone assignment was expressed in M9 minimal medium containing 1 g/L [^{15}N]NH_4Cl and 2 g/L [U-^{13}C]glucose (incubation conditions: 2 L culture volume in 5 L baffled flasks, 37°C, 120 rpm). For

selective ^{15}N-cysteine labeling, a modified M9 medium containing [^{15}N]cysteine was used, as described by Muchmore et al., 1989 (incubation conditions: 1 L culture volume in 5 L baffled flasks, 37°C, 120 rpm). Expression cultures were incubated at 37°C until an OD_{600} of 0.8 was reached, at which point protein expression was induced by addition of 1 mM IPTG and allowed to continue for 3 h before harvesting the cells by centrifugation (15 min, 5000×g, 4°C). Per 1 L of culture volume, cell pellets were resuspended in 35 mL of lysis buffer (50 mM sodium phosphate, pH 8.0, 500 mM NaCl, 20 mM imidazole, complete EDTA-free protease inhibitor tablet (Roche)). For resuspension volumes smaller than 50 mL, cells were lysed by sonication (6 × 1 min, 5 min cooling on ice, 40% cycle time, 60% power, Bandelin Sonoplus HD 2070). For resuspension volumes greater than 50 mL, lysis was performed using an M-110P microfluidizer (3 cycles, 1000 Bar, Microfluidics). Using a divided reference expression, we demonstrated that no artificial structural changes in gamma-B crystallin were introduced by different lysis conditions (data not shown). Cell lysates were cleared by centrifugation (2 × 30 min, 20.000×g, 4°C) and supernatants were loaded onto a 5 mL HisTrap HP Ni-NTA column (GE Healthcare) that was pre-equilibrated with NTA washing buffer (50 mM sodium phosphate, pH 8.0, 500 mM NaCl, 20 mM imidazole). Columns loaded with lysate were washed to A_{280} baseline levels and 6xHis-tagged gamma-B crystallin was eluted with NTA elution buffer (50 mM sodium phosphate, pH 8.0, 500 mM NaCl, 250 mM imidazole). Protein fractions were pooled and concentrated to a volume of 10 mL using a Vivaspin 20 centrifugal concentrator (Sartorius), followed by SEC using a HiLoad 26/60 Superdex 75 prep grade gel-filtration column (GE Healthcare) in IEX buffer A (50 mM Tris, pH 9.0). Fractions corresponding to monomeric gamma-B crystallin were pooled and loaded onto a 5 mL HiTrap Q XL anion exchange column (GE Healthcare), which was subsequently washed with 5 column volumes of IEX buffer A. pH gradient elution over 30 column volumes to 100% IEX buffer B (50 mM Tris, pH 6.0, 50 mM NaCl) was performed and corresponding peak fractions were pooled. Concentration to a final volume of 600 µL and buffer exchange into NMR buffer (50 mM Tris, pH 8.0, 200 mM NaCl, 10% D_2O, 0.1% DSS) was performed using Vivaspin 20 and Vivaspin 2 centrifugal concentrators (Sartorius). All chromatographic steps were performed on an ÄKTA purifier liquid chromatography system (GE Healthcare) at 25°C. The combination of protein concentrations above 5 mg/ml and temperatures below 20°C was avoided at all times during purification in order to prevent potentially structurally relevant phase transitions and precipitation. 10 µM samples for SDS-page and native-PAGE were taken after each purification step. Samples for C-4 reverse-phase HPLC were taken before and after ion exchange chromatography.

Oxidation of U with catalytic amounts of Cu(II)

Gamma-B crystallin U variant (10 µM in NMR buffer) isolated for NMR analysis was incubated overnight with 2 µM $CuCl_2$ under air supply and continuous stirring, essentially as described by Cavallini et al., (1969). The protein was then re-purified by SEC in NMR buffer, concentrated to 1 mM and subjected to NMR analysis.

Analytical HPLC

Analytical HPLC runs were performed on a Jasco HPLC system using a PerfectSil 300 C4-RP column (250 × 4.6 mm). Gradient: 10-35% B (20 min), 35% B isocratic (10 min), 45-50% B (15 min). Buffer A: H_2O, 0.1% TFA; Buffer B: Acetonitrile, 0.1% TFA.

Supplemental References

Cavallini, D., De Marco, C., Duprè, S., and Rotilio, G. (1969). The copper catalyzed oxidation of cysteine to cystine. Arch. Biochem. Biophys. *130*, 354-361.

Deuerling, E., Schulze-Specking, A., Tomoyasu, T., Mogk, A., and Bukau, B. (1999). Trigger factor and DnaK cooperate in folding of newly synthesized proteins. Nature *400*, 693-696.

Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M.V. (2013). EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. Science *339*, 85-88.

Gubbens, J., Kim, S.J., Yang, Z., Johnson, A.E., and Skach, W.R. (2010). In vitro incorporation of nonnatural amino acids into protein using tRNA(Cys)-derived opal, ochre, and amber suppressor tRNAs. RNA 16, 1660-1672.

Mayr, E.M., Jaenicke, R., and Glockshuber R. (1994). Domain interactions and connecting peptides in lens crystallins. J Mol Biol. *235*, 84-88.

Milon, P., Konevega, A.L., Peske, F., Fabbretti, A., Gualerzi, C.O., and Rodnina, M.V. (2007). Transient kinetics, fluorescence, and FRET in studies of initiation of translation in bacteria. Methods Enzymol. *430*, 1-30.

Olejniczak, M., and Uhlenbeck, O.C. (2006). tRNA residues that have coevolved with their anticodon to ensure uniform and accurate codon recognition. Biochimie *88*, 943-950.

Rodnina, M.V., and Wintermeyer, W. (1995). GTP consumption of elongation factor Tu during translation of heteropolymeric mRNAs. Proc. Natl. Acad. Sci U. S. A. *92*, 1945-1949.

Rudorf, S., Thommen, M., Rodnina, M.V., and Lipowsky, R. (2014). Deducing the kinetics of protein synthesis in vivo from the transition rates measured in vitro. PLoS Comput. Biol. *10*, e1003909.

Schägger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal. Biochem. *166,* 368-379.

Slingsby, C., and Miller, L. (1983). Purification and crystallization of mammalian γ-crystallins. Exp. Eye Res. *37*, 517-530.

Wohlgemuth, I., Pohl, C., and Rodnina, M.V. (2010). Optimization of speed and accuracy of decoding in translation. EMBO J. *29*, 3701-3709.