A Fluorescent Indicator for Detecting Protein−Protein Interactions in Vivo Based on Protein Splicing

ARTICLE *in* ANALYTICAL CHEMISTRY · DECEMBER 2000

Impact Factor: 5.64 · DOI: 10.1021/ac000617z · Source: PubMed

5 AUTHORS, INCLUDING:

Moritoshi Sato
The University of Tokyo
71 PUBLICATIONS **2,303** CITATIONS

SEE PROFILE

Yoshikazu Ohya
The University of Tokyo
125 PUBLICATIONS **4,091** CITATIONS

SEE PROFILE

Anal. Chem. **2000,** *72,* 5151−5157

A Fluorescent Indicator for Detecting Protein−Protein Interactions in Vivo Based on Protein Splicing

Takeaki Ozawa,[†] Satoru Nogami,[‡] Moritoshi Sato,[†] Yoshikazu Ohya,[‡] and Yoshio Umezawa*,[†]

Department of Chemistry, School of Science, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan, and Japan Science and Technology Corporation (JST), Tokyo, Japan, and Department of Integrated Biosciences, Graduate School of Frontier Sciences, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan

We describe a new method with general applicability for monitoring any protein−protein interaction in vivo. The principle is based on a protein splicing system, which involves a self-catalyzed excision of protein splicing elements, or inteins, from flanking polypeptide sequences, or exteins, leading to formation of a new protein in which the exteins are linked directly by a peptide bond. As the exteins, split N- and C-terminal halves of enhanced green fluorescent protein (EGFP) were used. When a single peptide consisting of an intein derived from *Saccharomyces cerevisiae* intervening the split EGFP was expressed in *Escherichia coli*, the two external regions of EGFP were ligated, thereby forming the EGFP corresponding fluorophore. Genetic alteration of the intein, which involved large deletion of the central region encoding 104 amino acids, was performed. In the expression of the residual N- and C-terminal intein fragments each fused to the split EGFP exteins, the splicing *in trans* did not proceed. However, upon coexpression of calmodulin and its target peptide M13, each connected to the N- and C-terminal inteins, fluorescence of EGFP was observed. These results demonstrate that interaction of calmodulin and M13 triggers the refolding of intein, which induces the protein splicing, thereby folding the ligated extein correctly for yielding the EGFP fluorophore. This method opens a new way not only to screen protein−protein interactions but also to visualize the interaction in vivo in transgenic animals.

It is well known that protein−protein interactions play key roles in structural and functional organization of living cells. Many unsolved problems currently studied in molecular biology and biochemistry are related to the protein−protein interactions. Important examples are the identification and functional characterization of novel gene products, the dissection of proteins into structural or functional motifs, and the study of the physical basis of protein−protein complementarity in naturally occurring proteins or in designed products. These problems have been gradually solved by the development of protein library screening techniques such as the yeast two-hybrid strategy,[1,2] in which a library of proteins is screened for interaction with a "bait" protein. The two-hybrid method, a standard functional assay, facilitates the identification of potential protein−protein interactions and has been proposed as a method for the generation of protein interaction maps. The limitation of the two-hybrid technique is the set of detectable protein interactions to those that occur in the nucleus, in proximity to the reporter gene. Conceptually, the biological information generated by two-hybrid analysis is often questioned because of the inherent artificial nature of the assay. Therefore, this method has to be tested afterward in a model organism using groups of proteins for which functional data are available.[3]

To overcome this limitation, a transcription-based selectable version of "ubiquitin split protein sensor" (USPS) was developed,[4−6] in which upon protein interactions, N- and C-terminal ubiquitins are reconstituted, and by cleavage of a transcription factor, a nuclear localized reporter is activated. Another system was also described, in which interacting protein partners reconstitute guanine exchange factor (GEF) or Ras by bringing the catalytic domain together with the required membrane localization domain, which then complements yeast carrying a temperature-sensitive mutation in yeast GEF.[7,8] In a more general approach, split enzyme technologies have been described,[9−12] where protein interaction-

* To whom correspondence should be addressed: (phone) +81-3-5841-4351; (fax) +81-3-5841-8349; (e-mail) umezawa@chem.s.u-tokyo.ac.jp.

[†] The University of Tokyo, and Japan Science and Technology Corp. (JST).

[‡] Graduate School of Frontier Sciences.

(1) Chien, C. T.; Bartel, P. L.; Sternglanz, R.; Fields, S. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 9578−9582.
(2) Fields, S.; Song, O. *Nature* **1989**, *340*, 245−246.
(3) Walhout, A. J. M.; Sordella, R.; Lu, X.; Hartley, J. L.; Temple, G. F.; Brasch, M. A.; Thierry-Mieg, N.; Vidal, M. *Science* **2000**, *287*, 116−122.
(4) Dunnwald, M.; Varshavsky, A.; Johnsson, N. *Mol. Biol. Cell* **1999**, *10*, 329−344.
(5) Johnsson, N.; Varshavsky, A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 10340−10344.
(6) Stagljar, I.; Korostensky, C.; Johnsson, N.; Heesen, S. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5187−5192.
(7) Aronheim, A. *Nucleic Acids Res.* **1997**, *25*, 3373−3374.
(8) Aronheim, A.; Zandi, E.; Hennemann, H.; Elledge, S. J.; Karin, M. *Mol. Cell. Biol.* **1997**, *17*, 3094−3102.
(9) Rossi, F.; Charlton, C. A.; Blau, H. M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 8405−8410.
(10) Remy, I.; Michnick, S. W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5394−5399.
(11) Pelletier, J. N.; Arndt, K. M.; Plückthun, A.; Michnick, S. W. *Nature Biotechnol.* **1999**, *17*, 683−690.
(12) Karimova, G.; Pidoux, J.; Ullmann, A.; Ladant, D. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 5752−5756.

