Supramolecular Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713649759

Peptide Assemblies in Living Cells. Methods for Detecting Protein-Protein Interactions†

Takeaki Ozawa[ab]; Yoshio Umezawa[ab]
[a] Department of Chemistry, School of Science, The University of Tokyo, Tokyo, Japan [b] Japan Science and Technology Corporation, Tokyo, Japan

Online publication date: 29 October 2010

To cite this Article Ozawa, Takeaki and Umezawa, Yoshio(2002) 'Peptide Assemblies in Living Cells. Methods for Detecting Protein-Protein Interactions†', Supramolecular Chemistry, 14: 2, 271 — 280
To link to this Article: DOI: 10.1080/10610270290026185
URL: http://dx.doi.org/10.1080/10610270290026185

Supramolecular Chemistry, 2002 Vol. 14 (2-3), pp. 271–280

Invited Paper

Peptide Assemblies in Living Cells. Methods for Detecting Protein–Protein Interactions†

TAKEAKI OZAWA[a,b] and YOSHIO UMEZAWA[a,b,*]

[a]*Department of Chemistry, School of Science, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan;* [b]*Japan Science and Technology Corporation, Tokyo, Japan*

(Received 10 September 2001)

In mammalian cells, protein–protein interactions constitute essential regulatory steps that modulate the activity of signaling pathways. To understand the complicated mechanisms of the interactions in living cells, chemical crosslinking or immunoprecipitation has extensively been used. However, such biochemical methods require cell disruption and do not necessarily preserve all interactions intact. In recent years, several elegant approaches have been developed towards understanding the interactions. A common advantage of these new approaches is direct observation of the interaction in living cells without the need for disrupting the cells. We describe herein recent advances of those methods including our recent works based on protein splicing for detecting protein–protein interactions *in vivo* and highlight some potential applications of these techniques.

Keywords: Protein–protein interaction; Protein splicing; Green fluorescent protein (GFP); Luciferase

INTRODUCTION

Protein–protein interactions are critical in many biological processes in living cells, that are involved in the assembly of enzymes, peptide homodimers and heterodimers in the regulation of intracellular transport pathways, gene expression, receptor–ligand interactions, and in the therapeutic or toxic effects of administered drugs. Identification of these interactions and characterization of their physiological significance constitute one of the main goals of current research in different biological fields. Towards this goal, several technologies have been developed for detecting protein–protein interactions without the need for disrupting living cells. The technologies are categorized into two systems; a transcriptional activation resulting from specific protein–protein interactions, and the interaction-induced complementation or reconstitution of a readout protein. At deep bottom, these methods have similar logic: Interacting proteins are tagged directly to polypeptides with a particular function, which works to convert the interaction event into a detectable signal. But implementations, the nature of the readouts and the experimental flexibility differ greatly. Here, we will summarize the overview of those systems for monitoring protein–protein interactions and then focus on our recent works on how to detect interaction between a pair of peptides or proteins of interest based on protein splicing.

DETECTION OF PROTEIN INTERACTION BASED ON TRANSCRIPTIONAL ACTIVATION

In eukaryotic cells, transcription from gene to RNA requires both widely distributed transcription

† To commemorate Prof. Eiichi Kimura's retirement from Hiroshima University after 30 years, with seminal contributions on functionalized macrocyclic polyamines for frontier supramolecular and medicinal chemistries.

*Corresponding author. Tel.: +81-3-5841-4351. Fax: +81-3-5841-8349. E-mail: umezawa@chem.s.u-tokyo.ac.jp

ISSN 1061-0278 print/ISSN 1029-0478 online © 2002 Taylor & Francis Ltd
DOI: 10.1080/10610270290026185

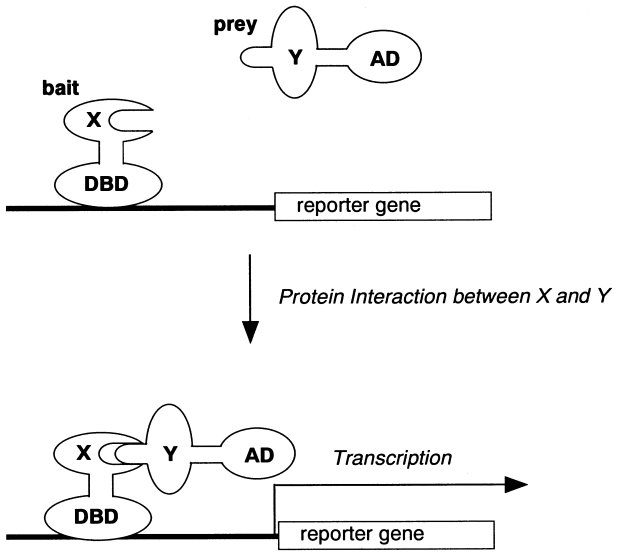

FIGURE 1 Principle of the yeast two-hybrid system. DBD-X is a chimeric protein consisting of a DNA-binding domain (DBD) fused to protein X, and AD-Y is a second chimeric protein consisting of a transcription activation domain (AD) and protein Y. Protein–protein interactions between X and Y brings AD into close proximity to the DNA-binding site that regulates the reporter gene and results in its transcriptional activation.

factors and other factors that have a limited, if not totally cell-specific, distribution. The transcription factors have two fundamental properties: First, eukaryotic transcription factors typically have a modular structure, with at lest two discrete domains, a DNA-binding domain and an activator domain. These two domains need not be present in the same polypeptide to give rise to an active transcription factor but just need to be in the vicinity of one another. Using this transcriptional machinery in eukaryotic cells, the yeast two-hybrid method pioneered the field of protein interactions, which enabled advances in the development of simple and automatic functional assays [1,2]. The two-hybrid screens require two chimeric proteins: A "bait" protein X connected to a DNA-binding domain is used to fish out a "prey" protein Y, which is connected to a transcription activation domain (Fig. 1). When the proteins interact, they bring together the DNA-binding domain and activation domain, which trigger the expression of a reporter gene. As the reporter gene, a *lacZ* gene, which encodes β-galactosidase, or a gene confers the ability to grow in selective media lacking specific amino acids, was used. If X interacts with Y in the yeast nucleus, activation domain will be in close proximity to the DNA-binding domain that regulates the reporter gene and results in its transcriptional activation. The activation brings to the production of a detectable protein such as β-galactosidase. This simple assay has been widely used to identify new interacting proteins, to map protein networks, and to address the functions of individual proteins and protein interactions.

