# *In vivo* and *in vitro* protein solubility assays using split GFP

Stéphanie Cabantous & Geoffrey S Waldo

[1]Bioscience Division, MS-M888, Los Alamos National Laboratory, P.O. Box 1663, Los Alamos, New Mexico 87545, USA. Correspondence should be addressed to G.S.W. (waldo@lanl.gov).



**The rapid assessment of protein solubility is essential for evaluating expressed proteins and protein variants for use as reagents for downstream studies. Solubility screens based on antibody blots are complex and have limited screening capacity[1–3]. Protein solubility screens using split β-galactosidase *in vivo*[4] and *in vitro*[5] can perturb protein folding. Split GFP used for monitoring protein interactions folds poorly[6], and to overcome this limitation, we recently developed a protein-tagging system based on self-complementing split GFP[7] derived from an exceptionally well folded variant of GFP termed 'superfolder GFP'[8]. Here we present the step-by-step procedure of the solubility assay using split GFP. A 15-amino-acid GFP fragment, GFP 11, is fused to a test protein. The GFP 1–10 detector fragment is expressed separately. These fragments associate spontaneously to form fluorescent GFP. The fragments are soluble, and the GFP 11 tag has minimal effect on protein solubility and folding[7]. We describe high-throughput protein solubility screens amenable both for *in vivo* and *in vitro* formats. The split-GFP system is composed of two vectors used in the same strain: pTET GFP 11 and pET GFP 1–10 (Fig. 1 and Supplementary Note online). The gene encoding the protein of interest is cloned into the pTET GFP 11 vector (resulting in an N-terminal fusion) and transformed into *Escherichia coli* BL21 (DE3) cells containing the pET GFP 1–10 plasmid. We also describe how this system can be used for selecting soluble proteins from a library of variants (Box 1). The large screening power of the *in vivo* assay combined with the high accuracy of the *in vitro* assay point to the efficiency of this two-step split-GFP tool for identifying soluble clones suitable for purification and downstream applications.**

## MATERIALS

### REAGENTS

Plasmid vectors pTET GFP 11 for expressing test proteins tagged with GFP strand 11, pET GFP 1–10 for detecting GFP strand 11 and pET SR–GFP 11 for expressing soluble sulfite reductase (SR) tagged with GFP strand 11 for use as a calibration standard (see **Supplementary Note**)

Restriction enzymes: *Nde*I, *Bam*HI (NEB)

Purification kit for PCR products (for example, QIAquick; Qiagen)

Purification kit for plasmid DNA (for example, QIAPrep; Qiagen)

Calf intestinal alkaline phosphatase (CIP; NEB)

T4 DNA ligase, 400 U/μl (NEB)

5× T4 DNA ligase buffer (Invitrogen; this ligase buffer contains PEG 4000, which increases ligation efficiency)

SOC recovery medium[9]

*E. coli* BL21 (DE3) strain made chemically competent (for example, BL21-Gold; Stratagene)

Electrocompetent DH10B cells (Invitrogen)

Luria-Bertani (LB) media and LB-agar plates used with appropriate antibiotics for selection (35 μg/ml kanamycin for the pET GFP 1–10 plasmid or pET SR–GFP 11, 75 μg/ml spectinomycin for the pTET GFP 11 plasmid)

Anhydrotetracycline (Antet; ACROS Organics or Fisher)

Isopropyl-β-D-thiogalactopyranoside (IPTG)

Nitrocellulose membranes (for example, 130 mm diameter supported nitrocellulose transfer membrane; GE Osmonics)

6His affinity resin (Talon; Clontech)

TNG buffer (100 mM Tris-HCl (pH 7.4), 150 mM NaCl, 10% (vol/vol) glycerol)

Elution buffer: 250 mM imidazole in TNG buffer

9 M urea prepared in water

1 M stock solution of dithiothreitol (DTT)

Protein extraction reagent (for example, Bugbuster; Novagen)

Bovine serum albumin (BSA)

### EQUIPMENT

Thermal cycler programmed with the desired amplification protocol

Microcentrifuge

Preparative centrifuge (for example, Beckman J2-21)

Probe sonicator (such as, Branson 450 with 1/8-inch probe tip)

Illumatool lighting system (LT-9500; Lightools Research) with 488 nm excitation filter (blue) and a colored glass 520 nm long pass filter

Microplate fluorescence reader (for example, FL600 BioTek)

Tube roller (for example, LabQuake, Barnstead)

PUBLISHED ONLINE 21 SEPTEMBER 2006; DOI:10.1038/NMETH932

## PROCEDURE

**Preparation of vectors encoding GFP 11 fusion proteins**

**1**| To prepare cloning vectors, transform the pTET GFP 11 plasmid into an *E. coli* strain (such as BL21-Gold or DH10B); and plate onto selective LB-agar medium containing 75 μg/ml of spectinomycin. Pick a single clone and prepare plasmid DNA using standard protocols. Concentration of purified plasmids should be ~0.1 μg/μl.