10.1021/ac000617z CCC: $19.00
Published on Web 09/29/2000

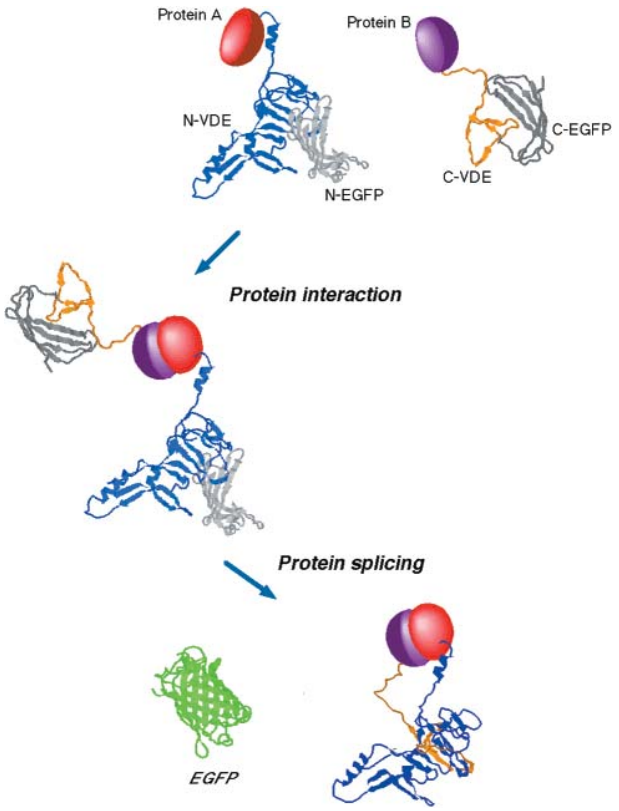

Figure 1. Split EGFP as a probe for protein−protein interaction. Ribbon diagrams of the N-terminal half of VDE (1−184 amino acids; blue) and the C-terminal half of VDE (389−454 amino acids; yellow) are each connected with the N-terminal half of EGFP (1−128 amino acids; light gray) and C-terminal half of EGFP (129−238 amino acids; dark gray), respectively. Interacted protein A and protein B are linked to opposite ends of the split VDE. Interaction between protein A and protein B accelerates the folding of N- and C-terminal VDE and protein splicing occurs. The N- and C-terminal halves of EGFP are linked together by a normal peptide bond to yield the β-can structure (green), in which the fluorophore is formed.

mediated reassembly of the split enzyme was found to restore its activity. The reconstituted enzymatic activity can be measured by the phenotype or fluorescent substrate analogues. Although these systems are well suited for assaying interactions between cytoplasm and membrane-proximal proteins, they can be utilized only in appropriately engineered cells and/or are prone to false positive signals.

We describe herewith a new method to detect protein−protein interaction based on the protein splicing of VMA1 intein, in which calmodulin−M13 interaction was used as a model system. The principle is schematically shown in Figure 1. The protein splicing is a posttranslational processing event involving precise excision of an internal protein segment, the intein, from a primary translation product with concomitant ligation of the flanking sequences, the exteins.[13] In the budding yeast *Saccharomyces cerevisiae*, the nascent 120-kDa *VMA1* translational product catalyzes protein splicing to yield a 70-kDa catalytic subunit of the vacuolar H^+-ATPase and a 50-kDa site-specific endonuclease (*VMA1*-derived endonuclease; VDE, also called PI−*Sce*I).[14,15] The deletion mutant of the endonuclease domain of VDE replaced with flexible dodecapeptide linker is known to have splicing activity.[16,17] In an attempt to monitor the protein−protein interaction in vivo, we used the N- and C-terminal halves of the intein without the endonuclease domain, which are termed as N_VDE and C_VDE, respectively. The N_VDE is fused to an N-terminal half of enhanced green fluorescent protein (EGFP) and the C_VDE to the rest of EGFP. Each of these fusion proteins is linked to a protein of interest (protein A) and its target protein (protein B). When interaction occurs between the two proteins, the N_ and C_terminal halves of VDE are brought in close proximity and undergo correct folding, which induces a splicing event and thereby N- and C-terminal fragments of EGFP directly linked to each other by a peptide bond. The matured EGFP forms the fluorophore with its emission maximum at 510 nm. The extent of the protein−protein interaction can be evaluated by measuring the magnitude of fluorescence intensity originated from the formation of EGFP.

EXPERIMENTAL PROCEDURES

Numbering of Residues in the VMA1 Intein Fusions. Amino acid numbers refer to the position in the 454-amino acid VDE intein, with Cys1 being the first residue to the intein and Asn454 the last. The numbering continues sequentially into the C_extein beginning with Cys455. The N_extein residue adjacent to Cys1 is referred to as −1, and the numbering continues sequentially into the N_extein.

