A New Protein-Protein Interaction Sensor Based on Tripartite Split-GFP Association

Stéphanie Cabantous¹, Hau B. Nguyen², Jean-Denis Pedelaq³, Faten Korachi¹, Anu Chaudhary⁴, Kumkum Ganguly², Meghan A. Locket⁵, Gilles Favre¹, Thomas C. Terwilliger² & Geoffrey S. Waldo²

1INSERM UMR1037, Cancer Research Center of Toulouse, Université de Toulouse, Institut Claudius Regaud, F-31052 Toulouse, France, ²Bioscience Division, MS-M888, Los Alamos National Laboratory, Los Alamos, NM 87545, USA, ³CNRS, IPBS (Institut de Pharmacologie et de Biologie Structurale), F-31077 Toulouse, France; Université de Toulouse; UPS, IPBS, F-31077 Toulouse, France, ⁴Department of Microbiology, University of Washington, Seattle, WA 98195, USA, ⁵Rockefeller University, New York, NY 10065, USA.

Monitoring protein-protein interactions in living cells is key to unraveling their roles in numerous cellular processes and various diseases. Previously described split-GFP based sensors suffer from poor folding and/or self-assembly background fluorescence. Here, we have engineered a micro-tagging system to monitor protein-protein interactions in vivo and in vitro. The assay is based on tripartite association between two twenty amino-acids long GFP tags, GFP10 and GFP11, fused to interacting protein partners, and the complementary GFP1-9 detector. When proteins interact, GFP10 and GFP11 self-associate with GFP1-9 to reconstitute a functional GFP. Using coiled-coils and FRB/FKBP12 model systems we characterize the minimal effect on fusion protein behavior and solubility should enable new experiments for monitoring protein-protein association by fluorescence.