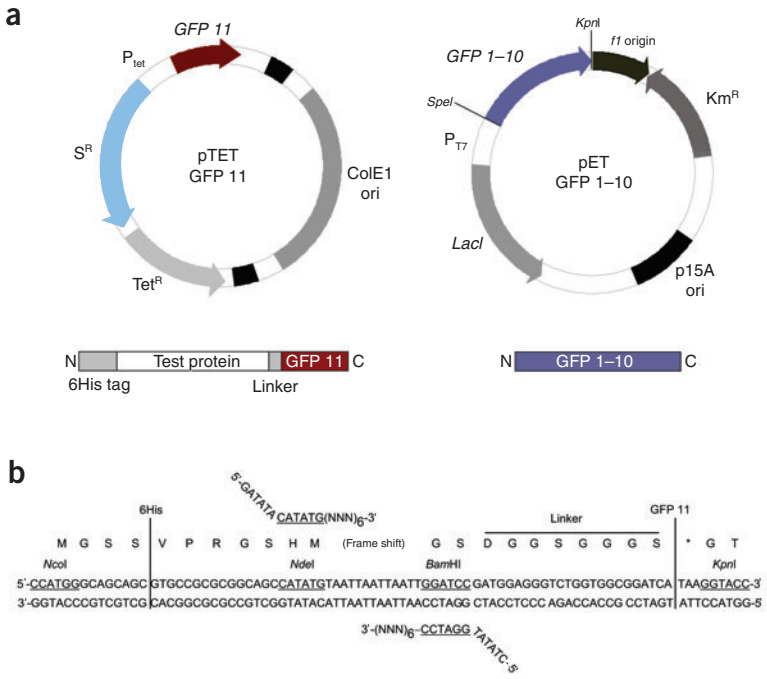


**Figure 1** | Split-GFP vectors. (**a**) The protein of interest is expressed, as N-terminal fusion with the GFP 11 tag, under the control of the tet promoter ($P_{tet}$) from a pTET plasmid, which contains the spectinomycin resistance marker ($S^R$), a ColE1 origin of replication and the gene encoding tetracycline repressor ($Tet^R$). GFP 1–10 is expressed under the control of the T7 promoter ($P_{T7}$) from a pET plasmid that contains a p15 origin of replication compatible for coexpression with the GFP 11 pTET plasmid, a kanamycin selection marker ($Km^R$) and the Lac repressor gene (*LacI*). (**b**) Cloning sites in the GFP 11 vector cassette. The insert protein is cloned using *Nde*I and *Bam*HI restrictions sites. A frame-shift stuffer with three translational stops (one in each frame) prevents false positives from a self-religated plasmid. A 6His tag followed by a thrombin cleavage site is located at the N terminus. An 8-amino-acid linker provides a spacer between the test protein and the C-terminal GFP 11 tag. Unique restriction sites in the plasmids are indicated above the 5′ sequence of the cassette.

**2**| Prepare inserts by PCR amplifying the gene encoding the protein of interest using primers that introduce *Nde*I and *Bam*HI restriction enzyme sites for subsequent cloning into the pTET GFP 11 vector (**Fig. 1**).

**3**| Digest the insert and the pTET GFP 11 vector (**Fig. 1**) with *Nde*I and *Bam*HI. Set up the following reactions for double digestion. Incubate the reactions at 37 °C for 2 h.

| **Reaction 1 (insert)** | | **Reaction 2 (vector)** | |
|---|---|---|---|
| PCR product | 50 μl | pTET GFP 11 plasmid | 50 μl |
| NEB buffer *Bam*HI | 7.2 μl | NEB buffer *Bam*HI | 7.2 μl |
| BSA 20× | 3.6 μl | BSA 20× | 3.6 μl |
| *Nde*I (20 U/μl) | 3.6 μl | *Nde*I (20 U/μl) | 3.6 μl |
| *Bam*HI (20 U/μl) | 3.6 μl | *Bam*HI (20 U/μl) | 3.6 μl |
| | | CIP (10 U/μl) | 0.2 μl |

*This step produces dephosphorylated vector, to reduce background of self-ligated or religated plasmid.*

**4**| Purify the insert by preparative electrophoresis. Clean up the reactions using a DNA gel purification kit (for example, QIAquick Gel Extraction Kit). Elute each reaction with 50 μl of 10 mM Tris-HCl (pH 8.3).

**5**| Ligate the insert into the pTET GFP 11 vector. Set up the following reaction and incubate at 30 °C for 1 h, or 16 °C for 12 h.

| **Reaction 3** | |
|---|---|
| Doubly digested insert DNA (Step 3) | 3.8 µl |
| Doubly digested, dephosphorylated vector DNA (Step 3) | 1.0 µl |
| 5× T4 DNA ligase buffer | 1.0 µl |
| T4 DNA ligase (400 U/µl) | 0.4 µl |

**Transformation into GFP 1–10 *E. coli* strain**

**6**| To prepare chemically competent cells containing the pET GFP 1–10 vector, transform the pET GFP 1–10 plasmid diluted 1:500 in 10 mM Tris-HCl (pH 8.3) into *E. coli* BL21 (DE3) to get single colonies on selective medium containing 35 µg/ml kanamycin. Grow a single clone and prepare chemically competent cells using a standard protocol.

**7**| Transform the ligated products by mixing 2 µl of ligation reaction (Step 5) with 40 µl of chemically competent BL21 (DE3) pET GFP 1–10 cells on ice. Heat-shock the cells for 1 min at 37 °C. Recover each reaction by immediately adding 1 ml of pre-warmed (37 °C) SOC medium and incubating for 1 h at 37 °C with shaking at ~250 r.p.m.
*Alternatively, a PCR machine may be used to perform chemical transformation following this program: cycle 1, –1 °C for 30 min; cycle 2, 37 °C for 1 min; cycle 3, –1 °C soak.*

**In vivo solubility screen**

**8**| Plate 300 µl of the 1 ml transformation reaction onto a nitrocellulose membrane on selective LB-agar containing 35 µg/ml kanamycin, 75 µg/ml spectinomycin. Dilute 200 µl of the remaining undiluted recovery reaction with 800 µl of SOC and plate onto a second nitrocellulose membrane on an LB-agar selective plate. Incubate overnight at 32 °C.
*Plating two dilutions helps ensure that at least one of the plates will have well-separated colonies.*
➨ **TROUBLESHOOTING**