The success of the yeast two-hybrid method has stimulated in a decade to develop further several approaches that facilitate non-destructive analysis of the association between two proteins. A bacterial two-hybrid system allows easy screening and selection of functional interactions between two proteins [3]. The principle is based on the reconstitution of a cAMP signal transduction pathway in a *Escherichia coli* (adenylate cyclase-deficient strain, Fig. 2). Adenylate cyclase consists of two complementary fragments. To these fragments, interacting polypeptides are genetically fused, respectively. Interaction between these two makes chimeric proteins resulted in functional complementation of the two fragments of adenylate cyclase to restore its enzymatic activity for the production of cAMP. This triggers the transcription of a reporter gene. A special feature of the bacterial two-hybrid system is that it is possible to analyze protein–protein interactions that occur either in the cytosol, at the inner membrane level, or on the DNA. In addition, the high efficiency of transformation that can be achieved in *E. coli* will allow screening of libraries of high complexity. These advantages rely on a signaling cascade that utilizes the diffusing regulatory molecule of cAMP.

In the same concept, it utilizes another regulatory molecule, where a split ubiquitin system exists for monitoring protein–protein interactions in mammalian cells [4]. Ubiquitin is a conserved protein of 76 amino acids, which is usually attached to the N terminus of intracellular proteins. The C- and N-terminal fragments of ubiquitin are connected with the proteins of interest (Fig. 3). The C-terminal fragment is expressed as a fusion with a reporter protein. When the interacting proteins bring the two parts of ubiquitin close enough, the reconstituted ubiquitin is recognized by ubiquitin-specific proteases (UBP), which cleave after the last residue of ubiquitin and release the reporter protein to activate a nuclear localized reporter. As a model system, specific interactions between Wbp1p and Ost1p, which are both subunits of oligosaccharyl transferase membrane proteins of the endoplasmic reticulum, are demonstrated. The interaction between Wbp1p and Ost1p induced cleaving and liberating the reporter protein into the nucleus, and thereby enabled to detect β-galactosidase activity. The success of split ubiquitin system led to a generally applicable screen for binding partners of specific proteins with a library of genomic DNA fragments [5], and detection of transient (short lived) interactions between a segment of a signal sequence-bearing proteins and Sec62p, a component of the translocation machinery on the endoplasmic reticulum [6].

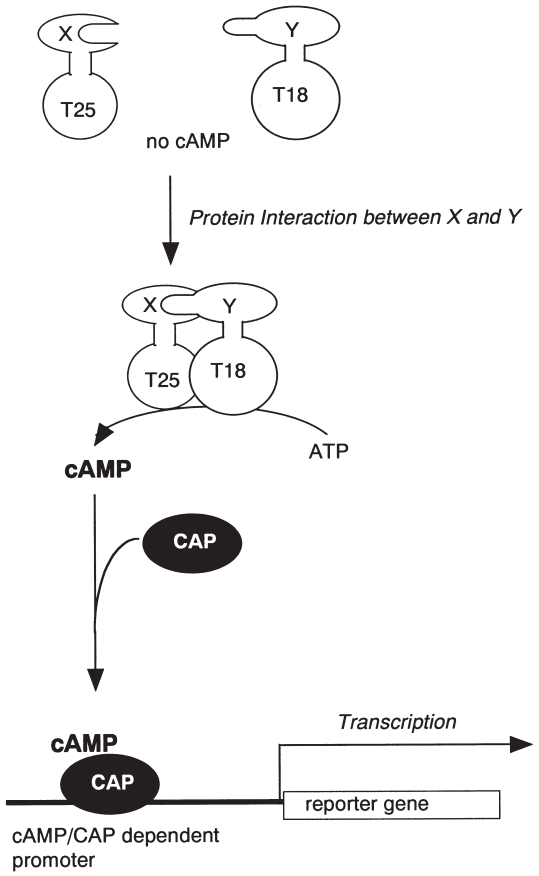

FIGURE 2 Principle of the bacterial two-hybrid system based on functional complementation of adenylate cyclase fragments. Two fragments, T25 and T18, represent amino acids 1–224 and 225–399 of the adenylate cyclase. The two fragments, fused to two interacting proteins, X and Y, are brought into close proximity, resulting in functional complementation followed by cAMP production. The cAMP binds to the catabolite gene activator protein CAP. The cAMP/CAP complex then can recognize specific promoters and switch on the transcription of the corresponding genes. These reporter genes can be either natural *E. coli* genes, such as *lacZ* or mal genes, or synthetic ones, such as antibiotic–resistance genes fused to a cAMP/CAP-dependent promoter.

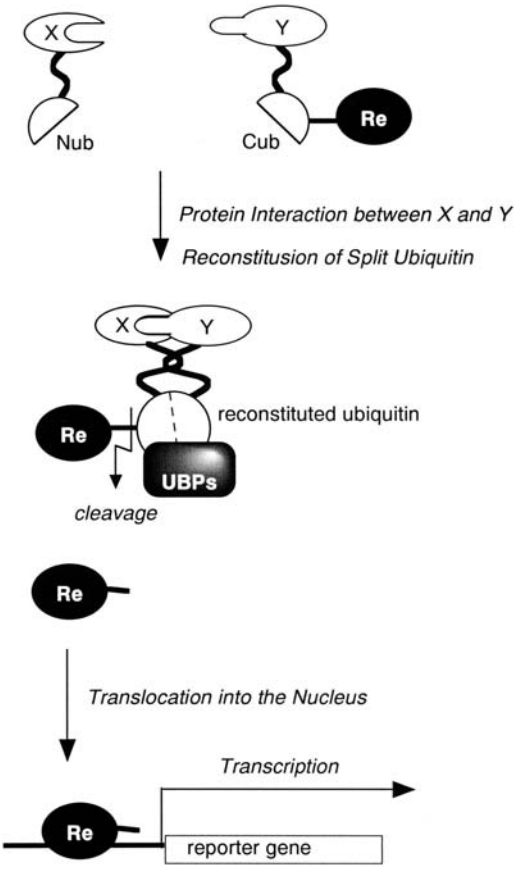

FIGURE 3 Principle of the split ubiquitin system. Proteins of interest, protein A and B, are linked to two ubiquitin fragments, Nub and Cub. Nub is an altered N-terminal fragment of ubiquitin that failed to reconstitute ubiquitin in the presence of Cub. Reporter protein (Re) is attached to the Cub by the peptide bond. Interaction between A and B results in formation of the split-ubiquitin heterodimer. The heterodimer is recognized and cleaved by ubiquitin specific proteases (UBP), liberating the reporter protein. The reporter protein can then enter the nucleus by diffusion and bind to its corresponding DNA-binding domain leading to activating transcription of a reporter gene.

