BioChip J. (2016) 10(4): 277-287
DOI 10.1007/s13206-016-0404-4

Review Article

Intein-mediated Protein Engineering for Biosensor Fabrication

Hyunjin Jeon[1], Minhyeong Lee[1], Wonhee Jang[2] & Youngeun Kwon[1,*]

Received: 14 October, 2016 / Accepted: 08 November, 2016 / Published online: 06 December, 2016
© The Korean BioChip Society and Springer 2016

Abstract Intein-mediated protein engineering has become a valuable tool for biosensing applications. Expressed protein ligation and protein trans-splicing were used to generate proteins with necessary tags or cyclic peptides that can be used to build various biosensing platforms such as biochips or sensor proteins. Especially the use of split-inteins that carry out conditional protein splicing reaction enabled the fabrication of various genetically-encoded biosensors that can monitor various signaling events *in vivo* including protein-protein interactions and protein translocalization. Biological targets activated split-inteins to generate functioning reporter molecules, such as luciferase and autofluorescent proteins. We here review various sensing platforms that utilize intein-mediated protein engineering technology mainly focusing on cell-based biosensors.

Keywords: Inteins, Biosensors, Cell-based sensors, Protein engineering, Protein trans-splicing

Introduction

Biosensors often refer to molecular sensors that can detect and monitor the interactions of biological molecules[1], such as proteins[2], DNAs[3], and small molecules[4]. Biosensors became a valuable tool not only for fundamental studies of biology but also for various applications including medical and environmental screenings. To widen the usage of biosensors, it is desirable to make biosensors with quick response time, high sensitivity, and selectivity. As such, many efforts have been paid to develop suitable *in vitro*[5-8] and/or *in vivo*[8-10] sensing platforms. Various platforms, such as chips, microorganisms, living cells, and animals, are currently utilized for fabricating biosensors. Among these, living cells offer unique opportunities as they are cheap and self-sustainable platforms which provide more biologically relevant information.

Sensors are comprised of a target-sensitive recognition unit, a transducer that changes stimuli to readable outputs, and mechanical signal amplification or processing system. Biosensors can be categorized based on the composition of target-recognition units such as molecular sensors, cellular sensors, and tissue sensors. Molecular sensors often utilize purified biomolecules, such as proteins and DNAs, as a recognition unit and are designed to detect target molecules *in vitro* with high specificity. On the other hand, cellular biosensors and tissue biosensors enable monitoring of the functions and interactions of target molecules in native environment where a plethora of biological molecules co-exist including biomolecules of similar functions.

In fabrication of cell-based biosensors, the native receptors or enzymes are often exploited, directly or indirectly, as the bio-recognition component. In order to detect the stimuli exerted to mammalian cells by the target, various electronic methods are developed such as measures of cell metabolism, impedance, intra- and extra-cellular potentials. Genetic modifications also provide useful methods, for example, by activating reporter proteins such as autofluorescent proteins (AFP) or bioluminescent proteins, as a response to external stimuli. These genetically-encoded cell-based sensors provide a suitable platform for high throughput screen-

[1]Department of Biomedical Engineering, Dongguk University, Pildong 3-ga, Seoul 04620, Republic of Korea
[2]Department of Chemical Biological Engineering, Dongguk University, Pildong 3-ga, Seoul 04620, Republic of Korea
*Correspondence and requests for materials should be addressed to Y. Kwon (✉ykwon@dongguk.edu)

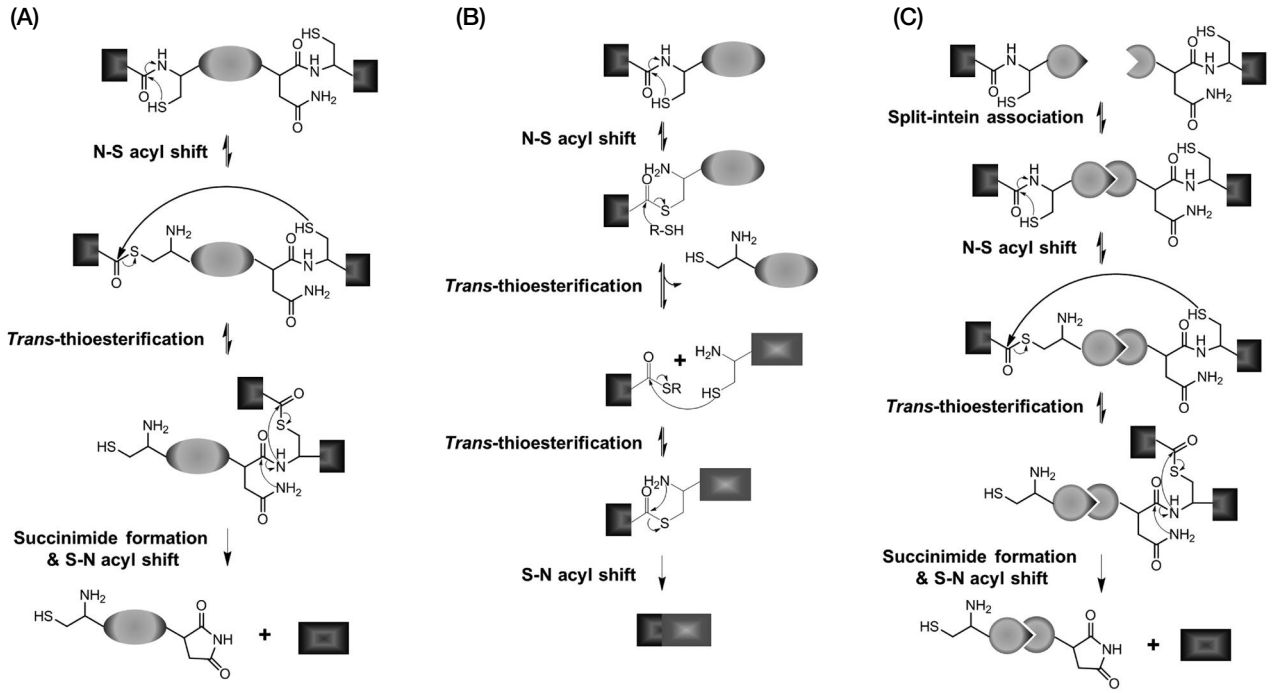

Figure 1. Curved arrow mechanisms of intein-mediated protein splicing (A), expressed protein ligation (EPL) (B), and protein trans-splicing (PTS) (C) reactions. ■, $_N$Extein; ■, $_C$Extein; ⬭, Intein; ▬, N-cys peptide; ◔, I_N; and ◔, I_C.

ing with minimal interferences as they allow non-invasive measurements of various stimuli. A self-processing protein, intein, and the intein-based protein engineering have been playing key roles in fabricating genetically-encoded cell-based sensors and provided solutions for many technical hurdles[11].

