Small molecule probes to quantify the functional fraction of a specific protein in a cell with minimal folding equilibrium shifts

Yu Liu[a,b,c,1], Yun Lei Tan[a,b,c,1], Xin Zhang[a,b,c,1,2], Gira Bhabha[d], Damian C. Ekiert[e], Joseph C. Genereux[a,b,c], Younhee Cho[a,b,c], Yakov Kipnis[f], Sinisa Bjelic[f], David Baker[f], and Jeffery W. Kelly[a,b,c,2]

Departments of [a]Molecular and Experimental Medicine and [b]Chemistry and [c]The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037; Departments of [d]Cellular and Molecular Pharmacology and [e]Microbiology and Immunology, University of California, San Francisco, CA 94158; and [f]Department of Biochemistry, University of Washington, Seattle, WA 98195

Edited by F. Ulrich Hartl, Max Planck Institute of Biochemistry, Martinsried, Germany, and approved February 3, 2014 (received for review December 16, 2013)

Although much is known about protein folding in buffers, it remains unclear how the cellular protein homeostasis network functions as a system to partition client proteins between folded and functional, soluble and misfolded, and aggregated conformations. Herein, we develop small molecule folding probes that specifically react with the folded and functional fraction of the protein of interest, enabling fluorescence-based quantification of this fraction in cell lysate at a time point of interest. Importantly, these probes minimally perturb a protein's folding equilibria within cells during and after cell lysis, because sufficient cellular chaperone/chaperonin holdase activity is created by rapid ATP depletion during cell lysis. The folding probe strategy and the faithful quantification of a particular protein's functional fraction are exemplified with retroaldolase, a de novo designed enzyme, and transthyretin, a nonenzyme protein. Our findings challenge the often invoked assumption that the soluble fraction of a client protein is fully folded in the cell. Moreover, our results reveal that the partitioning of destabilized retroaldolase and transthyretin mutants between the aforementioned conformational states is strongly influenced by cytosolic proteostasis network perturbations. Overall, our results suggest that applying a chemical folding probe strategy to other client proteins offers opportunities to reveal how the proteostasis network functions as a system to regulate the folding and function of individual client proteins in vivo.

chemical probes | pharmacologic chaperone | fluorescence labeling

All proteins are biosynthesized as linear chains, and most need to fold into 3D structures to function. Studies on protein folding in buffers have revealed that a kinetic competition typically exists between protein folding, misfolding, and aggregation. It is the role of the protein homeostasis or proteostasis network in each subcellular compartment to regulate this competition and keep the folded and functional proteome within the physiological concentration range, while minimizing misfolding and aggregation in the face of stresses (1–4). It remains a challenge to discern how the proteostasis network affects the folding of proteins into biologically active conformations required for function in vivo (5).

Current methodologies allow for quantification of the partitioning of a protein of interest (POI) between soluble and aggregated states but cannot determine the proportion of the soluble population that is properly folded and functional. Published folding probes have the potential to report on the folded fraction in cells or cell lysate (6–9); however, the extent to which they shift folding equilibria and quantify the folded and functional fraction faithfully has not been studied. Herein, we create POI folding probes by adapting the principle of activity-based protein profiling (10) to quantify the soluble folded and functional fraction of a particular protein in a cell lysate. We seek folding probes that bind to and selectively react with only the folded and functional state of a POI in a cell, leaving the nonfunctional states and other cellular proteins unmodified (Fig. 1*A*).

Fluorescent folding probes for the de novo-designed enzyme, retroaldolase (RA) (11), and fluorogenic folding probes (12) for the nonenzyme protein, transthyretin (TTR), were developed and scrutinized. We show that destabilized mutant RA and TTR proteins partition into folded and functional as well as misfolded soluble conformations and that this partitioning is sensitive to proteostasis network perturbations. Experiments show that a snapshot of the distribution between folded and functional vs. soluble and misfolded conformational states can be preserved during the small molecule folding probe labeling period, provided that the cellular chaperone holdase activity is sufficient, achieved by rapid ATP depletion in parallel with cell lysis. Sufficient chaperone/chaperonin holdase activity minimizes changes in the folded and functional concentration associated with probe binding and reaction with the POI and renders the relative folding and conjugation rates much less influential.

Results

Probe Strategy to Quantify the Folded and Functional Fraction. We exemplify the folding probe strategy (Fig. 1*A*) by, first, converting a substrate of RA into a folding probe and second, transforming a fluorogenic small molecule that covalently labels the

Significance

Historically, the folding of individual proteins in buffers has been studied spectroscopically. The majority of spectroscopic methods (NMR and fluorescence excluded) cannot be used in a cell, because the protein of interest (POI) cannot be distinguished from the background proteome. Herein, we introduce folding probes, which when used in cell lysates with sufficient holdase activity, faithfully quantify the folded and functional fraction of a POI at a time point of interest in a cell by selectively reacting with that state to afford a fluorescent signal. This work provides a blueprint for how to convert enzyme inhibitors, ligands for nonenzyme proteins, etc. into folding probes to efficiently and specifically investigate how intracellular function is controlled by the proteostasis network as a function of cellular perturbations.

Author contributions: Y.L., Y.L.T., X.Z., and J.W.K. designed research; Y.L., Y.L.T., X.Z., G.B., D.C.E., and J.C.G. performed research; Y.L., Y.L.T., X.Z., Y.C., Y.K., S.B., and D.B. contributed new reagents/analytic tools; Y.L., Y.L.T., X.Z., G.B., D.C.E., J.C.G., and J.W.K. analyzed data; and Y.L., Y.L.T., X.Z., and J.W.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The crystallographic data have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 4OU1).

[1]Y.L., Y.L.T., and X.Z. contributed equally to this work and are listed alphabetically.

[2]To whom correspondence may be addressed. E-mail: xinzhang@scripps.edu or jkelly@scripps.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1323268111/-/DCSupplemental.