**9**| For the sequential induction experiment (**Fig. 2**), prepare three selective plates each containing 35 µg/ml kanamycin, 75 µg/ml spectinomycin and the following induction reagents:

| | | |
|---|---|---|
| Plate I | 0.3 µg/ml Antet | For induction of GFP 11–tagged protein |
| Plate II | | 'Resting plate' |
| Plate III | 1 mM IPTG | For induction of GFP 1–10 detector fragment |

▲ **CRITICAL STEP**

**10**| Move the membrane (colony side up) from the overnight plate (Step 8) to a prewarmed Plate I. Incubate at 37 °C for 2 h. At this stage, only the GFP 11–tagged protein is induced, and clones are not fluorescent.
▲ **CRITICAL STEP**

**11**| Move the membrane (colony side up) from Plate I to Plate II (you may use the overnight plate as the resting plate). Incubate at 37 °C for 1 h to remove Antet from the colonies by diffusion to stop expression of the GFP 11–tagged protein.
➨ **TROUBLESHOOTING**

**12**| Move the membrane (colony side up) from Plate II (resting plate) to Plate III (IPTG plate). Incubate at 37 °C for at least 1 h to induce expression of GFP 1–10 from the T7 promoter. Clones expressing soluble GFP 11 fusion protein become fluorescent upon complementation with GFP 1–10 (**Fig. 3a** and **Box 1**). Illuminate the plate with the Illumatool to visualize the fluorescent clones.
➨ **TROUBLESHOOTING**

**13**| Pick consensus clones from the IPTG plate. Propagate in 3 ml of LB medium containing 35 µg/ml kanamycin and 75 µg/ml spectinomycin. Make a freezer stock of the cells by 1:10 dilution of the overnight growth in LB medium containing 20% glycerol, and freeze it at –80 °C.
*Larger numbers of clones can be grown in 96-well tissue culture plates (175 µl of selective medium) then copied to freezer stocks using a multichannel pipettor or replicator tool.*
■ **PAUSE POINT** The clones may be frozen indefinitely at –80 °C.

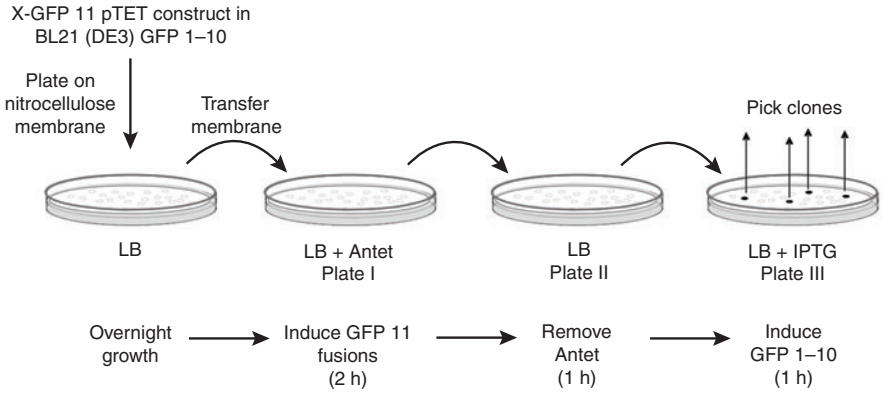


**Figure 2** | *In vivo* solubility screening using sequential induction. After sequential induction, soluble GFP 11 fusions spontaneously bind GFP 1–10, and the resulting fluorescence is proportional to the amount of soluble, nonaggregated GFP 11–tagged protein. Desired clones are then picked for propagation.



**Preparation of GFP 1–10 fragment for *in vitro* assays**

**14**| To purify GFP 1–10, start a 3-ml overnight culture of a GFP 1–10 competent strain. Inoculate 3 ml of the overnight culture into 500 ml of LB containing 35 μg/ml kanamycin. Induce the GFP 1–10 expression in the exponential phase (2 h after inoculation) with 1 mM IPTG for 5 h at 37 °C (this forces the GFP 1–10 to aggregate into inclusion bodies). Collect the cells by centrifugation for 15 min at 3,500*g* (5,000 r.p.m. in large flat-bottom Nalgene tubes in a J-6M Beckman centrifuge (rotor TY. JS-4.2) or at 3,500 r.p.m. in 250 ml Corning conical tubes in Allegra 6KR Beckman centrifuge (rotor GH 3.8)). Discard the medium and resuspend cell pellets in 15 ml of 100 mM Tris-HCl (pH 7.4), 150 mM NaCl, 10% glycerol vol/vol (TNG buffer). Sonicate three times for 2 min on ice at 50% duty cycle. Pellet cell debris containing crude inclusion bodies by centrifugation for 30 min at 30,000*g* (15,262 r.p.m. using Beckman JA-17 rotor) and discard the supernatant. *The engineered GFP 1–10 is about 50% soluble, and the inclusion body fraction is processed to take advantage of the enrichment and partial purification afforded by using inclusion bodies.*

**15**| Add 5 ml of Bugbuster (or another protein extraction reagent) and sonicate to resuspend the crude inclusion bodies. Centrifuge for 10 min at 30,000*g* (15,262 r.p.m. using Beckman JA-17 rotor) and discard the supernatant. Repeat this wash process three times with Bugbuster solution, then twice with 5 ml of TNG buffer (this step removes residual detergent from the pellet mass). Weigh the pellet and resuspend in TNG buffer to obtain a concentration of ~75 mg/ml of washed and packed pellet. Resuspend the inclusion bodies by briefly sonicating, transfer 1-ml aliquots to 1.7 ml microcentrifuge tubes. Centrifuge 10 min at 16,000*g* (13,000 r.p.m. in an Eppendorf centrifuge 5414D with F45-24-11 rotor), and remove and discard the supernatant. Store the inclusion bodies in the Eppendorf tubes at –80 °C.
■ **PAUSE POINT** These pellets may be stored at –80 °C for up to 6 months.

