

## TECHNIQUES FOR MOLECULAR ANALYSIS

# Molecular and cellular approaches for the detection of protein–protein interactions: latest techniques and current limitations

Sylvie Lalonde\*, David W. Ehrhardt, Dominique Loqué, Jin Chen, Seung Y. Rhee and Wolf B. Frommer  
Carnegie Institution, 260 Panama Street, Stanford, CA 94305, USA

Received 15 June 2007; revised 18 September 2007; accepted 20 September 2007.

\*For correspondence (fax +1 650 325 6857; e-mail slalonde@stanford.edu).

## Summary

Homotypic and heterotypic protein interactions are crucial for all levels of cellular function, including architecture, regulation, metabolism, and signaling. Therefore, protein interaction maps represent essential components of post-genomic toolkits needed for understanding biological processes at a systems level. Over the past decade, a wide variety of methods have been developed to detect, analyze, and quantify protein interactions, including surface plasmon resonance spectroscopy, NMR, yeast two-hybrid screens, peptide tagging combined with mass spectrometry and fluorescence-based technologies. Fluorescence techniques range from co-localization of tags, which may be limited by the optical resolution of the microscope, to fluorescence resonance energy transfer-based methods that have molecular resolution and can also report on the dynamics and localization of the interactions within a cell. Proteins interact via highly evolved complementary surfaces with affinities that can vary over many orders of magnitude. Some of the techniques described in this review, such as surface plasmon resonance, provide detailed information on physical properties of these interactions, while others, such as two-hybrid techniques and mass spectrometry, are amenable to high-throughput analysis using robotics. In addition to providing an overview of these methods, this review emphasizes techniques that can be applied to determine interactions involving membrane proteins, including the split ubiquitin system and fluorescence-based technologies for characterizing hits obtained with high-throughput approaches. Mass spectrometry-based methods are covered by a review by Miernyk and Thelen (2008; this issue, pp. 597–609). In addition, we discuss the use of interaction data to construct interaction networks and as the basis for the exciting possibility of using to predict interaction surfaces.

**Keywords:** protein interaction, fluorescence resonance energy transfer, fluorescence correlation spectroscopy, split GFP, split ubiquitin, co-localization.

## Introduction

Macromolecular interactions such as protein–protein interactions (PPI) are fundamental for all biological processes, ranging from the formation of cellular structures and enzymatic complexes to the regulation of signaling pathways. Proteins frequently function as stable or transient complexes with other proteins (Alberts, 1998; Grigoriev, 2003; Kerrien *et al.*, 2007). Interactions between proteins can serve diverse functions such as conferring specificity to interactions between enzymes and substrates in signal transduction events, protection of proteins from their environment, facilitation of substrate channeling, or build-

ing molecular machines such as the cytoskeleton (Huang *et al.*, 2001; Islam *et al.*, 2007; Kozer *et al.*, 2007). Some proteins function as obligatory oligomers, for example the *Escherichia coli* tryptophan repressor that forms two symmetric tryptophan-binding sites at the dimer interface (Schevitz *et al.*, 1985). Similarly, K<sup>+</sup> channels form a single pore from four similar or identical subunits (MacKinnon, 1991; Long *et al.*, 2005). Such oligomeric interactions provide the possibility of novel properties, as exemplified by the Ca<sup>2+</sup>/calmodulin-dependent protein kinase II (CaMKII). This consists of 12 identical subunits, each of which is able