Strains. *Escherichia coli* strain DH5α was used to produce the recombinant glutathione *S*-transferase (GST) fusion protein. *E. coli* strain BL21(DE3)pLysS was used to produce recombinant fusion proteins attached to the His tag in the N_terminals.

General Procedures. The crude extracts from *E. coli* were electrophoresed on 12−15% SDS−PAGE gels with a protein marker, 10−225 kDa (Novagen) and visualized by Coomassie Brilliant Blue staining. Western blotting data were probed with anti-VDE polyclonal, anti-His tag polyclonal (Santa Cruz Biotechnology) or anti-GFP monoclonal antibodies (Boehringer Mannheim). Protein concentrations were determined by the Bradford assay (BioRad). All cloning enzymes were from Takara Biomedicals and were used according to the manufacturer's instructions. All PCR fragments were sequenced by ABI310 genetic analyzer.

Protein Expression and Affinity Purification. . Detailed protocols of the present plasmid construction, all of which were shown in Figure 2, are available upon request. *E. coli* strain DH5α carrying plasmids pGEX series (Amersham Pharmacia Biotech) were grown to an OD_{600} of 0.6−0.8 in 10 mL of liquid Luria Broth (LB) medium containing 100 μg/mL ampicillin, at which time they were induced with 1 mM isopropyl β-D-thiogalactosidase (IPTG). The bacteria were allowed to express recombinant protein for 3 h at 25 °C and then overnight at 4 °C. Total cell extracts were prepared by pelleting the cells for 5 min at 8000*g* followed by lysis in 50 μL of SDS−PAGE sampling buffer (62.5 mM Tris-HCl, pH6.8, 2% SDS, 1% 2-mercaptoethanol, 10% glycerol, 0.01% bromophenol blue). Purification of GST fusion protein was performed using a

(13) Gimble, F. S. *Sci. Biol.* **1998**, *5*, R251−256.
(14) Hirata, R.; Ohsumi, A.; Nakano, H.; Kawasaki, K.; Suzuki, K.; Anraku, Y. *J. Biol. Chem.* **1990**, *265*, 6726−6733.
(15) Kane, P. M.; Yamashiro, C. T.; Wolczyk, D. F.; Neff, N.; Goebl, M.; Stevens, T. H. *Science* **1990**, *250*, 651−657.
(16) Cooper, A. A.; Chen, Y. J.; Lindorfer, M. A.; Stevens, T. H. *EMBO J.* **1993**, *12*, 2575−2583.
(17) Chong, S.; Xu, M.-Q. *J. Biol. Chem.* **1997**, *272*, 15587−15590.

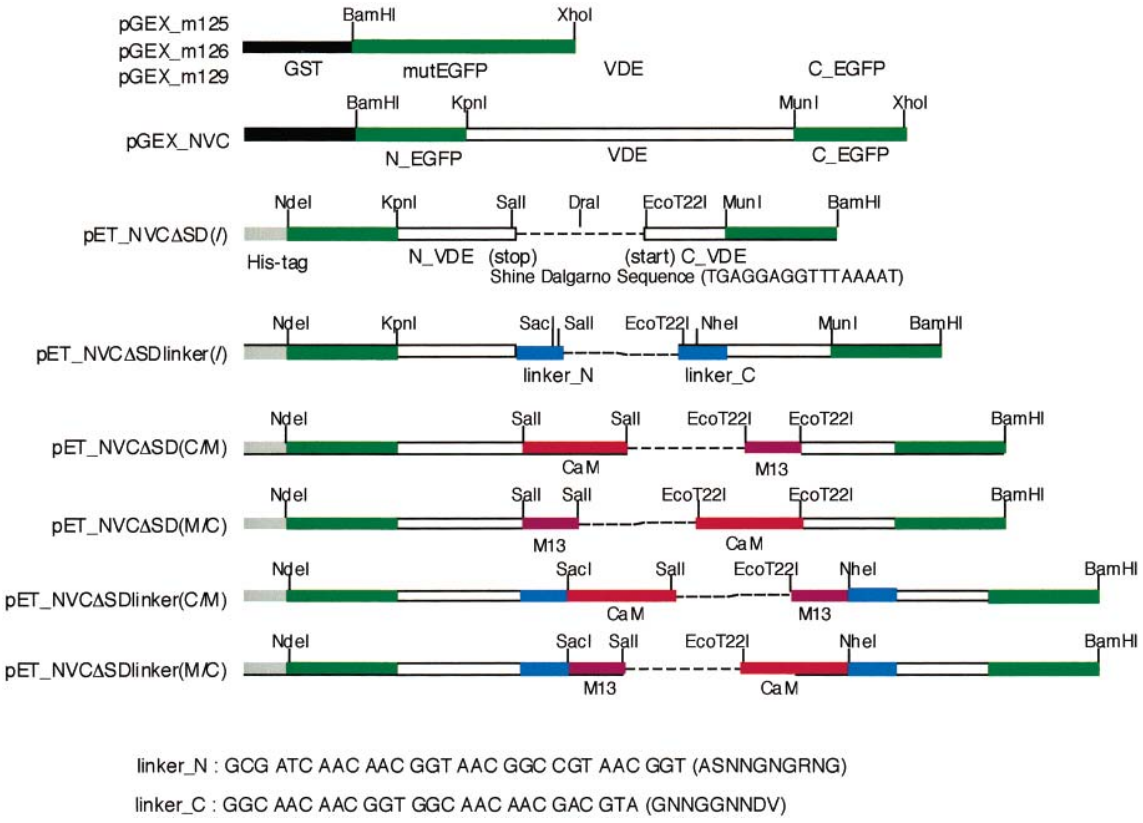

Figure 2. Schematic representation of the major plasmids used in this work. Shown at the top of each bar are the restriction sites. Dashed lines indicate a endonuclease domain in the intein VDE. The intein lacking the endonuclease domain indicates open bars with the N- and C-terminal halves of EGFP in green, calmodulin (CaM) in red, and M13 in purple. Linker_N and linker_C are blue, and their precise sequences are shown at the bottom. Tags of GST and His are shown in black and gray, respectively. Stop and start refer to translation termination and initiation codons.