In this review, we discuss how inteins have been used to design various sensing platforms especially cell-based sensors. There are several advantages of intein-mediated protein engineering. First, inteins mediate highly specific splicing reaction in physiological conditions without a need for external source of energy or cofactors[12,13]. Second, intein-mediated protein engineering allows modular approaches to enable easy exchanges of sensing and reporting elements[14,15]. Third, intein-mediated protein splicing and protein trans-splicing (PTS) both generate seamless target proteins conjugated by a stable amide bond in order to support fabrication of biosensors with high target sensitivity and bio-stability[12].

Inteins and Intein-mediated Protein Engineering

Inteins and the Mechanism of Protein-splicing Reaction

Inteins are self-processing proteins that conjugate two concomitant flanking sequences (exteins, attached to N- and C-termini of an intein) by an amide bond to make a seamless mature protein (Figure 1A)[16]. Intein itself is excised off from the host protein during the process and the whole process is called as an intein-mediated protein-splicing reaction[17]. Protein-splicing reactions are similar to intron-splicing but do not require other proteins or an external source of energy[12,13]. Inteins offer rare opportunities to modify proteins of interest (POI) through the formation of covalent amide-linkage and have emerged as a valuable tool for protein engineering in diverse biotechnological applications[18]. The mechanism of intein-mediated protein-splicing reaction is comprised of four nucleophilic displacement reactions led by functional groups on intein and the first Cys residue of C-extein. The inteins fold to bring together the splice junctions, provide nucleophiles, and assist the reaction by making the amide bond, between the last residue of N-extein and the first residue of intein, more scissile[19].

There are hundreds of inteins to choose from depending on the purpose of applications (Table 1). Subtyping of contiguous- and split-inteins is one way of categorizing inteins. While contiguous inteins are more common, there also are split-inteins which exist as two separate polypeptides, namely, the N-intein and C-intein. N- and C-inteins are not active individually but

Table 1. Numerous inteins of various characteristics are available for intein-mediated protein engineering and biosensing applications.

Intein	Type	Temp. (°C)	k_{splice} (s^{-1})	$t_{1/2}$	% Yield	Note	Ref.
Mxe GyrA	Contiguous	25	1.9×10^{-5}	10 h	>90		[20]
Pab PolIII	Contiguous	70	1.6×10^{-5}	12 h	74		[21]
Mtu SufB	Contiguous					Redox-sensitivity	[22]
Ssp DnaE	Naturally split	37	1.5×10^{-4}	76 m	<50		[23]
Npu DnaE	Naturally split	37	3.7×10^{-2}	19 s	>90	Fast reacting	[23]
Ssp DnaB	Artificially split	25	9.9×10^{-4}	12 m	32-56		[24]
Sce VMA	Artificially split	25	1.2×10^{-3}	10 m	67-73	Low binding affinity	[24]
Mtu RecA	Artificially split	30	Not reported	60-120 m		Requires refolding for activation	[25]

the splicing activity can be restored to carry out protein splicing *in trans* when folded together[26]. Inteins can also be grouped into spontaneously and conditionally functioning inteins. There are a few inteins regain the splicing activity when appropriate conditions are met and mediate conditional protein splicing (CPS) reaction[11].

Expressed Protein Ligation (EPL) and Protein Trans-splicing (PTS) Reaction

Both intein-mediated expressed protein ligation (EPL) and split-intein mediated PTS enable the introduction of various chemical modifications to POIs in a site-specific manner (Figure 1B and 1C). EPL is often used to modify the C-terminus of a recombinant protein. In EPL, the POI, which is recombinantly fused to an intein at its C-terminus, is converted to a labile protein-thioester via spontaneous protein splicing reaction (Figure 1B). The protein-thioester can then react with a synthetic peptide containing an N-terminal Cys to form an amide bond[20]. Split-intein mediated PTS reaction is also widely used for protein engineering (Figure 1C). The discovery of naturally split-inteins had a major impact in biotechnology applications. Split-inteins allow modular approaches to incorporate synthetic residues into POIs, thus expanding the pool of synthetic tags to be incorporated, from small molecules to solid supports or nanoparticles[27]. As the introduction of synthetic moiety to POI via PTS takes advantage of the affinity between the intein fragments which is in mid to high nanomolar concentration range, compared with EPL where the ligating segments have no natural affinity to each other and therefore have to overcome an entropic barrier. In addition, PTS reaction can be used to rearrange the backbone of polypeptide to generate backbone cyclized peptides via simple process.

Fabrication of Biosensors Using Intein-mediated Protein Engineering

Inteins have been utilized to advance biosensor technologies in different aspects. Here, we discuss several features of sensor technologies that have benefited from intein-mediated protein splicing reaction. We briefly discuss on protein chip fabrication as an example of *in vitro* sensor technology and generation of chimeric proteins functioning as target-recognition units using intein technology. We then discuss the fabrication of genetically-encoded cell-based biosensors using CPS as well as the reconstitution of various reporters.

Biochip Fabrication by Intein-mediated Protein Immobilization

Many of biosensors utilize proteins immobilized on solid support[28,29]. Especially, protein microarrays are receiving more and more attention as a tool for exploring the functions and relationships of proteins in proteomic context and as a platform for *in vitro* diagnostics[30]. Protein microarrays present a collection of probe proteins on a solid support in an immobilized fashion. Protein immobilization on solid support is often carried out by non-specific adsorption of proteins onto sticky substrates or by covalent bond formation between various functional groups on the proteins and the activated substrates. While these approaches immobilize purified proteins without a need for genetic modifications, the immobilized proteins often suffer from denaturation or random orientation, yielding a range of activities. The use of recombinant affinity tags (His-6, GST, streptavidin) addresses these issues. However, the interactions of affinity tags are often reversible so that proteins can be lost during the course of assays. Additionally, the presence of large linker proteins could be troublesome as they can mediate unwanted interactions and yield false positive results[31].

