

Profluorescent protein fragments for fast bimolecular fluorescence complementation *in vitro*

Vadim V Demidov & Natalia E Broude

Center for Advanced Biotechnology, Boston University, 36 Cummington Street, Boston, Massachusetts 02215, USA. Correspondence should be addressed to V.V.D. (vvd@bu.edu) or N.E.B. (nebroude@bu.edu).

Published online 6 July 2006; doi:10.1038/nprot.2006.114

Here, we present a protocol for isolating the large N-terminal fragment of enhanced green fluorescent protein (EGFP) with a preformed chromophore. By itself, the chromophore-containing EGFP fragment exhibits very weak fluorescence, but it rapidly becomes brightly fluorescent upon complementation with the corresponding small, C-terminal EGFP fragment. Each EGFP fragment is cloned and overexpressed in *E. coli* as a fusion with self-splitting intein. After solubilizing and refolding these fusions from inclusion bodies, both EGFP fragments are cleaved from intein and purified using chitin columns. When these EGFP fragments are linked with the two complementary oligonucleotides and combined in equimolar amounts, fluorescence develops within a few minutes. The isolation of profluorescent protein fragments from recombinant *E. coli* cells requires ~3 d, and their conjugation to oligonucleotides requires 1–4 h.

INTRODUCTION

Bimolecular complementation of fragments of fluorescent proteins is a robust assay for directly monitoring conformational changes or interactions of proteins, and for detecting specific nucleic acid sequences^{1–4}. The approach is based on the reassembly of a fluorescent protein from two nonfluorescent fragments driven by additional biomolecular interactions and resulting in restoration of fluorescence. In all prior studies, fluorescent chromophore formed *de novo* within the reconstituted full-size protein (Fig. 1a). Therefore, it took hours to restore the fluorescence of a split fluorescent protein, as formation of a protein chromophore is generally known to be a slow process^{4–7}.

In the course of our recent study on protein fluorescence complementation triggered by nucleic acid complementary interactions, we unexpectedly observed rapid kinetics of fluorescence development⁸. These results suggested that the large N-terminal fragment of split EGFP used in this study contained a preformed chromophore. Computer simulations verified that the large EGFP fragment may develop a compact structure in which arrangement of the chromophore-forming amino acids is essentially the same as in the full-size EGFP⁸. These data indicated that, under certain conditions, the protein chromophore may spontaneously form within the large N-terminal EGFP fragment, which will instantly become brightly fluorescent upon complementation with the C-terminal EGFP fragment (Fig. 1b).

We describe here the protocol that enabled us to isolate the large EGFP fragment with a preformed chromophore. Using the chromophore-containing large EGFP fragment yields a fluorescent response within 1 min, when reassembly of a split protein is supported by DNA hybridization⁸. The proposed methodology therefore represents a promising alternative for the *in vitro* monitoring of biomolecular interactions in real time, when compared to competing fast-responding techniques: protein-based fluorescent resonance energy transfer (FRET)⁹ and nucleic acid-based molecular beacons^{10,11}. Indeed, the reassembled split EGFP generates a much stronger signal over background than the protein-based FRET. Another potential advantage is the expected reduced sensitivity of a protein-shielded chromophore inside the reassembled split EGFP to environmental changes, such as pH, solvent and salt concentration, as compared to the solvent-exposed molecular beacons.