DETECTION OF PROTEIN INTERACTIONS BASED ON INTRACELLULAR SIGNAL TRANSDUCTION

As mentioned above, the yeast two-hybrid system is an excellent research tool, and is commonly used to identify and characterize novel and known interaction partners for proteins of interest. The yeast two-hybrid system, however, exhibits several limitations and inherent problems, because the two-hybrid system is based on a transcriptional readout. The two-hybrid system cannot be used with transcription activators or proteins with transcription repression activity. In addition, a kind of proteins, although not considered transcription factors, exhibits intrinsic transcription activity and, therefore, cannot be used in this assay. Moreover, the fact that protein interactions underlying this method occur in the yeast nucleus may result in problems of toxicity due to nuclear expression of DNA-binding proteins and cell cycle regulators. Finally, the two-hybrid system, used for almost a decade with different baits, has generated data regarding interacting pairs of proteins as a result of a library screening approach. While numerous proteins may pass bait-specificity tests, eventually they turn out to be "false positives" after all, resulting in wasted effort and confusion.

To overcome some of the problems and limitations, a cytoplasmic protein recruitment system was developed [7–10]. The system takes advantage of on/off switching mechanism in signal transduction (Fig. 4). Generation of local high concentrations of a signaling intermediate causes a dramatic increase in signaling activity. In yeast, cell proliferation requires an activation of the Ras signaling pathway. The yeast Ras protein attached to an inner cell membrane takes a GDP-bound inactive form. A protein of interest, protein X, is fused in frame with a membrane localization signal. This signal was either

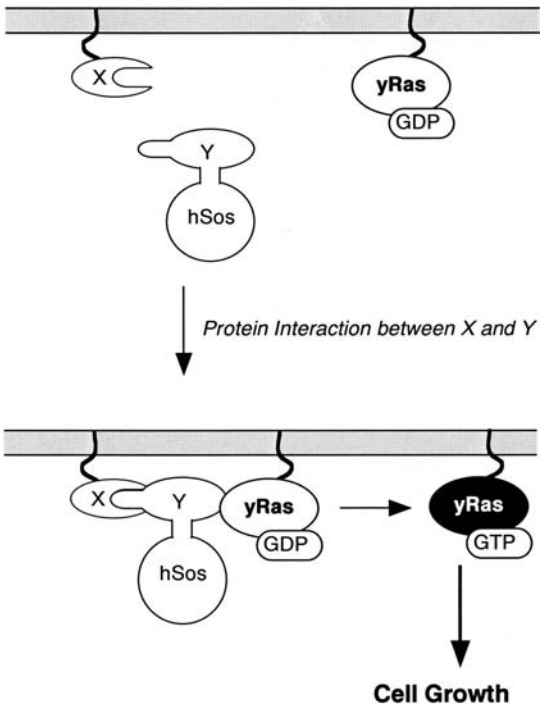

FIGURE 4 Schematic diagram describing the Sos recruitment system. The cDNA encoding an interacting protein X is fused to a membrane localization signal, such as a myristoylation sequence. Interacting partner, protein Y, is fused to hSos. The localization of hSos to the plasma membrane via protein–protein interactions results in exchange of GDP-bound yeast Ras for GTP-bound yeast Ras and thereby activation of the Ras viability pathway at the restrictive temperature conferring efficient cell growth.

a myristoylation or farnesylation sequence to be fused at the N- or C-terminal fragments, respectively. On the other hand, a protein partner, protein Y, is fused in frame with hSos. When protein–protein interactions occur between X and Y, the yeast Ras exchanges GDP for GTP and allows growth of the yeast cells. This system was shown to be suitable for identification and isolation of known and novel protein interactions. The approach has been improved by replacing the hSos effector molecule with Ras [11,12]. The both systems overcome some of the limitations and problems of the two-hybrid assay.

DETECTION OF PROTEIN INTERACTIONS BASED ON PROTEIN FRAGMENT COMPLEMENTATION ASSAY

Protein fragment complementation assay (PCA) is a more general approach to detect protein–protein interactions *in vivo*. The concept is quite simple: The gene coding an enzyme is rationally dissected into two pieces. Each fragment is connected with two test proteins that are thought to bind to each other. Refolding of the dissected enzyme from its fragments is initiated by the binding of the test proteins to each other, and is detected as recovering of the enzyme activity. Up to now, two practical applications for PCAs were demonstrated; dihydrofolate reductase (DHFR) [13–17] and β-galactosidase-fragment complementation assays [18,19].

Dihydrofolate reductase is central to cellular one-carbon metabolisms in both prokaryotes and eukaryotes and is absolutely required for their cell survival. It catalyzes the reduction of dihydrofolate to tetrahydrofolate for use in transfer of one-carbon units required for biosynthesis of serine, methionine, pantothenate (in prokaryotes), purines and thymidylate. To construct PCA with DHFR, it is split into two fragments with several mutations. The mutations render split DHFR, incapable of spontaneous assembly and folding from its fragments (Fig. 5). The complementary fragments of DHFR, when fused to interacting proteins and expressed in the DHFR-negative cell, folds and reassembles into the complete three-dimensional structure of the enzyme. Reconstitution of DHFR can be monitored by surviving cells grown in the medium in absence of nucleotides [13,14,20]. Only cells containing a pair of interacting proteins can undergo normal cell division and colony formation. The second approach is a fluorescence assay based on the detection of fluorescein-conjugated methotrexate (fMTX) binding to reconstituted DHFR. The reconstituted enzyme is able to bind fMTX with a high-affinity ($K_d = 540$ pM) in a 1:1 complex. Fluorescein-conjugated methotrexate is retained in cells by this complexation, while the unbound fMTX is actively and rapidly transported out of the cells. In addition, binding of fMTX to DHFR results in a 4.5-fold increase in the fluorescence quantum yield. Bound fMTX to complementary DHFR can then be monitored by fluorescence microscopy, fluorescence-activated cell sorting (FACS), or spectroscopy [15–17].