Signal transduction and gene expression often involves protein-protein interactions. Optical tracking of these interactions can help to illuminate regulatory mechanisms and identify aberrant processes in diseases. Fluorescent protein biosensors have been developed to measure protein complexes in living cells. For example, bioluminescence resonance energy transfer (FRET and BRET) enable dynamic observations of protein-protein interactions. More recent developments include protein-fragment complementation assays (PCA) that monitor protein-protein interactions by reconstitution of fragments of various enzymes split into two pieces (as fused tags on passenger proteins) including fragments of dihydrofolate reductase, β-galactosidase, β-lactamase, and the firefly and Gaussia luciferases. Though sensitive, these have the limitation that the observed products diffuse away from the protein interaction site. PCA based on split green fluorescent protein (GFP) and color variants is termed bimolecular fluorescence complementation (BiFC). BiFC relies on (i) interactions between bait and prey proteins that bring together two non-fluorescent split protein domains and (ii) subsequent co-folding into the β-barrel structure to form the chromophore. Because assembly of the GFP fragments is irreversible, the stability of the BiFC enables integration, accumulation, and subsequent detection even of transient interactions and low affinity complexes. These remain attached to the interacting proteins, enabling tracking of the complex. However, improving existing BiFCs based on large and bulky fragments is challenging as increasing BiFC fragment solubility and folding can increase background signals from spontaneous assembly of the fluorescent protein fragments. For example, BiFC fragments obtained by fragmentation of folding-reporter GFP (FR-GFP) and superfolder GFP (sfGFP) at permissive sites 156 and 173 were aggregation-prone and had high backgrounds from self-assembly (Supplementary Fig. S1). Using a smaller tag can reduce aggregation and folding interference. For example, one split GFP uses a very small 15 amino acid tag for quantification of soluble protein and tracking proteins in vivo. However the large size of the remaining piece and self-assembly of the fragments makes this system less suited for protein-protein interaction studies. We had previously developed a spontaneously assembling split GFP based on GFP 10–11 (residues 194–233, the tag) and GFP1-9 (residues 1–193, the detector) (Waldo lab, unpublished results). Here we describe the further engineering of this split GFP to yield an entirely new protein interaction assay based on the association of three fragments of the GFP: two short peptides GFP10 (residues 194–212) and GFP11 (residues 213–233) each tagged to one of the interacting partners, and a third large GFP1-9 (residues 1–193) detector fragment (Fig. 1). We present the characterization of the assay using attractive and repulsive pairs of charged coiled-coils peptides and the rapamycin mediated heterodimerization
Results
Engineering a three-body split-GFP system for improved solubility and complementation. We previously identified self-assembling split-GFP fragments corresponding to the C-terminal β-hairpin GFP10–11 of sfGFP14 (residues 194–238) and the large fragment GFP1–9 (residues 1–193) (Fig. 2). In order to improve GFP 1–9 folding efficiency, we performed two rounds of directed evolution and obtained a new variant of GFP1–9 named GFP1–9 M1 that contained five additional mutations relative to sfGFP1–9 (Fig. 2, Supplementary Note). We converted this bimolecular pair into a three-body split-GFP in a stepwise manner. First, in attempt to destabilize the GFP10–11 β-hairpin self-assembly, we inserted a long flexible linker that included a cloning site between GFP10 and GFP11 and evolved the whole cassette for improved solubility and complementation efficiency with GFP1–9 (Supplementary Note). In the final step, we evolved GFP1–9 M1 to recognize this “sandwich” bait topology. When complemented with the optimized GFP10 and GFP11 tags attached to the N and C termini of the HPS15, the resulting variant named GFP1–9 OPT, exhibited a 40-fold improvement in the rate of formation of fluorescence (Fig. 3a, black line) relative to GFP 1–9 M1 (Fig. 3a, gray line). Differences between the GFP1–9 M1 were less obvious when complemented with the GFP10–11 hairpin (displayed on permissive loop of a cherry fluorescent protein, Waldo lab, manuscript under review) (Fig. 3a, right panel), presumably due to the reduced entropy of the GFP 10–11 hairpin relative to the GFP10-HPS-GFP11 “sandwich” format. GFP1–9 OPT amino-acid sequence revealed the presence of four additional mutations relative to GFP1–9 M1 (Fig. 2). We then
examined whether optimized GFP10 and GFP11 would affect the solubility of fused proteins. We cloned set of eighteen *P. aerophilum* test proteins into a modified pTET vector as GFP10-POI-GFP11 “sandwiches” (POI = protein of interest) to assay solubility levels as previously described. We had previously measured the solubility of these eighteen controls without tags and shown that the GFP11 M3 tag engineered in the earlier study did not perturb passenger solubility. We first assayed the sandwich fusions using GFP1–10 to detect only the GFP11 tag (Fig. 3b). Co-expressing GFP1–10 and the sandwich resulted in bright colonies reflecting total protein expression (Fig. 3b, top). On the other hand, expressing the sandwich first, shutting off expression to allow the proteins to remain soluble or aggregate, then expressing the GFP1–10 detector (sequential induction) gave weak fluorescence for the protein expressed in insoluble form (proteins #6,8,15,16) as expected. This pattern was very similar to that previously seen for the GFP11 M3 tag study indicating that the GFP10-POI-GFP11 sandwich did not strongly perturb the solubility of the test proteins. Note that the version of GFPS11 developed in this manuscript differs slightly from the GFP11 M3 tag described earlier. We first assayed the sandwich fusions using GFP1–10 to detect only the GFP11 tag (Fig. 3b). Co-expressing GFP1–10 and the sandwich resulted in bright colonies reflecting total protein expression (Fig. 3b, top). On the other hand, expressing the sandwich first, shutting off expression to allow the proteins to remain soluble or aggregate, then expressing the GFP1–10 detector (sequential induction) gave weak fluorescence for the protein expressed in insoluble form (proteins #6,8,15,16) as expected. This pattern was very similar to that previously seen for the GFP11 M3 tag study indicating that the GFP10-POI-GFP11 sandwich did not strongly perturb the solubility of the test proteins. Note that the version of GFPS11 developed in this manuscript differs slightly from the GFP11 M3 tag described earlier. We next studied the complementation of the sandwich with GFP1–9 OPT (subsequently referred to as GFP1–9) (Fig. 3b, bottom). Interestingly, whether the sandwich and GFP1–9 were co-expressed or sequentially expressed, only the soluble proteins gave fluorescent colonies (Fig. 3b, bottom). When GFP1–10 is co-expressed, insoluble protein can be labeled (Fig. 3b, top), ostensibly because the available GFP1–10 can capture the GFP11 M3 tag early on, committing the fluorophore to form regardless of the subsequent fate of the fusion. In contrast, even when co-expressed with GFP1–9, one or both tags of insoluble proteins must become inaccessible prior to productive binding with GFP1–9. This suggests the GFP 1–9 could be used to stringently assess protein solubility, without the concern of false-positives caused by capturing transiently soluble proteins as with GFP1–10. Validating a protein-protein interaction sensor using coiled-coil heterodimerization. To test whether GFP1–9 is capable of detecting GFP10 and GFP11 fused to separate interacting protein molecules, we selected two pairs of coiled-coils developed by Hodges and coworkers based on their ability to direct preferential interaction between oppositely charged K1/E1 coiled-coils or repulsion for the negatively charged E1/E1 pair (Supplementary Fig. S2). This model system features rapid association of the K/E heterodimer and a lack of E coil homodimerization. Pair-wise combinations of E1-GFP11 and GFP10-E1, or GFP10-K1 and E1-GFP11 were expressed from an anhydrotetracycline (Antet) inducible bicistronic pTET vector (Supplementary Fig. S3), and transformed into *E. coli* cells expressing GFP1–9 from an isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible pET vector (Fig. 4a). After 3 h *E. coli* GFP1–9 cells co-expressing GFP10-K1 and E1-GFP11 turned brightly fluorescent, thus demonstrating efficient K/E coiled-coil heterodimerization. In contrast, co-expression of GFP10-E1 and E1-GFP11 produced residual background fluorescence, comparable with mock induction after either Antet or IPTG induction.
in *E. coli* (Fig. 4b). These results suggest minimal GFP10 and GFP11 self-association in the absence of interacting fused proteins. Using the K1/E1 model, we also chose the GFP10 construct (residue 194–212) as the optimal ‘break point’ for detecting protein-protein interactions (Supplementary Fig. S4). To assess the sensitivity of the assay, GFP10-K1, GFP10-E1, and E1-GFP11 fusion proteins were individually expressed from pET vectors in *E. coli* and purified using immobilized metal affinity chromatography (Supplementary Fig. S5a). Quantified GFP10-K1 or GFP10-E1 were mixed with E1-GFP11 at varying concentrations, and supplemented with a molar excess of the GFP1–9 fragment (8 μM) to initiate rapid complementation and ensure that kinetic rate is independent of GFP1–9 concentration. Fluorescence of the GFP10-K1/E1-GFP11 interaction assay was linearly related to the concentration of E1-GFP11 substrate, as previously observed with self-assembling split-GFP assays15 (Figure 4c). The assay could detect K1/E1 coil interaction down to 1 nM which corresponds roughly to the dissociation constant of the heterodimer18. As expected, the E1/E1 pair interaction signal remained at a basal level of fluorescence even at very high concentrations (1 mM) (Fig. 4c). These results are completely consistent with a model in which tripartite split-GFP complementation occurs only when GFP10 and GFP11 are brought into close proximity by interacting protein partners, presumably by reduction in entropy. To study how different concentrations of the detecting reagent (GFP1–9) affected the assay, two-fold successive dilutions of GFP1–9 were added to the protein mixtures containing GFP10–K1 and E1–GFP11 at the same final concentration (1 μM each). We observed a slower kinetic complementation rate as the reagent concentration decreased, completely consistent with simple bimolecular reaction kinetics (Fig. 4d). For example, the initial rate for appearance of fluorescence drops by a factor of two each time the concentration of GFP1–9 is halved.