There are several examples of protein chips fabricated by intein-based approaches. First, the inteins are used to introduce an affinity tag to a probe protein. Both PTS and EPL are used to introduce chemical tags to POIs, thus generating active probe proteins with proper immobilization handles. One of the most popular tags

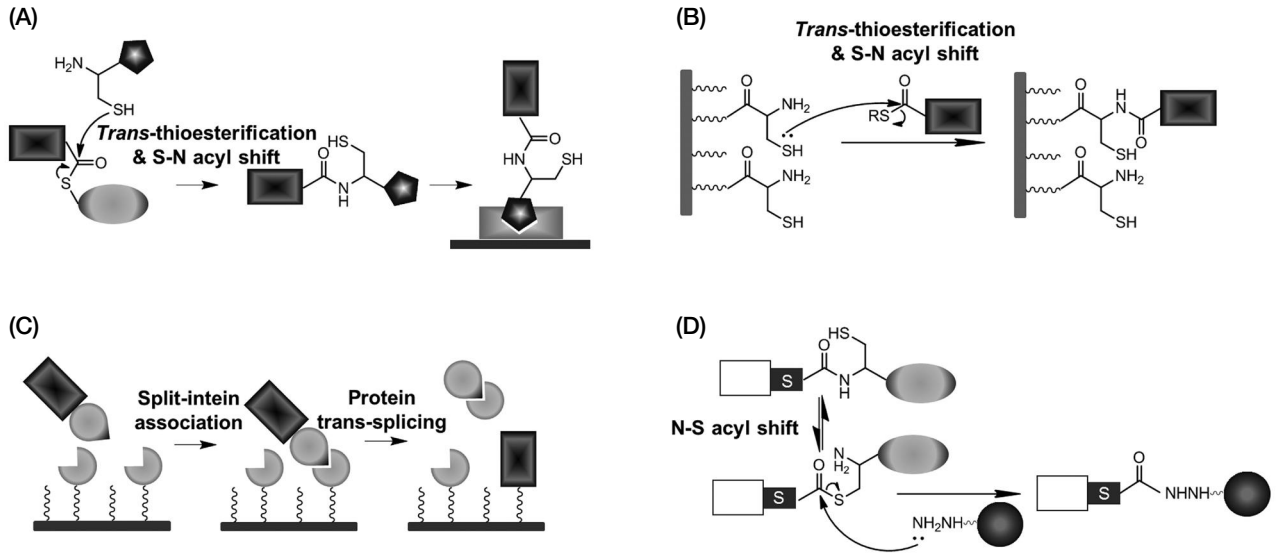

Figure 2. Intein-mediated protein immobilization approaches. (A) Site-specific introduction of affinity tags for protein immobilization. (B) Immobilization of reactive protein thioester onto Cys-presenting glass substrate. (C) Effective and traceless immobilization of protein via split-intein mediated PTS. (D) Fabrication of Q-dot based biosensors by selective immobilization of protein thioester. , Biotin; , Target; , Intein; , Avidin; , I_N; , I_C; , Protease substrate; , QD; , Modified glass; and , Luciferase.

is biotin as it binds to streptavidin with a high affinity. Lesaicherre and coworkers utilized this approach to make a protein array compatible with various biological assays to monitor the functions of proteins (Figure 2A)[32]. Biotin-tagged proteins were immobilized on streptavidin-coated substrates to generate well-oriented and site-specifically immobilized protein arrays. While this approach is useful, the presence of a large linker protein, streptavidin, is a drawback as it can mediate unwanted interactions to yield false positive results.

Alternatively, Kwon and Camarero utilized a protein-thioester intermediate to immobilize probe proteins to a glass substrate presenting N-terminal Cys via a chemoselective ligation reaction (Figure 2B). In this approach, EPL reaction was used to purify target proteins and to generate protein-thioester *in situ*. The protein thioester undergoes native chemical ligation reaction with a Cys-presenting substrate to form a stable amide bond[27]. By this method, the proteins are immobilized in biologically active form through a stable covalent bond without a linker protein. While this was a break through, the chemical interaction between the cysteine and the thioester was very slow when the concentrations of reagents are low. Thus, the amount of time needed for protein-immobilization remained as a hurdle to overcome. Kwon and Camarero also designed an entropically activated immobilization method by adopting split-intein mediated approach to accelerate the PTS reaction (Figure 2C)[33]. In this approach, the recombinant POI was fused to an N-intein. The C-intein is chemically synthesized and attached to an activated glass substrate. The N- and C-inteins bind with each other with high nanomolar affinity to lower the entropic barrier and the reconstituted intein assists the break and formation of chemical bond to ligate the POI to the substrate. During the PTS, the split-intein is spliced out into solution thus accomplishing the traceless immobilization[34].

Additionally, an intein-mediated conjugation strategy was also utilized to fabricate a Q-Dot nano-sensor (Figure 2D)[35]. POI was site-specifically conjugated to a number of quantum dots to generate a series of QD nano-sensors. Xia and coworkers utilized the developed sensors to detect protease matrix metalloproteinase (MMP) activity with high sensitivity. Since MMPs are valuable cancer markers, this assay holds a potential to make a useful cancer sensor.

Generation of Target-recognition Units Based on Intein-mediated Protein Engineering

Among various components required for fabrication of biosensors, the most important part is generating target-recognition modules. While native receptors function to monitor various stimuli in living cells, recently genetically-modified sensor proteins are often used to replace native receptors. There are several examples of proteins engineered to possess sensing capabilities based on split-intein mediated PTS reaction. These