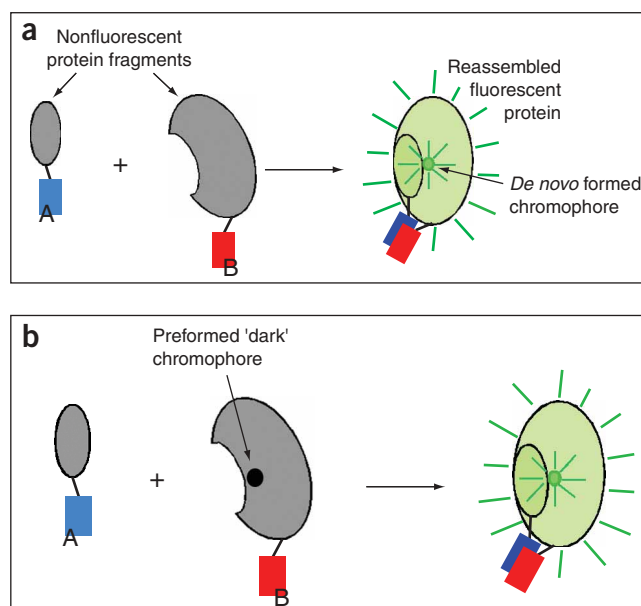


Figure 1 | The principle of protein fluorescence complementation. Fluorescent protein is split into two nonfluorescent protein fragments, which are tagged by interacting molecules A and B (A and B could be proteins or, as in experiments shown in Figure 4, oligonucleotides). Binding of A to B brings the nonfluorescent protein fragments in close proximity that reconstitutes split protein and restores fluorescence. (a) Traditional approach with slow development of fluorescence. Protein fragments do not have a fluorogenic chromophore prior their complementation; it slowly (up to 100 min) forms afterwards. (b) An unconventional approach presented here and yielding a rapid signal response. One of the protein fragments forms a fluorogenic chromophore before reassembly. This profluorescent polypeptide should rapidly develop fluorescence upon complementation with another protein fragment.

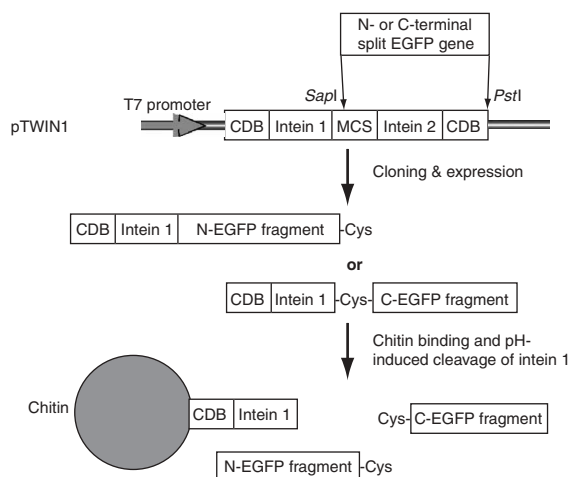


Figure 2 | Experimental design used here for cloning and isolating the two fragments of split EGFP. CBD, chitin-binding domain; MCS, multiple cloning site.

Experimental design

Our experimental design is shown schematically in **Figure 2**. Specifically, the DNA sequences encoding the large (N-terminal) and small (C-terminal) EGFP fragments were cloned in the pTWIN-1 vector to yield the C-terminal fusions with the *Ssp* DNAB intein. The N- and C-terminal EGFP-coding sequences were obtained by PCR amplification of the corresponding fragments of the pEGFP-1 vector containing the gene encoding EGFP. Each PCR product was ligated into the *Sap*I/*Pst*I-digested pTWIN-1 vector using adaptors; this cloning eliminated a second intein. With the chosen PCR primers (see MATERIALS), the large EGFP fragment contained the 158 N-terminal amino acids of EGFP plus a C-terminal cysteine, whereas the smaller fragment contained the remaining 81, C-terminal amino acids of EGFP plus an N-terminal cysteine. Terminal cysteines were added to the protein fragments to facilitate their conjugation to oligonucleotides (see ANTICIPATED RESULTS). The N- and C-terminal locations of added cysteines in the small and large EGFP fragments, respectively, warrant their reassembly similar to native conformation.

Protein fusions were expressed in *E. coli*, followed by their isolation from bacterial cells as inclusion bodies. After solubilizing and refolding these fusions from inclusion bodies, both EGFP fragments were cleaved from intein and purified using chitin columns. Intein strongly binds to chitin beads to anchor the protein fusion to a column, and it remains on the column after slow self-splitting from the fused protein triggered by a change of the buffer pH from 8.5 to 7.0. The flowchart of the entire procedure is shown in **Figure 3**. It should be noted that expression of the N-terminal,