Inducible FRB/FKBP interactions. To further test the specificity and dynamic range of the method, we used the well-studied FKBP12-FRB rapamycin inducible protein interaction19. We fused GFP10 to the rapamycin-binding domain of the mammalian target of rapamycin (mTOR) kinase (FRB) and GFP11 to the FK506-binding protein 12 (FKBP) and tested their association in vitro upon addition of recombinant GFP 1–9 protein (Fig. 5a). GFP10–FRB and FKBP–GFP11 fusion proteins were co-expressed from bicastronic pET GFP10–IRBS–GFP11 vector (Supplementary Fig. S3). In earlier studies of interacting proteins in cell extracts using BiFC, misfolding and aggregation issues interfered with complementation unless the
Figure 5 | Application of the tripartite split-GFP to study the rapamycin inducible FRB/FKBP interaction. (a) GFP10 and GFP11 tags were fused to FRB and FKBP proteins. Rapamycin ligand binding brings both protein fusions into proximity, permitting GFP fluorescence reconstitution upon addition of GFP1–9. (b) Raw fluorescence progress curves for GFP1–9 complementation with soluble extracts of GFP10-FRB and FKBP-GFP11 fusions in presence (+RAP) or absence (−RAP) of rapamycin (starting time marked by addition of rapamycin to initiate complementation) (c) Fluorescence levels of GFP10-FRB and FKBP-GFP11 assayed at various concentrations (320, 160, and 40 nM in each) initiated by addition of rapamycin to 150 nM final concentration (black bar) and no rapamycin (gray bar). (d) Rapamycin dose curve (0.6 to 300 nM) for FRB/FKBP binding in vitro measured as final fluorescence after 1 h.
proteins had been co-expressed or were refolded after denaturing. In contrast, our soluble crude E. coli cell extracts could be mixed with a 4 µM solution of purified GFP1–9 and incubated for several minutes prior to addition of 150 nM rapamycin, after which the rapid increase in fluorescence indicated soluble interacting GFP10-FRB and FKBP-GFP11 protein. A two-fold increase in fluorescence was easily detectable after only 10 minutes incubation (Fig. 5b, inset), with a half-life of 60 minutes. Without added rapamycin fluorescence remained near blank levels (Fig. 5b). To further examine the behavior of the protein interaction reporter, we studied purified 6-His-GFP10-FRB and 6-His-FKBP-GFP11 fusion proteins (Supplementary Fig. S5b). In the presence of 150 nM rapamycin (much greater than the reported Kd of ca. 12 nM), endpoint fluorescence was directly proportional to the amount of added protein complex as expected (Fig. 5c). To further evaluate reaction kinetics and equilibria of rapamycin-FKBP/FRB protein association, we tested the effect of increasing concentrations of rapamycin (0.05 to 300 nM). The rapamycin dose response curve presents a half maximal value of 13.5 nM, in accordance with previously published data (Fig. 5d).