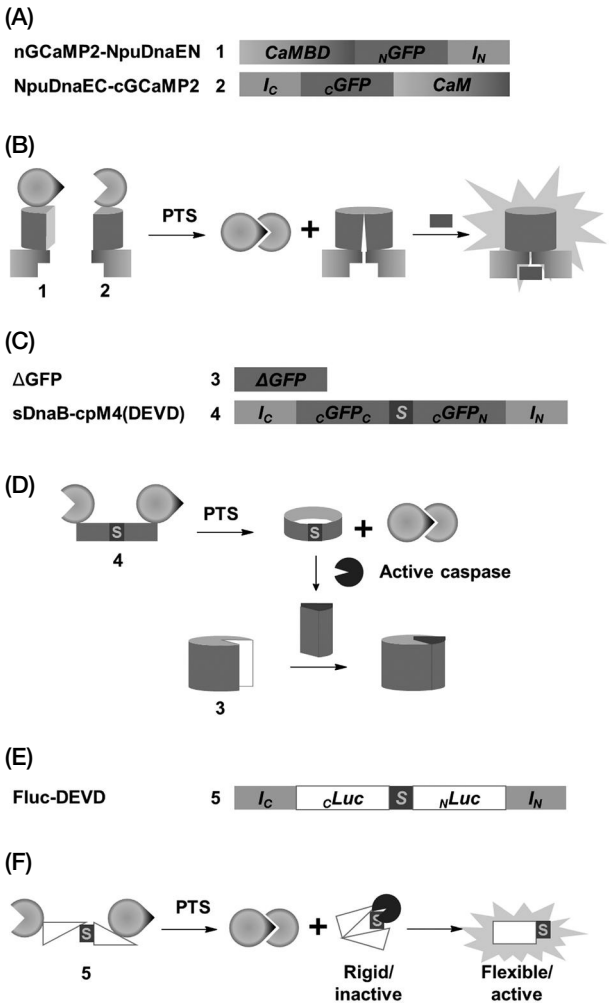

Figure 3. Fabrication of target-recognition units using intein-mediated protein engineering. (A) Schematics of two protein sequences for *in situ* generation of a Ca^{2+} indicator, GCaMP2. (B) Intein-mediated *in situ* re-assembly of GCaMP2 and detection of Ca^{2+}. (C) Schematics of two protein sequences to monitor the caspase activity using split-intein catalyzed ligation of proteins and peptides (SICLOPPS). (D) Active caspase cleaves caspase substrate to linearize cyclic GFP fragment which then bind to ΔGFP to yield a fluorescent signal. (E) A schematic drawing of a protein sequence which generates a cyclic caspase activity biosensor via SICLOPPS. (F) Active caspase processes and linearizes the cyclic and inactive fusion of luciferase and caspase substrate. The linearized fusion protein folds into an active luciferase to generate bioluminescent signals. , I_N; , I_C; , $_N$GFP; , $_C$GFP; , CaMBD; , CaM; , Ca^{2+}; , GFP; , Activation; , Deleted(Δ)-GFP; , C-fragment GFP; , Caspase; , Caspase substrate; , $_N$Luc; and , $_C$Luc.

modified sensor proteins are often expressed in living cells and function to detect biologically active molecules and to report induced protein-protein interactions in living cells.

Split-intein Mediated in situ Re-assembly of Genetically Encoded Ca^{2+} Indicators

Ca^{2+} is an important signaling molecule that controls various cellular events. Many different types of Ca^{2+} sensors are generated using Ca^{2+}-binding proteins. Rogers and colleagues showed an example of *in situ* re-assembly of genetically encoded Ca^{2+} indicators utilizing a pair of split-inteins. The highly efficient naturally-split intein of the DNA polymerase III (DnaE) from Nostoc punctiforme PCC73102 (*Npu*) is utilized for this work. *Npu* DnaE split-intein has a higher rate of protein-splicing compared with other artificial or natural split-inteins[36].

One of the representative Ca^{2+} indicators, GCaMP2, was previously used for the detection of Ca^{2+} *in vivo*[37]. GCaMP2 is a commonly used fluorescence turn-on Ca^{2+} indicator consisting of a tandem fusion of the calmodulin binding domain (CaMBD) from myosin light chain kinase, a circularly permutated EGFP (cpEGFP), and a calmodulin (CaM). The Ca^{2+} indicator was split into two fragments and each fragment was fused to N- and C-inteins, respectively (Figure 3A). The fusion proteins, nGCaMP2-NpuDnaEN **1** and NpuDnaEC-cGCaMP2 **2** were co-expressed in mammalian cells and the Ca^{2+} indicator was reconstituted *in situ* to monitor Ca^{2+} signaling (Figure 3B). This genetically-encoded Ca^{2+} indicator (GECI) system was also introduced to C. elegans and used to image Ca^{2+} signaling *in vivo*. The developed split-indicators can be encoded to two separate cell-lines to monitor the communication of two types of cells and increase cell-type targetability of GECI.

Split-intein Mediated Cyclization of Proteins for Sensing of Caspase Activity

Considerable efforts have been paid to prepare conformationally constrained cyclic peptides[38]. The increased stability and selectivity of cyclic peptides made them good substrates for various enzymatic assays, thus good drug candidates. In generating libraries of cyclic peptides, circularly permuted inteins are utilized to self-catalyze target peptides and this method is termed split-intein catalyzed ligation of proteins and peptides (SICLOPPS)[39,40].

Sakamoto and coworkers utilized this method to generate a genetically encodable sensing system to monitor caspase-3 activity[41]. Caspases are members of cysteine-aspartic acid proteases family and mediate a variety of biological processes in multicellular organisms, including cell cycle regulation, apoptosis and signal transduction. As such, development of a highly sensitive and specific *in vivo* monitoring of caspase activities will provides new insights into the roles of pro-

teases in the biological events. The sensor protein was built, first by splitting a reporter GFP to an inactivated ΔGFP **3** and a small complimentary GFP fragment sDnaB-cpM4(DEVD) **4** (Figure 3C). Then the complimentary GFP fragment was fused with a caspase-3 substrate and cyclized using split-intein mediated PTS to restrain the structural flexibility[42] (Figure 3D). The processing of the cyclic substrate by caspase-3 yielded a linearized and flexible complimentary GFP fragment which, then, can be inserted to an inactivated ΔGFP to recover the fluorescent signal.

The cyclized peptide substrates generated by SICLOPPS has also been utilized in consolidation with split enzymes to activate or amplify enzymatic activities[42]. As an example, Umezawa and coworkers made a cyclic luciferase for real-time sensing of caspase-3 activities (Figure 3E)[43]. They utilized a naturally split DnaE intein derived from Synechocystis sp. PCC6803 (*Ssp*.) to generate cyclic luciferases from Fluc-DEVD **5** via SCILOPPS (Figure 3F). The conformation of the cyclic luciferase was rigidly locked in inactive form. The processing of DEVD by caspase-3 yielded structural flexibility to alter the tertiary structure of the luciferase and recovered bioluminescent activity. This system was introduced to a mouse to monitor the caspase-3 activity *in vivo* based on luminescent signals.

Biosensors for Monitoring Nuclear Translocalization of Proteins

The main group of intein-based biosensors is usually built by constructing fusions of sensor domain, split-intein, and reporter domain. The stimuli is recognized by sensor domain, which then activates the PTS reaction mediated by fused split-intein to reconstitute the reporter domain. Then the reporter domain generates measurable signal. As one example, the translocalization sensors are built through nullifying the activities of spontaneously reacting split-inteins by separating each fragments in different compartment of cells[44,45]. For this system, the external stimuli function to transfer intein-fused sensor domain to different compartment, where the complementary fragment of inteins exists, so that the split-intein reacts and generates a reporter protein.