The first test system for the mammalian DHFR PCA was the pharmacologically well-characterized rapamycin-induced association of FK506-binding protein (FKBP) to its target, the FKBP-rapamycin-binding (FRB) domain of the FKBP12-rapamycin-associating protein (FRAP) [14,20]. The DHFR negative cells stably cotransfected with FRB and FKBP fused to one of the two DHFR complementary fragments, were selected for survival in nucleotide-free medium and in the presence of rapamycin. Only cells grown in the presence of rapamycin underwent normal cell division and colony formation. Formation of the FKBP-rapamycin-FRB complex was also detected with the fluorescence assay described above, based on stoichiometric binding of fluorescein-conjugated methotrexate to reconstituted DHFR. Fluorescence microscopy of unfixed cotransfected cells that had been incubated with fMTX showed high levels of fluorescence when cells were treated with rapamycin at its saturated concentration. The fluorescent response of cell population was quantified

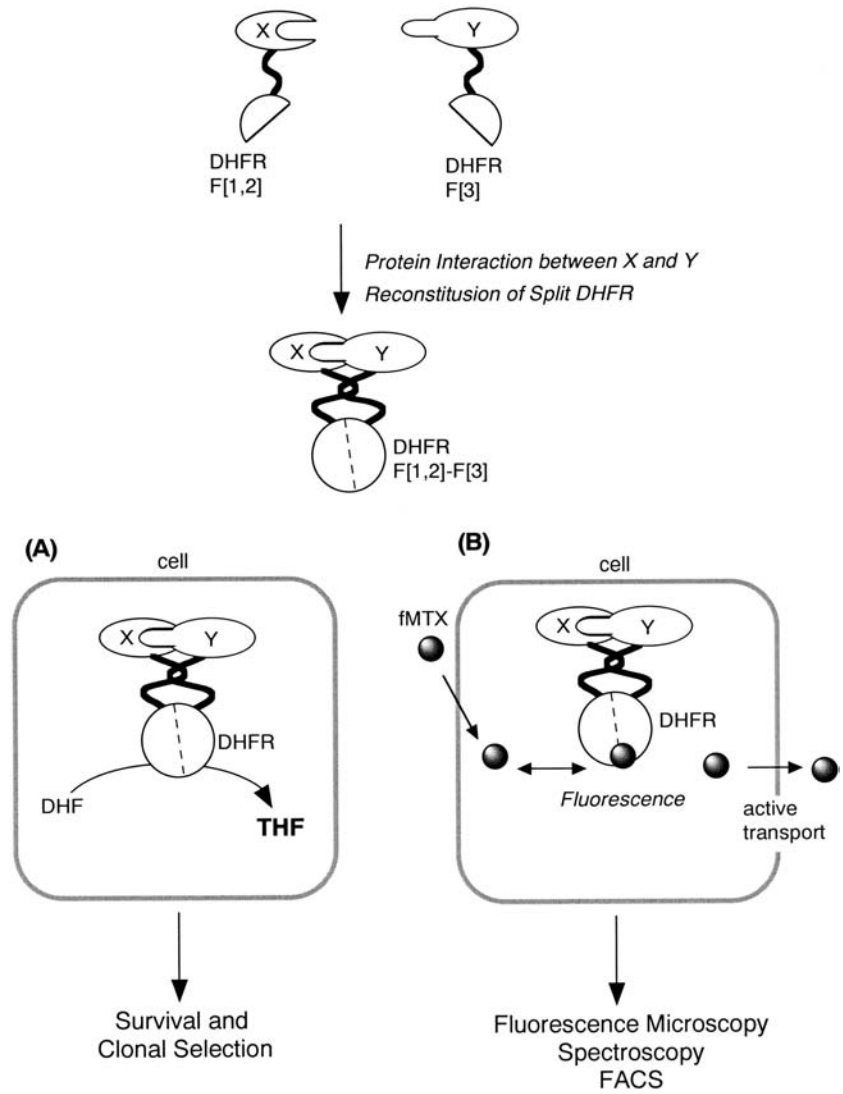

FIGURE 5 Principle of the DHFR PCA based on functional complementation of DHFR fragments. Interacting proteins, X and Y, are fused to one of two complementary fragments of murine DHFR (F[1,2] and F[3]). Association of X and Y drives the reconstitution of DHFR (F[1,2]-F[3]). (A) The reconstituted DHFR catalyzes the reaction of dihydrofolate to tetrahydrofolate required for biosynthesis of serine, methionine, purine and thymidylate, allowing DHFR-negative cells to grow in medium lacking nucleotides. (B) The fluorescence assay is based on high-affinity binding of the specific DHFR inhibitor fluorescein-methotrexate (fMTX) to reconstituted DHFR. Fluorescein-methotrexate passively crosses the cell membrane and binds to reconstituted DHFR and is thus retained in the cell. Unbound fMTX is rapidly exported from the cells by active transport. Bound and retained fMTX can then be detected by fluorescence microscopy, FACS, or fluorescence spectroscopy.

by FACS. Several applications of DHFR PCA strategy have also been demonstrated like quantitative detection of protein–protein interactions among protein pairs and of allosteric transitions in membrane-associated receptors in living mammalian cells [15], qualitative detection of the interactions in plant cells [17], and mapping a signal transduction network that controls initiation of translation in eukaryotes [16].

The second approach of PCA strategy is a β-galactosidase-fragment complementation assay. The β-galactosidase is composed of intracistronic complementation of α and ω domains. Pairs of inactive β-galactosidase deletion mutants are capable of complementing one another in trans and assembled to form an active enzyme. When two different, weakly complementing deletion mutants of β-galactosidase, Δα and Δω, are fused to two interacting proteins and expressed in a cell, the interaction of the non-β-galactosidase portions of the chimeric proteins drives β-galactosidase complementation, and the resulting β-galactosidase activity serves as a measure of that interaction (Fig. 6). In order to detect the β-galactosidase activity, Flour X-Gal was used, that consists of an azo dye, Fast Red Violet LB, with either X-Gal or with 5-bromo-6-chloro-3-indolyl-β-D-galactopyranoside (5–6 X-Gal). A practical application of the β-galactosidase PCA was demonstrated by an well-characterized interaction between the FKBP12 and the FKBP12-rapamycin-binding domain (FRB domain) of the

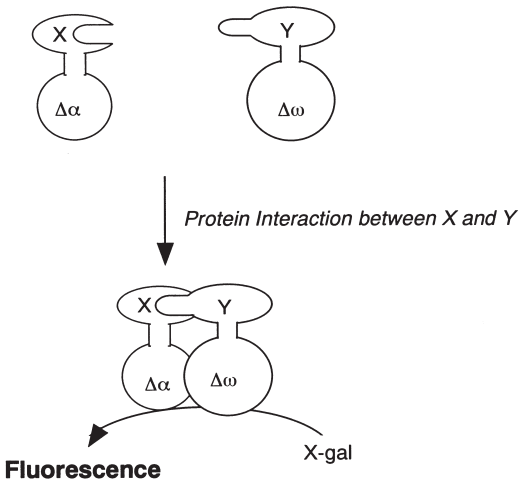

FIGURE 6 Schematic representation of β-gal PCA. When the Δα and Δω β-Gal mutants are fused to proteins of interest, X and Y, their association induces activation of the enzyme. The activity is measured by a chemiluminescence or fluorescence substrates.

FKBP12-rapamycin-associated protein (FRAP) [18], and dimerization of epidermal growth factor receptors in the cell surface [19].