Cloning of EGFP fragment-intein fusions

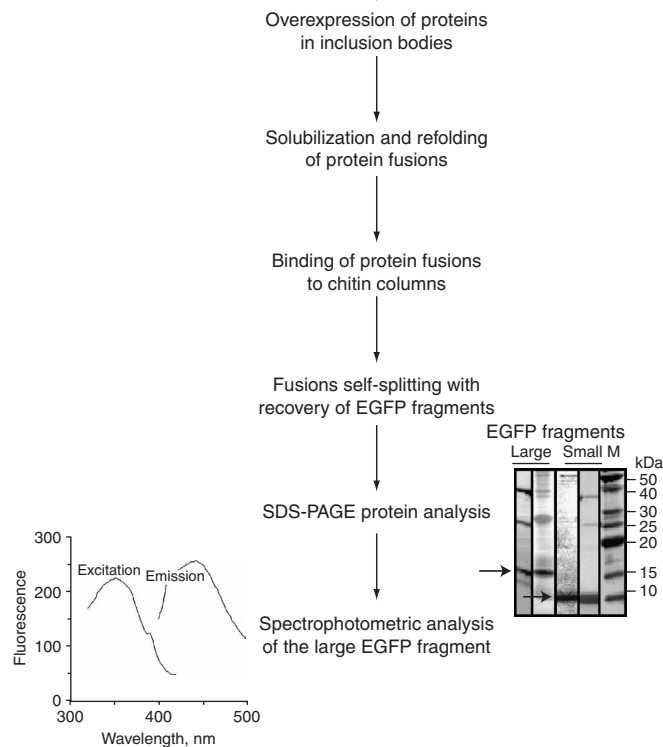


Figure 3 | Flowchart of the major procedure described in this protocol.

large EGFP fragment fused to intein is usually less efficient than expression of the C-terminal, small EGFP fragment-intein fusion. Therefore, larger cell-culture volumes are recommended for growing cells expressing the N-terminal EGFP fragment for obtaining an amount comparable with the amount of C-terminal fragment.

This procedure yields the large EGFP fragment with a preformed chromophore, which is revealed in: (i) a weak but characteristic fluorescence of the isolated N-terminal protein fragment, and (ii) the fast development of strong fluorescence upon addition of the C-terminal protein fragment, provided that both EGFP fragments are equipped with additional interacting molecules A and B (**Fig. 1**). Alternatively, A and B can interact with a third molecule, C, which will juxtapose profluorescent protein fragments (not shown). Here, we used a pair of complementary oligonucleotides for EGFP complementation. EGFP fragments can also be fused to the protein partners to detect their cross-interaction or conformational change induced by the binding of a small ligand.

MATERIALS

REAGENTS

- Primers used to PCR-amplify coding sequences of large and small EGFP fragments: Large EGFP fragment with C-terminal cysteine; Primer ALPHA_dir: 5'-AGTTTCTAGAAATGGTGAGCAAGGGCG-3'; Primer ALPHA-CYS_rev: 5'-ATCGCTCGAGTTAGCACTGCTTGTCCGCCATG-3'; Small EGFP fragment with N-terminal cysteine; Primer BETA-CYS_dir: 5'-ATCGGATATCATGTGCAAGAACGGCATCAAGGTG-3'; Primer BETA_rev: 5'-ATCGCTCGAGTTACTTGTACAGCTCGTCC-3';

- pEGFP-1 or other plasmid containing the EGFP coding sequence (Clontech)
- pTWIN-1 vector (New England Biolabs)
- BL21(DE3) pLys competent *E. coli* cells (Stratagene)
- Luria-Bertani (LB) cell-culture media
- Isopropyl thiogalactoside (IPTG, 1 M stock)
- Cell resuspension buffer: 50 mM Tris-HCl (pH 8.5), 25% sucrose, 1 mM EDTA, 10 mM 1,4-dithiothreitol (DTT)
- Inclusion bodies washing buffer: 50 mM Tris-HCl (pH 8.5), 0.5% Triton X-100, 100 mM NaCl, 1 mM DTT, 1 mM NaEDTA