**16**| Dissolve one GFP 1–10 inclusion body pellet in 1 ml of 9 M urea containing 5 mM DTT. Break up the pellet with a pipette tip. Incubate the tube in a 37 °C water bath to help dissolution of the inclusion bodies, or use a roller for mixing. After complete dissolution, centrifuge the microcentrifuge tube for 1 min at 16,000*g* to remove any aggregated material. To a 50 ml Falcon tube, add 25 ml of TNG buffer and the 1 ml of the soluble urea suspension. Mix gently by inversion. Filter the soluble solution through a 0.2 μm syringe filter. This solution is ready to use for *in vitro* protein quantification assays.
■ **PAUSE POINT** The remaining solution may be stored up to two weeks at –20 °C.

**Preparation of standard protein for assay calibration**

**17**| Transform the GFP 11 plasmid encoding the soluble protein SR cloned as an N-terminal fusion with GFP 11 in an N-terminal 6His pET vector[7] into an BL21 (DE3) *E. coli* strain and plate onto selective medium (35 μg/ml kanamycin). Pick a single clone and start a 3-ml overnight culture in selective LB medium.

**18**| Inoculate 500 ml of LB with 3 ml of the overnight culture. When cells reach an $OD_{600}$ of ~0.5, induce with 1 mM IPTG for 4 h at 37 °C. Pellet by centrifugation and resuspend in 2 ml of TNG buffer. Sonicate as indicated in Step 14.

**19**| Remove ethanol storage buffer from 20 ml of a 50% (vol/vol) slurry of Talon 6His affinity resin beads by washing three times with 40 ml of TNG buffer by centrifugation for 5 min at 500*g* (for example, 1,500 r.p.m. in Allegra 6KR Beckman centrifuge, rotor GH 3.8) in a 50 ml Falcon tube. Pour 10 ml of a 50% (vol/vol) slurry of beads in TNG buffer into a fresh 50 ml Falcon tube. Centrifuge and load the soluble extract onto the beads. Incubate on a roller for 30 min at 4 °C. Remove the unbound fraction by centrifugation and decanting. After two washes with two volumes of TNG loading buffer supplemented with 10 mM imidazole, elute the 6His-tagged SR–GFP 11 protein with 5 ml of 150 mM imidazole in TNG buffer. To remove imidazole, dialyze for 1 h at 4 °C against 500 ml of TNG buffer, then replace the buffer and dialyze overnight at 4 °C.
*The protocol gives around 25 mg/l protein of ~95% purity as determined by SDS-PAGE.*

**20**| Quantify the standard protein (for example, using the BioRad Protein assay). Dilute the solution to 1.625 mg/ml in TNG buffer, aliquot into 1.7-ml microcentrifuge tubes, and freeze at −20 °C. The protein is ready to set up calibration curves for *in vitro* protein quantification assays (Step 29).
■ **PAUSE POINT** The TNG buffer–diluted standard can be stored up to 1 year at −20 °C.

**Preparation of test protein**

**21**| Propagate an overnight growth of GFP 11 fusion–GFP 1–10 strain frozen stocks (Step 13) in LB medium containing 35 μg/ml kanamycin and 75 μg/ml spectinomycin. You may use 3 ml of medium in a 10 ml glass culture tube (Steps 22–27) or a 96-well tissue culture format (**Box 1**).

**22**| Inoculate 3.5 ml of selective LB with 35 μl of the overnight starter culture and shake at 350 r.p.m., 37 °C until $OD_{600}$ reaches 0.5 (~1.5 h).
▲ **CRITICAL STEP**

**23**| Induce cells in exponential phase by adding 3.5 μl of Antet 0.3 mg/ml (0.3 μg/ml Antet final concentration). Perform induction of GFP 11 fusions for 2 h at 37 °C or 3 h at 27 °C. At this stage, cell pellets are not fluorescent.
➨ **TROUBLESHOOTING**
▲ **CRITICAL STEP**

**24**| Collect the cells expressing GFP 11 fusions into 1.7 ml microcentrifuge tubes by centrifugation for 1 min at 16,000*g*. Remove the supernatant and resuspend the pellet in 150 μl of TNG buffer.

**25**| Lyse the cells using sonication or chemical lysis. Protein extraction reagents such as Bugbuster can be used following the manufacturer instructions. For sonication, keep the samples on ice and sonicate with three sequences of 10 pulses at 50% duty cycle using a sonicator equipped with 1/8-inch tip. Centrifuge the 1.7 ml microcentrifuge tubes for 1 min at 16,000*g* between sonication sequences. After the third pulse sequence, centrifuge 15 min at 16,000*g* to separate the soluble lysate from cell pellets. Transfer the soluble fraction to a fresh 1.7 ml microcentrifuge tube. Save this sample as 'soluble sample' for use in the *in vitro* assay.
▲ **CRITICAL STEP**

**26**| Wash the pellet by adding 150 μl of TNG buffer and mixing. Centrifuge for 5 min at 16,000*g*. Remove supernatant and resuspend the insoluble fraction in 110 μl of TNG buffer. Transfer 50 μl of the resuspended insoluble fraction to a fresh microcentrifuge tube ('pellet assay sample') and centrifuge for 5 min at 16,000*g*. Carefully pipette and discard the supernatant. Centrifuge again for 1 min and remove any remaining liquid by pipette.