**Monitoring the association of protein complexes in E. coli.** Protein signaling modules often involve more than two interacting partners. To examine the ability of our tagging system to detect multimeric complexes, we selected two polycistrons from the E. coli genome encoding the heterotrimer Tus BCD complex (YheNML) and a prototypical alphanate hydrolase constituted by the dimeric assembly YBGJ/YBGK. We also studied an alternate version of the Tus complex where YheN subunit was omitted, thus preventing stable association of the complex as previously reported. These polycistrons were PCR amplified from genomic DNA including their internal ribosome binding sites. Translation of individual complex subunits was dependent on natural IRBS present in the polycistron, so the first subunit had an N-terminal GFP10 and the last subunit had a C-terminal GFP11 furnished by the pTET SpecR vector (Fig. 6a). To ensure sufficient flexibility between complex subunits, especially for the YheNML whose subunits have nearly buried N-termini in the assembled complex, 30 and 25-mer linkers were inserted between the polycistron and the split-GFP tags (Fig. 6b and Supplementary Fig. S3). These constructs were transformed into E. coli cells expressing either the GFP1–9 fragment for complex detection (requiring both the GFP10 and GFP11 tag), or the GFP1–10 fragment to assess the expression level and solubility by monitoring the GFP11 tag (Fig. 6b and 6c). All the complexes were well expressed and highly soluble as shown in Fig. 6c, column A, B. Cells expressing the trimeric YheNML complex became fluorescent after 4h complementation with GFP1–9, as expected (Fig. 6c, 1D). In contrast, weak fluorescence was seen for colonies expressing only YheM and YheL (missing the YheN subunit) with GFP1–9 (Fig. 6c, 2D). Colonies expressing the YBGJ/YBGK control complex were brightly fluorescent after

![Figure 6](https://www.nature.com/scientificreports/srep02854/6)

**Figure 6** | Monitoring complex formation and stability in E. coli. (a) Analysis of YheNML heterotrimeric complex formation using split-GFP. GFP1–9 complementation provides information on interaction of GFP10 and GFP11 tagged domains, while GFP1–10 binds to GFP11 tagged protein, thus providing expression and solubility levels of the GFP11 tagged protein or complex. (b) E. coli polycistrons YheNML, YheML and YBGJ/K were subcloned between GFP10 and GFP11 strands with extended linkers (25 and 30) into pTET GFP10/11 tet inducible plasmids. (c) Split-GFP assay in E. coli cells expressing either GFP1–9 or GFP1–10. Sequential or co-induction was performed as previously described. Corresponding semi-quantitative measure of colony fluorescence using NIH Image (right).
Figure 7 | Visualization of complex formation in mammalian cells. (a) Leucine zipper and Ku70/80 heterodimerization. Cells transiently expressing GFP1–9 with GCN4 zipper (Z-11) (left panel) or Ku70-GFP11 (Ku70-11) (right panel) display no background fluorescence. Heterodimerization is visualized as fluorescence with interacting leucine zippers GFP10-Z and Z-GFP11 in CHO cells expressing GFP1–9. Complementation of GFP11-Z with GFP1–10 confirms localization of the zipper alone (bottom). Ku70/80 complex formation is visualized in cell nuclei (right panel). Expression of one Ku component tagged with GFP11 is monitored with GFP1–10. FACS analysis of HEK 293_GFP 1–9 cell lines transfected with corresponding constructs (24 h after transfection). Percentage of fluorescent cells (black bars); mean fluorescence intensity of the positive cell population (gray bars). Self-associating GFP10-Z-GFP11 domain is used as positive control of transfection and complementation with GFP1–9 (mean ± SD; N = 3). (b) Rapamycin induced FRB/FKBP interaction in mammalian cells. Stable HEK 293 cells expressing GFP1–9 were co-transfected with GFP10-FRB and FKBP-GFP11 constructs and stimulated with increasing concentrations of rapamycin (RAP) (0, 10, 20, 50 and 100 nM). Bipartite complementation of FKBP-11 and GFP1–10 is shown in the most left image; Green fluorescence at 488 nm excitation (GFP), DAPI nuclear staining (cyan). Scale bars = 10 μm. Bottom panel: FACS quantification of rapamycin induced association with or without addition of competitive inhibitor FK-506 (1 μM). Right graph: addition of increasing concentrations of FK-506 (20 nM rapamycin; 0.01, 0.1 and 1 μM FK-506) (mean ± SD; N = 3).
complementation with GFP1–9 as expected (Fig. 6c, 3D). These results are in good agreement with those obtained in our previous bead-based assays using GFP1–10 complementation17.

