



# G-quadruplexes rescuing protein folding

Ahyun Son<sup>a,1</sup> , Veronica Huizar Cabral<sup>a,1</sup>, Zijue Huang<sup>a</sup> , Theodore J. Litberg<sup>a</sup> , and Scott Horowitz<sup>a,2</sup>

Edited by Ken Dill, Stony Brook University, Stony Brook, NY; received September 23, 2022; accepted April 12, 2023

Maintaining the health of the proteome is a critical cellular task. Recently, we found G-quadruplex (G4) nucleic acids are especially potent at preventing protein aggregation in vitro and could at least indirectly improve the protein folding environment of *Escherichia coli*. However, the roles of G4s in protein folding were not yet explored. Here, through in vitro protein folding experiments, we discover that G4s can accelerate protein folding by rescuing kinetically trapped intermediates to both native and near-native folded states. Time-course folding experiments in *E. coli* further demonstrate that these G4s primarily improve protein folding quality in *E. coli* as opposed to preventing protein aggregation. The ability of a short nucleic acid to rescue protein folding opens up the possibility of nucleic acids and ATP-independent chaperones to play considerable roles in dictating the ultimate folding fate of proteins.

quadruplex | G4 | proteostasis | protein folding

Proteostasis is maintained through the action of a variety of proteins, particularly molecular chaperones and degradation machineries. Chaperones interact with unfolded, partially folded, or misfolded proteins to prevent aggregation, assist in protein folding, and/or target proteins for degradation (1, 2). ATP-independent chaperones (holdases) act as a first line of defense against toxic aggregation by binding proteins as they misfold or unfold (3, 4). At the cessation of stress, holdases transfer their clients to ATP-dependent chaperones (foldases) that can both prevent aggregation and actively assist the protein folding process (2, 4). In addition to traditional protein chaperones, several other molecules have been shown to play major roles in protein folding and aggregation (5, 6).

More recently, the role of nucleic acids in proteostasis has begun to receive attention; for example, ~10% of the proteome requires RNA for maintaining solubility, including many proteins that are not considered to be RNA-binding proteins (7). Moreover, stress granules, which are liquid-liquid phase separated condensates where proteins retreat under stress for protection, contain high levels of RNA (8–10). Nucleic acids maintain fluidity and solubility of stress granules (9, 10) and, in some cases, are responsible for scaffolding and guiding proteins to these condensates under stress (11, 12). These and other studies strongly suggest that RNAs constitute important players of multiple mechanisms in proteostasis. However, the precise roles of nucleic acids in proteostasis are still unclear.

We recently demonstrated that certain nucleic acids serve as powerful modulators of protein aggregation, oligomerization, and protein folding (13–16). Intriguingly, we discovered that hybrid (3+1) and parallel G4-containing sequences (Fig. 1A) are especially potent as protein chaperones. We found that G4s prevent aggregation in vitro more efficiently on a weight basis than known chaperone proteins, and an order of magnitude more efficiently than bulk DNA (14). Moreover, we demonstrated that many G4s could also promote a healthy protein folding environment in cells, although whether this was due to affecting protein folding directly was unclear at the time. Of note, unfolded nucleic acids did not show these effects in vitro or in cells, suggesting these properties were special to G4s (14).

G4s are nucleic acids that contain polyG stretches that form base tetrads (Fig. 1A). G4s can combine with other nucleic acid structural elements to create highly complex nucleic acid structures (17, 18) and have previously been characterized to have roles in regulating transcription, translation, and splicing (19). Although the full extent of G4 formation is not yet known in the cell, they have been identified in DNA, mRNA, and in the large rRNA extensions of ribosomes as well as in tRNA that is cleaved under stress conditions to create G4-forming tRNA (20–23).

Foldases are typically considered to aid proteins in folding in three major ways. In the case of Hsp60s (such as *Escherichia coli* GroEL), there is significant evidence that these chaperones can directly catalyze protein folding, which is hypothesized to occur through a combination of enclosure and electrostatic effects (24), as well as to increase apparent folding speed by isolating the protein from off-pathway aggregation (25). In Hsp60s as well as other foldases, there is also evidence of unfolding mechanisms in which the chaperone helps to unfold kinetically trapped misfolded states to allow another chance at folding (26). With ATP-independent holdases, effects on folding

## Significance

Correctly managing protein folding and aggregation is critical to cellular health, with many diseases resulting from mistakes in these processes. While nucleic acids have long been implicated as playing a role in protein folding and aggregation, whether nucleic acids could directly impact protein folding remained unclear. Here, we show that unlike a typical ATP-independent “holdase” mechanism, G-quadruplexes can trigger protein folding. These results broaden the set of strategies that can be used by the cell to guide protein folding.

Author affiliations: <sup>a</sup>Department of Chemistry & Biochemistry, Knobel Institute for Healthy Aging, University of Denver, Denver, CO 80208

Author contributions: A.S., T.J.L., and S.H. designed research; A.S., V.H.C., Z.H., and T.J.L. performed research; A.S., V.H.C., Z.H., T.J.L., and S.H. analyzed data; and A.S. and S.H. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

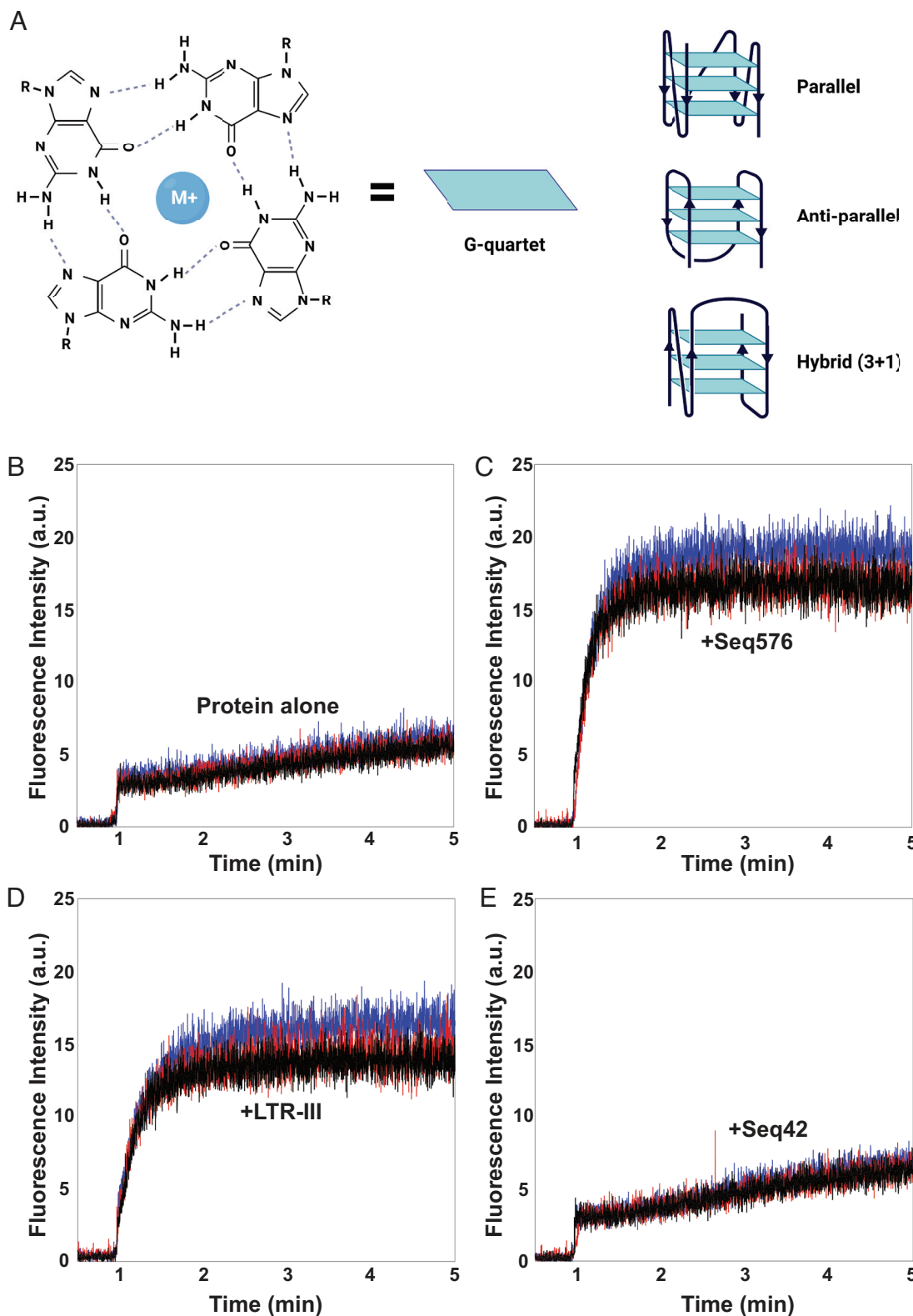
Copyright © 2023 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](#).