**27**| After the pellet assay sample is totally dry, add 50 μl of 9 M urea. Incubate at 37 °C to help dissolve the pellet, and mix by pipetting. Centrifuge at room temperature (critical to not obtain urea crystals) for 1 min at 16,000*g* to bring liquid at the bottom of the tube. Save this sample ('unfolded pellet') for the *in vitro* assay.
▲ **CRITICAL STEP**

## BOX 1 *IN VIVO* SELECTION OF SOLUBLE PROTEIN VARIANTS FROM A LIBRARY

The split GFP can be used to identify soluble proteins from a library of protein variants (for example, random mutants or fragments of a given protein). The basic principles for *in vivo* and *in vitro* selections are identical whether a single construct or library is assayed. The following procedure describes the generation of an ~$5 \times 10^6$ member library expressing random fragments of a multidomain enzyme in the pTET GFP 11 vector and subsequent screening for solubility using *in vivo* and *in vitro* assays:

**(a)** Create a library of desired variants of the target-protein gene using DNA mutagenesis, DNA terminal deletion protocol or DNA fragmentation technique.

**(b)** Ligate inserts (library) into the pTET GFP 11 vector by following the insert/vector ratio indicated in Step 5 and scaling up the reaction volumes tenfold. Incubate reaction at 16 °C for 12 h.

**(c)** Concentrate the large ligation reaction by ethanol precipitation and resuspend in 10 µl of 10 mM Tris-HCl (pH 8.3). Combine 100 µl of electrocompetent *E. coli* DH10B and 50 µl prechilled sterile $H_2O$. Mix with DNA and incubate on ice for 5 min. Separate into three 75 µl aliquots and transform each by electroporation. Recover the cells in 1 ml SOC for 1.5 h at 37 °C with shaking at ~250 r.p.m. Pool the recovery reactions and plate on selective LB-agar medium (75 µg/ml spectinomycin). Incubate overnight at 32 °C.

**(d)** Wash the cells off the plate using LB, and prepare plasmid DNA from an aliquot corresponding to a 3-ml culture cell mass pellet using a standard plasmid DNA miniprep procedure. Perform three transformations in parallel using 4 µl of the plasmid with 40 µl of chemically competent *E. coli* BL21 (DE3) GFP 1–10 cells. Pool the 1 ml SOC recovery cultures and plate them on selective medium (35 µg/ml kanamycin, 75 µg/ml spectinomycin). Incubate overnight at 32 °C.

**(e)** The next day wash the cells off the plate with 12 ml of SOC. Mix thoroughly and measure the $OD_{600}$. Dilute cells to $OD_{600}$ of 1.0 in LB, 20% glycerol in 160-µl aliquots in PCR tubes, and freeze at –80 °C.

**(f)** Dilute cells from the thawed 1.0 $OD_{600}$ freezer stock using two 400-fold serial dilutions in 1 ml LB. Plate 1 ml on 130-mm diameter nitrocellulose membranes on selective LB-agar Bauer plates containing 35 µg/ml kanamycin and 75 µg/ml spectinomycin. This procedure yields about 5,000 clones per plate. Use as many plates as needed to sample the expected diversity (we usually screen four plates per library in a random mutagenesis directed evolution experiment of an 800-bp gene). Incubate overnight at 32 °C.

**(g)** Perform an *in vivo* solubility screen using a sequential induction protocol as described in Steps 9–12.

**(h)** After sequential induction, the plate should show a diverse pool of colonies of differing brightness (**Fig. 3a**). The fluorescence is proportional to the amount of soluble protein expressed for each clone. Pick the desired clones from the IPTG plate, and grow them individually in 175 µl of selective medium (35 µg/ml kanamycin, 75 µg/ml spectinomycin; 96-well tissue culture plate) overnight at 32 °C.

**(i)** Dilute 20 µl of the overnight growth in 180 µl of LB containing 20% glycerol for a freezer stock in a new tissue-culture plate. In parallel, inoculate 1 ml of selective medium with 10 µl of the overnight growth (for example, in narrow wall 2 ml square-bottom deep-well plates; Marsh Bioproduct). Induce expression of the GFP 11 fusions by addition of 100 µl of a 3 µg/ml Antet solution diluted in selective medium to the 1 ml culture using a multichannel pipette. Agitate using a titer plate shaker. Induce the cells as indicated in Step 23.

**(j)** Collect the cells by centrifugation for 30 min at 1,900*g* in a rotor for multiwell plates (for example, 2,900 r.p.m in a Microplus carrier for GH3.8 rotor, Allegra 6KR Beckman). Discard supernatant. The plate may be stored at –20 °C.

**(k)** Resuspend the cell pellets in 110 µl of TNG buffer using a plate shaker or a multichannel pipette. Transfer the total volume (130 µl) into a 96-well PCR plate (for example, Thermowell plate; Fisher Scientific). Cover with a tape pad (Qiagen). Centrifuge for 10 min at 1,900*g*.

**(l)** Separate soluble and insoluble fractions using a plate sonicator (for example, XL-2020; Misonix, Inc.) or a lysis method (for example, BugBuster; Step 25). For sonication, we perform 3 cycles of 3 min 70% duty cycle, each followed by centrifugation for 5 min at 1,900*g* . Perform the last centrifugation step for 30 min at 2,900*g* to pellet all insoluble material. Transfer the soluble fraction to a new PCR plate (soluble plate). We recommend using an automatic pipettor (for example, Robbins Scientific Hydra 96 liquid handling system) for 96-well plate formats to avoid transferring material from insoluble to soluble fraction.

**(m)** Wash and unfold the pellet fraction as indicated in the main protocol (Steps 26–27), centrifuging the plates for 10 min at 1,900*g* and resuspending by sonication.

**(n)** Perform an *in vitro* complementation assay (Step 28) on both soluble and pellet fractions of the clones selected *in vivo* (**Fig. 3b**). Variants may be subcloned if desired (**Fig. 3d**).

