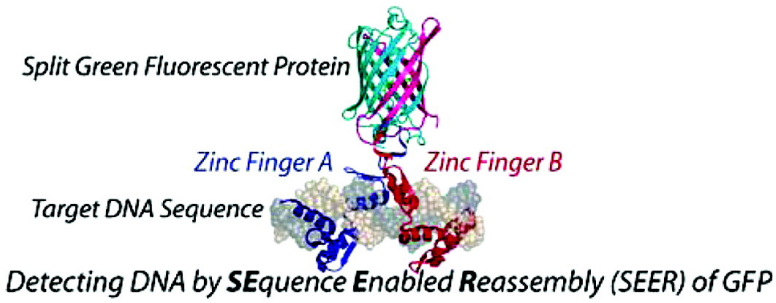
Communication

DNA Sequence-Enabled Reassembly of the Green Fluorescent Protein

Cliff I. Stains, Jason R. Porter, Aik T. Ooi, David J. Segal, and Indraneel Ghosh

J. Am. Chem. Soc., **2005**, 127 (31), 10782-10783• DOI: 10.1021/ja051969w • Publication Date (Web): 14 July 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 8 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 07/14/2005

DNA Sequence-Enabled Reassembly of the Green Fluorescent Protein

Cliff I. Stains,[†] Jason R. Porter,[†] Aik T. Ooi,[‡] David J. Segal,*[,‡] and Indraneel Ghosh*[,†]

Department of Chemistry and Department of Pharmacology and Toxicology, University of Arizona, Tucson, Arizona 85721

Received March 28, 2005; E-mail: ghosh@email.arizona.edu; segal@pharmacy.arizona.edu

Nature utilizes DNA binding proteins that activate or repress the transcription of desired genes. Many of these DNA binding proteins require the recognition of two adjacent DNA recognition sites for productive binding and subsequent signaling.[1] We utilize this principle of sequence-enabled activity in the rational design of a split protein system[2] that reassembles only in the presence of a specific DNA sequence (Figure 1) to allow for direct detection of DNA sequences. This approach called **SE**quence **E**nabled **R**eassembly of proteins (SEER) builds upon the ability to rationally dissect proteins to construct oligomerization-dependent protein reassembly systems and the ready availability of DNA binding Cys_2-His_2 zinc finger motifs[3] for targeting desired sequences of double-stranded DNA.

Oligomerization-assisted protein reassembly is possible when a protein can be fragmented into two halves that do not reassemble until appended to suitable protein oligomerization domains. This approach has been successfully utilized for the detection of oligomerizing proteins utilizing fragmented ubiquitin,[2a] beta-galactosidase,[2b] beta-lactamase,[2c] dihydrofolate reductase,[2d] green fluorescent protein (GFP),[2e,f] luciferase,[2g] and PH domains[2h] among others. However split protein reassembly has not been utilized for the direct detection of specific DNA sequences by ternary complexation. In order for ternary complexation in the presence of DNA, we have chosen to employ the ubiquitous Cys_2-His_2 family of zinc fingers that are the most widely used DNA binding motif in the human genome. Each of the zinc finger domains is capable of recognizing a 3-base pair tract in the major groove utilizing an α-helix.[4] Thus, a 3-finger protein can recognize a tract of 9 base pairs with picomolar to nanomolar affinity.[3] Moreover, recent experiments have resulted in the identification of a recognition code for nearly all-possible 3 base-pair DNA recognition sites, allowing for the design of unique zinc fingers for any DNA target of interest.[3] We envisioned that appending sequence-specific zinc fingers to appropriately fragmented proteins should, in principle, allow for protein reassembly only in the presence of the correct DNA sequence (Figure 1).