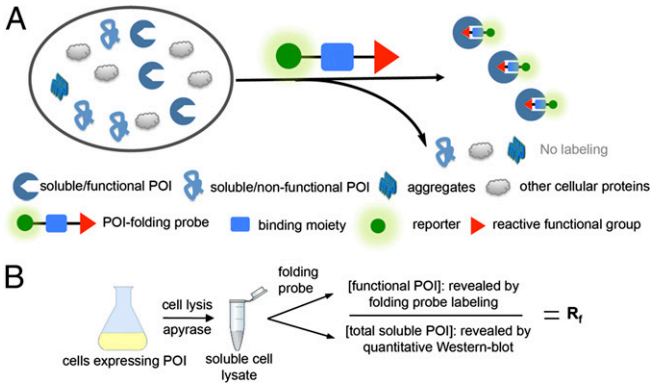

Fig. 1. A small molecule folding probe strategy to quantify the soluble folded and functional fraction of a POI in a cell lysate. (*A*) Overview of the general strategy to selectively covalently label a folded and functional POI without labeling its nonfunctional conformations and other cellular proteins. (*B*) The experimental scheme to quantify the ratio of the soluble POI that is functional (R_f).

TTR native state into a folding probe (13). A question to be addressed using small molecule folding probes is what fraction of the soluble POI is correctly folded and functional vs. soluble but nonfunctional (Fig. 1*B*) at any point in time in a cell. The POIs in this study were expressed in *Escherichia coli*, and the resulting cells were lysed in the presence of apyrase to deplete ATP. The lysates were centrifuged to remove insoluble and large soluble aggregates of the POI, and the concentration of the total soluble POI remaining (including folded, misfolded, and misassembled) was determined by quantitative immunoblotting in the linear range by LI-COR. To quantify the concentration of the folded and functional POIs at each time point of interest in the cell before lysis, the ATP-depleted soluble lysate was incubated with a folding probe that reacts only with the folded and functional POI, achieving complete labeling before quantification of the fluorescent signal of the probe. These two concentrations afford the fraction of the POI that is soluble, folded, and functional (i.e., R_f in Fig. 1*B*).

Design of Folding Probes for RA. We chose the designed RA enzyme because it is not found in nature. Thus, it can be incorporated into cells to assess proteostasis network influences on RA folding without concerns about loss- or gain-of-function cellular phenotypes. The RA of focus, RA114.3 (14), adopts an eight-stranded α/β barrel fold (Fig. S1*A*), a common natural enzyme scaffold (15). A pK_a-perturbed active site lysine residue [K210 (16), pK_a of 6.6] (Fig. S1*B*) is used for catalysis (Fig. S1*C*), akin to aldolase enzymes (17). Because nucleophilic amino acid side chains are frequently found in proteins as either the catalytic residue(s) or near the substrate binding pocket, RA provides a prototype to test the POI folding probe strategy.

The reactive small molecule folding probes for RA (probes **1**–**3** in Fig. 2*A*) that we designed and synthesized are derived from its substrate (Fig. S1*C*) and incorporate three essential components: the methoxy-naphthalene substructure conferring substrate binding affinity and structural complementarity to the active site of RA (Figs. 1*A* and 2*A*, substructure colored blue), the epoxide electrophile on the naphthalene ring enabling catalytic K210 labeling but not general proteome reactivity (Figs. 1*A* and 2*A*, substructure colored red), and a fluorescence reporter or affinity tag covalently attached by a linker to what was the methoxy group on the naphthalene (Figs. 1*A* and 2*A*, substructure colored green).

The ε-amino group of K210 is the nucleophile to be targeted by the RA folding probes. To identify an electrophile that reacted with only K210 of the folded and functional state of RA but not other nucleophiles in proteome, we synthesized probe **1** comprising an α,β-epoxyalcohol (Fig. 2*A*). A 1.25-Å resolution crystal structure of the RA–probe **1** conjugate showed that K210 was covalently modified by probe **1** (Fig. 2*B* and Table S1). The chemoselectivity of probe **1** for the K210 residue of RA was confirmed by an MS analysis that detected no reaction between probe **1** and the K210A RA mutant (Fig. 2*C*). To engineer a reporter or affinity function into the folding probes, the methoxy end of probe **1** was modified by attaching either a biotin affinity tag (Fig. 2*A*, probe **2**) or a fluorescein fluorophore (Fig. 2*A*, probe **3**). The conjugation kinetics of probe **3** reacting with RA were measured using electrophoresis and single turnover kinetic experiments using fluorescein fluorescence (Fig. S1 *D*–*G*). Probe **3** efficiently labeled RA with a maximum rate constant (k_{inact}) of 0.75 min^{-1}, the rate of the chemical step (Fig. S1*H*). The second-order association rate constant (k_{inact}/K_i) between probe **3** and RA is 5.51×10^3 M^{-1} min^{-1}.

We next evaluated the ability of fluorescent probe **3** to distinguish functional RA from its nonfunctional soluble conformations, using a pH titration. Previous studies revealed that acidic conditions protonate the K210 lysine, thus reducing the concentration of functional RA (16). We observed that the k_{cat}/K_m values decreased on buffer acidification (Fig. S1*B*), indicating reduced RA catalysis without significant changes to its structure, which was ascertained by far-UV CD, tryptophan fluorescence, and ^{1}H-NMR spectroscopies (Fig. S2 *A*–*C*, respectively). Accordingly, the amount of RA that was labeled by probe **3** also declined on buffer acidification (Fig. 2*D*). Notably, the decrease in labeling efficiency of probe **3** closely mirrored the decrease in the catalytic activity of RA (Fig. 2*E*), showing that probe **3** only reacts with functional RA.