Nuclear translocalization of functional cytosolic proteins plays an important role in regulating gene expressions in response to external signals. Umezawa and coworkers created several examples of high-throughput sensing and imaging of protein nuclear transport based on intein-mediated reconstitution of split luciferase[44]. In one example, they developed a genetically encoded indicator for detecting nuclear transport of a hormonal

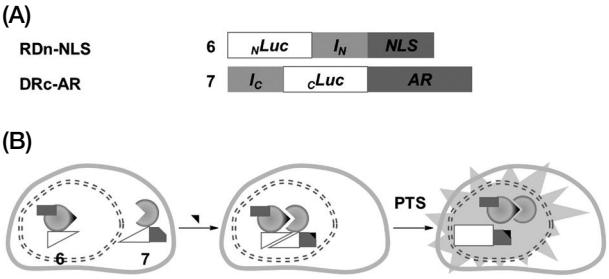

Figure 4. Biosensors for monitoring nuclear translocalization. (A) Schematics of two split-intein fusion proteins to generate a translocalization sensor using living cells. (B) Binding of androgen to androgen receptor translocalizes DRc-AR **6** into nucleus where the RDn-NLS **7** are located. PTS between two fusion proteins **6** and **7** yields a reconstituted luciferase to activate bioluminescent signals in nucleus. ◔, I_N; ◔, I_C; ◺, $_N$Luc; ◺, $_C$Luc; ■, NLS; ◣, AR, and ◥, Androgen.

receptor (Figure 4A). In their approach, the indicator consisted of two fusion proteins of a fragment of split-luciferase and a fragment of *Ssp*. DnaE intein. Each fragment of split-luciferase did not have bioluminescent activity. The fusion protein of N-terminal fragments RDn-NLS **6** was localized into the nucleus via nuclear localization signal (NLS), whereas the fusion protein of C-terminal fragment DRc-AR **7** joined to a well-known nuclear receptor (androgen receptor, AR) was localized in the cytosol. Translocation of DRc-AR **7** into the nucleus was induced by binding of 5-dihydrotestosterone to an AR and allowed the assembly of N- and C-inteins, thereby activating PTS to produce a functional full-length luciferase (Figure 4B). The developed indicator was also modified by using glucocorticoid receptor and was genetically-encoded in an animal model to detect the secretion of endogenous stress hormone, corticosterone. The authors demonstrated that the encoded indicator enables non-invasive imaging against different types of imposed stress[45].

Genetically Encoded Biosensors Using CPS

There are a number of inteins with controllable activities which can be regulated by external factors such as temperature, redox environment, pH, light, or small molecules. PTS reactions mediated by the controllable inteins are called conditional protein-splicing and are used for fabricating biosensors. Especially, the artificially-split inteins with reduced activities are often used for monitoring of protein-protein interactions of sensor domains. The interaction between sensor domains brings fused intein fragments to a close proximity to activate the PTS reaction to generate reporter module that produce measurable signals, such as en-

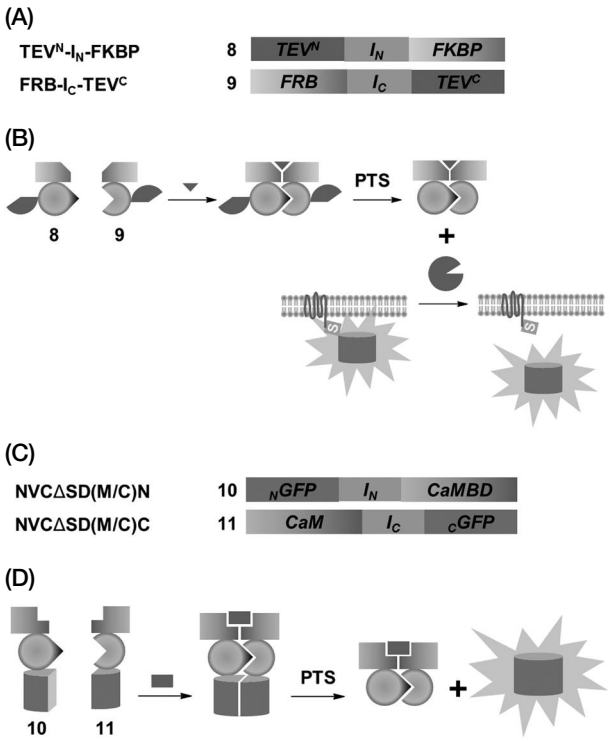

Figure 5. Live-cell based biosensors to monitor small molecule triggered protein-protein interactions via conditional protein splicing (CPS) reaction. (A) Schematics of two protein sequences to monitor rapamycin-triggered binding of FKBP and FRB. (B) The presence of rapamycin induces heterodimerization of two fusion proteins TEV^N-I_N-FKBP **8** and FRB-I_C-TEV^C **9**. As a result, split-fragments of VMA intein bind to each other to carry out PTS to generate an active TEV protease. TEV protease releases membrane bound reporter GFP into cytosol by cleaving the TEV substrate sequence. (C) Schematics of two protein sequences to monitor Ca^{2+}-triggered binding of CaM and CaMBD. (D) The presence of Ca^{2+} induces heterodimerization of two fusion proteins NVCΔSD(M/C)N **10** and NVCΔSD(M/C)C **11**. As a result, split-fragments of VMA intein bind to each other to carry out PTS to generate an active GFP. , FKBP; , FRB; , TEV^N; , TEV^C; , I_N; , I_C; , Rapamycin; , TEV substrate; , GFP; , Activation; , $_N$GFP; , $_C$GFP; , CaMBD; , CaM; and , Ca^{2+}.

zymes, luciferase, or AFPs. Here, we discuss examples that utilize CPS reaction for generation of genetically-encoded biosensors and also the representative reporter proteins.

There are a few cases showing the activation of CPS by a small molecule trigger. Previously, vacuolar ATPase synthase catalytic subunit A of Saccharomyces cerevisiae (Sce VMA) intein was engineered to generate a pair of artificial split-inteins[23]. Split VMA inteins did not mediate PTS ordinarily due to the low binding affinity of two intein fragments. However, the PTS activity was regained by induced heterodimerization of split VMA inteins. Small molecules can function to induce heterodimerization of split-inteins and these systems can be used as sensors to detect target molecules.

The representative systems of small molecule-triggered dimerization include rapamycin- induced dimerization of FK506 binding protein (FKBP) and FKBP12-rapamycin binding domain (FRB) (Figure 5A and 5B)[46], Ca^{2+}-induced binding of CaM and CaMBD (Figure 5C and 5D)[47], and TNF-α mediated binding of myogenic differentiation protein (MyoD) and inhibitor of DNA-binding (Id) protein[48]. Each of these systems was interfaced with engineered low affinity split-VMA intein to generate small molecule sensors.