PROTOCOL

- Protein solubilization buffer: 25 mM MES (pH 8.5), 8 M urea, 10 mM NaEDTA, 0.1 mM DTT
- Protein refolding buffer: 50 mM Tris-HCl (pH 8.5), 500 mM NaCl, 1 mM DTT
- Protein refolding buffer (neutral version): same composition as the previous buffer but with pH 7.0
- Column equilibration-loading buffer: 50 mM Tris-HCl (pH 8.5), 500 mM NaCl, 1 mM EDTA, 1 mM PMSF and 0.1% Triton X100
- Column cleavage buffer: same composition as the previous buffer but with pH 7.0
- Precast 15% polyacrylamide gels for protein electrophoresis (PAGE, Tris-glycine type, Cambrex)
- Protein electrophoresis buffer: 25 mM Tris, 0.2 M glycine, 0.1% SDS (pH 8.3)

- Chitin beads (New England Biolabs, cat. no. S6651)

EQUIPMENT

- Thermal cycler (PTC-100 or PTC-200, MJ Research)
- Controlled-environment incubator shaker (model G25, New Brunswick Scientific)
- Superspeed centrifuge (SORVALL RC 5B, Thermo Electron Corporation)
- Sonifier cell disrupter (model W185c, Branson Sonic Power)
- Poly-prep chromatography columns (Bio-Rad, cat. no. 731-1550)
- Set of Pipetmans (P2 to P1000, Gilson)
- Vertical electrophoresis system (10 × 10 cm, Fisher Biotech)
- Spectrofluorometer (Hitachi F-2500) and, optionally, spectrophotometer (Hitachi U-3010) with quartz cuvette(s)

PROCEDURE

- 1| PCR-amplify DNA sequences coding for the large and small EGFP fragments using a common PCR practice (www.promega.com/guides/pcr_guide/default.htm), and using a plasmid with the EGFP gene (Clontech) as a template and the corresponding primers (see MATERIALS and Experimental design).
 - 2| Separately clone each of the corresponding DNA sequences into pTWIN-1 vector using traditional recombinant techniques¹² to yield the IPTG-inducible fusions of EGFP fragments to mini-intein.
▲ CRITICAL STEP Check the cloned plasmid sequences for correctness of all necessary protein-expression elements (promoter, initiation and stop codons, and protein-coding genes).
 - 3| With these plasmids, transform the protease-deficient BL21 (DE3) pLys competent *E. coli* cells.
 - 4| Grow the transformed *E. coli* cells in a shaking incubator overnight at 37 °C in 2 ml LB medium.
 - 5| By taking the corresponding aliquots, dilute the cell culture 500-fold in 200 ml to 1 liter LB (for submilligram to milligram protein amounts), and allow cells to grow at 37 °C for about 2–3 h up to $A_{600} \approx 0.6$.
 - 6| Induce cells with 0.35 mM IPTG and grow them overnight at 25 °C or for 4 h at 30 °C (in the latter case, somewhat better protein expression is normally observed).
▲ CRITICAL STEP Check the efficiency of IPTG induction by SDS-PAGE directly loading onto a 15% polyacrylamide gel a required amount of SDS-denatured cell suspension before and after induction, and analyzing the intensity of corresponding protein bands (~35 kDa and ~40 kDa for the intein-fused small and large EGFP fragments, respectively).
- ### ? TROUBLESHOOTING
- 7| Precipitate cells by centrifugation at 15,000 r.p.m. for 10 min at 4 °C. Resuspend cells in resuspension buffer (15 ml/1 l cell culture).
 - 8| Freeze 1-ml aliquots at -70 °C for 10 min, then thaw them at 37 °C for 5 min; repeat the freeze-thaw process three times.
 - 9| Keeping cells on ice, and break them up by sonication with three 30-s bursts, each followed by a 30-s interval.
 - 10| Combine the resulting samples and precipitate the inclusion bodies by centrifugation at 15,000 r.p.m. for 5 min at 4 °C.
 - 11| Remove the supernatant, resuspend the precipitate in washing buffer (10 ml/1 l cell culture) by vortexing and sonicate this suspension as before.
▲ CRITICAL STEP Keep this supernatant in a refrigerator until the end of the procedure to use it as an alternative source for isolation of EGFP fragments from a soluble cellular fraction (see the TROUBLESHOOTING TABLE, Step 14).
 - 12| Centrifuge for 5 min at 15,000 r.p.m. and repeat the washing step three times, each time discarding the supernatant.
 - 13| Resuspend pellets in solubilization buffer (5 ml/1 l cell culture) and incubate at 25 °C for 1 h. Clear the solution by centrifugation at 15,000 r.p.m. for 5 min.
 - 14| Use the supernatant containing solubilized denatured protein to refold it by adding this solution drop by drop to an excess of the refolding buffer with 1:100 dilution. For drop-by-drop dispensing (normally requires several minutes), use a 100- μ l Pipetman (P100) with ~1-s intervals between drops.
■ PAUSE POINT Leave the refolding solution overnight at 4 °C.
▲ CRITICAL STEP We believe that intein facilitates the proper folding of EGFP fragments⁸. Still, the success rate of correct refolding, and hence of chromophore formation, is ~50%: we got two fluorescently active samples of the large EGFP fragment out of four independent isolations from inclusion bodies (with a smaller yield, a fluorescently active preparation of split EGFP