<sup>1</sup>A.S. and V.H.C. contributed equally to this work.

<sup>2</sup>To whom correspondence may be addressed. Email: Scott.Horowitz@du.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2216308120/-DCSupplemental>.

Published May 8, 2023.



**Fig. 1.** Fluorescence recovery of TagRFP675 in the presence of DNAs containing G4 structures in vitro. (A, *Left*) A schematic illustration of a G-tetrad. Four guanine bases arranged in a square planar tetrad structure through Hoogsteen hydrogen bonding. Monovalent cations ( $M^+$ , shown as a blue sphere), such as  $K^+$ , are required as a stabilizer of G4s. (A, *Right*) Examples of intramolecular G4 complex with different structures and loop conformations. (B–E) TagRFP675 was denatured in 6M guanidine-hydrochloride (Gu-HCl) for 11 h at 23 °C. Denatured protein was diluted to 0.5  $\mu$ M into potassium phosphate and the increase in fluorescence intensity was measured over time in the absence of DNAs (B) and in the presence of 1  $\mu$ M of Seq576 (C), LTR-III (D), and Seq42 (E). The absence of DNAs and the addition of Seq42 served as the negative controls and LTR-III and Seq576 contain G4 structures. Three triplicate measurements are shown with the three separate colors.

have previously been found to be primarily passive (27), but it remains possible that there are direct effects on protein folding.

Before finding the potency of G4s as chaperones, we investigated the ability of disordered and unfolded ssRNA to affect protein folding. We examined the ability of polyU RNA to help the

folding of luciferase. In these experiments, polyU had no effect on the refolding of luciferase on its own, but could serve as a holdase during stress that transferred the partially folded protein to Hsp70 for refolding. As a result, in the presence of polyU RNA and Hsp70, a much greater amount of luciferase was refolded than with just Hsp70 (13). We thus hypothesized that all nucleic acids function as holdases, similar to known ATP-independent chaperones. Unexpectedly, our work here reveals this hypothesis was not accurate for G4s: G4s can accelerate protein folding by rescuing kinetically trapped intermediates, a property previously thought to be predominantly performed by ATP-dependent foldases. Continuing experiments in *E. coli* shows that the effects of chaperone nucleic acids are not protease, solubility, or expression level-dependent, but promote protein folding.

## Results

**G4-Containing DNA Rescues Protein Folding of TagRFP675 In Vitro.** To investigate whether G4s could not only prevent aggregation but also affect protein folding, we checked whether G4s we previously found that had good holdase activity (such as Seq576, see *Materials and Methods* for detailed sequence information) (14) could potentially affect protein refolding in vitro. To test this question, we purified TagRFP675, and performed in vitro refolding assays. In these assays, TagRFP675 is chemically denatured, and then rapidly diluted into a native buffer to allow refolding in the presence or absence of G4-containing chaperones. We monitored the refolding of TagRFP675 by the appearance of its fluorescence, which depends almost solely on its chromophore formation and being in the native state (28). At room temperature, we found that a small amount of protein renatured quickly, followed by a continued slow increase in folded protein (Fig. 1*B*). Adding Seq576 (14) caused a marked increase in the speed of protein refolding (Fig. 1*C*). A second G4 of known structure, LTR-III (29), also instigated substantial protein folding (Fig. 1*D*). Contrasting the effects of the G4s, using a ssDNA sequence of the same length but containing no G4 structure (Seq42, negative control) (Fig. 1*E*) caused no increase in the speed of refolding. This experiment suggests that the G4 aids in the folding of the protein, unlike an unfolded nucleic acid that has no effect on folding. The RNA equivalents showed nearly identical activities (*SI Appendix, Fig. S1*).

The difficulty of TagRFP675 to refold on its own in this experiment is consistent with previous literature on the folding of fluorescent proteins. It has been well established that fluorescent  $\beta$ -barrel proteins display a hysteresis in folding, in which returning to the native state after chemical denaturation occurs on the timescale of days (30–34). This very slow refolding rate is due to a rugged folding landscape, and particularly the formation of kinetically trapped intermediates that escape to the native state very slowly (30–33). These intermediates have  $\beta$ -barrel formation but appear to be structurally different primarily in the lid region in the case of superfolder GFP (30). We therefore hypothesized that TagRFP675 could similarly be trapped in intermediate states and unable to refold without the presence of a chaperone to assist transitioning from the kinetically trapped intermediate to the native state.

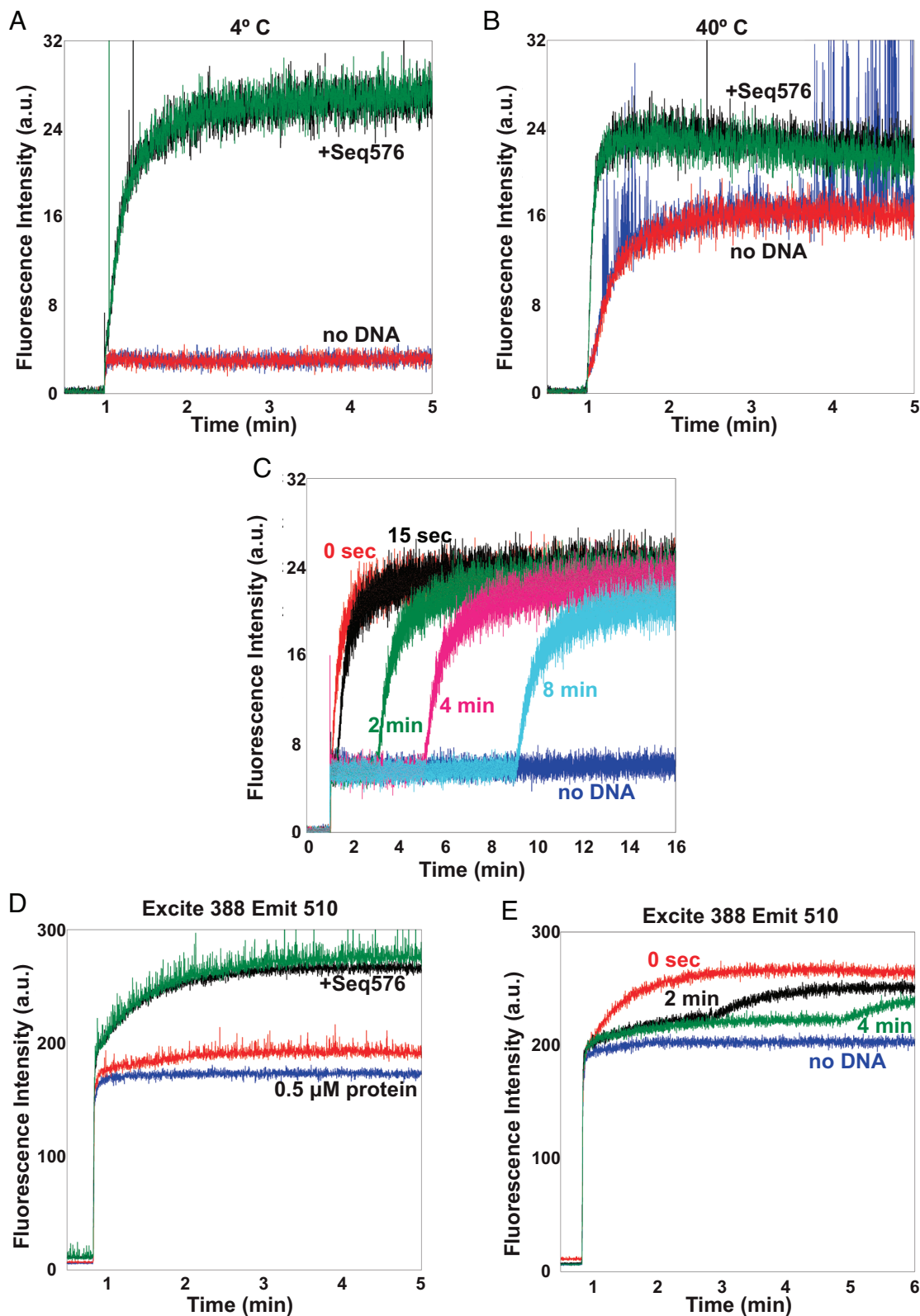
To directly test the kinetically-trapped intermediate hypothesis, we reasoned that by increasing the refolding temperature, TagRFP675 would overcome the kinetic folding barrier more quickly without a chaperone (35). In this case, we would observe an increase in the folding rate as a function of temperature. Indeed, our data showed that the rate of spontaneous refolding dramatically decreased at lower temperature and improved with

increasing temperature, making the presence of the chaperone G4 almost unnecessary at 40 °C (Fig. 2*A* and *B*). These results strongly support the hypothesis that some portion of TagRFP675 is in a kinetically trapped state that the G4 is able to rescue to the native state. Of note, we also are not able to exclude the effect of hyperfluorescence from near native states that could contribute to the structural fluctuations of the chromophore at 40 °C (36).