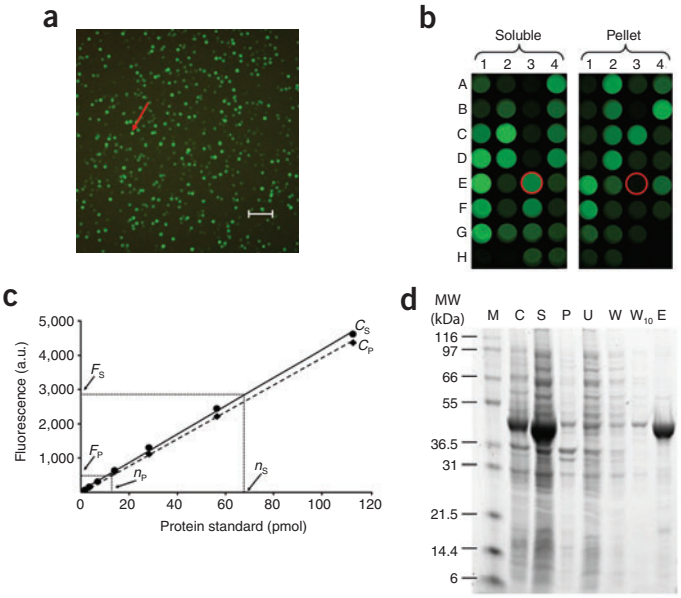


**Figure 3 |** Successive *in vivo* and *in vitro* protein solubility screens. (**a**) Fluorescence image of *E. coli* colonies expressing protein fragments fused to GFP 11 upon complementation with GFP 1–10 after sequential induction. The red arrow indicates the colony selected for *in vitro* assays. Scale bar, 1 cm. (**b**) *In vitro* complementation assay for 32 candidate clones isolated from *in vivo* solubility screen of a library. The image of the microplate shows the fluorescence after complementation of the assayed soluble and pellet fraction *in vitro*. The circled samples (96-well position E3) in soluble and pellet assays plates correspond to a soluble protein fragment chosen for purification. (**c**) Calibration curves for sulfite-reductase GFP 11 fusion standard protein in TNG buffer for quantifying soluble ($C_S$), and urea-denatured pellets ($C_P$). *In vitro* fluorescence of soluble ($F_S$) and pellet ($F_P$) fractions for variant E3 and the corresponding pmol of protein in the soluble fraction ($n_S$) and in the pellet ($n_p$). (**d**) SDS-PAGE image showing expression of variant E3 (see **b**) after subcloning into a pET N-6His expression vector without the GFP 11 tag (crude cell extract, C; soluble fraction, S; pellet fraction, P; unbound, U; wash, W; 10 mM imidazole wash, $W_{10}$; eluted fraction, E. Protein fragment E3 (apparent molecular weight ~42 kDa) could be purified to 95% homogeneity using Talon resin



***In vitro* solubility assay**

**28|** To set up the *in vitro* assay, in parallel add 20 µl of target protein soluble fractions and 10 µl of the unfolded samples to white 96-well assay plates. Use assay plates with low fluorescence background (for example, white Nunc-Immuno plates). Rapidly add the refolded GFP 1–10 fragment (Step 16) to reach a final volume of 200 µl (add 180 µl of GFP 1–10 to the soluble clones, or 190 µl of the GFP 1–10 to the unfolded pellets). Immediately measure fluorescence (Step 30).

*The protocol for obtaining GFP 1–10 (Step 14–16) produces a 3 mM GFP 1–10 detector reagent stock solution. Lower amounts of detection fragment may be used, but to help ensure a linear response to analyte–GFP 11, keep the molar concentration of the GFP 1–10 at least fourfold in excess of the highest GFP 11 concentration expected.*

▲ **CRITICAL STEP**

**29|** Set up a calibration series that will span the dilution range of 230–1.8 pmol of standard protein (SR–GFP 11). Block a 96-well white assay plate using 200 µl of a solution of 0.5% bovine serum albumin (BSA) in TNG buffer. Incubate for 10 min, and then discard the solution. Prepare 8 twofold serial dilutions of the soluble standard protein (Step 20). Transfer 20 µl of each dilution to the blocked assay plate. Rapidly add 180 µl of refolded GFP 1–10 (Step 16).

*To set up the pellet fraction calibration for quantifying urea-denatured pellets, repeat the calibration as above, spiking each calibration well in the assay plate with 10 µl of 9 M urea before adding the GFP 1–10 assay solution. Adding the urea to the pellet fraction calibration compensates for the urea present in the sample pellet in Steps 27-28.*

▲ **CRITICAL STEP**

**30|** Immediately after setting up the sample assay and the soluble and pellet calibrations, measure the fluorescence values ($\lambda_{exc}$= 488 nm /$\lambda_{em}$= 530 nm) with a fluorescence microplate reader to obtain the initial fluorescence. For samples with greater than 2 pmol of tagged protein, the fluorescence can be measured again after 15 min of complementation (**Fig. 3b**), or after 4 h for samples containing as little as 0.2 pmol of tagged protein.

*The response is linear over four orders of magnitude from 0.2 to 200 pmol of tagged protein. The plate can be incubated overnight at 4 °C before measuring the final fluorescence. If desired, initial and final fluorescence values can also be measured for a blank sample (for example, 20 µl of E. coli lysate expressing a protein without the GFP 11 tag, and 180 µl of refolded GFP 1–10). The blank typically yields a fluorescence change less than half that of an authentic 0.2 pmol tagged protein sample, and can often be neglected for typical working samples containing 2–200 pmol tagged protein.*

➨ **TROUBLESHOOTING**

**31**| The GFP fluorescence after complementation is proportional to number of moles of tagged protein. Complementation fluorescence $\Delta F$ is given by equation 1,

$$\Delta F = (Fs_{final} - Fb_{final}) - (Fs_{initial} - Fb_{initial}), \quad (1)$$

where $Fs_{final}$ and $Fs_{initial}$ are the final and initial fluorescence values for a sample, and $Fb_{final}$ and $Fb_{initial}$ are the final and initial fluorescence values for the blank.