Our design entailed choosing appropriate zinc fingers and a suitable disassembled protein that could generate a readily detectable optical signal upon successful reassembly. For our disassembled protein, we chose fragments of a GFP variant that have been previously demonstrated to be capable of functional reassembly only when appended to oligomerizing protein or peptide partners.[2e,5] For our DNA binding domains, two well-characterized 3-domain containing zinc fingers Zif268 and PBSII, with low nanomolar affinity for unique 9-base pair sequences were chosen.[3,4] Protein constructs were designed such that the C-terminus of the GFP fragment (1−157) was fused to the N-terminus of Zif268 by means of a 15-residue linker, and the N-terminus of the GFP fragment (158−236) was fused to the C-terminus of PBSII through a 15-residue linker. Both protein constructs were incorporated together or separately in the PetDuet vector and verified by DNA sequencing (Supporting Information). Protein expression profiles revealed that NGFP−ZnFingerA was expressed in the insoluble fractions, whereas the smaller CGFP−ZnFingerB was expressed in both soluble and insoluble fractions as had been previously observed for coiled−coil peptides appended to similar GFP fragments.[2e] Importantly, no detectable fluorescence was observed in cells expressing each protein alone or together, over a period of 1 week, indicating lack of any detectable nonspecific reassembly in the presence of native *Escherichia coli* DNA. Both proteins, NGFP−ZnFingerA and CGFP−ZnFingerB were separately purified under denaturing conditions utilizing affinity chromatography and characterized by SDS-gel electrophoresis and mass spectrometry (Supporting Information). As proof of concept for SEER, we designed a double-stranded oligonucleotide that contained the two 9-base pair recognition sites for Zif268 and PBSII separated by a 10-nucleotide spacer, *Zif268-10-PBSII*.[3] The 10-nucleotide spacer was designed to allow for both halves of GFP to be juxtaposed on the same face of the target DNA but avoid steric crowding. Equimolar mixtures (15 μM) of the two purified proteins were refolded into 10 mM Tris-HCl, 100 mM NaCl, 1 mM DTT, and 100 μM $ZnCl_2$ at pH7.5 (buffer A) in the presence or absence of the target oligonucleotide (4 μM). Under these unoptimized conditions the concentration of DNA was 4-fold lower than that of the protein halves, such that the zinc finger-tagged GFP halves

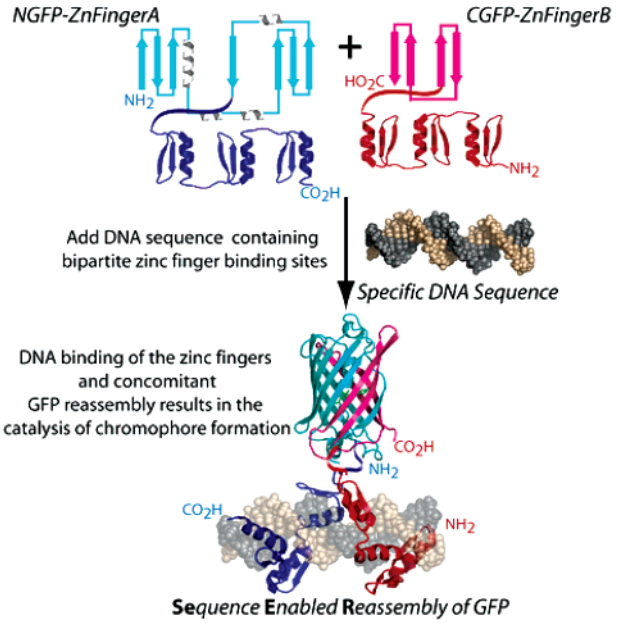


Figure 1. SEER strategy. NGFP−ZnFingerA (cyan and blue) comprises residues 1−157 of GFP fused by a 15-residue linker to the DNA binding zinc finger Zif268. CGFP−ZnFingerB (pink and red) comprises residues 158−238 of GFP fused by a 15-residue linker to the zinc finger PBSII.

[†] Department of Chemistry.
[‡] Department of Pharmacology and Toxicology.