We next investigated whether kinetically trapped states were being rescued by changing when the G4 was added. If kinetically trapped states were stable, we would observe the refolding regardless of when the G4 was added to the kinetically trapped protein. We therefore performed the refolding experiment as before but added Seq576 at defined times after the protein was diluted into the refolding buffer (Fig. 2*C*). The results showed that regardless of whether the G4 was added before or after the addition of protein, the addition of the G4 immediately prompted the refolding of TagRFP675 (Fig. 2*C*) to nearly the same extent. This observation suggests that there is a subpool of trapped intermediates that are able to overcome the kinetic barrier to the folded state by association with the G4. We also varied the concentration of Seq576 to detect the effect on protein refolding. Increasing the amount of G4 did lead to an increase in the level of protein refolding, but this effect leveled off at a small fraction of total protein folding (*SI Appendix, Fig. S2*).

TagRFP675's chromophore forms an intricate set of hydrogen bonds that affect its emission spectrum. Specifically, the hydroxyl group forms several hydrogen bonds that shift the emission maximum from 510 nm to 675 nm (37). By also monitoring 510 nm during refolding, we could therefore separately observe the protein in an intermediate state that is distinct from the native state. Monitoring at 510 nm in the absence of Seq576, we observed roughly an order of magnitude increase in fluorescence relative to that observed at 675 nm, suggesting that a considerably larger fraction of the protein collapses into a non-native fluorescent intermediate than to the native state (Fig. 2*D*). Performing this same experiment in the presence of Seq576 caused a considerable increase in the fast formation of this non-native intermediate over the case without G4 (Fig. 2*D*). This observation suggests that the G4 is able to promote the formation of both fluorescent non-native intermediates as well as to rescue a smaller portion of kinetically trapped intermediates to the native state. Furthermore, in both cases, the amount of protein in this partially folded state observed at 510 nm decreases slowly over time that roughly corresponds with the slow increase in fluorescence at 675 nm, suggesting that there is a slow conversion from this intermediate state to the native state (Fig. 2*D*). Like with the native state, delayed addition of the G4 also rescues folding to the intermediate state (Fig. 2*E*). Overall, these results show that G4s can rescue protein folding from intermediate states.

**G4 Enhances Proper Folding of TagRFP675 in Cells.** Having seen an effect of promoting protein folding in vitro, we next examined the effects of a G4 on folding in cells. Previously, we showed that fluorescence of TagRFP675 was increased in the presence of G4s in *E. coli* but not by non-G4 sequences (14) using a tightly controlled two expression vector system (Fig. 3*A*). Here, we used this system for a fluorescence time-course assay to observe folding, solubility, and proteolysis after translation inhibition. Briefly, we co-expressed TagRFP675 with Seq576 RNA under heat stress. As controls, we compared the fluorescence in the presence of Seq576 both against empty vector (negative control) as well as GroEL (positive control). The expression of TagRFP675 was induced for 3 h (Fig. 3*B*), and then protein synthesis was stopped by the



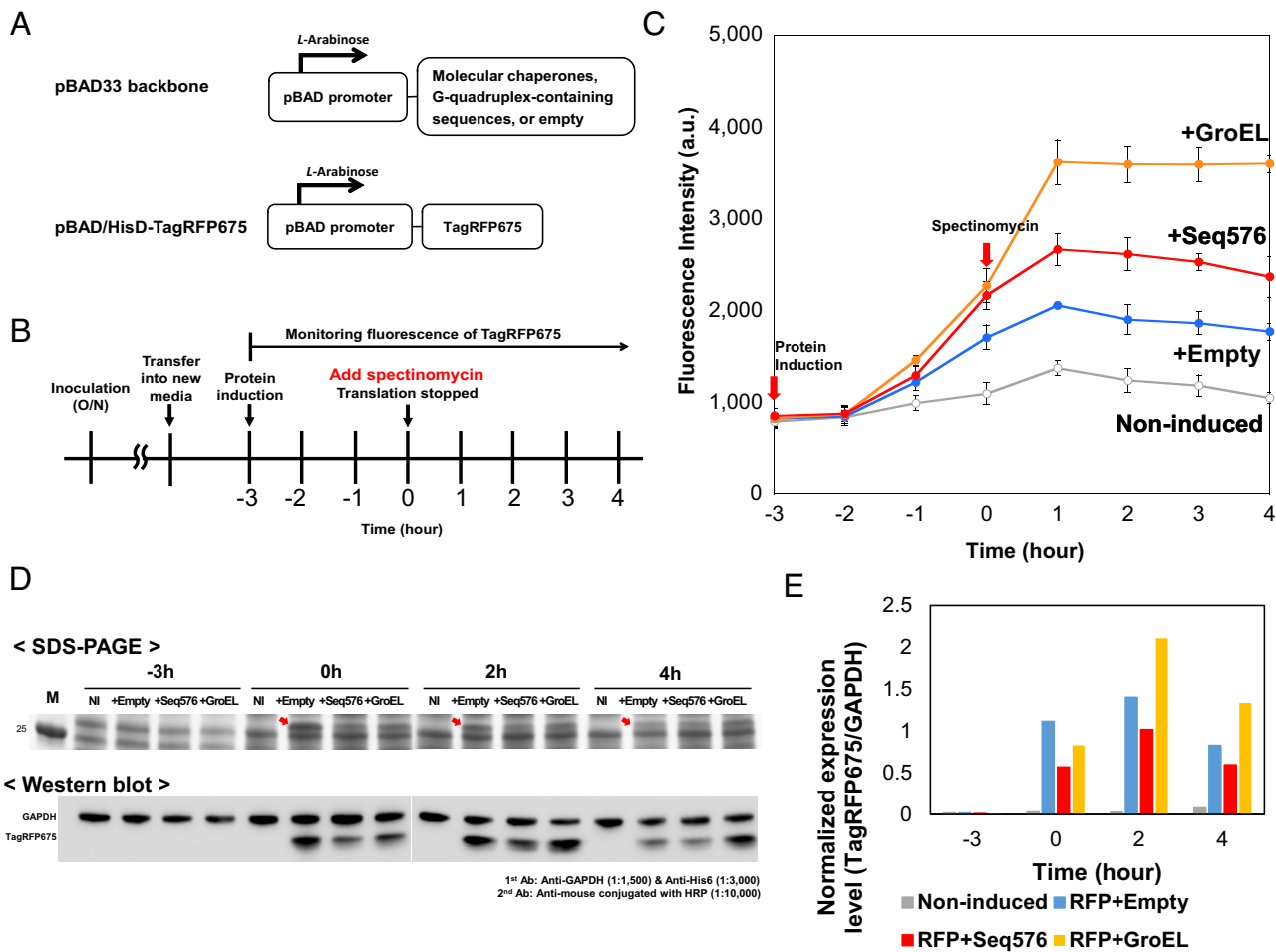
**Fig. 2.** G4-containing DNA rescues protein folding of TagRFP675 in vitro. TagRFP675 was diluted from denaturing buffer (6 M Gu-HCl) to 0.5  $\mu$ M protein into refolding buffer. Addition of 1  $\mu$ M Seq576 at 4°C (A) and 40°C (B). Replicates shown in separate colors. Experiments at 23°C can be seen in Fig. 1. (C) Addition of G4-containing DNA triggers refolding after dilution. Timed addition of Seq576, showing folding rescue at timepoints after TagRFP675 dilution into refolding buffer at 4°C. TagRFP675 was diluted in the presence of 1  $\mu$ M Seq576 at 23°C and observed at 510 nm emission at 23°C either before dilution (D) or with timed addition of Seq576 (E).

addition of spectinomycin, which has previously been used to halt protein translation for protein fate analyses (38–40).