**Visualization of protein-protein interactions in mammalian cells.** We adapted the split-GFP fragments for optimized expression in mammalian cells and tested the formation of several eukaryotic complexes by fluorescence microscopy. As proof of principle, we first fused the yeast GCN4 leucine zipper25 to each of the GFP10 and GFP11 tags and co-transfected both constructs along with the GFP1–9 detector fragment into CHO cells to study the formation of GCN4 heterodimer (Fig. 7a, left panel). Nuclear fluorescence corresponding to leucine zipper heterodimerization was observed in cells expressing GFP1–9 after 24 h of transfection (Fig. 7a, lane 2). As expected, co-expression of GFP1–9 and one of the leucine zipper domain C-terminally fused to GFP11 did not produce fluorescence (Fig. 7a, lane 1). In a separate experiment we observed the fluorescence complementation of the GFP11 tagged proteins with GFP1–10, which measures soluble protein levels in cells13. As expected, the leucine zipper domain alone was strongly localized in the nucleus and only weakly in the cytoplasm (Fig. 7a, lane 3). We were also able to detect much larger protein complexes, such as the Ku70–Ku80 complex27 that is required in the non-homologous end-joining (NHEJ) pathway in DNA repair. We successfully detected heterodimerization between GFP10–Ku80 and Ku70–GFP11 in HEK 293 cells expressing stably the GFP1–9 fragment (Fig. 7a, right panel). To shed additional light on possible background from the tripartite assay in living cells, we co-expressed leucine-zipper and Ku proteins that localize in the same subcellular compartment. Quantification of the fluorescence levels by flow activated cell sorting (FACS) indicated basal levels of fluorescence from individual split-GFP controls and between non-interacting proteins tested. Pairwise leucine zipper and Ku subunit heterodimerization led to fluorescence levels comparable to the bipartite split-GFP assay, which titrates soluble protein levels in cells13 (Fig. 7a). In parallel, we co-expressed GFP10–FRB and FKBP-GFP11 protein fusions in HEK 293 GFP1–9 cells, stimulated with increasing concentrations of rapamycin. Fluorescence levels for the unstimulated cells remain basal, whereas bright fluorescent cells could be visualized with as little as 10 nM rapamycin, in the range of the EC50 value measured for the ternary complex in vitro24 (Fig. 7b). Moreover, pre-addition of an excess of competitive inhibitor FK-506 totally prevented FRB/FKBP association by rapamycin in an FK-506 dose dependent manner (Fig. 7b, bottom panel), thus demonstrating that tripartite split-GFP complementation can be used to monitor inhibitors of protein-protein interactions by small molecule compounds prior to association of interacting partners.

**Discussion**

Here we describe the first protein-protein interaction reporter based on tripartite split-GFP association. Instead of the bulky and poorly folded BiFC fragments of GFP12,25, our tags based on small engineered β-strands of GFP (~20 amino acids long) minimize protein interference and aggregation. Interaction assays using the E-coil/K-coil model14 have a sensitivity limit in the picomole range. Chemically induced interactions of the FRB/FKBP complex are detectable using the split-GFP within a few minutes after addition of rapamycin in vitro. Bulky BiFC fragments from various fluorescent proteins are expressed largely in E. coli in inclusion bodies, and the effect of the fragments from other enzyme-based PCA assays on the folding of fused proteins is poorly understood22. Here we show that our tagging system is highly soluble, permitting production of high yield of fusion proteins at 37 °C in E. coli. The assay is therefore not temperature dependent, unlike other BiFC for which co-expression and decreasing growth temperature, or refolding denatured extracts are the only ways to improve assembly of split-GFP fragments22. We have shown that the method is sensitive enough to detect association of proteins expressed from E. coli polycistrionic mRNAs, and fluorescence is correlated with independently measured protein complex formation and stability. As shown for the TusABC (YheNML) studies above, we could monitor the formation of larger trimeric complexes by extending the length of the linkers between the GFP tag and the protein of interest. This suggests that precise geometry of the GFP10 and GFP11 is not critical for interaction with GFP1–9, but rather that reduction of entropy is important for triggering assembly of the GFP10, GFP 11, and GFP1–9.