10.1021/ja051969w CCC: $30.25

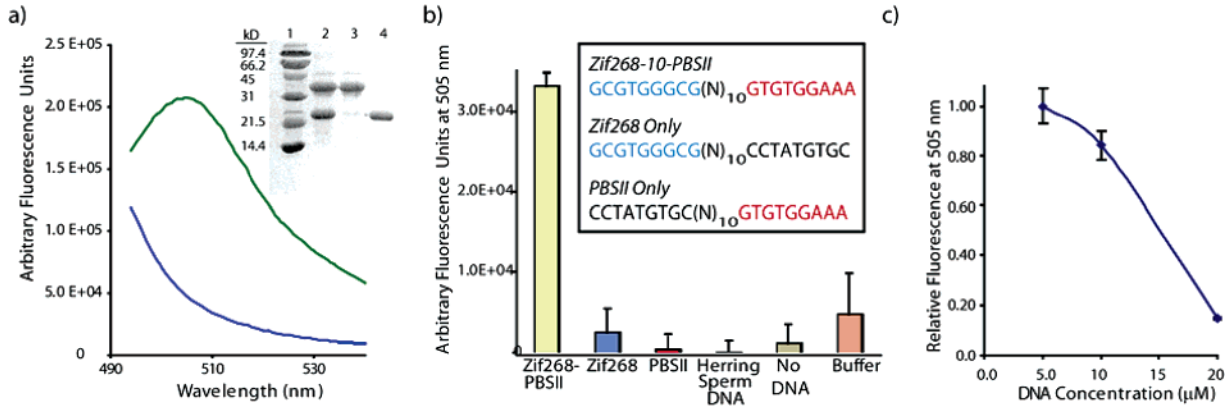


Figure 2. (a) Fluorescence emission spectra of NGFP−ZnFingerA (15 μM) + CGFP−ZnFingerB (15 μM) in the presence (green) and the absence (blue) of 4 μM target DNA (*Zif268-10-PBSII*) excited at 468 nm. (Inset) SDS-gel with MW standards (lane 1); equimolar mixture of NGFP−ZnFingerA and CGFP−ZnFingerB used in the SEER experiments (lane 2); NGFP−ZnFingerA (lane 3); and CGFP−ZnFingerB (lane 4). (b) Fluorescence emission at 505 nm of NGFP−ZnFingerA (5 μM) + CGFP−ZnFingerB (5 μM) in the presence of indicated double-stranded DNA controls (2.5 μM each). (c) Relative fluorescence emission at 505 nm of NGFP−ZnFingerA (5 μM) + CGFP−ZnFingerB (5 μM) as a function of increasing concentrations of target DNA (*Zif268-10-PBSII*).

would not localize to different DNA strands. Fluorescence spectra were acquired 48 h post-refolding by excitation at 468 nm. Fluorescence emission due to GFP chromophore formation was only observed for samples containing both halves of GFP−zinc-finger fusions in the presence of target oligonucleotide (Figure 2a), thus strongly supporting our SEER approach.

To further test the sequence specificity of reassembly and subsequent chromophore catalysis, several control experiments were designed. DNA sequences to determine specificity of reassembly consisted of the two-half-sites, *Zif268* alone, *PBSII* alone, and nonspecific herring sperm DNA. Equimolar mixtures of the two proteins, NGFP−ZnFingerA and CGFP−ZnFingerB, were allowed to refold in the presence of the control and target DNA sequences. No fluorescence was observed in the presence of any of the controls (Figure 2b), strongly confirming that the reassembly of the two halves of GFP requires the presence of both the zinc finger target sites on a single double-stranded DNA template. A final control experiment entailed varying the concentration of the target DNA, with the hypothesis that high molar ratios of the target DNA: dissected proteins would not allow for GFP reassembly as the two halves of GFP would statistically localize to different oligonucleotides with increasing DNA concentrations. The results of this experiment (Figure 2c) clearly demonstrated that only 4-fold excess of the *Zif268-10-PBSII* target DNA (20 μM) strongly inhibits GFP (5 μM) reassembly. A first attempt at gauging the effect of spacing of the two DNA target sites also revealed that our designed 10-bp separation between binding sites was substantially better than a 3-bp separation (Supporting Information). Future experiments will aim to optimize the minimum and maximum allowable separation between the two target DNA sequences as well as to optimize the amino acid linker separating the zinc fingers and GFP halves.