Folding probes must be selective to their target in a complex cellular proteome to be useful in vivo. To test for selectivity, probe **3** (200 μM) was incubated with concentrated *E. coli* cell lysate with or without the addition of fully functional RA. Importantly, no significant off-target fluorescent conjugate bands

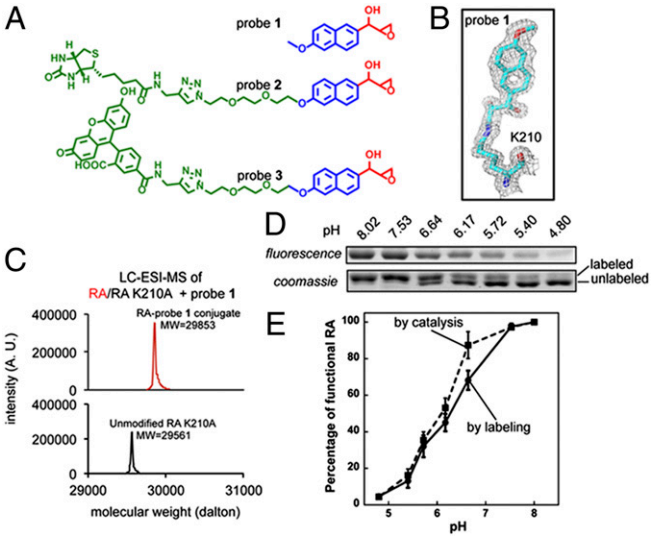

Fig. 2. The RA folding probes selectively label folded and functional RA. (*A*) Structures of the substrate-based RA folding probes using a reactive functional group (red), a binding substructure based on the substrate (blue), and a fluorophore or an affinity tag attached by a flexible linker (green). (*B*) Probe **1** reacts with K210 of RA to afford a covalent conjugate, which is shown in the $2F_o - F_c$ omit map (σ = 1.0) of the K210–probe **1** structure (Protein Data Bank ID code 4OU1). (*C*) Probe **1** reacts with RA but not the K210A RA mutant as shown by liquid chromatography–electrospray ionization–MS (LC-ESI-MS). MW, molecular weight. (*D*) The labeling efficiency of RA by probe **3** declines with acidification. (*E*) The relative labeling efficiency correlates with retroaldolase catalytic activity as a function of pH, indicating that the RA labeling reaction and RA retro-aldol catalysis take place by similar chemistries.

were observed (Fig. S2*D*). Moreover, no bands were detected when the RA added to *E. coli* lysate was denatured by boiling before incubation with probe **3** (Fig. S2*D*, lane 5). These results strongly suggest that probe **3** is highly selective to functional RA conjugation (Fig. S2*D*, lane 4), allowing it to be used in vivo. High-affinity binding and a fast probe reaction maximize selectivity.

Soluble Nonfunctional and Folded RA in Cellular Lysate. To discern whether RA adopts both soluble functional and soluble nonfunctional conformations in *E. coli* cytosol, we studied two mutants of RA, RAm1 (E10K:D120V:N124S:L225P) and RAm2 (K135I), that are as catalytically active as the parent RA when folded (Fig. 3*A*, *Right*). The parent RA is the most thermodynamically stable, exhibiting a urea denaturation midpoint (C_m) of 6.5 M, the concentration affording equal folded and unfolded populations (Fig. 3*A*, *Left*). The RAm1 and RAm2 mutants are thermodynamically destabilized, reflected by their C_m values of 3.8 and 5.6 M urea, respectively. Interestingly, RAm1, unlike RAm2, populates a partially denatured intermediate. When these energetically distinct RAs were expressed in *E. coli*, they produced roughly equal amounts of total protein (Fig. 3*B*). The parent RA protein is largely soluble, whereas the mutants RAm1 and RAm2 are three- and sevenfold less soluble in *E. coli* (Fig. 3*B*). The lower concentration of the soluble fraction and enhanced aggregate formation (I in Fig. 3*B*) suggest that both RA mutants were defective in folding.

To derive the ratio of soluble RA that is functional (R_f), the concentration of total soluble RA in lysate was determined by quantitative immunoblotting calibrated by a standard curve (Fig. S3*A*) after RA aggregates were removed (*SI Results*, *Note S1*). The concentration of folded and functional RA was determined by incubating the soluble ATP-depleted cell lysate with probe **3** for 1 h, achieving complete labeling before electrophoresis (Fig. 3*C*) and quantification of the fluorescent RA–probe **3** conjugate band (Fig. 3 *D* and *E*) by comparison with a standard curve (Fig. S3*B*). The soluble parent RA was predominantly (93%) functional in *E. coli* lysate (Fig. 3 *C*–*E*). Of the three- and sevenfold less soluble RAm1 and RAm2 found in *E. coli* lysate relative to parent RA, only 65% and 55% was correctly folded and functional, respectively (Fig. 3 *C*–*E*). These results were validated by two additional approaches. First, a small (~700 Da) size increase of probe **3**-labeled RA was detectable by electrophoresis/immunoblotting analysis (Fig. 3*C*, row 2). Second, by labeling RA with probe **2**, we could visualize the folded and functional populations using an antibody to biotin (Fig. 3*C*, rows 3–5 and Fig. S3*C*). Using a standard curve, we quantified the concentration of folded and functional RA (as labeled by probe **2**), which afforded very similar R_f values to the R_f values measured by probe **3** (Fig. 3*E*). These data collectively suggest that destabilized RA mutants, when biosynthesized in a cellular context, can be soluble but in misfolded and nonfunctional conformations, calling into question the often used assumption that solubility equates to proper folding.