glutathione Sepharose-affinity column. A GST fused protein was eluted by removing the GST tag with PreScission protease cleavage (Pharmacia Biotech).

Visualization of *E. coli* on Plates and Fluorescence Measurements. *E. coli* strain DH5α carrying plasmids pGEX series and BL21(DE3)pLysS carrying plasmids pET series (Novagen) was grown in the same conditions as described above. In the case of pGEX series, the bacteria were allowed to express recombinant protein with IPTG for 3 h at 25 °C and 10 μL of LB medium containing the bacteria was streaked onto LB-agarose plates and incubated overnight at 25 °C. The bacteria on the plates were illuminated with blue-LED at 470 nm (LAS-1000plus, Fujifilm), emitting maximum at 510 nm, which was detected by a cooled CCD equipped with an emission filter (530DF30).

To visualize the fluorescence from bacteria carrying pET series, the bacteria were allowed to express recombinant protein with IPTG for 12 h at 25 °C and stored for more than 24 h at 4 °C. Ten microliters of LB medium including the bacteria was streaked onto M9-agarose clear plates containing 20% glucose, 25 mM NH_4Cl, 8.6 mM NaCl, 20 mM KH_2PO_4, and 48 mM $Na_2HPO_4 \cdot 7H_2O$. The bacteria were illuminated with Ar laser at 488 nm, and the fluorescence was detected by PMT with an emission filter (530DF30) (Molecular Imager FX, BioRad).

Fluorescence spectra of crude extracts from *E. coli* were obtained using a spectrofluorometer (FP-750, Jasco Co., Tokyo, Japan). The growth bacteria in 10 mL of LB medium were collected by centrifugation at 8000*g* for 5 min and redissolved in 1 mL of PBS buffer. The cells were lysed with a tip sonicator, and the lysate was subjected to centrifugation at 15000*g* for 10 min at 4 °C. A portion of 100 μL of the supernatant was diluted into a 600 μL of PBS buffer, which was allowed for the measurement of fluorescence spectra.

RESULTS AND DISCUSSION

For efficient splicing to occur, the protein precursor has to fold properly to bring the two splice junctions into close proximity and precisely align,[18] both of which were derived from C-terminal ends of N_extein and VDE. At the splice junctions of *S. cerevisiae*, one Cys residue, another Gly residue at its upstream of the position of −1, and three hydrophobic amino acid residues at their positions of −5, −4, and −3 are required for the N-terminal half of EGFP to form β-strand sheet with the C-terminal end of VDE (Figure 3A).[19,20] To meet these demands in EGFP, at least two amino acid residues at a position from I124 to I129 have to be replaced as described below. The reason that the position is chosen from the total 238 amino acids in EGFP is that it locates at the end of the sixth β-sheet strand of EGFP, at which the folding of its N-terminal half is just finished, and therefore, the structure

(18) Duan, X.; Gimble, F. S.; Quiocho, F. A. *Cell* **1997**, *89*, 555−564.
(19) Kawasaki, M.; Makino, S.-I.; Matsuzawa, H.; Satow, Y.; Ohya, Y.; Anraku, Y. *Biochem. Biophys. Res. Commun.* **1996**, *222*, 827−832.
(20) Nogami, S.; Satow, Y.; Ohya, Y.; Anraku, Y. *Genetics* **1997**, *147*, 73−85.