DETECTION OF PROTEIN INTERACTIONS BASED ON PROTEIN SPLICING

Protein Splicing as Useful Engineering Tools

Protein biosynthesis was initially thought to be a simple process in which the genetic information in DNA was directly copied into messenger RNAs, which in turn directed the biosynthesis of proteins. But an unexpected discovery was made by two groups independently in 1990 that in *Saccharomyces cerevisiae*, a nascent 120 kDa translational product of VMA1 gene autocatalytically excised out a 50 kDa site-specific endonuclease (VMA1-derived endonuclease; VDE, also called PI-*Sce* I) and spliced the two external polypeptides to form a 70 kDa catalytic subunit of vacuolar H^+-ATPase [21,22]. This discovery led to conclude that posttranslational removal of polypeptide segments can occur by protein splicing (Fig. 7). The protein splicing is a multi-step processing event involving precise excision of an internal protein segment (intein), from a primary translation product with concomitant ligation of the flanking sequences (extein) [23,24]. Since the initial discovery of the VMA1 intein, nearly 100 putative inteins have been identified in eubacteria, archea and eukaryotic unicellular organisms (InBase, the Intein registry Web site at http://www.neb.com/neb/inteins.html; New England Biolabs, Beverly, MA). These inteins range in size from 134 to 608 amino acid residues, of which more than 20 have been shown to undergo protein splicing. A common important feature of the protein splicing is a

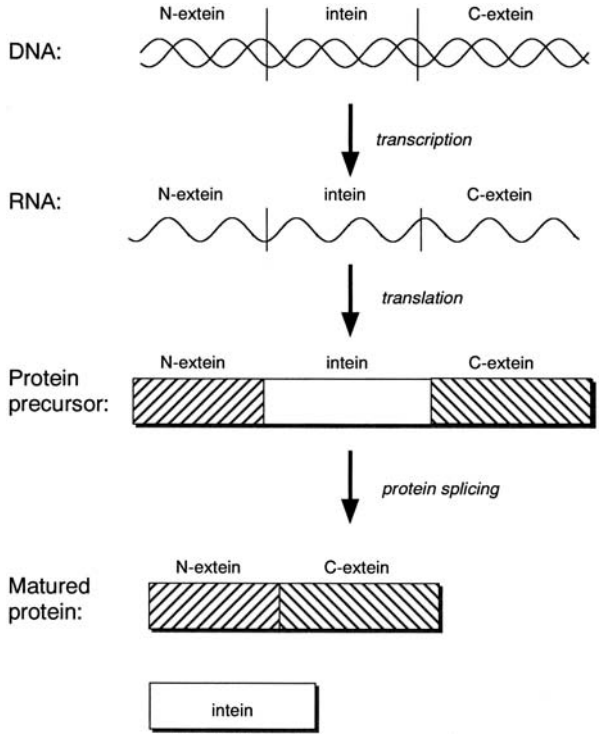

FIGURE 7 Posttranslational modification by protein splicing. Genetic information in DNA is directly copied into messenger RNAs (transcription), which in turn directs the biosynthesis of proteins (translation). A specific type of intervening sequence, termed an intein, is excised from an internal site in a precursor protein and the surrounding polypeptide (exteins) are ligated to form the matured protein.

self-catalyzed excision of the intein and ligation of concomitant exteins without any external enzymes. This special feature has led to develop a number of applications, including incorporation of synthetic peptides containing fluorescence and phosphorylation [25–27], self-cleaving affinity tags for protein purification [28], a novel polypeptide ligation system for protein semisynthesis [29–32], segmental labeling of proteins for NMR analysis [33,34], fluorescent biosensors [35], and protein–protein interactions [36–38].

Principle of the Detection Methods Using Protein Splicing

The principle for detecting protein–protein interactions is based on reconstitution of a functional protein by the protein splicing reaction. As to the functional protein, split-enhanced green fluorescent protein (EGFP) was chosen. The structure of EGFP is composed of eleven strands of β-sheet that form an anti-parallel barrel with short α-helixes forming lids on each end [39]. The fluorescent active center of EGFP is located inside the barrel. In using EGFP, we were initially not sure whether split EGFPs really connect to each other by protein splicing and the ligated EGFP reconstitutes the barrel structure to

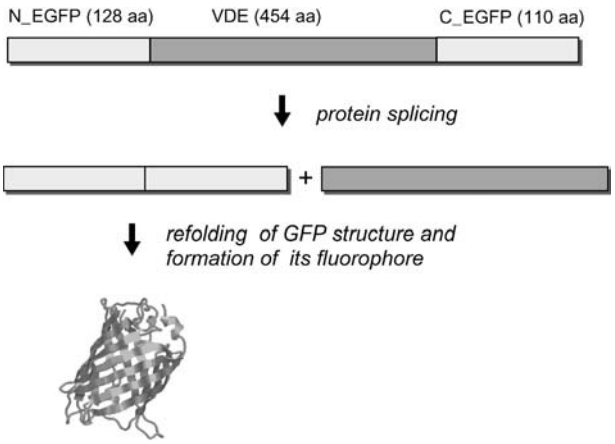

FIGURE 8 Reconstitution of split EGFP by protein splicing. A single polypeptide, that is composed of 128 from the N-terminal half of EGFP, 454 residues from VDE and 110 residues from the C-terminal half of EGFP, undergoes protein splicing and thereby N- and C-terminal fragments of EGFP ligate by a peptide bond. The matured EGFP thus formed folds correctly and its fluorophore is formed inside the barrel.

form its fluorophore. To ascertain this, we performed the following pilot experiment before demonstrating the utilization of protein splicing as a tool for detecting protein–protein interactions. The EGFP was dissected at the position between 128 and 129, that locates at the end of the sixth β-sheet strand of EGFP, and VDE was inserted into the position. A single polypeptide encoding the VDE intervening the N- and C-terminal halves of EGFP was expressed in *E. coli* and analyzed its splicing products (Fig. 8). The results revealed that the two external regions of the N- and C-terminal halves of EGFP were ligated with a peptide bond by protein splicing and the ligated EGFP folded correctly to form its fluorophore [38].

To explore this finding further, the single polypeptide was dissected into two functional splicing parts. This dissection was achieved by deletion of a functionally unrelated endonuclease motif from a 185–389 amino acid region. The resulting peptide fragments, termed as optical probes, are linked to a protein of interest (protein A) and its target protein (protein B) (Fig. 9). When an 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 the splicing and thereby N-and C-terminal fragments of EGFP directly link to each other by a peptide bond. This reconstitution of EGFP is monitored by its fluorescence at 510 nm. The intensity of the fluorescence is proportional to the number of interacting protein pairs. In our proof of this principle, we attached calmodulin (CaM) and its target peptide, known as M13 derived from skeletal muscle myosine light-chain kinase. Upon coexpression of CaM and M13 connected with the optical probes in *E. coli*, their interaction induced protein splicing, by which split EGFP underwent correct reconstitution and its fluorescence was monitored.