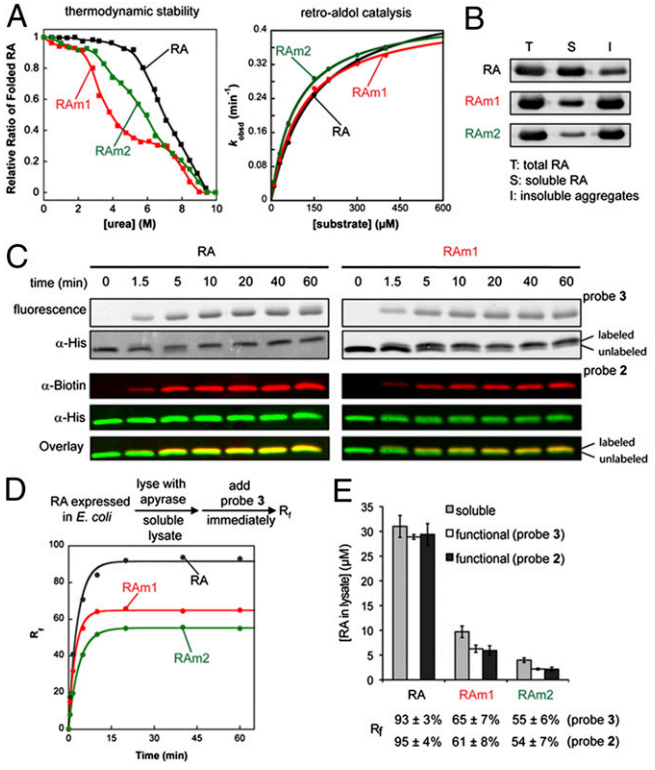

Fig. 3. Soluble RA mutants are only partially functional in *E. coli* cell lysate. (*A*) Metastable RAm1 (E10K:D120V:N124S:L225P) and RAm2 (K135I) are thermodynamically destabilized mutants of RA (*Left*) but are not defective in their ability to catalyze the retro-aldol reaction after folding (*Right*). (*B*) RAm1 and RAm2 form insoluble aggregates when expressed in *E. coli* to a greater extent than the parent RA sequence. (*C*) Time course of probes **2** and **3** labeling of folded and functional RA (*Left*) and RAm1 (*Right*). Probe **3** labeling was quantified by fluorescence (rows 1 and 2), whereas probe **2** labeling was visualized by a two-color Western blot analysis (rows 3–5). (*D*) Probe **3** (200 μM) rapidly labels folded and functional RA and its mutants in *E. coli* lysates depleted of ATP ($k_{obsd} = 0.41\ min^{-1}$), with complete labeling of the functional RA or RA mutants occurring within 15 min. The bacteria express roughly equivalent amounts of total RA. (*E*) Soluble RA, and especially the destabilized mutants, are not fully functional in lysate. Concentrations of total soluble (gray) RA, RAm1, and RAm2 in lysate quantified by immunoblot and soluble functional RA and RAm1 and RAm2 quantified by probes **3** (white) and **2** (black) were used to derive the R_f values. Data are represented as mean ± SD.

Fluorogenic Folding Probe for Transthyretin, a Nonenzyme. As another example, we show that probe **4** (Fig. 4*A*), previously used to quantify the native TTR concentration in *E. coli* lysate (13), could be used as a folding probe. Probe **4**, comprising a stilbene binding motif and a vinyl sulfonamide electrophile, chemoselectively alkylates the pK_a-perturbed K15 residue on the periphery of the two thyroxine binding sites of the TTR tetramer. Probe **4** is fluorogenic (12) [i.e., it is dark in *E. coli* lysate lacking TTR, remains dark after binding TTR, and only becomes fluorescent after reacting with properly folded tetrameric TTR (13)]. The excellent selectivity of probe **4** for covalently modifying the TTR tetramer in *E. coli* lysate was previously shown (13).

The selectivity and fluorogenicity of probe **4** makes the determination of an R_f value very convenient, because a separation step is not required like it is with RA probe **3**. The concentration of folded TTR in *E. coli* lysate was quantified in a fluorescent plate reader by comparison of the fluorescent signal with a standard curve after probe **4** labeling was complete (30 min) (13). Using the experimental strategy outlined in Fig. 1*B*, we discovered that WT-TTR and its destabilized mutant, A25T-TTR (18), partitioned into soluble and nonfunctional states as well as a folded state in cell lysate (Fig. 4*B*). Although soluble WT-TTR was mostly tetrameric and functional in *E. coli* lysate (80%), only 32% of the soluble destabilized A25T-TTR protein was a functional tetramer, despite the fact that the concentration of soluble A25T-TTR was nearly twofold higher than the concentration of the WT-TTR (Fig. 4*C*). These data show that destabilized proteins can be soluble but in misfolded and nonfunctional conformations in cells.

Probes Minimally Shift Equilibria in ATP-Depleted Lysate. There is a general perception that folding probes, such as **3** and **4**, will substantially overestimate the folded and functional fraction of the POI in cell lysate, because after that fraction binds to and reacts with a folding probe, the resulting conjugate is stabilized; therefore, the population of the folded and functional fraction will increase. Thus, the folding equilibrium or equilibria will be

BIOCHEMISTRY

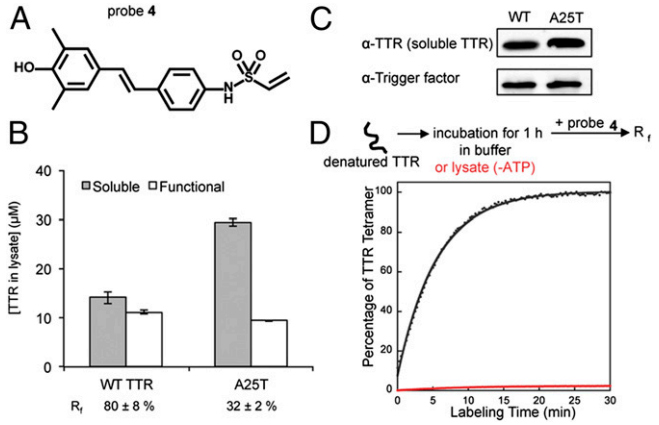

Fig. 4. Probe 4 reveals soluble but nonfunctional TTR in *E. coli* cell lysate. (*A*) Structure of folding probe **4**. (*B* and *C*) Not all soluble WT-TTR and its destabilized A25-TTR mutant in lysate (expressed at similar levels) are functional. Concentrations of total soluble (gray) and soluble folded and functional tetrameric (white) WT-TTR and A25T-TTR in lysate were quantified by (*C*) quantitative Western blot and (*B*) probe **4** reactivity, respectively. These concentrations were used to derive the R_f values. Data are represented as mean ± SD. (*D*) Denatured WT-TTR refolds to predominantly folded and functional TTR tetramer in buffer, whereas folding is 98% prevented by the cellular holdase activity in *E. coli* lysate created by ATP depletion.

shifted to the stabilized folded and functional population (19). In buffer, the magnitude of this so-called pharmacologic chaperoning effect depends on the probe labeling reaction rate, the folding rate of the POI, and the conjugation period (Fig. S4) (19, 20).