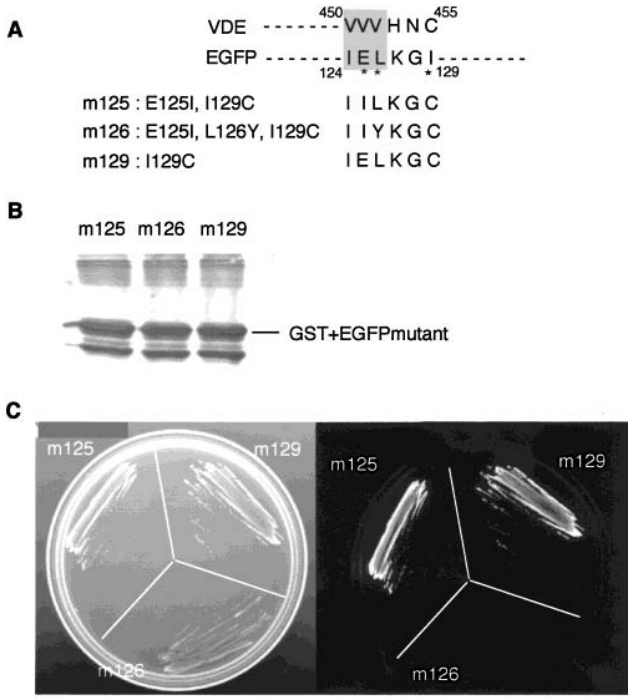

Figure 3. Mutagenesis analysis of EGFP. (A) Schematic representation of spatial interaction between EGFP residues 124−129 and VDE residues 450−455 near the splicing junctions. Hydrophobic residues required for efficient protein splicing are in the shaded box. Amino acid residues of each mutant m125, m126, and m129 EGFP are aligned parallel to those of EGFP. (B) Expression of EGFP mutants in *E. coli*. Western blots stained for anti-GFP monoclonal antibody. (C) Visual appearance of *E. coli* expressing three different mutants of EGFP. The bacteria were streaked onto LB agarose, transmitted with white light (left panel), or illuminated with blue-LED at 470 nm (right panel). The fluorescence was detected by a cooled CCD through an emission filter of 530DF30.

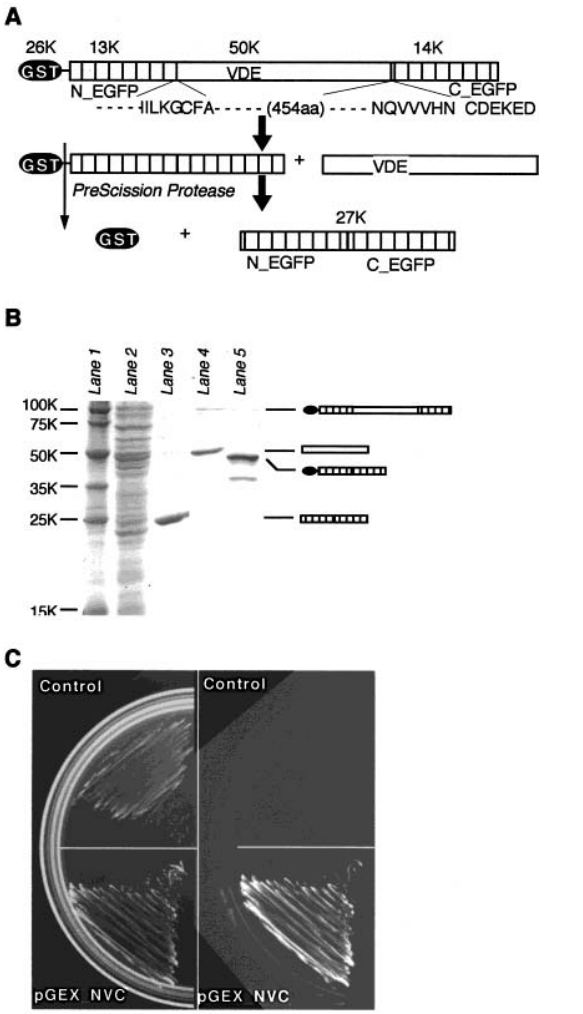

Figure 4. Construction and expression of a GST-fused EGFP−VDE protein in *E. coli*. (A) Model of the protein splicing of the pGEX_NVC gene product. (B) SDS−PAGE analysis of proteins expressed in *E. coli* transformed with pGEX_NVC. Lanes 1−3, Coomassie Blue-stained SDS−PAGE. Lane 1, protein molecular mass standards (Novagen) with their molecular masses (kDa) shown on the left; lane 2, crude extract before loaded onto a GST-affinity column; lane 3, a GST-affinity purified sample. Lanes 4 and 5, Western blotting analysis of crude cell extract using antibodies specific for VDE (lane 4) and GFP (lane 5). (C) Visual appearance of *E. coli* carrying pGEX_NVC (lower panels) and pGEX6p-2 (upper panels). The bacteria were streaked onto LB agarose, transmitted with white light (left panel), or illuminated with blue-LED at 470 nm (right panel). The fluorescence was detected by a cooled CCD through an emission filter of 530DF30.

of the N-terminal seems to be of relatively stable conformation.[21] Figure 3C shows the mutational analysis of EGFP. Mutations of I129C and E125I generated fluorescence, of which excitation and emission peaks were 488 and 510 nm, respectively, which was consistent with those of EGFP (data not shown). However, the additional mutation of L126Y to m125 lost fluorescence though the expression levels of those three mutants were the same (Figure 3B), indicating that the L126Y mutation caused incorrect folding of the m129EGFP mutant and/or became incapable of forming the fluorophore. Hereafter, the m125 EGFP mutant is simply called as EGFP mutant.

To examine whether protein splicing occurs in a single polypeptide fusing VDE intervening between N- and C-terminal halves of the EGFP mutant, expression of pGEX_NVC was studied in *E. coli* at 25 °C. The plasmid that covers the entire VDE region and N- and C-terminal halves of EGFP mutant was fused to the GST gene under control of *tac* promoter. This construct expresses a chimeric protein that is composed of GST (26 kDa), 125 residues from the N_terminal half of EGFP mutant (13 kDa), VDE (50 kDa), and C-terminal half of the mutant (14 kDa). To analyze the products after protein expression, crude and affinity-purified extracts from *E. coli* were subjected to SDS−PAGE. As shown in Figure 4B, a major component of the crude extract (lane 2) was a ~50-kDa protein, the size being expected for VDE (50 kDa) and the ligated extein, i.e., GST fusion with the N-terminal half of EGFP mutant (26 kDa plus 13 kDa) and the C-terminal half of the mutant (14 kDa). This result indicates that a 103-kDa precursor of the GST fusion protein was processed into a 50-kDa VDE and 53-kDa GST−EGFP mutant fusion protein. The expected sizes of VDE and GST−EGFP mutant fusion protein were also detected by Western blotting. Anti-VDE and anti-GFP antibodies reacted with the excised 50-kDa intein (lane 4) and 53-kDa GST-EGFP mutant fusion protein (lane 5), respectively. In addition, a 100-kDa minor component of unspliced precursor was also