was also obtained from a soluble cellular fraction by an alternative protocol without refolding). So, refolding of several samples in parallel is recommended at this step of the protocol. It is also recommended to estimate the concentration of solubilized proteins by SDS-PAGE using protein standards. The protein concentration > 100 mg/l would be optimal for subsequent protein refolding and column purification.

? TROUBLESHOOTING

15| Fill an empty 9-cm-high, conical 0.8 × 4 cm polypropylene chromatography column with 2 ml of the chitin bead suspension (supplied as a slurry in 20% ethanol with a bead content of ~50% (vol/vol)). Equilibrate the column with 10 ml of the column equilibration-loading buffer.

▲ **CRITICAL STEP** For the intein-fused protein purification using the chitin column procedure, see http://t7l.bimcore.emory.edu/manuals/IMPACT_I_Sys_man.pdf#search='chitin%20column%20new%20england%20biolabs' and <http://www.neb.com/nebecomm/ManualFiles/manualE6950.pdf>

16| Load a refolded protein solution onto the column and wash it with 20 ml of the column equilibration-loading buffer.

17| Wash the column with 10 ml of the cleavage buffer and close the bottom tip with the tip closure.

18| Load the column with 10 ml of the cleavage buffer, then close the top with the end cap.

■ **PAUSE POINT** Leave the column at room temperature overnight for intein self-cleavage.

19| Open the column's top end and the bottom tip to elute EGFP fragment split from intein. Collect ~300-µl fractions and analyze them by SDS-PAGE for purity and amounts of the large and small EGFP fragments (~18 kDa and ~10 kDa, respectively; see Fig. 3). Close and save the column for possible additional elution. Store both protein solutions at 4 °C.

? TROUBLESHOOTING

20| Verify the presence of the chromophore within the large EGFP fragment by spectrophotometry. The chromophore-containing protein fragment should show significant absorbance in the range of 300–400 nm. The presence of a chromophore in the large EGFP fragment should be more evident by recording its fluorescence with excitation maximum near 350 nm and emission maximum near 450 nm (see Fig. 3). The fluorescence intensity of this protein fragment should be ~100 times weaker than the peak fluorescence of the intact EGFP⁸.

● **TIMING**

Steps 1–3: 2–3 d

Steps 4–6: 20–30 h

Steps 7 and 8: 1 h

Steps 9–13: 2 h

Step 14: 15 h

Steps 15–18: 16–17 h

Step 19: 3 h

Step 20: 30 min

The protein-oligonucleotide or protein-protein conjugation reactions normally require 1–4 h, depending on the particular chemistry used.

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 1**.

TABLE 1 | Troubleshooting table.