We tested the hypothesis that depleting ATP during cell lysis (apyrase treatment) to increase the cellular chaperone/chaperonin holdase activity would make folding probe labeling largely independent of the conjugation rate of the probe, the conjugation period, and the folding rate of the POI. To this end, we added denatured RAm1 to *E. coli* lysate supplemented with ATP (5 mM) or after apyrase-mediated ATP depletion. After a 1-h incubation period, only 5 ± 2% denatured RAm1 could fold to functional conformations in *E. coli* lysate depleted of ATP during the incubation and labeling periods (Fig. 5*A*). In other words, under conditions of ATP depletion, <5% of the RA escaped the considerable cellular chaperone/chaperonin holdase activity and folded. In contrast, the presence of ATP enables cellular foldase function, allowing 15 ± 3% of the denatured RAm1 to fold to a functional conformation during the incubation and labeling periods (Fig. 5*A*). This interpretation was substantiated by an in vitro reconstitution experiment, wherein denatured RAm1 was added to buffer with GroEL lacking ATP for 1 h. Then, probe **3** was added, revealing <2.5% refolding in the presence of GroEL, an established holdase chaperone (Fig. S5*A*) (21).

To scrutinize whether cellular holdase activity prevents nonfunctional RAm1 from converting into a functional conformation during the labeling period, we allowed denatured RAm1 to refold for 1 h in buffer and used probe **3** to label folded RAm1 for the indicated period without or with added GroEL lacking ATP (a holdase) (Fig. 5*B*). Without GroEL, two labeling kinetic phases are apparent: the faster phase deriving from the labeling of prefolded RAm1 and the slower phase reflecting a pharmacologic chaperoning phase, which was discerned by the agreement between the fit of the experimental data (Fig. 5*B*, black curve) and the curve predicted by the pharmacologic chaperoning model (Fig. 5*B*, green curve and Fig. S4 *A* and *B* showing the pharmacologic chaperone model). In contrast, if probe **3** plus GroEL lacking ATP is added to RAm1 after 1 h of folding in buffer, labeling kinetics (Fig. 5*B*, blue curve) are accurately recapitulated by the holdase trapping model (i.e., there is no second kinetic phase associated with pharmacologic chaperoning or probe-associated folding equilibrium shifts) (Fig. 5*B*, compare blue and red curves and Fig. S4*C* showing the holdase trapping model). The folded and functional fraction of RA quantified by probe **3** in the labeling period exhibits the expected conjugation rate constant (Fig. S5*B*) in the presence of the GroEL holdase, identical to the probe **3** conjugation reaction derived from fully folded and functional RAm1 in buffer (Fig. S5*B*).

Based on these results, it is clear that ATP depletion during cell lysis creates sufficient cellular holdase activity (22) that soluble misfolded and/or misassembled RA is effectively held in that state, allowing quantification of the folded and functional fraction in the cell lysate without or with only a minimal pharmacologic chaperoning effect (<5%). If RA is biosynthesized in *E. coli* and the cells are lysed with ATP depletion, the kinetics of probe **3** labeling (Fig. 5*C*, black curve) faithfully fit the time course predicted by the holdase trapping model (Fig. 5*C*, red curve and Fig. S4 *A* and *C* showing the model). We also showed that folded and functional RA purified from *E. coli* retained an $R_f = 1$ in buffer (Fig. S5*C*) and ATP-depleted *E. coli* lysate (Fig. S5*D*), indicating that the cellular holdase activity affected by ATP depletion did not lead to measurable unfolding of RA after folding. If a 1:1 mixture of denatured RAm1 and fully folded RAm1 was added to *E. coli* lysate depleted of ATP and incubated for 1 h before adding probe **3**, only one kinetic phase was observed (Fig. S5*E*), with a rate constant that was identical to the probe **3** conjugation reaction derived from folded and functional RAm1 (Fig. S5*B*). Again, less than 5% of RA escaped

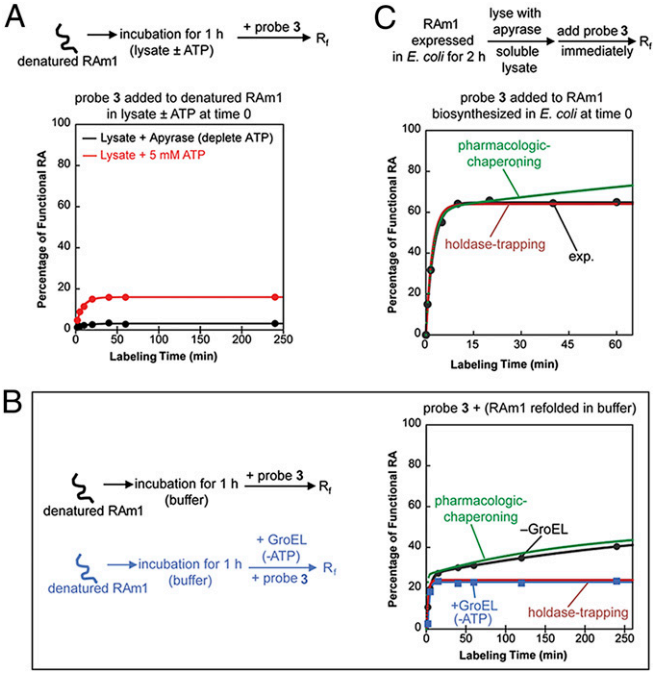

Fig. 5. Cellular holdase chaperone activity traps nonfunctional RAs to prevent probe-mediated refolding. (*A*) When fully denatured RAm1 (6 μM) was refolded in *E. coli* lysate without (black) or with ATP (5 mM; red), only 5 ± 2% and 16 ± 3% of functional RA, respectively, were formed after a 1-h folding period followed by the probe **3** (200 μM) labeling period shown. (*B*) Covalent labeling by probe **3** (black curve) perturbs the refolding equilibrium of RAm1 in buffer in the absence of GroEL (i.e., without holdase activity), which was predicted by the pharmacologic chaperoning model (green curve) (Fig. S4*B*). The presence of GroEL without ATP effectively prevents the pharmacologic chaperoning effect (blue curve), which was predicted by the holdase trapping model (red curve) (Fig. S4*C*). (*C*) The presence of cellular chaperone holdase activity in the lysate inhibits biologically expressed nonfunctional RAm1 from converting to a folded and functional conformation (black curve) in the presence of probe **3**, which was predicted by the holdase trapping model (red curve) (Fig. S4*C*).

the cellular holdase activity and folded during the folding and labeling periods.