Small Molecule Biosensors Based on CPS

Mootz's group demonstrated the activation of VMA split-intein mediated PTS in the presence of rapamycin, using Tacacco etch virus (TEV) protease as a reporter protein. Two tandem fusion proteins were prepared as shown in Figure 5A. N-terminal fusion, TEV^N-I_N-FKBP **8**, was comprised of N-terminal fragments of TEV protease, split N-intein, and FKBP and C-terminal fusion, FRB-I_C-TEV^C **9**, was comprised of FRB, split C-intein, and C-terminal TEV protease.

A small molecule rapamycin bound to both FKBP and FRB domains and brought two fusion proteins **8** and **9** in closer proximity thereby inducing the folding and activation of the split VMA intein (Figure 5B). As a result, a splicing product, TEV protease, was generated and cleaved a membrane anchored GFP tag to release the fluorescent signal into the cytosol[46].

Umezawa and coworkers designed a Ca^{2+} sensor utilizing Ca^{2+}-induced binding between CaM and its target peptide M13. Reconstruction and following activation of GFP based on PTS was used as a reporting system. Two tandem fusion proteins were prepared as shown in Figure 5C. N-terminal fusion, NVCΔSD(M/C)N **10**, was comprised of N-EGFP, split N-intein and CaM, and C-terminal fusion, NVCΔSD(M/C)C **11**, was comprised of CaMBD, split C-intein, and C-EGFP. Two fusion proteins heterodimerized in the presence of Ca^{2+} to yield an active EGFP reporter (Figure 5D)[47]. The developed system was introduced to a transgenic mouse to monitor Ca^{2+}-mediated protein-protein interactions and offered a new opportunity to image Ca^{2+}-induced signaling. However, the drawbacks of the system include low splicing efficiency of intein as well as low sensitivity. High background signal coming from spontaneously reconstituted AFP is also a problem to overcome.

Reconstitution of Split-reporters

Biosensors include biological recognition elements

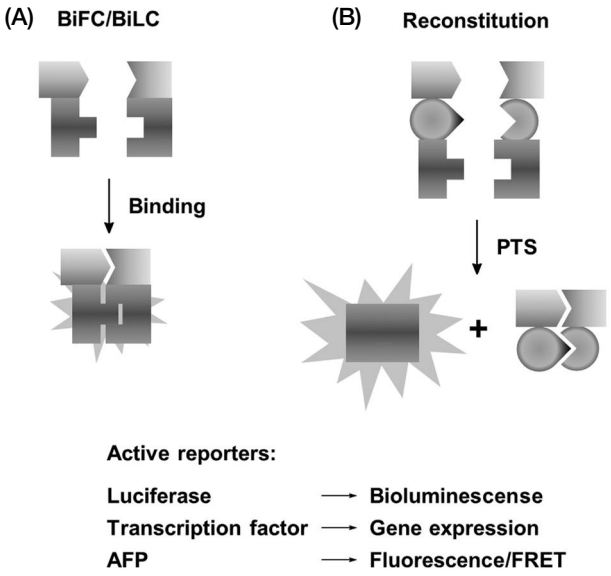

Figure 6. Generation of reporter proteins in biosensor via complementation (A) or reconstitution (B) of split-reporters. Popular reporters include luciferases, transcription factors, and autofluorescent proteins (AFPs). , Target; , Bait; , $_N$Reporter; , $_C$Reportor; , Activation; , I_N; and , I_C.

and reporter units. When specific targets or events are detected by the recognition unit, the sensor generates a measurable signal through a reporter unit. Various approaches to build target recognition units based on intein-mediated protein splicing reaction were discussed in the preceding sections of this review. Here, the design strategies of various reporter modules will be discussed. The most common reporters in cell-based biosensors include AFP and enzymes that generate colored or luminescent products. These proteins are useful as they enable non-destructive spectroscopic monitoring, which are sensitive and suitable for high-throughput analysis.

One of most common strategies to build a reporter system is a bimolecular complementation approach interfaced with luminescent[49] or fluorescent proteins[50]. This method is based on complementation and subsequent activation of split luminescent or fluorescent proteins. The presence of target molecules induce heterodimerization of target recognition elements as previously discussed. While bimolecular luminescence complementation (BiLC) and bimolecular fluorescence complementation (BiFC) (Figure 6A) offer useful strategies, one of the major drawbacks of these strategies are that the split-reporters can be activated in the absence of stimuli to yield false positive signals when they exist in high concentration. Intein-mediated PTS reaction offers an alternative approach by enabling reconstitution of split-reporter molecules instead of reversible complementation (Figure 6B). In CPS, the N- and C-termini of an intein are brought together in close proximity when stimuli are present. Then they undergo a splicing event resulting in reconstitution of active reporter protein by conjugating two fragments through a peptide bond[51]. The reconstituted seamless reporter proteins show significantly enhanced signal intensities with low background signals compared with complemented split-reporters. Additionally, intein-based sensors can easily be modified due to the modular design of these sensors, allowing the easy exchange of sensing and reporter domains[11].

Over the years, many groups have used split-enzymes for the main reporter module. Some of the examples of enzymes used are β-galactosidase and Firefly and Renila luciferases, which enables background-free and highly sensitive bioluminescent detection[52]. One of the major advantage of using enzymes as a reporter module in biosensors is natural amplification of signals as one enzyme can catalyze a number of reactions to generate signals. Also enzyme reporters are genetically encodable, thus can be used to build self-sustainable cell-based sensors as well as *in vivo* sensors. On the other hands, enzymes often require substrates and cofactors to generate signals and can be unstable when isolated.

AFPs are useful reporter proteins in building biosensors. Monomeric AFPs such as GFP, mCherry, and mRFP are frequently used. Intein-mediated approaches can be interfaced with various AFP-based reporting systems. Aforementioned, reconstitution and activation of inactivated split-AFP upon stimulation is a representative example[53]. Alternatively, a Förster resonance energy transfer (FRET) pairs which are introduced to N- and C-termini of split N- and C-inteins, respectively, are used to build a reporting unit. In the presence of a small molecule, heterodimerization of tandem fusion and consequent PTS reaction bring two fluorescent proteins in close proximity to activate a FRET signal[26].

Other groups used transcription factors as a reporter module. A transcription factor, a sequence-specific DNA-binding protein, regulates the rate of transcription often by forming a complex with other proteins. The specific transcription-factor-based reporters generate semi-quantitative results with reduced background signal.

Summary

Following the discovery of inteins in 1990s, a variety of inteins were identified from the nature and engineered to obtain diverse and useful characteristics.