STEP	PROBLEM	SOLUTION
Step 6	Little or no protein expression was detected after induction.	Vary the IPTG concentration and/or cell growth temperature to find the optimal conditions for the particular cell clone. If this does not help, grow a single clone of the transformed <i>E. coli</i> cells first on the agar plate, then repeat Steps 4–6.
Step 14	Concentration of the refolded large EGFP-intein fusion is low (< 1 mg/l ⁻¹).	In this case, subsequent column purification of the large EGFP fragment will be inefficient. As an alternative to the column-splitting procedure, cleave the fusion directly in the refolding buffer with pH 7.0 overnight at 4 °C. In some applications, there is no need for separation of the large EGFP fragment from the intein, as it will not interfere with the downstream procedures. If the intein removal is necessary, the 25-kDa split intein can be separated from solution of the 18-kDa large EGFP fragment by size- selective ultrafiltration using Centricon centrifugal filter microdevices (YM-50, Millipore). Alternatively, fluorescently active EGFP fragments



TABLE 1 | Troubleshooting table (continued).

<p>Step 19 Although the concentration of the large EGFP fragment was high enough, no or very small amounts of this protein were eluted, because most of it adhered to the column.</p>	<p>can be obtained, though with lower yield, from soluble cellular fractions. Using the supernatant from Step 12 instead of a refolded protein solution, go directly to Step 15, thus avoiding protein refolding and extensive dilution. The likelihood of obtaining the large EGFP fragment with preformed chromophore using this alternative protocol may be higher, though at the expense of protein yield.</p> <p>Wash the column with additional 10 ml of the cleavage buffer to remove the protein that remained bound to the chitin beads.</p>
---	---

ANTICIPATED RESULTS

Profluorescent protein fragments with mature chromophore should be useful for detecting many types of fast pairwise interactions. To demonstrate this potential, we obtained the chromophore-containing large and the complementary small EGFP fragments from the inclusion bodies (see Fig. 3), and used them in the protein fluorescence complementation assay triggered by the DNA-DNA duplex formation⁸. The DNA duplexes normally form very quickly^{13,14}. Given that refolding and fluorescence restoration of the chromophore-containing full-size EGFP from totally unfolded state requires just a few minutes⁵, we expected that reassembly of the split EGFP from fragments, one of which already contains the chromophore, should result in the fast development of protein fluorescence, if supported by DNA hybridization.

In our *in vitro* system⁸, large and small EGFP fragments isolated using the above procedure were first biotinylated using sulfhydryl-based chemistry, then coupled with complementary biotinylated oligonucleotides using streptavidin as a linker. The oligonucleotide-conjugated large EGFP fragment was only weakly fluorescent, and the small EGFP fragment-oligonucleotide conjugate did not show any detectable fluorescence. However, when the two EGFP fragments were brought together by nucleic acid complementary interactions, a strong, up to 100-fold increase in fluorescence was detected with excitation-emission spectra resembling EGFP and with a $t_{1/2} \leq 1$ min (Fig. 4)⁸. The fluorescence intensity of the reassembled complexes varied from experiment to experiment, with maximal response close to that of the intact EGFP. In contrast, control experiments (i.e., mixing streptavidin-bound protein fragments without complementary oligonucleotides) did not show any appreciable increase in fluorescence. These proof-of-principle data demonstrate that protein fluorescence complementation with the preformed chromophore occurs quickly. We believe that this approach could be extended to other fluorescent proteins^{4,15,16}, to other conjugation chemistries and to detecting various complexes *in vivo* by injecting the corresponding conjugates into different cellular compartments.