The newly described tripartite split-GFP assay is a promising tool to study protein-protein interactions in vitro and in living cells. Its chief advantage over the bulky aggregation-prone fragments of existing BiFC is the small sizes of the GFP10 and GFP11 tagging peptides. It is therefore well suited to study interactions of unstable protein complexes that are difficult to detect with larger GFP tags. Our system greatly expands the prospect for protein-interaction screening and the design of new biosensors for protein complex assembly and association23. Although association of the 3-body GFP fragments is currently irreversible, this system can be exploited to turn on the detection of protein complex formation by simple addition of GFP1–9 reagent. The method is particularly well adapted to high-throughput interaction screens of libraries of protein-protein interfaces and domains29, the study of complex formation stability, and the robust detection of soluble proteins and protein complexes by flow cytometry. The technology should prove useful for screening small molecule compounds, for example inhibitors for the interfaces of macromolecular protein complexes. Further work might yield split tripartite GFPs whose assembly can be regulated with light30.

**Methods**

**Cloning**

K1-coil and E1-coil DNA, FRB (NM, 09,0458) and GFP11 (NM, 00,0801) sequences were amplified by PCR using synthetic oligonucleotides (SI note) and cloned into a pTET ColE1 SpecR GFP10/11 via Nde1Kpn1 (GFP10 fusion) or Spe1:BamH1 (GFP11 fusion). Linker extension of pTET GFP10/11 was performed using inverse PCR with synthetic oligonucleotides (Supplementary Fig. S3). GFP1–9 was cloned into a pET28a p15 Kan vector via Nde1:BamH1 sites. Proteins expressed with GFP10 (E1 and K1 coils, FRB) and proteins expressed with GFP10 (E1 coil, P28a p15 Kan) were expressed as C-terminally fused to GFP10 or GFP11 (E1-E1 or C-terminal FRB), subcloned into a pET vector bearing a C-terminal or N-terminal 6-His tag respectively, prior to purification and characterization. For mammalian expression, all the constructs used in the study were derived from pcDNA 3.1 Zeo backbone vector encoding GCN4-hGluc and hLuc2-GCN4 kindly provided by Stephen Michnick. Split-lucerase domains were replaced with GFP10 and GFP11 fragments. Ku80, FRB were cloned into Bspe:Xhol sites of pcDNA, GFP10, Ku70 and FKBP were amplified from cDNA and inserted into NotI:ClaI cloning sites of pcDNA, GFP11 vector.