$\Delta F$ is converted into moles of protein by the following procedure. First, obtain the slope and intercept of each calibration by a standard linear least squares fitting program (for example, Microsoft Excel) according to the simple linear model equation 2,

$$\Delta F_i = a(n_i) + b, \quad (2)$$

where the dependent variable is $\Delta F$, the independent variable $n$ is the number of moles in the calibration well, $i$ is the sample index from 1 to $N$ where $N$ is the number of calibration points, a is the slope, and b the intercept. It is helpful to plot the calibration data, which should be linear with analyte concentration over the range 0.2–200 pmol for a 200-µl assay volume (**Fig. 3c**).

Calculate the number of moles of tagged protein in each unknown protein test assay ($n_t$) using equation 3,

$$n_t = (\Delta F - b) / a, \quad (3)$$

where $\Delta F$ represents the complementation fluorescence of the test protein sample calculated by Equation 1, and $a$ and $b$ the slope and intercept of the calibration curves coefficients determined by least squares analysis using equation 2.

The number of moles $n_t$ of tagged protein be converted to concentration (in mg/l) expressed protein using equation 4,

$$\text{Concentration} = (1{,}000\ \text{mg/g}) \times (n_t) \times (\text{FW})_t \times (V_s / V_a) \times (1{,}000 / V_c), \quad (4)$$

where $n_t$ is the number of moles of GFP 11 tagged protein calculated by equation 3, $(\text{FW})_t$ is the molecular weight of the test protein, $V_s$ is the sonicant volume in microliters, $V_a$ is the assay volume in microliters, and $V_c$ is the culture volume in milliliters.

Target proteins can be subcloned into a pET N-6His vector for high-level expression and purification without the GFP 11 tag (**Fig. 3d** and **Supplementary Fig. 1** online)

➨ **TROUBLESHOOTING**

**TROUBLESHOOTING TABLE**

| PROBLEM | SOLUTION |
|---|---|
| **Step 8** Yield of transformation of GFP 11 fusions into GFP 1–10 strain is low. | Include a sufficient clamp sequence on the 5′ end of each PCR primer for efficient restriction for cloning. Gel-purify the insert. |
| Yield of colonies too dense. | Perform an additional dilution of the cells before plating. |
| **Step 11** Some fluorescent clones are visible on plates before induction of GFP 1–10. | Do not let your cells rest too long after Antet induction; 1–2 h is sufficient. Leakage expression from GFP 1–10 pET vector may occur, resulting in early complementation with the GFP 11 expressed fusion. Ensure that no lactose is present in the medium. |
| **Step 12** All the clones are fluorescent, even those that are supposed to be insoluble. | Perform an *in vivo* sequential induction experiment if you want to screen for solubility. Fluorescence after simultaneous induction of tagged protein and GFP 1–10 reflects the total expressed tagged protein. |
| **Step 23** After Antet induction of GFP 11 fusions in liquid culture, cell pellets are fluorescent. | Leakage complementation with GFP 1–10 may have occurred. Decrease the duration of Antet induction. Do not over-inoculate, and make sure that no lactose is present in medium. |

**TROUBLESHOOTING TABLE (CONTINUED)**

| | |
|---|---|
| **Step 30** The calibration graph is not linear, falling off at low sample concentration. | The calibration assay plate may have not been blocked sufficiently. When assaying purified samples, blocking prevents sticking of protein molecules to the well surface.<br>Change tips between dilution steps. |
| **Step 31** No fluorescence is visible in the sample and calibration assay plates after overnight complementation. | GFP 1–10 may not be active. This may happen if the solution of refolded GFP 1–10 is stored more than 6 months at −20 °C. Prepare a fresh refolded solution from a new frozen pellet. Assay solutions of pH < 6.5, 0.1% SDS (wt/vol), guanidine HCl > 0.25 M or urea > 0.5 M may also inhibit complementation. |
| There is no fluorescence observed in the pellet assay even though insoluble material is visible after induction. | The pellet may have been unfolded improperly. Repeat the unfolding procedure (Step 27) and ensure pellet is totally dry before dissolving it in urea. |
| The fluorescence signal is very high but appears to be underestimated given the expected quantity of protein in the samples. | Dilute the test sample to ensure at least a fourfold molar excess of GFP 1–10 relative to the sample.<br>To reduce self-absorption, you may also use a different wavelength for exciting the GFP fluorescence (for example 420 instead of 488 nm). Be certain to measure the calibration and sample fluorescence at the same excitation wavelength. |
| Fluorescence is lower than expected in the presence of adjuvants. | Measure SR-GFP 11 calibration with and without adjuvants to test the effect of the adjuvant on the complementation reaction. Perform calibration under precisely the same conditions (chemical composition) as the sample.<br>Dilute your sample further in TNG buffer to reduce interference of adjuvants. |

## CRITICAL STEPS

**Step 9** Antet must be diluted in ethanol. Wait long enough for the LB-agar medium to cool down (<45 °C) before adding Antet. Mix thoroughly after addition of the inducer to obtain a homogeneous solid medium plate. Do not store Antet solution or LB-agar plates containing Antet in direct light. Cover with aluminum foil to avoid degradation of Antet.

**Step 10** Colony size is critical for proper induction on membranes. Do not let clones grow too large after the overnight incubation (between 0.3 and 0.5 mm). We usually perform a 14 h growth at 32 °C. Make sure that you have prewarmed the induction plate (Plate I) for at least 30 min at 37 °C if the plate was stored in the cold, before using it for induction. The membrane bearing the clones should not be allowed to become wet after colonies develop. Ensure that all plates have dried; use a ventilated hood.