Analogous experiments show a negligible pharmacologic chaperoning effect for probe **4** labeling of TTR with adequate holdase activity. When denatured WT-TTR was incubated in ATP-depleted *E. coli* lysate for 1 h and then subjected to probe **4** labeling, only 2% of TTR escaped the holdase activity and folded, whereas folding in buffer resulted in >95% folding during the incubation and labeling periods (Fig. 4*D* and Fig. S6*A* shows additional supportive data).

Collectively, these results and additional control experiments described in *SI Results*, *Note S2: No Significant Pharmacologic Chaperoning Effect by Probe 3 Labeling in the Presence of Cellular Chaperone Holdases* and *SI Results*, *Note S3: Estimation of the Correction for the Pharmacologic Chaperoning Effect of Probe 3 Labeling in Cellular Lysate Depleted of ATP* show that probes **3** and **4** exhibit a minimal pharmacologic chaperone effect (2–5%) when using 30-min or 1-h labeling periods (Fig. 4*D* and Fig. S6*C*) and carried out in the presence of adequate cellular holdase activity. Because we are generally making comparisons in this paper, these small corrections were not applied below.

Proteostasis Network Perturbations Influence RA R_f. We determined the fraction of folded and functional RA as a function of cellular proteostasis network perturbations (Fig. 6*A*). The Hsp70 (DnaK), Hsp40 (DnaJ), and nucleotide exchange factor (GrpE) pathway (23, 24) (the DnaKJE pathway) and the chaperonin (GroEL) cochaperonin (GroES) pathway (25) (the GroEL/ES pathway) are the two most important pathways in the bacterial cytosolic proteostasis network. They typically assist their client proteins to attain function, while minimizing misfolding and aggregation.

When DnaKJE was overexpressed (Fig. S7*A*), the concentration of soluble RAm1 did not change significantly; however, there was a marked decrease in the R_f from 65% to 30% (Fig. 6 *B* and *C* and Fig. S7*B*), suggesting that recruitment of RAm1 to the DnaKJE pathway produced more soluble and nonfunctional conformations through a holdase function, consistent with previous observations (23, 26, 27). The concentration of soluble RAm2 increased fourfold with overexpression of DnaKJE, but the R_f of RAm2 decreased dramatically to 3% (Fig. 6 *D* and *E*). Overexpressing GroEL/ES increased the soluble concentration of RAm1 by a factor of two, and almost all of the soluble RAm1 was functional with an R_f of ∼92% (Fig. 6 *B* and *C* and Fig. S7*B*), which was not surprising, because most eight-stranded α/β barrel proteins are bona fide GroEL/ES clients (21). Both the soluble and functional concentrations of RAm2 were enhanced by a factor of 1.5 with GroEL/ES overexpression; however, its R_f value remained the same (Fig. 6 *D* and *E*). The RA mutants share ∼98% sequence identity, but it seems that they have different chaperone/chaperonin dependencies.

Although DnaKJE and GroEL/ES cooverexpression (Fig. S7*A*) preserved the concentration of soluble RAm1 observed with GroEL/ES overexpression, the DnaKJE machinery decreased the R_f value relative to GroEL/ES overexpression alone (Fig. 6 *B* and *C* and Fig. S7*B*), confirming DnaKJE holdase function. When both pathways were overexpressed, RAm2 retained the increased concentration observed with DnaKJE pathway overexpression alone and most of the increased function observed with GroEL/ES pathway overexpression alone (Fig. 6 *D* and *E*), indicating that the GroEL/ES pathway appears to be important for the small fraction of RAm2 that is folded.

Overexpressing polyglutamine-containing proteins is established to reduce cytosolic proteostasis capacity, resulting in the misfolding and/or aggregation of metastable proteins not harboring a polyglutamine expansion (28, 29) (e.g., the firefly luciferase folding sensor). Thus, we investigated RAm1 folding and function on overexpression of the polyglutamine-expanded ataxin-3 protein Atx3(tr)-Q62 (62 CAG repeats), which is linked to Machado–Joseph disease (30). Overexpressed Atx3(tr)-Q62 formed aggregates in the *E. coli* cytosol (Fig. 6*F*). The soluble