We observed that the fluorescence of TagRFP675 continued to increase after inhibition of translation by about a factor of two

(Fig. 3C), suggesting either post-translational folding, or the tail end of co-translational folding of a slow-folding protein. We confirmed that this result is not due to insufficient translation inhibition by testing varying concentrations of spectinomycin and





**Fig. 3.** Fluorescence time course assay of TagRFP675 in MC4100(DE3). (A) A schematic illustration of expression vectors. Both the expression of G4s and TagRFP675 are under the control of pBAD promoter, which is induced by L-Arabinose. (B) An experimental workflow of the fluorescence time-course assay. Protein was induced when OD<sub>600</sub> reaches more than 0.7 (at time point -3 h) and expressed for 3 h at 42 °C, and then protein translation was stopped by adding spectinomycin (at time point 0 h). The fluorescence of TagRFP675 was continuously monitored after protein induction. (C) Time-course assay of TagRFP675 in the presence of G4 (Seq576) or molecular chaperone (GroEL). Protein translation was halted at time point 0 h after 3 h induction of protein. The experiment shown here were technical triplicates (n = 3) and performed more than 2 times. (D) SDS-PAGE (Top) and western blot assay (Bottom) from the cells at time points -3 h, 0 h, 2 h, and 4 h of the time-course assay. Red arrows indicate expressed TagRFP675. The concentrations of first anti-GAPDH, first anti-His6, and second anti-mouse antibodies are listed at bottom right. The full images are shown in *SI Appendix, Fig. S4*. (E) Normalization of the expression level of TagRFP675. Normalized ratio was calculated by dividing the expression level of TagRFP675 analyzed by three people using ImageJ by that of GAPDH. A replica western blot is shown in *SI Appendix, Fig. S4*.

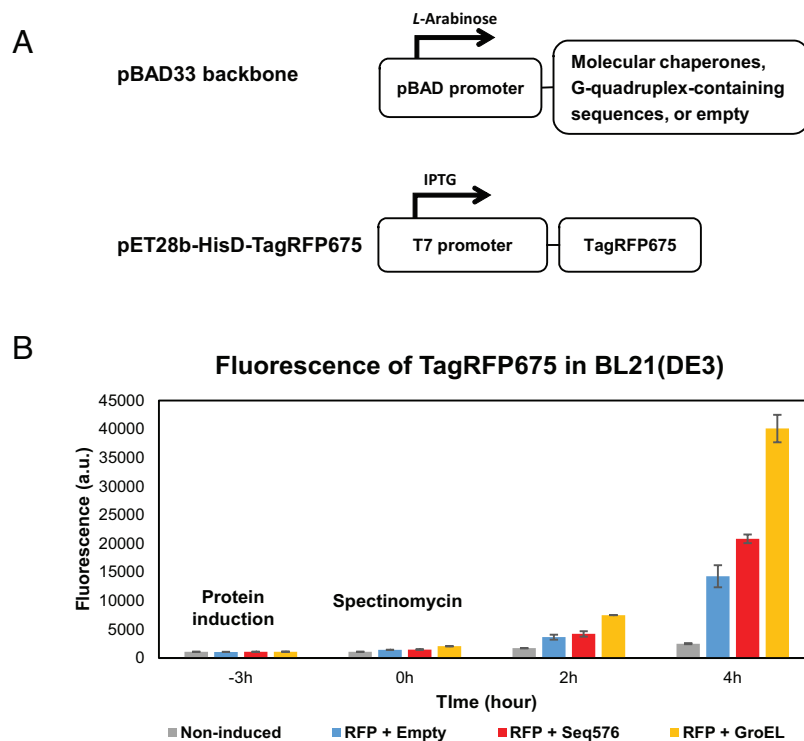
monitoring O.D. levels (*SI Appendix, Fig. S3 A and B*) and other inhibitors such as kanamycin (*SI Appendix, Fig. S3 C and D*). This behavior is instead consistent with the slow folding and maturation time of TagRFP675 (37). Consistent with in vitro refolding, co-expression of the G4-containing sequence Seq576 strongly increased the fluorescence of TagRFP675 above empty vector, indicating a higher level of properly folded TagRFP675 in cells (Fig. 3C).

To determine whether Seq576 has a role in the expression level of TagRFP675 in cells, we harvested and visualized the amount of TagRFP675 present at each time point with SDS-PAGE and western blot (Fig. 3D and E and *SI Appendix, Fig. S4*). The expression level of TagRFP675 was not significantly changed shortly after translation inhibition (0 h to 2 h), but it did decrease at later timepoints (4 h), suggesting proteolysis was slowly decreasing the protein level. This result correlated with the slight decrease of the fluorescence at the 4 h timepoint (Fig. 3C). Curiously, the expression level of TagRFP675 was not perfectly proportional to the fluorescence. We further confirmed that there were not significant differences between pH, the translation inhibition had not killed the cells within this timeframe (*SI Appendix, Fig. S3 E and F*), and

the fluorescence of expressed TagRFP675 did not affect overall absorbance spectra and OD<sub>600</sub> of cells for normalization (*SI Appendix, Fig. S5*). Taken together, these data suggest that not all translated TagRFP675 formed properly folded structure due to its slow folding and maturation time, and Seq576 promoted some portion of the protein population to fold.

Previous studies have reported that nucleic acid binding stimulates *E. coli* Lon protease activity, and Lon protease itself is a well-known nucleic acid-binding protein that binds a variety of nucleic acid sequences and structures (41–45). As we observed degradation of TagRFP675 at later time points, the Seq576 could be potentially affecting either the activity of Lon protease or protein folding, or both. We therefore performed the same experiment in a BL21(DE3), a strain that notably lacks Lon (Fig. 4A).

In BL21(DE3), we also observed that in the presence of Seq576, the fluorescence of TagRFP675 was significantly higher than that with empty vector control, consistent with the result in MC4100(DE3). This suggests that Seq576 mainly affects folding of the protein, and not in preventing degradation by Lon. Similar to MC4100(DE3), in BL21(DE3) the fluorescence continued to increase after translation inhibition, but for a much greater



**Fig. 4.** Fluorescence time course assay of TagRFP675 in BL21(DE3). (A) A schematic illustration of expression vectors. The expression of G4s is under the control of pBAD promoter, which is induced by *L*-Arabinose. TagRFP675 expression is under the control of T7 promoter, which is induced by IPTG. (B) Time-course assay of TagRFP675 in the presence of G4s (Seq576) or molecular chaperone (GroEL). The experiment shown here were technical triplicates ( $n = 3$ ) and performed more than two times; error bars are SE.

timespan than in MC4100(DE3) (Fig. 4B). Correlating in vitro proteolytic assays with Lon protease found a higher amount degradation of partially misfolded TagRFP675 than native state (SI Appendix, Fig. S6), similar to other fluorescent proteins (46). We again confirmed that in BL21(DE) cell growth, pH, and viability between samples are similar (SI Appendix, Fig. S7) and GFPwt fluorescence is unaffected by Seq576 (SI Appendix, Fig. S8). These data further suggest that slow folding and maturation of TagRFP675 also continue in BL21(DE3), but due to reduced degradation, a larger proportion of the TagRFP675 is chaperoned and folded to the native state.

Next, we investigated whether the expression level of TagRFP675 was affected by the presence or absence of Seq576 in BL21(DE3) (Fig. 5 and SI Appendix, Fig. S9). To minimize the discrepancy between expression level and fluorescence due to gel uncertainty, we performed SDS-PAGE multiple times and confirmed the results from different experiments were consistent (SI Appendix, Fig. S9 D and E). Even though the difference of fluorescence between co-expressed GroEL or Seq576 and empty vector was large, the solubility and protein levels of TagRFP675 between Empty and Seq576 were very similar (Fig. 5 A–C and SI Appendix, Fig. S10). These results indicate that the primary effect of Seq576 is on enhancing proper folding, not increasing the amount of protein or its solubility. Normalizing the fluorescence by expression level confirms this observation (Fig. 5D). Taken together with in vitro results, the mechanisms of G4s in assisting folding appear to enhance proper folding by helping protein intermediates to overcome kinetic traps in their protein folding landscapes.