**Step 22** Maintain sufficient oxygenation of the culture. Adjust the culture volume to 3 ml for 10-ml culture tubes, or 1 ml for 2-ml 96-well deep-well plates.

**Step 23** If the GFP 1–10 plasmid is also present in the expression strain, the GFP 11 fusion proteins should not be induced longer than 3 h. Otherwise leakage expression of GFP 1–10 can cause premature complementation with tagged GFP 11 proteins.

**Step 25** To avoid underestimating the amount of protein in the pellet fraction; avoid transferring pellet material when transferring soluble fraction.

**Step 27** For correct unfolding and quantification of the pellet fraction, make sure that no liquid remains after the pellet has been dried and centrifuged. Excess buffer can prevent dissolution of pellets in urea. If large pellets are present, use a smaller amount of cell mass for unfolding the pellet to ensure complete dissolution of the inclusion bodies in urea.

**Step 28** When quantifying purified samples, block the plate with a solution of 0.5% BSA in TNG buffer for 10 min to prevent protein adsorption on the well surface. This precaution is less of a concern for assaying crude soluble fractions or inclusion bodies, which contain exogenous cellular proteins. Assay no more than 10 µl of urea dissolved pellet in a 200 µl assay to ensure that the final concentration of urea is less than 0.5 M (urea interferes with the complementation assay)[7].

**Step 29** Before using a new reagent (for example, if testing several refolding conditions) always test the effect of the adjuvant on complementation by performing a calibration with and without the adjuvant. Keeping the starting concentration of the standard protein sufficiently high allows it to be conveniently diluted with adjuvants. It is important to measure test samples and calibrations under the same conditions (adjuvants and chemical composition).

## COMMENTS

Protein solubility assays using self-assembling split-GFP fragments system provide a powerful alternative to conventional approaches based on SDS-PAGE or antibody blots. The split RNase A (S-protein–S-peptide) protein assay (FretWorks, Novagen EMD) works well *in vitro* but not *in vivo*. The split β–galactosidase (*lacZ*) protein assay (QTag, Stratagene) works both *in vivo* and *in vitro*, but the 45-amino-acid β-gal–α tag (QTag) can interfere with passenger protein folding and solubility. The split-GFP can be used for screening a protein construct *in vivo* then *in vitro*, with no need to subclone between vectors. No exogenous reagents are required. The assay is amenable to low-throughput formats on individual genes or high-throughput arrays for screening a large number of genes or libraries of protein variants. The small 15-amino-acid GFP 11 tag is less likely to perturb protein folding and localization compared to more bulky protein labels. Once complemented, the split GFP is irreversible, which means the GFP-labeled protein can be used for a variety of downstream applications using GFP fusions. No external substrate or cofactor is required. The sensitivity (as little as 0.2 pmol in a 200-µl assay), and the linearity of the assay over four orders of magnitude (0.2–200 pmol in a 200-µl assay), provide an accurate quantification in a wide range of concentrations, allowing screening of small volumes of reagents. Visible fluorescence appears *in vivo* within 15 min, and *in vitro* protein quantification time is fast, as little as 15 min for >10 pmol in a 200-µl assay. The pTET plasmid can be subsequently transformed into a BL21 (DE3) strain without GFP 1–10 to entirely eliminate leakage of the GFP 1–10 if this is required for downstream applications. A pET version of the GFP 11 is also available (**Supplementary Fig. 1**). In addition to screening proteins for soluble mutants and domains, the system can be used to screen *in vitro* protein refolding conditions that reduce misfolding and aggregation. The versatility of the split-GFP protein tagging and detection platform should open new research opportunities for finding productive chaperones, protein folding partners, identifying environmental factors and adjuvants modulating protein folding and solubility, and engineering hosts for enhanced protein expression.

*Note: Supplementary information is available on the Nature Methods website.*

**COMPETING INTERESTS STATEMENT**

The authors declare competing financial interests (see the *Nature Methods* website for details).

**Published online at http://www.nature.com/naturemethods/**
**Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/**

1. Knaust, R.K. & Nordlund, P. Screening for soluble expression of recombinant proteins in a 96-well format. *Anal. Biochem.* **297**, 79–85 (2001).
2. Vincentelli, R., Canaan, S., Offant, J., Cambillau, C. & Bignon, C. Automated expression and solubility screening of His-tagged proteins in 96-well format. *Anal. Biochem.* **346**, 77–84 (2005).
3. Nguyen, H., Martinez, B., Oganesyan, N. & Kim, R. An automated small-scale protein expression and purification screening provides beneficial information for protein production. *J. Struct. Funct. Genomics* **5**, 23–27 (2004).
4. Wigley, W.C., Stidham, R.D., Smith, N.M., Hunt, J.F. & Thomas, P.J. Protein solubility and folding monitored *in vivo* by structural complementation of a genetic marker protein. *Nat. Biotechnol.* **19**, 131–136 (2001).
5. Horecka, J. *et al*. Antibody-free method for protein detection on blots using enzyme fragment complementation. *Biotechniques* **40**, 381–383 (2006).
6. Magliery, T.J. *et al*. Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. *J. Am. Chem. Soc.* **127**, 146–157 (2005).
7. Cabantous, S., Terwilliger, T.C. & Waldo, G.S. Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat. Biotechnol.* **23**, 102–107 (2005).
8. Pedelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C. & Waldo, G.S. Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* **24**, 79–88 (2006).
9. Sambrook, J., Fritsch, E.F. & Maniatis, T., eds. In *Molecular Cloning* 2nd edn. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989).