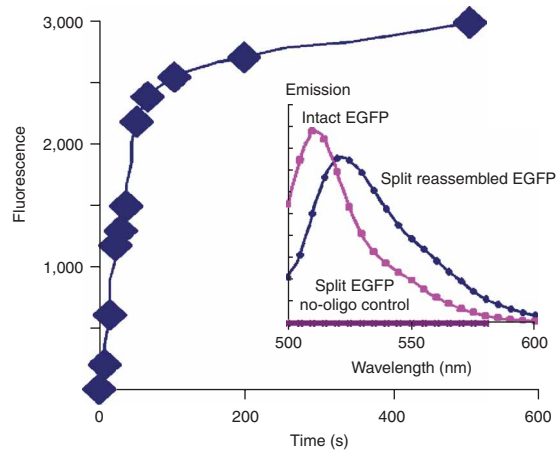


Figure 4 | Fluorescent response kinetics of the split EGFP system upon DNA hybridization recorded at 524 nm. The complementary EGFP fragments were obtained from the inclusion bodies following the described protocol. They were then biotinylated at the N- or C-terminal cysteines, and each protein fragment was conjugated via streptavidin to one of the two biotinylated 21-nt oligonucleotides with complementary sequences at equimolar ratio of all components⁸. Inset shows the fluorescence spectra of the intact EGFP and split EGFP reassembled by DNA hybridization (~200 nM concentrations in PBS buffer, pH 7.4; spectra recorded 20 min after mixing).

ACKNOWLEDGMENTS We acknowledge our graduate students P. Chalasani and H.-W. Yiu, as well as C. Witte-Hoffmann, I. Smolina and Y. Yu, who participated in the protocol development. We also thank C.R. Cantor for inspiration and encouragement in this project.

COMPETING INTERESTS STATEMENT The authors declare that they have no competing financial interests.

Published online at <http://www.natureprotocols.com>
 Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions>

- Ghosh, I., Hamilton, A.D. & Regan, L. Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein. *J. Am. Chem. Soc.* **122**, 5658–5659 (2000).
- Stains, C.I., Porter, J.R., Ooi, A.T., Segal, D.J. & Ghosh, I. DNA sequence-enabled reassembly of the green fluorescent protein. *J. Am. Chem. Soc.* **127**, 10782–10783 (2005).
- Jeong, J. *et al.* Monitoring of conformational change in maltose binding protein using split green fluorescent protein. *Biochem. Biophys. Res. Commun.* **339**, 647–651 (2006).
- Shyu, Y.J., Liu, H., Deng, X. & Hu, C.D. Identification of new fluorescent protein fragments for bimolecular fluorescence complementation analysis under physiological conditions. *BioTechniques* **40**, 61–66 (2006).

5. Reid, B.G. & Flynn, G.C. Chromophore formation in green fluorescent protein. *Biochemistry* **36**, 6786–6791 (1997).
6. Tsien, R.Y. The green fluorescent protein. *Annu. Rev. Biochem.* **67**, 509–544 (1998).
7. Zimmer, M. Green fluorescent protein (GFP): applications, structure, and related photophysical behavior. *Chem. Rev.* **102**, 759–781 (2002).
8. Demidov, V.V. *et al.* Fast complementation of split fluorescent protein triggered by DNA hybridization. *Proc. Natl. Acad. Sci. USA* **103**, 2052–2056 (2006).
9. Endoh, T., Funabashi, H., Mie, M. & Kobatake, E. Method for detection of specific nucleic acids by recombinant protein with fluorescent resonance energy transfer. *Anal. Chem.* **77**, 4308–4314 (2005).
10. Marras, S.A., Kramer, F.R. & Tyagi, S. Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acids Res.* **30**, e122 (2002).
11. Tan, W., Wang, K. & Drake, T.J. Molecular beacons. *Curr. Opin. Chem. Biol.* **8**, 547–553 (2004).
12. Sambrook, J.F. & Russell, D.W. (eds.) *Molecular Cloning: A Laboratory Manual* 3rd edn. (Cold Spring Harbor Laboratory Press, 2001).
13. Hamad-Schifferli, K., Schwartz, J.J., Santos, A.T., Zhang, S. & Jacobson, J.M. Remote electronic control of DNA hybridization through inductive coupling to an attached metal nanocrystal antenna. *Nature* **415**, 152–155 (2002).
14. Sekar, M.M., Bloch, W. & St. John, P.M. Comparative study of sequence-dependent hybridization kinetics in solution and on microspheres. *Nucleic Acids Res.* **33**, 366–375 (2005).
15. Hu, C.D. & Kerppola, T.K. Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. *Nat. Biotechnol.* **21**, 539–545 (2003).
16. Chudakov, D.M., Lukyanov, S. & Lukyanov, K.A. Fluorescent proteins as a toolkit for *in vivo* imaging. *Trends Biotechnol.* **23**, 605–613 (2005).