(21) Ormö, M.; Cubitt, A.; Kallio, K.; Gross, L.; Tsien, R.; Remington, S. *Science* **1996**, *273*, 1392−1395.

detected by both antibodies. The identities of the GST-EGFP mutant fusion protein was further verified by purification with a GST-affinity column. The crude extract was applied to the affinity column and bound proteins to the resin were excised with PreScission protease, which subjected to SDS−PAGE (lane 3). A single 25-kDa band was observed, which was almost the same as the calculated mass of EGFP mutant. Fluorescence spectra of the affinity-purified protein were further measured, of which excitation and emission maximums were found to be 488 and 510 nm, respectively, which were consistent with those of EGFP. These results demonstrate that spliced intervening VDE was excised out and the two external regions of the N- and C-terminal halves of EGFP mutant were ligated with a peptide bond, and remarkably, the ligated protein of EGFP mutant folded correctly and the fluorophore was formed in the antiparallel β-sheet can structure.

Split EGFP Works as a Probe for Protein Interaction. In testing for protein−protein interactions, plasmid pET_NVCΔSD-(/) was constructed by substituting GST of pGEX_NVC for a His tag of a pET16-b vector, followed by splitting functional splicing parts. This splitting was achieved by replacing a functionally unrelated endonuclease motif from a 185−389-amino acid region with a cassette consisting of (translation termination codon)−(Shine-Dalgarno sequence)−(translation initiation codon).[22] The resulting plasmid pET_NVCΔSD(/) is essentially a two-gene operon, with the first gene (gene I) encoding the N-terminal halves of EGFP and VDE (N_EGFP−VDE) and with the second gene (gene II) encoding the C-terminal halves of VDE and EGFP (C_VDE−EGFP). To ensure that N_ and C_terminal halves of VDE could be spatially proximal when protein interactions occurred, flexible peptide linkers containing Gly-Asn repeat were inserted in frame just before the translation termination codon and just after the initiation codon (pET_NVCΔSDlinker(/); see Figure 2).[17]

To examine whether some particular protein−protein interaction in *E. coli* facilitates the splicing event to yield matured EGFP, calmodulin (CaM) and its target peptide, M13, were chosen as a model system of the protein interaction. The structure of the CaM and M13 complex has been well resolved by NMR,[23] which gave information of the distances from a given amino acid in CaM to the one in M13. The binding of CaM to M13 is known to be Ca^{2+} dependent; thus, it seems uncertain whether the interaction between CaM and M13 occurs at the physiological concentration range in *E. coli*. However, it has been revealed that the dissociation constant between CaM and Ca^{2+} in the presence of M13 is 2 orders of magnitude lower than that in the absence of M13,[24] and it has been shown by surface plasmon resonance studies that the Ca^{2+}−CaM−M13 ternary complex is formed in the Ca^{2+} concentrations higher than 3×10^{-8} M.[25,26] Considering these facts, CaM is assumed to bind to M13 in the physiological concentration ranges of Ca^{2+} in *E. coli*.

Recombinant plasmid, pET_NVCΔSD(C/M),pET_NVCΔSD-linker(C/M), and pET_NVCΔSDlinker(/) were introduced into *E. coli* cells to produce the corresponding fusion proteins. To begin the splicing event in the bacteria, the protein expressions were performed for 12 h at 25 °C,[20] and the bacteria were stored 1−2 days at 4 °C. The expressed protein products of the first gene (gene I) of the two-gene operon were recognized by an anti-His tag antibody (Figure 5A). Major components specifically recognized by the antibody were 55-kDa proteins (lanes 1 and 4), the size expected for unspliced precursors of N_EGFP−VDE (36-kDa) plus CaM (17-kDa) or its connected linker (1-kDa). As its control, the lysate of *E. coli* carrying the plasmid pET_NVCΔSDlinker(/) was subjected to SDS−PAGE (lane 6), which gave a single band of N_EGFP−VDE as expected. The expression levels of unspliced precursor proteins from those three plasmids were almost the same. The protein products of the second gene operon (gene II) were identified by anti-GFP monoclonal antibody, which was confirmed to recognize not the N-terminal but the C-terminal half of EGFP (data not shown). In cells containing pET_NVCΔSD-(C/M) and pET_NVCΔSDlinker(C/M), the expressed protein size from the second gene operon corresponded well with the predicted sizes of precursor protein of N_VDE-EGFP (20 kDa) plus M13 (3 kDa) (Figure 5B, lane 1) or linker-including M13 (4-kDa) (lane 4). Similarly, cells containing the control plasmid produced an expected protein (20 kDa, lane 6). The amounts of those three proteins originating from operon II were all of the same expression level.