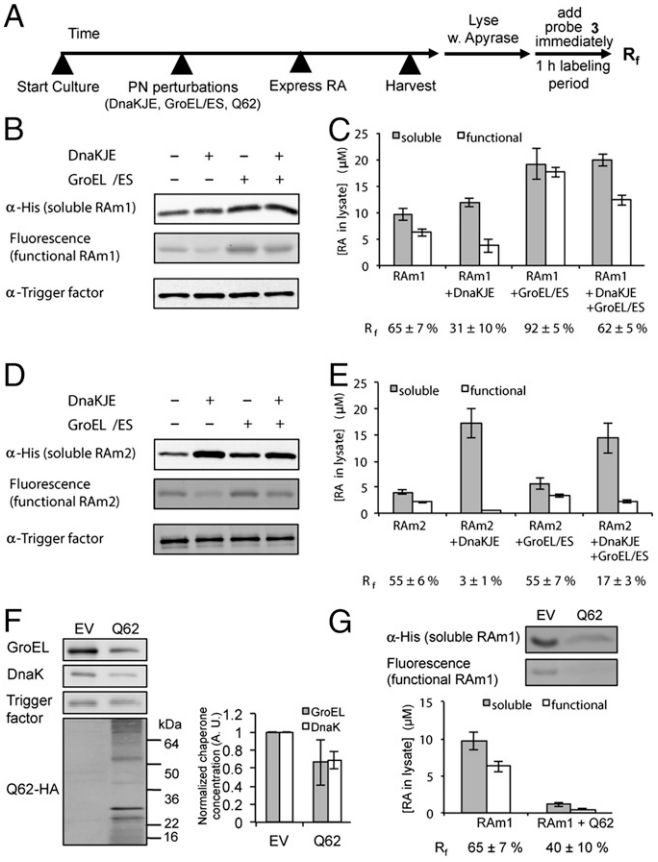

Fig. 6. The folded and functional fraction of RAm1 and RAm2 is dependent on proteostasis network composition and capacity. (*A*) The experimental scheme by which RA mutants are expressed in the context of a perturbed proteostasis network [resulting from the expression of DnaKJE and/or GroEL/ES, or Atx3(tr)-Q62]. (*B* and *D*) The R_f of RAm1 and RAm2 depends differentially on the DnaKJE and GroEL/ES pathways. The total soluble proteins were analyzed by Western blot, and the functional proteins were quantified by fluorescence signal from probe **3** labeling in lysate. Trigger factor serves as the loading control. (*C* and *E*) Quantification of soluble (gray) and folded and functional (white) concentrations of RA mutants as a function of proteostasis network perturbation. Data are represented as mean ± SD. (*F*) Atx3(tr)-Q62 forms soluble oligomers when overexpressed in *E. coli*. As a result, the soluble DnaK and GroEL levels were reduced as shown by Western blot. (*G*) The concentrations of total soluble and folded and functional RAm1 dramatically decrease when cooverexpressed with Atx3(tr)-Q62, resulting in a reduced R_f value. Data are represented as mean ± SD. EV, empty vector.

cellular concentrations of both DnaK and GroEL decreased by ∼40% (Fig. 6*F*), likely because of a transcriptional effect reported previously that reduces proteostasis capacity (31). When RAm1 was overexpressed after overexpression of Atx3(tr)-Q62 (Fig. 6*A*), the concentration of soluble RAm1 decreased by ∼7.5-fold, and the R_f value decreased to ∼40% (Fig. 6*G*), presumably owing to a transcriptional effect and chaperone/chaperonin sequestration.

Discussion

Fluorescent or fluorogenic folding probes that react selectively and chemoselectively with a folded and functional POI in a cell can be prepared. With sufficient cellular holdase activity, we expect generally negligible shifts in folding equilibria, even for rapidly folding proteins and slowly reacting folding probes. Historically, the soluble fraction has been considered to be folded and functional; however, the data present herein reveal that soluble proteins may not always be functional, especially for folding compromised proteins. These folding probes allow us to

study how the cellular proteostasis network, when altered, influences the folding and function of metastable proteins. It is clear that the presence of an aggregation-prone protein can dramatically influence the folding and function of metastable proteins, potentially leading to complex degenerative phenotypes.

The data herein suggest that the folding probe strategy can be generally applied to enzymes and nonenzyme proteins that contain nucleophilic residues critical for function. The availability of numerous enzyme substrates, activity-based probes, and suicide inhibitors serves as starting points for developing folding probes for many POIs (Fig. S8 and *SI Results, Note S4*). Fluorogenic folding probes like **4** are highly desirable, because they do not require a separation step to determine an R_f value for a POI in a cell lysate. This strategy should be applicable to low-abundance proteins, provided that the probe exhibits high selectivity for the POI and that the fluorophore exhibits a high quantum yield. Using probes **3** and **4**, we were able to detect RA at a concentration of 10 nM and TTR at a concentration of 1 nM in *E. coli* lysate (Fig. S3*D*).

In the future, we aspire to use folding probes in living cells, but quantitative corrections for pharmacologic chaperoning in cells will need to be applied (Figs. 4*D* and 5*A*). Folding probes used in living cells have to be cell-permeable, not readily metabolized, etc.

The blueprint that we provide enables the quantification of intracellular protein folding and function in a cellular context. We envision that POI folding probes will be very important for quantitatively discerning how the cellular proteostasis network influences individual proteins in the background of inherited proteins that can misfold or aggregate and lead to pathology.

Materials and Methods

Design and Synthesis of Folding Probes. The design, synthesis, purification, and characterization of folding probes **1**–**3** are described in *SI Materials and Methods*.

Quantification of Functional RA and TTR in Lysates. Cells expressing RA or TTR were lysed in the presence of 50 U/mL apyrase to deplete ATP. Soluble cell lysates were subsequently treated with folding probes for chemical labeling until completion of reaction. The fluorescence signal was analyzed by electrophoresis or a fluorescence spectrometer and was used to determine the concentration of the soluble functional protein using a standard curve generated with purified, fully functional, and labeled RA or TTR. More details are described in *SI Materials and Methods*.

Experimental Scheme to Determine R_f. The concentrations of total soluble protein (functional + nonfunctional fractions) and functional protein were quantified as detailed in *SI Materials and Methods* to derive the ratio (R_f) of the soluble protein that was functional. In brief, soluble cell lysates were split into two aliquots. One aliquot was labeled by folding probes for up to 1 h and analyzed by electrophoresis or a fluorescence spectrometer. The fluorescence signal was used to determine the concentration of the soluble functional protein. The other aliquot was analyzed directly by electrophoresis to determine the concentration of the total soluble protein by comparison with a standard curve. These two concentrations were used to arrive at R_f. More details can be found in *SI Materials and Methods*.

Control Experiments to Evaluate the Pharmacological Chaperoning Effect of Probes 3 and 4. The extent to which folding probes **3** and **4** perturb the RA and TTR folded–misfolded or folded–unfolded state equilibria was evaluated in buffer with or without a specific chaperone holdase activity and in *E. coli* lysate with or without cellular chaperone holdase activity. Experimental measurements were rationalized with kinetic models that describe how folding probes perturb the folding–misfolding–unfolding equilibria in the absence (the pharmacologic chaperoning mechanism) and presence (the holdase trapping mechanism) of chaperone holdase activity. More details can be found in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank M. Saure, G. Dendle, G. J. Kroon, X. Han, and W. Chen for technical assistance. The authors also thank C. S. Parker at the California Institute of Technology for providing the full-length Atx3-Q62 plasmid. Y.L.T. is supported by a predoctoral fellowship from the Agency of Science, Technology and Research (A*STAR). X.Z. is the Howard Hughes Medical Institute Fellow of the Helen Hay Whitney Foundation. G.B. is the Merck Fellow of the Damon Runyon Cancer Research Foundation (DRG-2136-12). D.C.E. is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (DRG-2140-12). This work was supported by National Institutes of Health Grant NS05636 (to J.W.K.), the Skaggs Institute for Chemical Biology, and the Lita Annenberg Hazen Foundation.