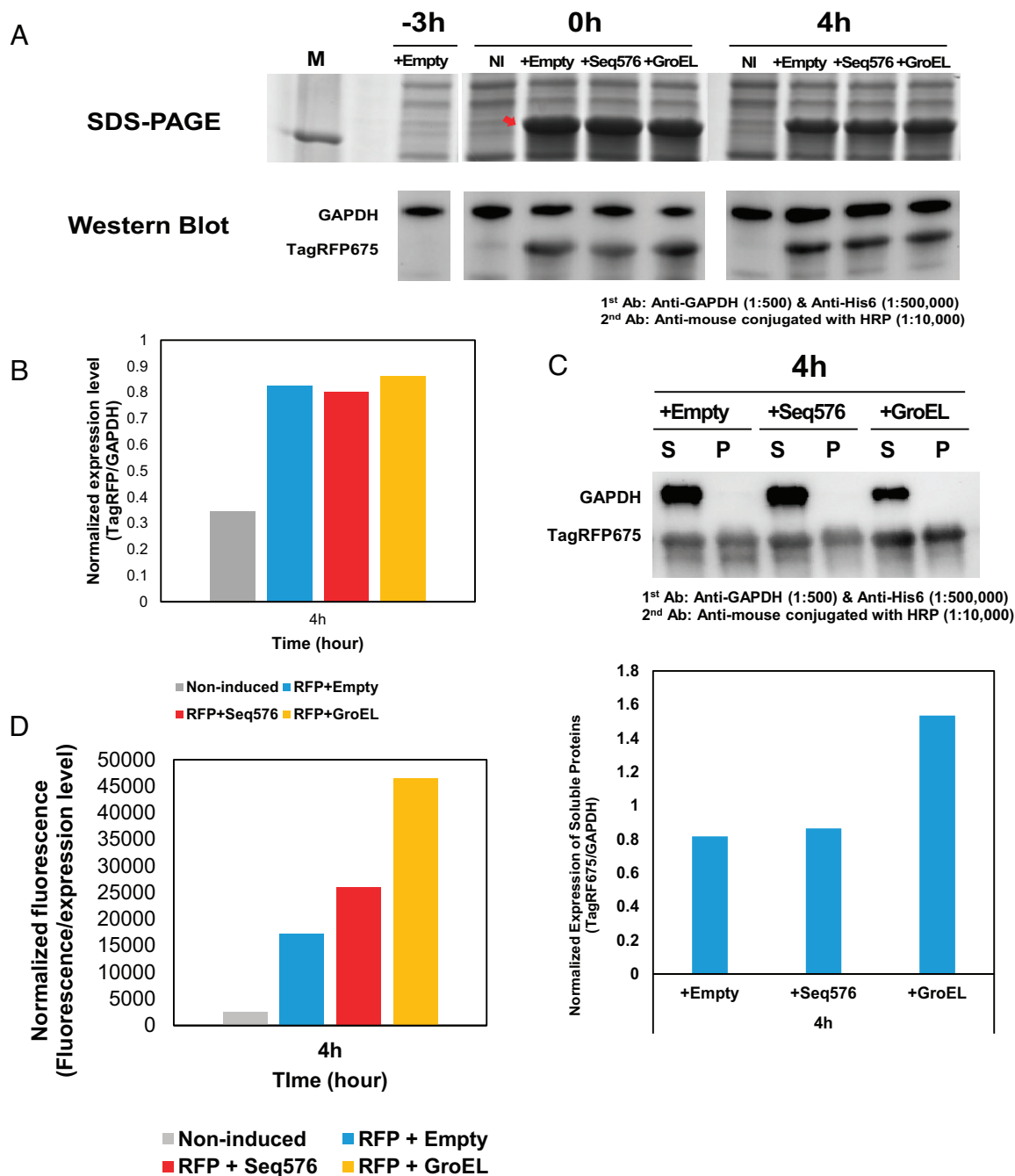
To further test whether Seq576 affects Lon activity, we added Lon protease back into BL21(DE3) cells (SI Appendix, Fig. S9). With Lon expression, both fluorescence and expression level of TagRFP675 in BL21(DE3) were highly decreased and were more

similar to that in MC4100(DE3). Of note, we found that the increase in fluorescence by Seq576 compared to Empty vector was proportional both with and without Lon protease. This result suggests that the primary effect of Seq576 is not to prevent degradation by Lon, and it is consistent with the increase in fluorescence in the presence of Seq576 not relying on an increased amount of protein.

## Discussion

In this study, we examined the effects of G4s with chaperone activity on the folding of an unstable biosensor protein, TagRFP675. In vitro, we unexpectedly found that the G4s accelerate protein folding to native and near-native states by rescuing kinetically trapped intermediates. In *E. coli*, we found that the G4 Seq576 improves protein folding of TagRFP675 as opposed to affecting protein solubility or proteolysis. Together, these results illuminate the potential power of G4s to modulate protein folding.

The ability for G4s to accelerate protein folding in vitro is somewhat surprising based on literature on protein folding rates and chaperone activity. Several ATP-independent chaperones have been observed to improve protein folding yield, such as Spy and SecB (27, 47). However, in these cases the rate of folding slows down due to the affinity of partially folded intermediates for the chaperone (27, 48). That said, ATP-independent “holdase” chaperones have been shown to have the capacity to potentially reshape folding landscapes (47, 49, 50). The observation that the G4 rescues kinetically trapped intermediates to both native and fluorescent non-native states suggests that the G4 is acting in a fashion that is general, but this still could have several different molecular explanations, including modification of multimeric state. These effects could come from destabilization or catalysis-like effects

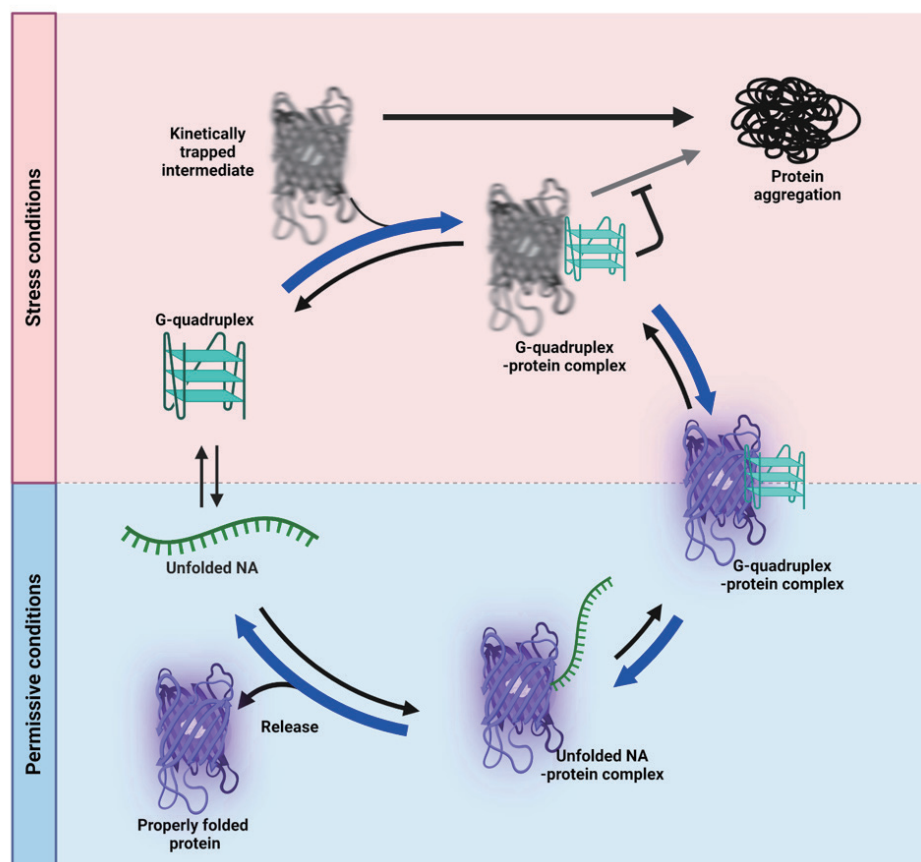
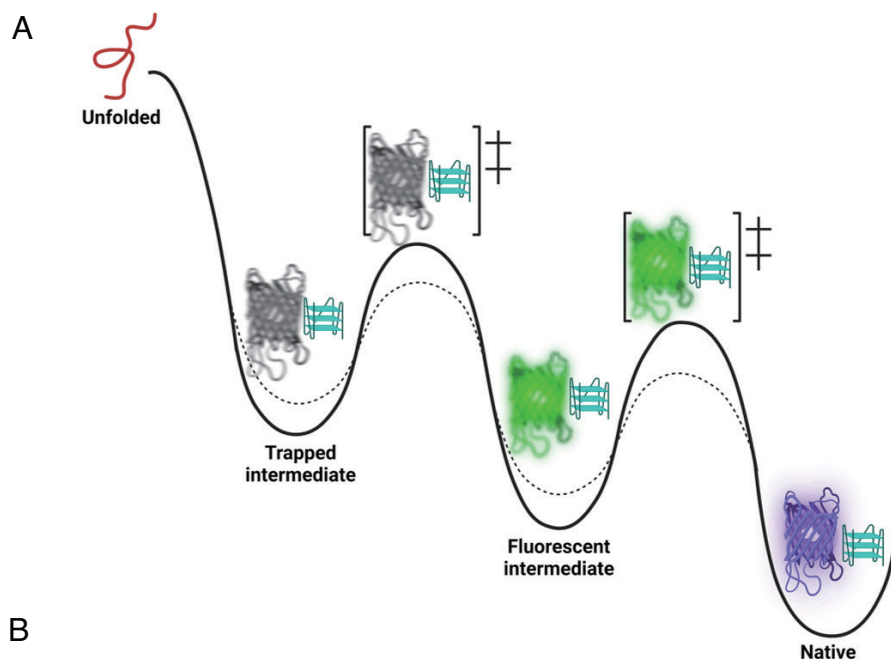