In conclusion we have clearly demonstrated the successful DNA-templated reassembly of the two fragments of GFP appended to the zinc fingers, Zif268 and PBSII. This to our knowledge is the first example of DNA-dependent reassembly of protein fragments, which can be applied to other split-protein enzymes, such as β-lactamase[2c] or luciferase,[2f] which can further amplify signal by substrate turnover. Future redesigns of the GFP SEER system will aim to modulate the relative DNA affinity of the substrates (protein halves) and product (reassembled protein) to allow for DNA-templated multiple turnover of reassembled GFP.[5] SEER could prove to be a valuable approach toward the direct in vivo detection of specific sequences of double-stranded DNA as well as for conditional responses to specific genetic mutations by reassembling proteins that act as cellular toxins. More generally, DNA-templated protein assembly also has bearing on the prebiotic co-evolution of oligonucleotides, proteins, and small-molecule metabolism.[6]

Acknowledgment. C.I.S. was supported by an NIH training grant. D.J.S. was supported by the American Cancer Society IRG7400125. We thank Carlos F. Barbas, III for helpful discussions.

Supporting Information Available: Details of cloning, sequencing, purification, characterization, and spectroscopy. This material is available free of charge via the Internet at http://pubs.acs.org

References

(1) Patikoglou, G.; Burley, S. K. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 289−325.
(2) (a) Johnsson, N.; Varshavsky, A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 10340−10344. (b) Rossi, F.; Charlton, C. A.; Blau, H. M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 8405−8410. (c) Galarneau, A.; Primeau, M.; Trudeau, L. E.; Michnick, S. W. *Nat. Biotechnol.* **2002**, *20*, 619−622. (d) Pelletier, J. N.; Campbell-Valois, F. X.; Michnick, S. W. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12141−12146. (e) Ghosh, I.; Hamilton, A. D.; Regan, L. *J. Am. Chem. Soc.* **2000**, *122*, 5658−5659. (f) Hu, C. D.; Kerppola, T. K. *Nat. Biotechnol.* **2003**, *21*, 539−545. (g) Paulmurugan, R.; Gambhir, S. S. *Anal. Chem.* **2003**, *75*, 1584−1589. (h) Sugimoto, K.; Mori, Y.; Makino, K.; Ohkubo, K.; Morii, T. *J. Am. Chem. Soc.* **2003**, *125*, 5000−5004.
(3) (a) Blancafort, P.; Segal, D. J.; Barbas, C. F., III. *Mol. Pharmacol.* **2004**, *66*, 1361−1371. (b) Segal, D. J. *Methods* **2002**, *26*, 76−83. (c) Pabo, C. O.; Peisach, E.; Grant, R. A. *Annu. Rev. Biochem.* **2001**, *70*, 313−340.
(4) Pavletich, N. P.; Pabo, C. O. *Science* **1991**, *252*, 809−817.
(5) Magliery, T. J.; Wilson, C. G. M.; Pan, W. L.; Mishler, D.; Ghosh, I.; Hamilton, A. D.; Regan, L. *J. Am. Chem. Soc.* **2005**, *127*, 146−157.
(6) (a) Paul, N.; Joyce, G. F. *Curr. Opin. Chem. Biol.* **2004**, *8*, 634−639. (b) Ghosh, I.; Chmielewski, J. *Curr. Opin. Chem. Biol.* **2004**, *8*, 640−644. (c) Calderone, C. T.; Liu, D. R. *Curr. Opin. Chem. Biol.* **2004**, *8*, 645−653.

JA051969W