Fluorescence spectra of *E. coli* lysates carrying each plasmid are shown in Figure 5C. In the case of *E. coli* containing plasmid pET_NVCΔSDlinker(/), no change in the spectrum was observed. Coexpression of CaM and M13 in the bacteria carrying the plasmid pET_NVCΔSD(C/M) resulted in a slight change in the fluorescence emission maximum at 510 nm. In contrast, fluorescence at 510 nm increased to a large extent upon expression of CaM and M13, both of which were connected with flexible peptide linkers to the N_EGFP−VDE and C_VDE−EGFP, respectively. The fluorescence intensity of crude extract from *E. coli* carrying pET_NVCΔSDlinker(C/M) was strong enough to discriminate the cells containing control plasmid pET_NVCΔSDlinker(/) or the plasmid not coding flexible linkers pET_NVCΔSD(C/M) (Figure 5C). These results demonstrate that interaction between CaM and M13 facilitated splicing *in trans* and the two external regions of N- and C-terminals of the EGFP mutant were ligated to yield the EGFP fluorophore.

For the splicing event to occur, the N- and C-terminal halves of VDE have to correctly fold each other. This folding is achieved by the C-terminal end of N_VDE proximal enough to the N-terminal end of C_VDE. In the case of CaM and M13 complex, a distance between the N-terminal end in CaM and the C-terminal end in M13, both of which are in contact respectively to N_EGFP−VDE and C_VDE−EGFP, is ∼50 Å (retrieved from Brookhaven Protein Data Bank). The distances may be too large for N_VDE to be adjacent to C_VDE, thereby disturbing the correct folding in *E. coli* carrying the plasmid pET_NVCΔSD(C/M), albeit splicing occurred to some extent. In contrast, an increase in the fluorescence intensity at 510 nm for the plasmid pET_NVCΔSD-linker(C/M) indicates that 10- and 9-amino acid linkers each attached to N_ and C_VDE, respectively, gave flexibility to the

(22) Wu, H.; Xu, M.-Q.; Liu, X.-Q. *Biochim. Biophys. Acta* **1998**, *1387*, 422−432.
(23) Ikura, M.; Clore, G. M.; Gronenborn, A. M.; Zhu, G.; Klee, C. B.; Bax, A. *Science* **1992**, *256*, 632−638.
(24) Porumb, T.; Yau, P.; Harvey, T. S.; Ikura, M. *Protein Eng.* **1994**, *7*, 109−115.
(25) Ozawa, T.; Sasaki, K.; Umezawa, Y. *Biochim. Biophys. Acta* **1999**, *1434*, 211−220.
(26) Ozawa, T.; Kakuta, M.; Sugawara, M.; Umezawa, Y. *Anal. Chem.* **1997**, *69*, 3081−3085.