**Fig. 5.** Normalization of expression level, fluorescence, and solubility of TagRFP675 from the fluorescence time course assay in Fig. 4 in BL21(DE3). (A) SDS-PAGE (Top) and western blot assay (Bottom) from the cells at time points  $-3$  h,  $0$  h, and  $4$  h of the time-course assay. A red arrow indicates expressed TagRFP675. The concentrations of first anti-GAPDH, first anti-His6, and second anti-mouse antibodies are listed at bottom right. The full images of gels are shown in *SI Appendix, Fig. S9C*. (B) Normalization of the expression level of TagRFP675 at  $4$  h timepoint. The expression ratio was calculated via ImageJ. Normalized ratio was calculated by dividing the expression level of TagRFP675 by that of GAPDH. (C) Solubility of TagRFP675 at  $4$  h timepoint. Cell lysates taken from  $4$  h timepoints were separated into Soluble (S) and Pellet (P) fractions by centrifugation and visualized by western blot (Top). The expression level of soluble proteins was measured using ImageJ, normalized based on the expression level of GAPDH, and shown in the Bottom panel. (D) Normalization of the fluorescence of TagRFP675 at  $4$  h timepoint. The ratio was calculated by dividing the fluorescence of induced TagRFP675 by the normalized expression level from B. The full images of C and D are shown in *SI Appendix, Fig. S10*.

(Fig. 6A). Of note, it is also possible that similar effects could occur with ATP-independent chaperone proteins.

From the *E. coli* experiments, we found that the G4s affected protein folding primarily, and not protein level or protein solubility. A common assumption made in biological experiments is that soluble protein is typically well-folded and active. However, several studies have pointed out that this is an oversimplification (51–54), as appears to be the case for TagRFP675 when not enough chaperone capacity is available to aid its folding.

In combination with recent work on the stability of G4s under stress conditions, we can now present a potential model for the cycle of G4s in protein folding (Fig. 6B). Under stress conditions, G4s preferentially fold (55). The protein binds to the G4 under stress, and interaction with the G4 rescues kinetically trapped intermediates. It is possible that this could cause a release of the G4, but for those proteins that stay bound, the cessation of stress causes helicase-mediated unfolding of the G4s (55) and release of bound proteins. This cycle could then be repeated if



**Fig. 6.** Schematic view of the effect and mechanism of G4 on protein folding of TagRFP675. (A) Simplified model of the effects of G4 on the folding of TagRFP675. The dashed line represents the two possible effects of the G4 to rescue protein folding. The G4 rescues protein folding either by stabilizing the transition state between the trapped intermediate, later intermediate (green), and native states (purple), or partially destabilizing the trapped intermediate state, or both. Trapped intermediate and transition state are depicted here as fuzzy protein states, while the G4 is shown in cyan. (B) Mechanistic folding cycle of G4 clients. Under stress condition, unfolded single-stranded G4-sequence-containing nucleic acid (NA) tends to form G4 sequences. Kinetically trapped or misfolded intermediates are aggregation-prone without aid of G4s. G4s bind to the intermediates, enhancing proper protein folding. Under nonstress conditions, G4s are likely to unfold, which favors release of properly folded protein from NA-protein complex.

stress recurs. It is also possible that the G4s could still play other roles in proteostasis in addition to protein folding. Of note, this model still leaves room for more nuanced regulation and

mechanistic detail to be added, as well as includes steps that extrapolate between results demonstrated on different model systems.



Concluding, the role of G4s in protein folding and aggregation is still an evolving area of study, but ATP-independent chaperones such as G4s can play potentially substantial roles in dictating the folding state of proteins through rescuing trapped intermediates. In addition to the power of G4s, it is also possible that other nucleic acid structures could have not-yet-identified properties of modifying protein folding.

## Materials and Methods

**Construction of expression vectors and nucleic acid sourcing.** Strains and plasmids are listed in *SI Appendix, Table S1*. pBAD/HisD-TagRFP675, pET28b-His6-TagRFP675 and pBAD18-wtGFP were used for the expression of folding biosensors. pBAD/HisD-TagRFP675 was a gift from Vladislav Verkhusha (Addgene plasmid # 44274; <http://n2t.net/addgene:44274>; RRID:Addgene\_44274) (37). pET28b-His6-TagRFP675 was generated from pET28b vector, using NcoI/EcoRI restriction enzymes. pBAD24-Lon WT was given by Igor Konieczny's laboratory (43). pBAD/HisD-wtGFP was obtained by using NcoI/EcoRI restriction enzymes by using pBAD/HisD-TagRFP675 vector as a mother vector. wtGFP sequence was from pBAD18-wtGFP vector given by Jonathan S. Weismann's laboratory (56). pBAD24-empty vector was generated by using EcoRI and HindIII restriction enzymes, and then blunting and ligating the cleaved vector. pBAD33 vector was used to generate expression vectors of protein folding enhancing factors, i.e., pBAD33-GroEL, pBAD33mut-Seq576, and pBAD33mut-Empty. pBAD33-GroEL was constructed using the SacI and HindIII sites and the GroEL gene was obtained by PCR using the *E. coli* genome. pBAD33mut-empty vector was generated from pBAD33 vector by changing MCS into SacI and SpeI, using MluI and BglII restriction enzymes. The gene of Seq576 was synthesized (Genscript) and inserted into pBAD33mut vector using the SacI and SpeI sites. The structure of RNA expression vector using this paper and expected RNA sequences from the vectors are shown in *SI Appendix, Table S2*.

**DNA and RNA Preparation.** DNAs and RNAs for in vitro refolding assays were ordered from Integrated DNA Technology (IDT) using their standard desalting and purification procedures. Following is the entire sequence information. 1) Seq576, parallel G4 (14): TGT CGG GCG GGG AGG GGG GG, 2) Seq42, non-G4 (14): AAC GAA AGA ACA TAA TCT CG, and 3) LTR-III, hybrid 3+1 (29): GGG AGG CGT GGC CTG GGC GGG ACT GGG G. Sequences were ordered lyophilized and normalized to guaranteed molar weights by IDT, and samples were resuspended in 10 mM potassium phosphate buffer pH 7.5. Prior to every experiment, the stocks were annealed by heating to 95 °C for 2 min and allowed to cool to 25 °C by a controlled temperature ramp of 1 °C/min with 550 rpm shaking in a heating block (Eppendorf ThermoMixer C). For all experiments, the freshly annealed nucleic acid was diluted into the assay/experiment, and the experiment began within 30 min of the annealed samples reaching 25 °C. All RNA samples used here were freshly resuspended and consumed immediately after resuspending/annealing. Seq576 had previously been found to have a high melting temperature and similar structure across the range of temperatures used here (14).

**Protein Expression and Purification.** TagRFP675 was expressed in BL21(DE3) bacterial host cells with plasmids pBAD/HisD-TagRFP675 and co-expressing pBAD33-GroEL. Cells were grown in Luria-Bertani (LB) medium with ampicillin and chloramphenicol for 24 h at 37 °C while shaking at 180 rpm. Afterwards, a starter culture was transferred to fresh LB medium containing the previous antibiotics at 37 °C while shaking at 250 rpm, until an OD<sub>600</sub> of 0.7 was achieved. Cells were induced with 0.2% L-arabinose and allowed to incubate at 250 rpm for 45 min or until spun down aliquots had a distinct red/purple pellet. Cells were collected by centrifugation at 6,000 rpm for 20 min. The supernatant was discarded while the pellet was resuspended in 50 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, 10 mM PMSF, and autoclaved DI water. Cells were then lysed by sonication on ice. The lysate was centrifuged for 30 min at 15,000 g and the supernatant was subsequently filtered using 0.45 µm micron syringe filters. The recombinant protein was purified using Bio-Scale Mini Nuvia IMAC Ni-Charged column. The protein was eluted through a gradient elution with an elution buffer made of 50 mM sodium phosphate, 300 mM NaCl, and 300 mM imidazole. The loading buffer consisted of 50 mM sodium phosphate, 300 mM

NaCl, and 10 mM imidazole. Purification was followed by dialysis overnight in 10 mM sodium phosphate, pH 7.5, and then concentrated using a Amicon Ultra-15 Centrifugal Filter Unit spun at intervals of 20 min at 20,000 rpm. After the 20-min intervals, the sample was pipetted to reach a homogenous solution, and this was done until the desired concentration and volume was achieved. The protein was then flash frozen with 15% glycerol for storage. Purity was assessed through an SDS-PAGE at greater than 95% purity. *SI Appendix, Fig. S11* shows the gel portraying the purified protein sample.