The emergence of intein-mediated protein engineering technology have expanded the capabilities in various biotechnology applications, including development of biosensors, drug screening platforms, as well as protein purification methods. The development of inteins with controllable activities along with the appearance of fast-functioning split-inteins provided a new opportunity in intein-mediated protein engineering. In this review, we discussed on sensor technologies that take advantages of intein-mediated protein engineering. First we discussed on *in vitro* sensing platforms which are fabricated in tailored manner using intein-mediated protein engineering for enhanced sensitivities and selectivities. We also discussed on the use of structurally-strained cyclic peptides in fabrication of sensing modules as well as *in situ* generation of previously reported sensor proteins for an enhanced applicability, such as monitoring cell-cell communication. Various types of genetically-encoded sensor that can monitor induced protein-protein interactions based on CPS are also discussed including a brief overview on reconstitution of split-reporter domains. The developed cell-based sensors have many advantages compared to conventional sensors built on native receptors and hold a large potential for *in vivo* applications. Even though there are still many hurdles to overcome for use of these systems for practical applications, we believe that intein-based protein engineering will greatly contribute to the field in biomedicine and environments.

Acknowledgements This work was supported by Business for Cooperative R&D between Industry, Academy, and Research Institute funded by KSMBA in 2015 (Grant No. C0349391) and Gachon university Gil medical center (Grant No. FRD2014-07-02).

References

1. Liu, Q. *et al*. Cell-Based Biosensors and Their Application in Biomedicine. *Chem. Rev.* **114**, 6423-6461 (2014).
2. Ramachandran, N., Larson, D.N., Stark, P.R.H., Hainsworth, E. & LaBaer, J. Emerging tools for real-time label-free detection of interactions on functional protein microarrays. *FEEB Journal* **272**, 5412-5425 (2005).
3. Liu, Y. & Wilson, W.D. Quantitative Analysis of Small Molecule-Nucleic Acid Interactions With a Biosensor Surface and Surface Plasmon Resonance Detection. *Methods. Mol. Biol.* **613**, 1-23 (2010).
4. Brian, R.B. *et al*. An Electronic, Aptamer-Based Small-Molecule Sensor for the Rapid, Label-Free Detection of Cocaine in Adulterated Samples and Biological Fluids. *J. Am. Chem. Soc.* **128**, 3138-3139 (2006).
5. Zhou, W., Zhang, Y., Ding, J. & Liu, J. *In Vitro* Selection in Serum: RNA-Cleaving DNAzymes for Measuring Ca^{2+} and Mg^{2+}. *ACS Sens.* **1**, 600-606 (2016).
6. Kim, B.K. *et al*. Sensitivity response to coating material thickness for an optical resonant reflective biosensor based on a guided mode resonance filter. *BioChip J.* **8**, 35-41 (2014).
7. Zhang, C.Q., Park, Y.M., Yang, D., Yoo, T.H. & Yoon, H.C. Development of a matrix metalloproteinase-2 (MMP-2) biosensing system by integrating an enzyme-mediated color development reaction into a common electronics components setup. *BioChip J.* **10**, 198-207 (2016).
8. Liu, Z., Tabakman, S., Welsher, K. & Dai, H. Carbon Nanotubes in Biology and Medicine: *in vitro* and in vivo Detection, Imaging and Drug Delivery. *Nano Res.* **2**, 85-120 (2009).
9. Park, M., Tsai, S.L. & Chen, W. Microbial Biosensors: Engineered Microorganisms as the Sensing Machinery. *Sensors* **13**, 5777-5795 (2013).
10. Choi, M. *et al*. Light-guiding hydrogels for cell-based sensing and optogenetic synthesis *in vivo*. *Nat. Photonics* **7**, 987-994 (2013).
11. Wood, D.W. & Camarero, J.A. Intein applications: from protein purification and labeling to metabolic control methods. *J. Biol. Chem.* **289**, 14512-14519 (2014).
12. Elleuche, S. & Pöggeler, S. Inteins, valuable genetic elements in molecular biology and biotechnology. *Appl. Microbiol. Biotechnol.* **87**, 479-489 (2010).
13. Mills, K. V., Johnson, M. A. & Perler, F. B. Protein splicing: how inteins escape from precursor proteins. *J. Biol. Chem.* **289**, 14498-14505 (2014).
14. Jung, D., Min, K., Jung, J., Jang, W. & Kwon, Y. Chemical biology-based approaches on fluorescent labeling of proteins in live cells. *Mol. Biosyst.* **9**, 862-872 (2013).
15. Jung, D. *et al*. Photo-triggered fluorescent labelling of recombinant proteins in live cells. *Chem. Commun.* **51**, 9670-9673 (2015).
16. Vila-Perelloó, M. *et al*. Streamlined Expressed Protein Ligation Using Split Inteins. *J. Am. Chem. Soc.* **135**, 286-292 (2013).
17. Carvajal-Vallejos, P., Pallisse, R., Mootz, H.D. & Schmidt, S.R. Unprecedented rates and efficiencies revealed for new natural split inteins from metagenomic sources. *J. Biol. Chem.* **287**, 28686-28696 (2012).
18. Eryilmaz, E., Shah, N.H., Muir, T.W. & Cowburn, D. Structural and dynamical features of inteins and implications on protein splicing. *J. Biol. Chem.* **289**, 14506-14511 (2014).
19. Topilina, N.I. & Mills, K.V. Recent advances in *in vivo* applications of intein-mediated protein splicing. *Mob DNA* **5**, 5 (2014).
20. Frutos, S., Goger, M., Giovani, B., Cowburn, D. & Muir, T.W. Branched intermediate formation stimulates peptide bond cleavage in protein splicing. *Nat. Chem. Biol.* **6**, 527-533 (2010).