With same concept as the EGFP system, we have recently developed a split luciferase system for detecting protein–protein interactions in mammalian cells [36]. Firefly luciferase is known to be folded into two compact domains, one is the large N-terminal globular domain, while the C-terminal portion of the enzyme is separated from the N-terminal by a wide cleft, which is the location of the active site of the enzyme (Fig. 10) [40]. This luciferase is split into the N- and C-terminal fragments. As for the intein, a dnaE intein from *Synechosystis* is used that possesses an ability to ligate

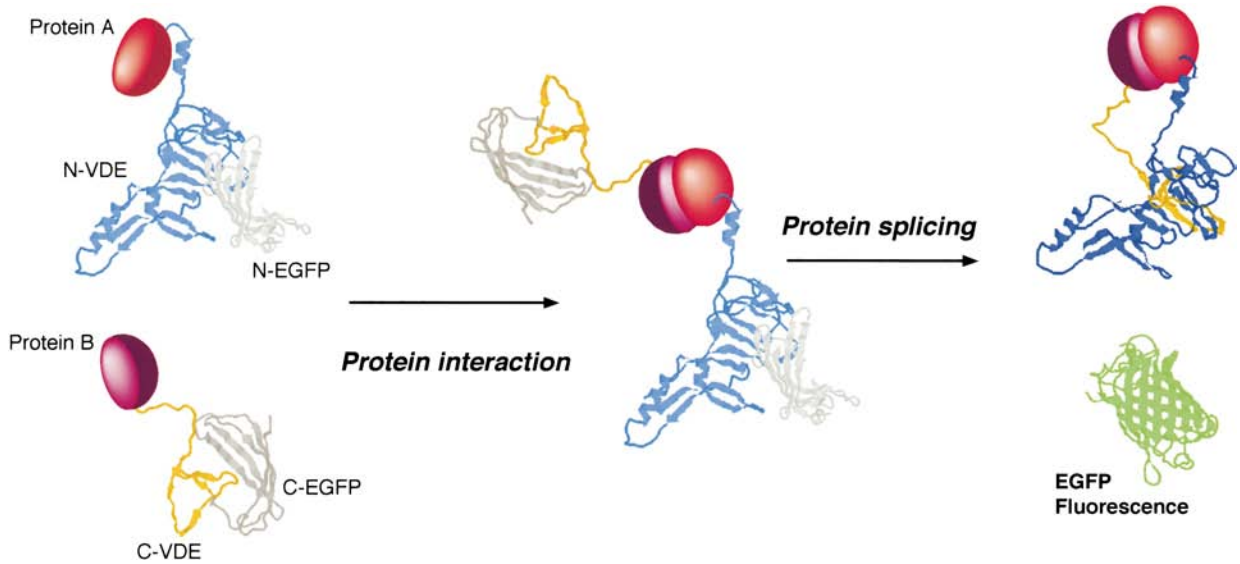

FIGURE 9 Proteins of which interactions are being monitored are attached to the N- (blue) and C- (orange) terminal portions of VDE and split EGFP (gray). When the proteins interact, the two portions of VDE come close enough to fold together and initiate protein splicing. The two halves of EGFP are ligated and released.

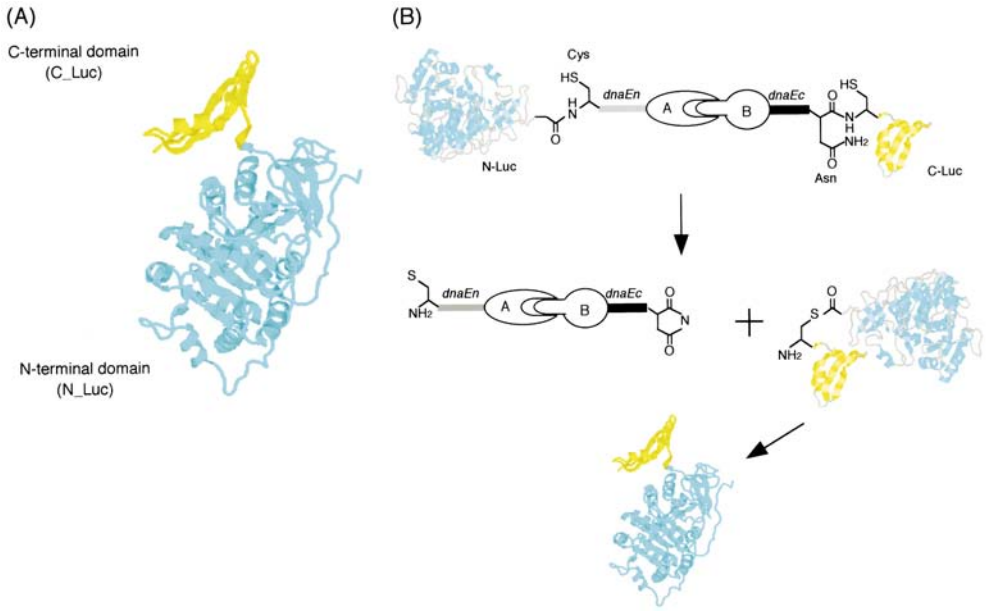

FIGURE 10 (A) 3D structure of firefly luciferase. N- (cyan, 1–437 amino acids) and C- (yellow, 438–544 amino acids) terminal halves of luciferase are shown as cyan and yellow, respectively. (B) Principle of the split-luciferase system. N- and C-dnaE are connected to the N-and C-terminal halves of luciferase, respectively. Partner proteins A and B are linked to opposite ends of those dnaE. Interactions between the two proteins accelerate the folding of N- and C-dnaE and protein splicing occurs. The N- and C-terminal halves of luciferase are linked together by a peptide bond to recover its bioluminescent activity.

N- and C-exteins [41]. Each N- and C-dnaE intein is connected with N- and C-terminal halves of the firefly luciferase, respectively. The opposite ends of dnaE thus formed are further bonded respectively with a pair of proteins of interest and expressed in mammalian cells. Upon interactions between the two proteins, the two dnaE fragments are brought close enough to fold together and initiate splicing and linking of the two halves of luciferase with a peptide bond. Reconstitution of luciferase is monitored by its bioluminescence, of which intensity is again proportional to the number of interacting protein pairs.