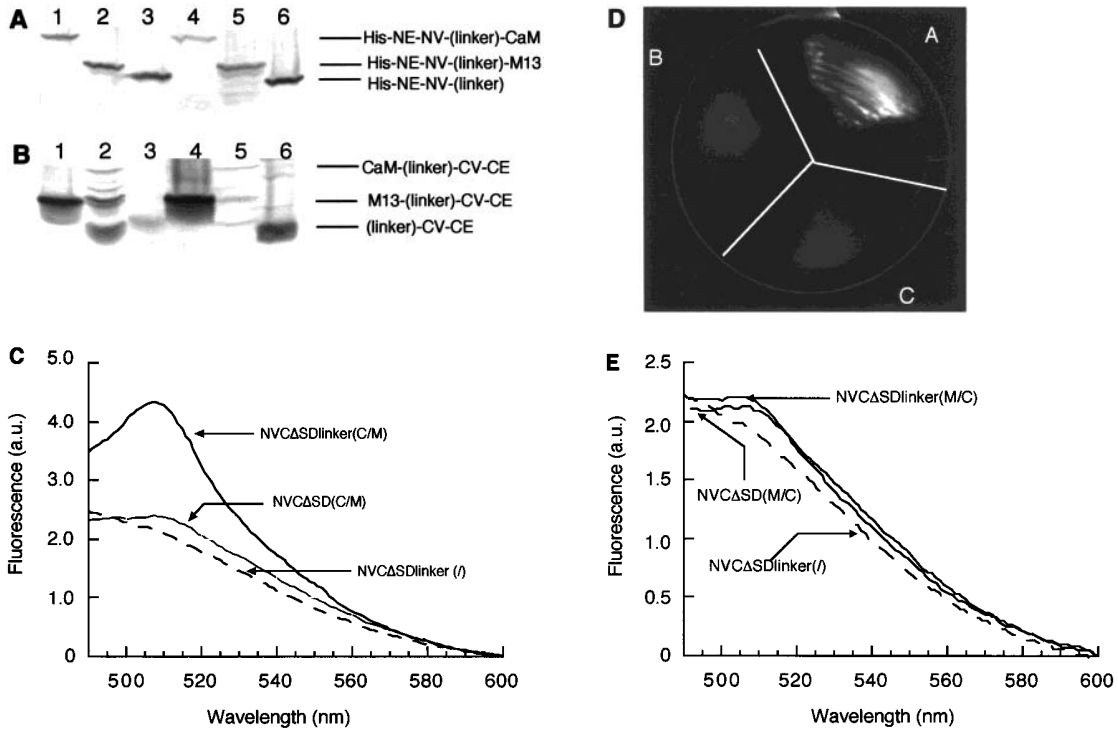

Figure 5. Protein splicing induced by CaM-M13 interaction. SDS−PAGE analysis of proteins expressed in *E. coli* transformed with pET_NVCΔSD(C/M) (lane 1), pET_NVCΔSD(M/C) (lane 2), pET_NVCΔSD(/) (lane 3), pET_NVCΔSDlinker(C/M) (lane 4), pET_NVCΔSDlinker-(M/C) (lane 5), and pET_NVCΔSDlinker(/) (lane 6). Western blotting was done by anti-His tag (A) and anti-GFP (B) antibodies. (C) Fluorescence emission spectra of crude extracts of *E. coli* carrying the plasmid pET_NVCΔSDlinker(C/M), pET_NVCΔSD(C/M), and pET__NVCΔSDlinker-(/). Excitation was 470 nm with 5.0-nm bandwidth. The emission bandwidth was 5.0 nm. (D) Visual appearance of *E. coli* containing the plasmid pET_NVCΔSDlinker(C/M) (A), pET_NVCΔSD(C/M) (B), and pET_NVCΔSDlinker(/) (C). The bacteria were streaked onto M9 agarose, illuminated with Ar laser at 488 nm. The fluorescence was detected by a PMT with an emission filter (530DF30). (E) Fluorescence emission spectra of crude extracts of *E. coli* carrying the plasmid pET_NVCΔSDlinker(M/C), pET_NVCΔSD(M/C), and pET_NVCΔSDlinker(/). Excitation was 470 nm with a 5.0-nm bandwidth. The emission bandwidth was 5.0 nm.

folding of VDE, thereby allowing VDE to have proper conformation for efficient splicing.

We next examined the splicing efficiency of *E. coli* each containing the plasmid, pET_NVCΔSDlinker(M/C) and pET_NVCΔSD(M/C), in which CaM and M13 were inserted in the opposite positions. To analyze the products, crude extracts from the bacteria were subjected to SDS−PAGE. As shown in Figure 5A, 23-kDa unspliced precursors were detected by anti-His tag antibody (lanes 2 and 5). The protein expression levels of the unspliced precursors for each *E. coli* carrying pET_NVCΔSD-linker(M/C) and pET_NVCΔSD(M/C) were almost the same as that for the control plasmid pET_NVCΔSDlinker(/) (lane 6). However, the expression of proteins derived from operon II detected by the anti-GFP antibody gave several bands (Figure 5B, lanes 2 and 5): 23-kDa and 20-kDa unexpected proteins plus the expected 37-kDa unspliced precursors, the expression levels of which were lower than the control. The fluorescence spectra of crude extracts from *E. coli* carrying both plasmids showed relatively weak emission maximums at 510 nm (Figure 5E), indicating that a very small quantity of matured EGFP was produced by CaM and M13 interaction.

In pET_NVCΔSDlinker(M/C) and pET_NVCΔSD(M/C) constructs, a distance from the C-terminal end of N_VDE−EGFP to the N-terminal end of C_VDE−EGFP, both connected to M13 and CaM, respectively, is estimated to be ∼20 Å. The distance is shorter than that of the former case, and the splicing efficiency was expected to increase as a result. However, very little mature EGFP was obtained from each construct. The reason for this may be that expression level of expected 37-kDa unspliced precursor derived from operon II was very low in comparison to the one of the plasmid pET_NVCΔSD(C/M) and of pET_NVCΔSDlinker-(C/M) (Figure 5B, lanes 1 and 4) and thereby reduced the splicing efficiency of the two separate VDE.

The present method is based on protein splicing to produce matured EGFP, which distinguishes it from the two-hybrid, USPS, or split enzyme methods. Contrary to the earlier methods, neither reporter genes nor substrates for enzymes are needed. This indicates that using the present method, protein−protein interactions can also be measured on the cell membrane or in intercellular events such as activation of receptors by hormones, intracellular membrane-associated proteins that mediate signaling cascade or adhesion of cell membranes. Another advantage of the present splicing-based method is that EGFP accumulates in a target cell until it degradates, and therefore, information of the protein−protein interaction is integrated in the cell. In addition, unless the interaction occurs, a splicing event does not proceed and no fluorophore is formed. This feature is a strong advantage in the highly sensitive detection of a low amount of EGFP. All of these features should be useful in vivo analysis of protein−protein interactions not only in living cells as shown in the present study but also in cell−cell interactions of interest in organisms or, importantly, in living animals. The use of GFP in transgenic

vertebrates has been limited, in most cases, as a reporter for gene or protein expression. The concept addressed in the present study will open a way to visualize any protein−protein interaction of interest in living animals and thereby make it possible to analyze the localization or the interaction in the tissues or its roles in the development stage. The limitation of the present method may arise when interactions between large proteins or giant oligomeric proteins are detected. In this case, steric hindrance of the large proteins can possibly inhibit the interactions between the two separate VDE sequences that form the splicing complex of EGFP. These conceivable problems could be overcome by correct design of the fusion constructs, and it is now under examination.

ACKNOWLEDGMENT

This work was financially supported by Core Research for Evolutional Science and Technology (CREST) of Japan Science and Technology (JST) and by grants for Scientific Research by the Ministry of Education, Science and Culture, Japan.

Received for review May 31, 2000. Accepted August 21, 2000.

AC000617Z