**Directed evolution of split-GFP fragments for tripartite complementation.** The DNA construct encoding (GFP10–11)D(VGSGGGGGS)-Nde1:GGGSGGSGG:Ku80-HPS:BamHI–12 (GGSGGGGSG)-GFP11 (GFP11), which GFP10 and GFP11 sequences were derived from superfolder GFP (Supplementary Note), was evolved by DNA shuffling. Libraries of GFP10–11 variants expressed from the pTET SpecR vector were screened for improved solubility using a sequential induction protocol21 in E. coli cells containing GFP1–9 M1 on a pET p15 vector. At each round, protein solubility of selected optima was verified by complementation with GFP1–10 in vitro. From the six brightest clones sequenced after three rounds of evolution, one best mutant (#5) was identified, termed GFP11 M4 (Supplementary Note). Upstream GFP10 fragment (GFP10 M1) was further evolved by DNA shuffling and primer doping mutagenesis with a pool of fourteen synthetic oligonucleotide primers. Each primer was centered at one of the fourteen amino acids of the GFP10 M1 domain, containing an NNN coding degeneracy at the central target amino acid and flanking homology to the GFP10 M1 domain. The newly described tripartite split-GFP assay is a promising tool to study protein-protein interactions in vitro and in living cells. Its chief advantage over the bulky aggregation-prone fragments of existing BiFC is the small sizes of the GFP10 and GFP11 tagging peptides. It is therefore well suited to study interactions of unstable protein complexes that are difficult to detect with larger GFP tags. Our system greatly expands the prospect for protein-interaction screening and the design of new biosensors for protein complex assembly and association23. Although association of the 3-body GFP fragments is currently irreversible, this system can be exploited to turn on the detection of protein complex formation by simple addition of GFP1–9 reagent. The method is particularly well adapted to high-throughput interaction screens of libraries of protein-protein interfaces and domains29, the study of complex formation stability, and the robust detection of soluble proteins and protein complexes by flow cytometry. The technology should prove useful for screening small molecule compounds, for example inhibitors for the interfaces of macromolecular protein complexes. Further work might yield split tripartite GFPs whose assembly can be regulated with light30.
GFP1–9 in vitro assays. Production of recombinant GFP1–9 OPT protein fragment was performed according to previous protocol described for split GFP 1–10. GFP 1–9 was expressed from a pET15b vector without a 6His tag and produced after induction 1 nM IPTG for 4 hours at 37 °C. The protein was refolded from washes inclusions bodies (see GFP1–10 protocol3) and solubilized in TNG buffer. To perform kinetic characterization of the newly engineered GFP1–9 OPT and GFP 1–9 M1, 180 μl of equal amounts of GFP1–9 refolded pellet fractions (ca. 0.5 mg/ml) were mixed with 20 μl of a soluble protein control sulfito reductase (SR) in fusion with sandwich GFP10 and GFP11 (GFP10-SR-GFP11) or GFP10–11 peptide (3.5 μM each). For K1/E1 and E1/E2 kinetic saturation studies, 25 μl of E1-GFP11 (6.25 μM) were mixed with 50 μl of GFP10-K1 or GFP10-E1 at various concentrations (800 nM down to 0.4 nM). For ligand induced interactions, rapamycin (LC laboratories) was diluted in DMSO and added in 10 μl aliquots to the final FRB/FKBP protein assay mix 20 μl of GFP10–FRB + 20 μl of FKBP–GFP11 (0.5 mg/ml). Assay was repeated by diluting FRB/FKBP samples in TNG buffer. Complementation was induced by addition of a large excess of GFP1–9 OPT (150 μl, 0.25 mg/ml). Fluorescence kinetic measurements were performed using a BD FACSCaliburTM cytometer. Data analysis was performed using detector array). Flow cytometry measurements were performed using a BD FACSCaliburTM cytometer. Data analysis was performed using detector array). Imaging of HEK cells was performed using a Zeiss (Carl Zeiss) M410 confocal laser scanning inverted microscope (LSM 710 NLO with Quazar spectral detector array).

Mammalian cell interactions assays. CHO cells were grown in Ham’s F-12 medium (Gibco, Invitrogen Co) and 10% (v/v) fetal bovine serum (FBS); HEK 293 cells were grown in Dulbecco’s Modified Eagle Medium (DMEM) and 10% (v/v) FBS. CHO cells were co-transfected with Lipofectamine 2000 (Gibco, Invitrogen Co.) with plasmids encoding for GFP1–9 OPT, GCN4-GFP11, GFP10–GCN4 or GCN4-GFP11 M4+ GFP1–10. Stable GFP1–9 cell lines were produced by lentiviral transduction of HEK 293. HEK 293_GFP1–9 cells were co-transfected with jetPRIME reagent (Polyplus transfection) with 0.5 μg of each plasmid (GFP10 and GFP11 fusions); Twenty-four hours after transfection, cells were washed with PBS and mounted in DPI-containing ProLong Gold antifading reagent (Molecular Probes). Imaging was performed using a LEICA DMI-FL microscope with a 40× and 100× oil immersion objective to visualize the stained cells. Images from CHO cells were acquired with a Photometric CoolSNAP HQ camera and analysed with Metamorph or ImageJ softwares. Imaging of HEK cells was performed using a Zeiss (Carl Zeiss) confocal laser scanning inverted microscope (LSM 710 NLO with Quazar spectral detector array). Flow cytometry measurements were performed using a BD Biosciences FACS CaliburTM cytometer. Data analysis was performed using CellQuest® software (BD Biosciences).