**Protein Denaturation and Fluorescence Recovery Assay.** The fluorescence intensity of renaturing TagRFP675 was measured with a fluorimeter in the presence and absence of G4 containing nucleic acids. TagRFP675 was denatured chemically with 6 M guanidine-HCl (Gu-HCl), 40 mM HEPES, pH 7.5 (KOH) for approximately 11 h at 23 °C in tear drop vials at 20 mM protein. Denatured protein was then diluted into 10 mM potassium phosphate, pH 7.5, in a clear cuvette and with constant stirring at 23 °C. Fluorescence was measured by the fluorimeter set at an emission wavelength of 675 nm, an excitation wavelength of 598 nm, high detector voltage (800 V), and 10-nm slits on excitation and emission. When measuring the fluorescence recovery in the presence of nucleic acid, the nucleic acid was first added to the 10 mM potassium buffer in the cuvette, followed by the dilution of the denatured protein into the cuvette after 1 min of starting the readings in the fluorimeter. All samples and conditions were run in triplicates and graphed in an intensity vs. time plot.

**Fluorescence Recovery: Temperature.** Measurements were recorded at 4 °C and 40 °C. The same procedure was followed as the fluorescence recovery assay in the presence of nucleic acids. Seq576 was present in the potassium buffer prior to the dilution of the denatured TagRFP675. The concentration ratio of TagRFP675 to Seq576 was 1:2. For all temperatures, the buffer was allowed to rest in the cuvette holder for approximately 2 min before diluting protein in buffer.

**Fluorescence Recovery: Delayed Addition of Protein.** Seq576 was added to the sample in the cuvette before and after the dilution of the denatured TagRFP675. Seq576 was added before dilution (0 s), or at later time points indicated in figure after TagRFP675 was diluted in the native buffer. The fluorescence recovery was taken at 23 °C.

**Fluorescence Assay in *Escherichia coli*.** Each expression vector of protein folding enhancing factors and pBAD/HisD-TagRFP675 and pET28b-His6-TagRFP675 were co-transformed into the *E. coli* strain MC4100(DE3) or BL21(DE3) by heat shock. To perform time-course fluorescence assay, the same expression vector set mentioned above was used (Figs. 3A and 4A). Briefly, each transformant was inoculated into 3 mL LB media with appropriate antibiotics and incubated at 37 °C overnight with vigorous shaking. Next day, 200 µL of each culture was transferred into fresh 3 mL LB with the same concentration of antibiotics used above. When the OD<sub>600</sub> reached > 0.7, 0.1% L-Arabinose was added to induce protein and RNA expression in MC4100(DE3), and 0.1% L-Arabinose and 10 µM IPTG were used for the expression in BL21(DE3) (referred as –3 h). After 3 h induction (referred to as 0 h), translation was stopped by adding excess amounts of Spectinomycin (0.5 mg/mL). As high concentrations of spectinomycin ranged from 250 µg/mL to 1 mg/mL were used from previous studies (38–40), we first performed a time-course assay at the various concentrations of spectinomycin to establish an optimum concentration of spectinomycin for our assay (*SI Appendix, Fig. S3A*). The fluorescence of TagRFP675 with 0.1 mg/mL of spectinomycin continuously increased, while more than 0.2 mg/mL of spectinomycin halted the increase in fluorescence. Thus, 0.5 mg/mL was finally chosen for our assay to monitor the degradation of protein in bacteria. Fluorescence was measured by microplate reader (Infinite M200 Pro, Tecan) using a black 96-well plate (Corning Black NBS 3991). Fluorescence emission of TagRFP675 at 675 nm was measured upon excitation at 598 nm.

**SDS-PAGE and Western blot.** Each sample from the time-course fluorescence assay was harvested at certain time points by centrifugation at 10,000 g for 2 min at 4 °C. Cells were lysed with EZLys™ Bacterial Protein Extraction Reagent (BioVision) by following the manufacture's method. Lysed cells were determined by dividing the OD<sub>600</sub> by three to normalize the number of cells per mL for all samples. To examine the solubility of TagRFP675, the lysates were separated into soluble and pellet fractions by centrifugation at 16,000 × g for 15 min at 4 °C. The samples were denatured with the same volume of SDS loading

buffer at 98 °C for more than 20 min and visualized on an SDS-PAGE gel. To perform western blots, each gel was transferred to 0.2 µm PVDF membrane using Trans-Blot Turbo Transfer system (Bio-Rad). To detect TagRFP675, blots were incubated with 6x-His Tag Monoclonal Antibody (Thermo Fisher Scientific). As a control, GAPDH Loading Control Monoclonal Antibody (GA1R) (Thermo Fisher Scientific) was obtained. Rabbit anti Mouse IgG (H+L) Secondary Antibody, HRP (Invitrogen) was used for visualization of bands. The dilution ratio of three antibodies for the samples in MC4100(DE3) and BL21(DE3) are listed in the figures. The protein bands were then visualized by Clarity Max™ Western ECL Substrate (Bio-Rad) using Imaging System Chemidoc (Bio-Rad). The intensity of

each band was quantified by ImageJ software (n = 3) (57), and the expression level of TagRFP675 was calculated based on the expression level of GAPDH (TagRFP675/GAPDH).

**Data, Materials, and Software Availability.** All study data are included in the article and/or *SI Appendix*.

**ACKNOWLEDGMENTS.** We would like to thank J. Bardwell and U. Jakob for helpful discussions and F. Stull and K. Ghosh for commentary on the manuscript. This work was funded by NIH R35GM142442.