Applications of the Protein Splicing Systems

The split luciferase system is easily adapted to work in mammalian cells. The most straightforward and immediate applications of this method are likely to be in high-throughput screening of chemical compounds that can either block or induce a particular protein–protein interaction. For example, in the physiologically relevant insulin signaling, active insulin receptor phosphorylates its substrate, IRS-1, which interacts with phosphatidyl inositol 3-kinase (PI3-kinase). The peptide fragment of IRS-1 (Y941 peptide) and the N-terminal SH2 domain (SH2N) of PI3-kinase were connected respectively to the luciferase-based flanking protein fragments. These molecules were expressed in the Chinese hamster ovary cells overexpressing human insulin receptor (CHO-HIR) cells; stimulation of the cells with insulin caused intracellular protein interactions between Y941 peptide and SH2N in the cells.

The interaction induced protein splicing and thereby matured luciferase thus formed recovered its bioluminescent activity. This system allows a selective and quantitative detection *in vivo* of insulin-stimulated protein phosphorylation and protein–protein interactions in the insulin signaling pathways. Using this system, screening and targeting of agonist-like drugs for the insulin-signaling pathway are possible for hundreds or thousands of pharmaceutical candidates. Recently, a new lead compound has been identified as a possible insulin mimetic that directly increases insulin receptor kinase activity [42]. If thousands of the pharmaceutical samples are derived from the lead compound, they can be screened out to find a better pharmacological effect.

In the split-EGFP system, there exist several advantages; (i) substrates of enzymes are not needed, (ii) EGFP accumulates in a target cell until it degrades and information of the interaction is thereby integrated in the cell, (iii) the interaction can be monitored on the cell membrane or adhesion of cell membranes, because this system is irrelevant to reporter genes. On the basis of these advantages, we expect wider applications such as a bacterial screening and selection system (Fig. 11) [43]. Several bacterial one-and two-hybrid systems have been proposed, in which there is a common principle that when the proteins interact, they trigger a transcriptional activation of a reporter gene and produce a signal protein that is accumulated in the bacteria.

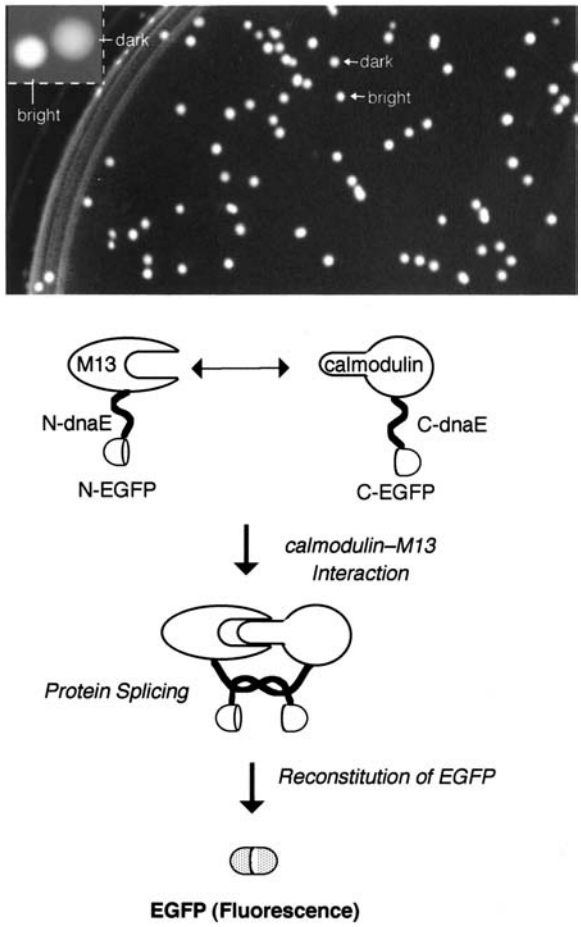

FIGURE 11 Fluorescent images of bacterial colonies on Luria Broth (LB) agar plates. As interaction partner, calmodulin and its interacting peptide, M13, were used. Each of the bacterial colonies includes probe molecules including split EGFPs, among which around 50% of the colonies contain calmodulin and M13 directly linked to the optical probes. The half of the colonies in the presence of the interacting pairs showed strong fluorescence and the rest of the colonies fluoresced weakly (inset: an expanded image of bright and dark colonies of the bacteria). The strong fluorescence demonstrates that the calmodulin-M13 interaction induces protein splicing to produce mature EGFP.

Unlike these earlier protein interaction assays, the split-EGFP system involves the reconstitution of EGFP, and does not require that the protein–protein interactions take place near the cell nucleus and reporter gene or that an enzyme substrate be present. This will make the method more generally useful and allow the interactions be screened in the cytosol or at the inner-membrane level. Moreover, faster growth rate of bacteria than eukaryotic cells is a potentially significant advantage for screening times. The selection of bacterial clones also permits a single-step isolation of the candidates in an *in vivo* context. This advantage helps to greatly facilitate the screening speed and identification of the interaction partner.

Other applications that we can think of are visualization of the interaction in eukaryotic cells and mapping protein interactions *in vivo* in transgenic animals. The fusion point between test proteins and dnaE connected with split EGFP is possible to be reconstituted *in vivo* by knock-in in frame with the endogenous gene. Animals carrying both transgenes should develop parent EGFP only in those cells in which the endogenous proteins are expressed and actually interacted; histochemical detections of EGFP therein is expected to extend these finding in animal models by allowing the mapping of specific protein interactions in single cells both during embryogenesis and in specific disease states.

CONCLUSION

To date, the yeast two-hybrid system and its related techniques have provided predominant *in vivo* methods to study protein–protein interactions. The systems are suitable for automation and high-throughput screening to map the interactions in a living cell, but the prey-bait pairs are likely to be limited to constitutive protein interactions. Other methods to detect protein–protein interactions have been developed over the past several years, many of which are based on intracellular signal transduction, complementation of protein fragments, or protein splicing. With these systems, protein interactions on a specific cellular compartment, organelle or membrane surface, can be detected in living mammalian cells, which enabled to study ligand-induced interactions between two proteins. These systems complement the limitations and inherent problems of the conventional two-hybrid approach and may cover a wide range of protein–protein interactions to be explored. Although the protein splicing based systems are still in the early stages of their developments, they have the potential advantages in the nature of the readouts and the experimental flexibility. Improvements will certainly be forthcoming as well as a deeper understanding of the underlying molecular mechanisms.

Acknowledgements

This work was financially supported by Core Research for Evolutional Science and Technology (CREST) of Japan Science and Technology. (JST).

References

[1] Fields, S. and Song, O. (1989), *Nature* **340**, 245–246.
[2] Chien, C.T., Bartel, P.L., Sternglanz, R. and Fields, S. (1991), *Proc. Natl Acad. Sci. USA* **88**, 9578–9582.
[3] Karimova, G., Pidoux, J., Ullmann, A. and Ladant, D. (1998), *Proc. Natl Acad. Sci. USA* **95**, 5752–5756.
[4] Stagljar, I., Korostensky, C., Johnsson, N. and Heesen, S.T. (1998), *Proc. Natl Acad. Sci. USA* **95**, 5187–5192.