1. Day, R. N. & Davidson, M. W. The fluorescent protein palette: tools for cellular imaging. Chem Soc Rev 38, 2887–2921 (2009).
2. Pfleger, K. D. & Edna, K. A. Illuminating insights into protein–protein interactions using fluorescence resonance energy transfer (FRET). Nat Methods 3, 165–174 (2006).
3. Pelletier, J. N., Arndt, K. M., Pluchthun, A. & Michnick, S. W. An in vivo library-versus-library selection of optimized protein–protein interactions. Nat Biotechnol 17, 683–690 (1999).
4. Ross, Z., Charlton, C. A. & Blau, H. M. Monitoring protein–protein interactions in intact eukaryotic cells by beta-galactosidase complementation. Proc Natl Acad Sci USA 94, 8405–8410 (1997).
5. Galarneau, A., Primeau, M., Trudeau, L. E. & Michnick, S. W. Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein–protein interactions. Nat Biotechnol 20, 619–622 (2002).
6. Shekhawat, S. S. & Ghosh, I. Split protein systems: beyond binary protein–protein interactions. Curr Opin Chem Biol 15, 789–797 (2011).
7. Hu, C. D. & Kerpolla, T. K. Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. Nat Biotechnol 21, 539–545 (2003).
8. Maguire, 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).
9. Morrell, M., Espargaro, A., Aviles, F. X. & Ventura, S. Detection of transient protein–protein interactions by bimolecular fluorescence complementation: the Abh-Sh3 case. Proteomics 7, 1023–1036 (2007).
10. MacDonald, M. L. et al. Identifying off-target effects and hidden phenotypes of drugs in human cells. Nat Chem Biol 2, 329–337 (2006).
11. Kodama, Y. & Hu, C. D. An improved bimolecular fluorescence complementation assay with a high signal-to-noise ratio. Biotechniques 49, 793–805 (2010).
12. Hu, C. D., Chinenov, Y. & Kerpolla, T. K. Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. Mol Cell 9, 789–798 (2002).
13. Waldo, G. S., Standish, B. M., Berendzen, J. & Terwilliger, T. C. Rapid protein–fold assay using green fluorescent protein. Nat Biotechnol 17, 691–695 (1999).
14. Pedelaqc, 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).
15. 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).
16. Cabantous, S. & Waldo, G. S. In vivo and in vitro protein solubility assays using split GFP. Nat Methods 3, 845–854 (2006).
17. Kaddoum, L., Magdevleine, E., Waldo, G. S., Joly, E. & Cabantous, S. One-step split GFP staining for sensitive protein detection and localization in mammalian cells. Biotechniques 49, 727–730, 732, 732 passim (2010).
18. Tripet, B. et al. Engineering a de novo designed coiled-coil heterodimerization domain for the rapid detection, purification and characterization of recombinantly expressed peptides and proteins. Protein Eng Des Sel 10, 299 (1997).
19. Banaszynski, L. A., Liu, C. W. & Wandless, T. J. Characterization of the FKBP, rapamycin. FRB ternary complex. J Am Chem Soc 127, 4715–4721 (2005).
20. Numata, T., Fukui, S., Ikeuchi, Y., Suzuki, T. & Nureki, O. Structural basis for sulfur relay to RNA mediated by heterorexameric TuB3C complex. Structure 14, 357–366 (2006).
21. Lockard, M. A. et al. A high-throughput immobilized bead screen for stable proteins and multi-protein complexes. Protein Eng Des Sel 24, 565–578 (2011).
22. Remy, I. & Michnick, S. W. A highly sensitive protein–protein interaction assay based on Gausia luciferase. Nat Methods 3, 977–979 (2006).
23. Cary, R. B., Chen, F., Shen, Z. & Chen, D. A. Central region of KU80 mediates interaction with KU70 in vivo. Nucleic Acids Res 26, 974–979 (1998).
24. Marz, A. M., Fabian, A. K., Kozany, C., Bracher, A. & Hauß, F. Large FK506-binding proteins shape the pharmacology of rapamycin. Mol Cell Biol 33, 1357–1367 (2013).
25. Ghosh, I. Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein. J. Am. Chem. Soc. 122, 5658–5659 (2000).
26. Ozawa, T. Designing split reporter proteins for analytical tools. Anal Chem Acta 556, 58–68 (2006).
27. Robida, A. M. & Kerpolla, T. K. Bimolecular fluorescence complementation analysis of inducible protein interactions: effects of factors affecting protein folding on fluorescent protein fragment association. J Mol Biol 394, 391–409 (2009).
28. Kellermann, S. J., Rath, A. K. & Rentmeister, A. Tetramolecular fluorescence complementation for detection of specific RNAs in vitro. Chembiochem 14, 200–204 (2013).
29. Pedelaqc, J. D. et al. Experimental mapping of soluble protein domains using a hierarchical approach. Nucleic Acids Res (2011).
30. Do, K. & Boxer, S. G. Thermodynamics, kinetics, and photochemistry of beta-strand association and dissociation in a split-GFP system. J Am Chem Soc 133, 18078–18087 (2011).