- W. E. Balch, R. I. Morimoto, A. Dillin, J. W. Kelly, Adapting proteostasis for disease intervention. *Science* **319**, 916–919 (2008).
- F. U. Hartl, A. Bracher, M. Hayer-Hartl, Molecular chaperones in protein folding and proteostasis. *Nature* **475**, 324–332 (2011).
- M. Haslbeck, E. Vierling, A first line of stress defense: Small heat shock proteins and their function in protein homeostasis. *J. Mol. Biol.* **427**, 1537–1548 (2015).
- B. D. Eisenhardt, Small heat shock proteins: Recent developments. *Biomol. Concepts* **4**, 583–595 (2013).
- I. Baskakov, D. W. Bolen, Forcing thermodynamically unfolded proteins to fold. *J. Biol. Chem.* **273**, 4831–4834 (1998).
- M. J. Gray *et al.*, Polyphosphate is a primordial chaperone. *Mol. Cell* **53**, 689–699 (2014).
- J. Aarum *et al.*, Enzymatic degradation of RNA causes widespread protein aggregation in cell and tissue lysates. *EMBO Rep.* **21**, e49585 (2020), 10.15252/embr.201949585.
- S. Namkoong, A. Ho, Y. M. Woo, H. Kwak, J. H. Lee, Systematic characterization of stress-induced RNA granulation. *Mol. Cell* **70**, 175–187.e8 (2018).
- S. Maharana *et al.*, RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **360**, 918–921 (2018).
- D. Mateju *et al.*, An aberrant phase transition of stress granules triggered by misfolded protein and prevented by chaperone function. *Embo. J.* **36**, 1669–1687 (2017).
- M. Aliquet *et al.*, Assembly of proteins by free RNA during the early phase of proteostasis stress. *J. Proteome. Res.* **18**, 2835–2847 (2019).
- O. Bounedjah *et al.*, Free mRNA in excess upon polysome dissociation is a scaffold for protein multimerization to form stress granules. *Nucleic Acids Res.* **42**, 8678–8691 (2014).
- B. E. Docter, S. Horowitz, M. J. Gray, U. Jakob, J. C. A. Bardwell, Do nucleic acids moonlight as molecular chaperones? *Nucleic Acids Res.* **44**, 4835–4845 (2016).
- A. Begeman *et al.*, G-Quadruplexes act as sequence-dependent protein chaperones. *EMBO Rep.* **21**, e49735 (2020), 10.15252/embr.201949735.
- S. Horowitz, J. C. Bardwell, RNAs as chaperones. *RNA Biol.* **13**, 1228–1231 (2016).
- T. J. Litberg, B. Docter, M. P. Hughes, J. Bourne, S. Horowitz, DNA facilitates oligomerization and prevents aggregation via DNA networks. *Biophys. J.* **118**, 162–171 (2020).
- R. K. Reddy Sannapureddi, M. K. Mohanty, A. K. Gautam, B. Sathyaamoorthy, Characterization of DNA G-quadruplex topologies with NMR chemical shifts. *J. Phys. Chem. Lett.* **11**, 10016–10022 (2020).
- M. T. Banco, A. R. Ferre-D'Amare, The emerging structural complexity of G-quadruplex RNAs. *RNA* **27**, 390–402 (2021).
- D. Varshney, J. Spiegel, K. Zyner, D. Tannahill, S. Balasubramanian, The regulation and functions of DNA and RNA G-quadruplexes. *Nat. Rev. Mol. Cell Biol.* **21**, 459–474 (2020).
- P. Ivanov *et al.*, G-quadruplex structures contribute to the neuroprotective effects of angiogenin-induced tRNA fragments. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 18201–18206 (2014).
- X. Shao *et al.*, RNA G-quadruplex structures mediate gene regulation in bacteria. *mBio* **11**, e02926-19 (2020).
- S. Mestre-Fos, C. Ito, C. M. Moore, A. R. Reddi, L. D. Williams, Human ribosomal G-quadruplexes regulate heme bioavailability. *J. Biol. Chem.* **295**, 14855–14865 (2020).
- L. I. Jansson *et al.*, Telomere DNA G-quadruplex folding within actively extending human telomerase. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 9350–9359 (2019).
- F. Georgescauld *et al.*, GroEL/ES chaperonin modulates the mechanism and accelerates the rate of TIM-barrel domain folding. *Cell* **157**, 922–934 (2014).
- N. K. Tyagi, W. A. Fenton, A. A. Deniz, A. L. Horwich, Double mutant MBP refolds at same rate in free solution as inside the GroEL/GroES chaperonin chamber when aggregation in free solution is prevented. *FEBS Lett.* **585**, 1969–1972 (2011).
- D. Thirumalai, G. H. Lorimer, C. Hyeon, Iterative annealing mechanism explains the functions of the GroEL and RNA chaperones. *Protein Sci.* **29**, 360–377 (2020).
- S. Horowitz, P. Koldewey, F. Stull, J. C. Bardwell, Folding while bound to chaperones. *Curr. Opin. Struct. Biol.* **48**, 1–5 (2018).
- T. D. Craggs, Green fluorescent protein: Structure, folding and chromophore maturation. *Chem. Soc. Rev.* **38**, 2865–2875 (2009).
- E. Butovskaya, B. Heddi, B. Bakalar, S. N. Richter, A. T. Phan, Major G-quadruplex form of HIV-1 LTR reveals a (3 + 1) folding topology containing a stem-loop. *J. Am. Chem. Soc.* **140**, 13654–13662 (2018).
- B. T. Andrews, M. Roy, P. A. Jennings, Chromophore packing leads to hysteresis in GFP. *J. Mol. Biol.* **392**, 218–227 (2009).
- S. T. Hsu, G. Blaser, S. E. Jackson, The folding, stability and conformational dynamics of beta-barrel fluorescent proteins. *Chem. Soc. Rev.* **38**, 2951–2965 (2009).
- J. R. Huang, T. D. Craggs, J. Christodoulou, S. E. Jackson, Stable intermediate states and high energy barriers in the unfolding of GFP. *J. Mol. Biol.* **370**, 356–371 (2007).
- O. V. Stepanenko *et al.*, Comparative studies on the structure and stability of fluorescent proteins EGFP, zFP506, mRFP1, “dimer2”, and DsRed1. *Biochemistry* **43**, 14913–14923 (2004).
- H. Fukuda, M. Arai, K. Kuwajima, Folding of green fluorescent protein and the cycle3 mutant. *Biochemistry* **39**, 12025–12032 (2000).
- K. A. Dill, S. Bromberg, *Molecular Driving Forces: Statistical Thermodynamics in Chemistry and Biology* (Garland Science, New York, 2003), 666 p.
- J. Ervin, E. Larios, S. Osvath, K. Schulten, M. Gruebele, What causes hyperfluorescence: Folding intermediates or conformationally flexible native states? *Biophys. J.* **83**, 473–483 (2002).
- K. D. Piatkevich *et al.*, Extended Stokes shift in fluorescent proteins: Chromophore-protein interactions in a near-infrared TagRFP675 variant. *Sci. Rep.* **3**, 1847 (2013).
- C. Barembruch, R. Hengge, Cellular levels and activity of the flagellar sigma factor FliA of *Escherichia coli* are controlled by FlgM-modulated proteolysis. *Mol. Microbiol.* **65**, 76–89 (2007).
- Z. Granot *et al.*, Turnover of mitochondrial steroidogenic acute regulatory (StAR) protein by Lon protease: The unexpected effect of proteasome inhibitors. *Mol. Endocrinol.* **21**, 2164–2177 (2007).
- K. Yoshitani, Y. Hizukuri, Y. Akiyama, An in vivo protease activity assay for investigating the functions of the *Escherichia coli* membrane protease HtpX. *FEBS Lett.* **593**, 842–851 (2019).
- L. Ambro, V. Pevala, J. Bauer, E. Kutejova, The influence of ATP-dependent proteases on a variety of nucleoid-associated processes. *J. Struct. Biol.* **179**, 181–192 (2012).
- S. Kubik, K. Wegrzyn, M. Pieriechod, I. Konieczny, Opposing effects of DNA on proteolysis of a replication initiator. *Nucleic Acids Res.* **40**, 1148–1159 (2012).
- A. Karłowicz *et al.*, Defining the crucial domain and amino acid residues in bacterial Lon protease for DNA binding and processing of DNA-interacting substrates. *J. Biol. Chem.* **292**, 7507–7518 (2017).
- C. H. Chung, A. L. Goldberg, DNA stimulates ATP-dependent proteolysis and protein-dependent ATPase activity of protease La from *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 795–799 (1982).
- S. H. Chen, C. K. Suzuki, S. H. Wu, Thermodynamic characterization of specific interactions between the human Lon protease and G-quartet DNA. *Nucleic Acids Res.* **36**, 1273–1287 (2008).
- M. L. Wohlever, A. R. Nager, T. A. Baker, R. T. Sauer, Engineering fluorescent protein substrates for the AAA+ Lon protease. *Protein Eng. Des. Sel.* **26**, 299–305 (2013).
- L. Huang *et al.*, DAXX represents a new type of protein-folding enabler. *Nature* **597**, 132–137 (2021).
- K. Wu, F. Stull, C. H. Lee, J. C. A. Bardwell, Protein folding while chaperone bound is dependent on weak interactions. *Nat. Commun.* **10**, 4833 (2019).
- A. Mashaghi *et al.*, Reshaping of the conformational search of a protein by the chaperone trigger factor. *Nature* **500**, 98–101 (2013).
- F. Moayed *et al.*, The anti-aggregation holdase Hsp33 promotes the formation of folded protein structures. *Biophys. J.* **118**, 85–95 (2020).
- D. A. Nissley *et al.*, Universal protein misfolding intermediates can bypass the proteostasis network and remain soluble and less functional. *Nat. Commun.* **13**, 3081 (2022).
- J. Danielsson *et al.*, Thermodynamics of protein destabilization in live cells. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 12402–12407 (2015).
- S. Ebbinghaus, A. Dhar, D. McDonald, M. Gruebele, Protein folding stability and dynamics imaged in a living cell. *Nat. Methods* **7**, 319–323 (2010).
- E. Waztor, M. Rutkiewicz, M. S. Weiss, P. Wilk, Co-expression with chaperones can affect protein 3D structure as exemplified by loss-of-function variants of human prolidase. *FEBS Lett.* **594**, 3045–3056 (2020).
- P. Kharel *et al.*, Stress promotes RNA G-quadruplex folding in human cells. *Nat Commun* **14**, 205 (2023), <https://doi.org/10.1038/s41467-023-35811-x>.
- J. D. Wang, C. Herman, K. A. Tipton, C. A. Gross, J. S. Weissman, Directed evolution of substrate-optimized GroEL/S chaperonins. *Cell* **111**, 1027–1039 (2002).
- C. A. Schneider, W. S. Rasband, K. W. Eliceiri, NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).