[5] Laser, H., Bongards, C., Schuller, J., Heck, S., Johnsson, N. and Lehming, N. (2000), *Proc. Natl Acad. Sci. USA* **97**, 13732–13737.
[6] Dunnwald, M., Varshavsky, A. and Johnsson, N. (1999), *Mol. Biol. Cell* **10**, 329–344.
[7] Huang, W., Wang, S.L., Lozano, G. and Crombrugghe, B. (2001), *Biotechniques* **30**, 94–100.
[8] Aronheim, A. (1997), *Nucleic Acids Res.* **25**, 3373–3374.
[9] Aronheim, A., Zandi, E., Hennemann, H., Elledge, S.J. and Karin, M. (1997), *Mol. Cell Biol*. **17**, 3094–3102.
[10] Yu, X., Wu, L.C., Bowcock, A.M., Aronheim, A. and Baer, R. (1998), *J. Biol. Chem.* **273**, 25388–25392.
[11] Aronheim, A., Broder, Y.C., Cohen, A., Fritsch, A., Belisle, B. and Abo, A. (1998), *Curr. Biol.* **8**, 1125–1128.
[12] Broder, Y.C., Katz, S. and Aronheim, A. (1998), *Curr. Biol.* **8**, 1121–1124.
[13] Pelletier, J.N., Arndt, K.M., Pluckthun, A. and Michnick, S.W. (1999), *Nat. Biotechnol.* **17**, 683–690.
[14] Remy, I. and Michnick, S.W. (1999), *Proc. Natl Acad. Sci. USA* **96**, 5394–5399.
[15] Remy, I., Wilson, I.A. and Michnick, S.W. (1999), *Science* **283**, 990–993.
[16] Remy, I. and Michnick, S.W. (2001), *Proc. Natl Acad. Sci. USA* **98**, 7678–7683.
[17] Subramaniam, R., Desveaux, D., Spickler, C., Michnick, S.W. and Brisson, N. (2001), *Nat. Biotechnol.* **19**, 769–772.
[18] Rossi, F., Charlton, C.A. and Blau, H.M. (1997), *Proc. Natl Acad. Sci. USA* **94**, 8405–8410.
[19] Blakely, B.T., Rossi, F.M., Tillotson, B., Palmer, M., Estelles, A. and Blau, H.M. (2000), *Nat. Biotechnol.* **18**, 218–222.
[20] Pelletier, J.N., Campbell-Valois, F.-X. and Michnick, S.W. (1998), *Proc. Natl Acad. Sci. USA* **95**, 12141–12146.
[21] Hirata, R., Ohsumi, A., Nakano, H., Kawasaki, K., Suzuki, K. and Anraku, Y. (1990), *J. Biol. Chem.* **265**, 6726–6733.
[22] Kane, P.M., Yamashiro, C.T., Wolczyk, D.F., Neff, N., Goebl, M. and Stevens, T.H. (1990), *Science* **250**, 651–657.
[23] Noren, C.J., Wang, J. and Perler, F.B. (2000), *Angew. Chem. Int. Ed.* **39**, 450–466.
[24] Paulus, H. (2000), *Annu. Rev. Biochem.* **69**, 447–496.
[25] Evans, T.C., Martin, D., Kolly, R., Panne, D., Sun, L., Ghosh, I., Chen, L., Benner, J., Liu, X.-Q. and Xu, M.-Q. (2000), *J. Biol. Chem.* **275**, 9091–9094.
[26] Cotton, G.J., Ayers, B., Xu, R. and Muir, T.W. (1999), *J. Am. Chem. Soc.* **121**, 1100–1101.
[27] Severinov, K. and Muir, T.W. (1998), *J. Biol. Chem.* **273**, 16205–16209.
[28] Chong, S., Mersha, F.B., Comb, D.G., Scott, M.E., Landry, D., Vence, L.M., Perler, F.B., Benner, J., Ducera, R.B., Hirvonen, C.A., Pelletier, J.J., Paulus, H. and Xu, M.-Q. (1997), *Gene* **192**, 271–281.
[29] Iwai, H. and Pluckthun, A. (1999), *FEBS Lett.* **459**, 166–172.
[30] Evans, T.C.J., Benner, J. and Xu, M.-Q. (1999), *J. Biol. Chem.* **274**, 18359–18363.
[31] Camarero, J.A. and Muir, T.W. (1999), *J. Am. Chem. Soc.* **121**, 5597–5598.
[32] Scott, C.P., Abel-Santos, E., Wall, M., Wahnon, D.C. and Benkovic, S.J. (1999), *Proc. Natl Acad. Sci. USA* **96**, 13638–13643.
[33] Xu, R., Ayers, B., Cowburn, D. and Muir, T.W. (1999), *Proc. Natl Acad. Sci. USA* **96**, 388–393.
[34] Otomo, T., Ito, N., Kyogoku, Y. and Yamazaki, T. (1999), *Biochemistry* **38**, 16040–16044.
[35] Muir, T.W., Sondhi, D. and Cole, P.A. (1998), *Proc. Natl Acad. Sci. USA* **1998**, 6705–6710.
[36] Ozawa, T., Kaihara, A., Sato, M., Tachihara, K. and Umezawa, Y. (2001), *Anal. Chem.* **73**, 2516–2521.
[37] Ozawa, T. and Umezawa, Y. (2001), *Curr. Opin. Chem. Biol.* **5**, 1–7.
[38] Ozawa, T., Nogami, S., Sato, M., Ohya, Y. and Umezawa, Y. (2000), *Anal. Chem.* **72**, 5151–5157.
[39] Ormo, M., Cubitt, A., Kallio, K., Gross, L., Tsien, R. and Remington, S. (1996), *Science* **273**, 1392–1395.
[40] Conti, E., Franks, N.P. and Brick, P. (1996), *Structure* **4**, 287–298.
[41] Wu, H., Hu, Z. and Liu, X.-Q. (1998), *Proc. Natl Acad. Sci. USA* **95**, 9226–9231.
[42] Zhang, B., Salituro, G., Szalkowski, D., Li, Z., Zhang, Y., Royo, I., Vilella, D., Diez, M.T., Pelaez, F., Ruby, C., Kendall, R.L., Mao, X., Griffin, P., Calaycay, J., Zierath, J.R., Heck, J.V., Smith, R.G. and Moller, D.E. (1999), *Science* **284**, 974–977.
[43] Ozawa, T., Takeuchi, M., Kaihara, A., Sato, M., Umezawa, Y., (2001), *Anal. Chem.* **73**, 5866–5874