21. Mills, K.V., Manning, J.S., Garcia, A.M. & Wuerdeman, L.A. Protein splicing of a Pyrococcus abyssi intein with a C-terminal glutamine. *J. Biol. Chem.* **279**, 20685-20691 (2004).
22. Topilina, N.I. *et al*. SufB intein of Mycobacterium tuberculosis as a sensor for oxidative and nitrosative stresses. *Proc. Natl. Acad. Sci. U S A* **112**, 10348-10353 (2015).
23. Shah, N.H., Dann, G.P., Vila-Perello, M., Liu, Z. & Muir, T.W. Ultrafast protein splicing is common among cyanobacterial split inteins: implications for protein engineering. *J. Am. Chem. Soc.* **134**, 11338-11341 (2012).
24. Brenzel, S., Kurpiers, T. & Mootz, H.D. Engineering artificially split inteins for applications in protein chemistry: biochemical characterization of the split Ssp DnaB intein and comparison to the split Sce VMA intein. *Biochemistry* **45**, 1571-1578 (2006).
25. Mills, K.V., Lew, B.M., Jiang, S. & Paulus, H. Protein splicing in trans by purified N- and C-terminal fragments of the Mycobacterium tuberculosis RecA intein. *Proc. Natl. Acad. Sci. U S A* **95**, 3543-3548 (1998).
26. Shah, N.H. & Muir, T.W. Inteins: Nature's Gift to Protein Chemists. *Chem. Sci.* **5**, 446-461 (2014).
27. Camarero, J.A., Kwon, Y. & Coleman, M.A. Chemoselective attachment of biologically active proteins to surfaces by expressed protein ligation and its application for "protein chip" fabrication. *J. Am. Chem. Soc.* **126**, 14730-14731 (2004).
28. Jung, D., Min, K., Jeon, Y., Jang, W. & Kwon, Y. Optimized magnetic bead-based immunoassay for automated detection of protein toxins. *BioChip J.* **6**, 293-298 (2012).
29. Yoo, R.J. & Choi, S.J. Identification of a peptide ligand for antibody immobilization on biosensor surfaces. *BioChip J.* **10**, 88-94 (2016).
30. Kwon, J.H. *et al*. Simultaneous subtyping and pathotyping of the novel reassortant influenza A (H5N8) virus from clinical samples using a diagnostic microarray. *BioChip J.* **10**, 167-173 (2016).
31. Min, K., Jung, D., Jeon, Y., Jeoung, E. & Kwon, Y. Site-specific and effective immobilization of proteins by Npu DnaE split-intein mediated protein trans-splicing reaction. *BioChip J.* **7**, 288-294 (2013).
32. Lesaicherre, M.L., Lue, R.Y., Chen, G.Y., Zhu, Q. & Yao, S.Q. Intein-mediated biotinylation of proteins and its application in a protein microarray. *J. Am. Chem. Soc.* **124**, 8768-8769 (2002).
33. Kwon, Y., Coleman, M.A. & Camarero, J.A. Selective immobilization of proteins onto solid supports through split-intein-mediated protein trans-splicing. *Angew. Chem.* **45**, 1726-1729 (2006).
34. Camarero, J.A. & Kwon, Y. Traceless and Site-specific Attachment of Proteins onto Solid Supports. *Int. J. Pept. Res. The.* **14**, 351-357 (2008).
35. Xia, Z. *et al*. Multiplex detection of protease activity with quantum dot nanosensors prepared by intein-mediated specific bioconjugation. *Anal. Chem.* **80**, 8649-8655 (2008).
36. Zettler, J., Schutz, V. & Mootz, H.D. The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS Lett.* **583**, 909-914 (2009).
37. Wong, S.S., Kotera, I., Mills, E., Suzuki, H. & Truong, K. Split-intein mediated re-assembly of genetically encoded Ca(2+) indicators. *Cell Calcium* **51**, 57-64 (2012).
38. Young, T.S. *et al*. Evolution of cyclic peptide protease inhibitors. *Proc. Natl. Acad. Sci. U S A* **108**, 11052-11056 (2011).
39. Naumann, T.A., Tavassoli, A. & Benkovic, S.J. Genetic selection of cyclic peptide Dam methyltransferase inhibitors. *Chembiochem.* **9**, 194-197 (2008).
40. Abel-Santos, E., Scott, C.P. & Benkovic, S.J. Use of inteins for the *in vivo* production of stable cyclic peptide libraries in E. coli. *Methods Mol. Biol.* **205**, 281-294 (2003).
41. Taylor, R.C., Cullen, S.P. & Martin, S.J. Apoptosis: controlled demolition at the cellular level. *Nat. Rev. Mol. Cell. Biol.* **9**, 231-241 (2008).
42. Sakamoto, S. *et al*. Creation of a caspase-3 sensing system using a combination of split-GFP and split-intein. *Chem. Commun.* **49**, 10323-10325 (2013).
43. Kanno, A., Yamanaka, Y., Hirano, H., Umezawa, Y. & Ozawa, T. Cyclic luciferase for real-time sensing of caspase-3 activities in living mammals. *Angew. Chem.* **46**, 7595-7599 (2007).
44. Kim, S.B., Ozawa, T., Watanabe, S. & Umezawa, Y. High-throughput sensing and noninvasive imaging of protein nuclear transport by using reconstitution of split Renilla luciferase. *Proc. Natl. Acad. Sci. U S A* **101**, 11542-11547 (2004).
45. Kim, S.B., Ozawa, T. & Umezawa, Y. Genetically encoded stress indicator for noninvasively imaging endogenous corticosterone in living mice. *Anal. Chem.* **77**, 6588-6593 (2005).
46. Sonntag, T. & Mootz, H.D. An intein-cassette integration approach used for the generation of a split TEV protease activated by conditional protein splicing. *Mol. Biosyst.* **7**, 2031-2039 (2011).
47. Ozawa, T., Nogami, S., Sato, M., Ohya, Y. & Umezawa, Y. A fluorescent indicator for detecting protein-protein interactions *in vivo* based on protein splicing. *Anal. Chem.* **72**, 5151-5157 (2000).
48. Paulmurugan, R., Umezawa, Y. & Gambhir, S.S. Noninvasive imaging of protein-protein interactions in living subjects by using reporter protein complementation and reconstitution strategies. *Proc. Natl. Acad. Sci. U S A* **99**, 15608-15613 (2002).
49. Paulmurugan, R. & Gambhir, S.S. Monitoring protein-protein interactions using split synthetic renilla luciferase protein-fragment-assisted complementation. *Anal. Chem.* **75**, 1584-1589 (2003).
50. Kerppola, T.K. Design implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. *Nat. Protoc.* **1**, 1278-1286 (2006).

51. Kanno, A., Ozawa, T. & Umezawa, Y. Intein-mediated reporter gene assay for detecting protein-protein interactions in living mammalian cells. *Anal. Chem.* **78**, 556-560 (2006).
52. Kanno, A., Ozawa, T. & Umezawa, Y. Genetically encoded optical probe for detecting release of proteins from mitochondria toward cytosol in living cells and mammals. *Anal. Chem.* **78**, 8076-8081 (2006).
53. Ozawa, T., Takeuchi, T.M., Kaihara, A., Sato, M. & Umezawa, Y. Protein splicing-based reconstitution of split green fluorescent protein for monitoring protein-protein interactions in bacteria: improved sensitivity and reduced screening time. *Anal. Chem.* **73**, 5866-5874 (2001).

Copyright for the published papers belongs to the Korean BioChip Society. pISSN 1976-0280. eISSN 2092-7843