1. Hartl FU, Hayer-Hartl M (2002) Molecular chaperones in the cytosol: From nascent chain to folded protein. *Science* 295(5561):1852–1858.
2. Kim YE, Hipp MS, Bracher A, Hayer-Hartl M, Hartl FU (2013) Molecular chaperone functions in protein folding and proteostasis. *Annu Rev Biochem* 82(2013):323–355.
3. Balch WE, Morimoto RI, Dillin A, Kelly JW (2008) Adapting proteostasis for disease intervention. *Science* 319(5865):916–919.
4. Chen B, Retzlaff M, Roos T, Frydman J (2011) Cellular strategies of protein quality control. *Cold Spring Harb Perspect Biol* 3(8):a004374.
5. Ebbinghaus S, Dhar A, McDonald JD, Gruebele M (2010) Protein folding stability and dynamics imaged in a living cell. *Nat Methods* 7(4):319–323.
6. Scheck RA, Schepartz A (2011) Surveying protein structure and function using bis-arsenical small molecules. *Acc Chem Res* 44(9):654–665.
7. Ignatova Z, Gierasch LM (2004) Monitoring protein stability and aggregation in vivo by real-time fluorescent labeling. *Proc Natl Acad Sci USA* 101(2):523–528.
8. Griffin BA, Adams SR, Tsien RY (1998) Specific covalent labeling of recombinant protein molecules inside live cells. *Science* 281(5374):269–272.
9. Myung N, et al. (2013) Bifunctional coumarin derivatives that inhibit transthyretin amyloidogenesis and serve as fluorescent transthyretin folding sensors. *Chem Commun (Camb)* 49(80):9188–9190.
10. Cravatt BF, Wright AT, Kozarich JW (2008) Activity-based protein profiling: From enzyme chemistry to proteomic chemistry. *Annu Rev Biochem* 77:383–414.
11. Jiang L, et al. (2008) De novo computational design of retro-aldol enzymes. *Science* 319(5868):1387–1391.
12. Jing C, Cornish VW (2011) Chemical tags for labeling proteins inside living cells. *Acc Chem Res* 44(9):784–792.
13. Suh EH, et al. (2013) Stilbene vinyl sulfonamides as fluorogenic sensors of and traceless covalent kinetic stabilizers of transthyretin that prevent amyloidogenesis. *J Am Chem Soc* 135(47):17869–17880.
14. Bjelic S, et al. (2014) Exploration of alternate catalytic mechanisms and optimization strategies for retroaldolase design. *J Mol Biol* 426(1):256–271.
15. Nagano N, Orengo CA, Thornton JM (2002) One fold with many functions: The evolutionary relationships between TIM barrel families based on their sequences, structures and functions. *J Mol Biol* 321(5):741–765.
16. Lassila JK, Baker D, Herschlag D (2010) Origins of catalysis by computationally designed retroaldolase enzymes. *Proc Natl Acad Sci USA* 107(11):4937–4942.
17. Heine A, Luz JG, Wong CH, Wilson IA (2004) Analysis of the class I aldolase binding site architecture based on the crystal structure of 2-deoxyribose-5-phosphate aldolase at 0.99A resolution. *J Mol Biol* 343(4):1019–1034.
18. Sekijima Y, et al. (2005) The biological and chemical basis for tissue-selective amyloid disease. *Cell* 121(1):73–85.
19. Banaszynski LA, Chen LC, Maynard-Smith LA, Ooi AG, Wandless TJ (2006) A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell* 126(5):995–1004.
20. Yu ZQ, Sawkar AR, Kelly JW (2007) Pharmacologic chaperoning as a strategy to treat Gaucher disease. *FEBS J* 274(19):4944–4950.
21. Kerner MJ, et al. (2005) Proteome-wide analysis of chaperonin-dependent protein folding in Escherichia coli. *Cell* 122(2):209–220.
22. Teter SA, et al. (1999) Polypeptide flux through bacterial Hsp70: DnaK cooperates with trigger factor in chaperoning nascent chains. *Cell* 97(6):755–765.
23. Schröder H, Langer T, Hartl FU, Bukau B (1993) DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. *EMBO J* 12(11):4137–4144.
24. Rüdiger S, Buchberger A, Bukau B (1997) Interaction of Hsp70 chaperones with substrates. *Nat Struct Biol* 4(5):342–349.
25. Martin J, et al. (1991) Chaperonin-mediated protein folding at the surface of groEL through a 'molten globule'-like intermediate. *Nature* 352(6330):36–42.
26. Chang L, Thompson AD, Ung P, Carlson HA, Gestwicki JE (2010) Mutagenesis reveals the complex relationships between ATPase rate and the chaperone activities of Escherichia coli heat shock protein 70 (Hsp70/DnaK). *J Biol Chem* 285(28):21282–21291.
27. Tiwari S, Kumar V, Jayaraj GG, Maiti S, Mapa K (2013) Unique structural modulation of a non-native substrate by cochaperone DnaJ. *Biochemistry* 52(6):1011–1018.
28. Gidalevitz T, Ben-Zvi A, Ho KH, Brignull HR, Morimoto RI (2006) Progressive disruption of cellular protein folding in models of polyglutamine diseases. *Science* 311(5766):1471–1474.
29. Gupta R, et al. (2011) Firefly luciferase mutants as sensors of proteome stress. *Nat Methods* 8(10):879–884.
30. Bettencourt C, Lima M (2011) Machado-Joseph Disease: From first descriptions to new perspectives. *Orphanet J Rare Dis* 6:35.
31. Schaffar G, et al. (2004) Cellular toxicity of polyglutamine expansion proteins: Mechanism of transcription factor deactivation. *Mol Cell* 15(1):95–105.