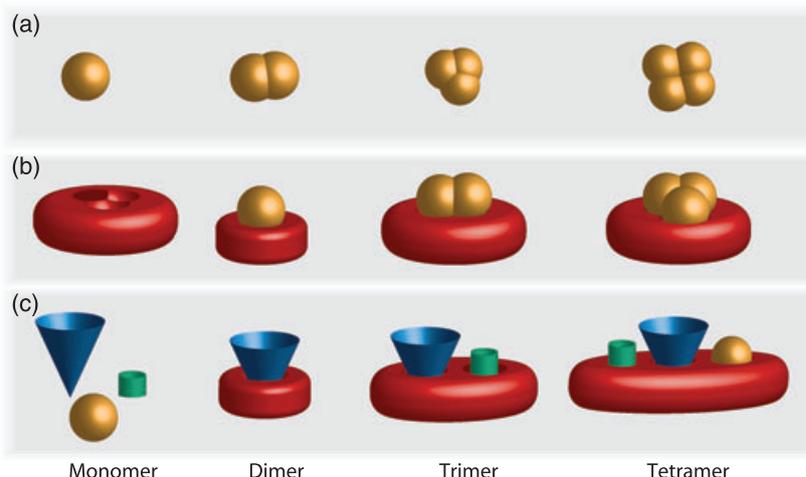
to phosphorylate its neighbor. Modeling showed that the interplay between the autocatalytic phosphorylation of CaMKII and removal of phosphate groups by protein phosphatase produces two stable states of CaMKII at basal free calcium levels, enabling it to act as a switch involved in memory and as a decoder of calcium spike patterns (Mitra *et al.*, 2004; Shen *et al.*, 2000). A second example of emergent properties due to oligomerization is provided by the ammonium transporter AtAMT1, where *trans*-activation among constituent subunits allows for rapid non-linear shutdown of transport activity and memory of activity state (Loqué *et al.*, 2007).

Protein interactions are characterized by kinetic and thermodynamic parameters. Some proteins interact with high affinity, forming stable interactions, such as TEM1- $\beta$ -lactamase and its inhibitor protein (Reichmann *et al.*, 2007). Other proteins interact more dynamically, requiring a lower-affinity binding. Typically, these are proteins that serve regulatory roles, including G-protein-coupled receptors, protein kinases, and cell surface receptors that are activated by dimerization (Pellicena and Kuriyan, 2006). The methods described in this review differ in their sensitivity, specificity, and ability to detect interactions of differing affinity; thus selection of a suitable method is crucial for a given investigation. The detection of weak or transient interactions may require special techniques such as cross-linking (Trakselis *et al.*, 2005). For example, data obtained in heterologous systems often overexpress the two partners and thus an interaction detected in such a system may not reflect a native or conditional occurrence *in situ*. It is also important to

decide whether the objective is to determine the interaction between a protein pair or to analyze for the existence of protein in larger complexes before choosing a method. Two-hybrid screens measure direct binary interactions, while immunoprecipitation-based methods and fluorescence resonance energy transfer (FRET) measure the presence of a bait protein in a complex or in the vicinity of the prey (Figure 1). Therefore, datasets derived from different methods each have their characteristics and thus are expected to share only partial overlap; extensive follow-up using different approaches is therefore required to generate a comprehensive interaction map (Rual *et al.*, 2005).

#### The structural basis of protein interactions

The binding of two proteins is determined by the shape and chemistry of the binding surface, i.e. the amino acid composition and tertiary structure of the proteins (Perozzo *et al.*, 2004; Reichmann *et al.*, 2007). Protein interactions can occur between identical or non-identical polypeptides (homo- and hetero-oligomers). Protein interactions can be further classified into obligate and non-obligate complexes or, depending on the lifetime of the complex, into transient or permanent complexes (Nooren and Thornton, 2003). Proteins that interact typically have complementary surfaces, and the forces that stabilize the interaction are identical to those that play a role in protein folding as well as in interactions of proteins with small molecules: ionic interactions, dipole interactions, hydrogen bonds, van der Waals forces, hydrophobic interactions, and also water-mediated



**Figure 1.** Types of protein complexes.

(a) A variety of proteins occur as monomers, homodimers, homotrimers, homotetramers, or even larger homomeric complexes (e.g. CaMKII as a dodecamer; Rosenberg *et al.*, 2006).

(b) Proteins may assemble as hetero-oligomers consisting of homo- or hetero-oligomers.

(c) Indirect interactions: for example scaffolding proteins such as INAD (Wang and Montell, 2007) can bind various monomeric or oligomeric proteins, as for example in the signalosome (Wei and Deng, 2003). The methods used to identify protein interactions differ with respect to the type of interaction they detect and thus yield non-overlapping datasets. The yeast two-hybrid or mating-based split ubiquitin systems identify binary interactions only (as in a or b); while affinity chromatography/mass spectrometry and fluorescence resonance energy transfer identify both binary complexes as well as complexes in which two proteins interact via a third partner (e.g. trimer or tetramer in c).

interactions between residues on the surface of the two partners (Reichmann *et al.*, 2007).

Stable protein–protein interfaces consist of a set of modules (Reichmann *et al.*, 2005). Well-studied examples include the nanomolar-affinity-interaction between TEM1  $\beta$ -lactamase and its protein inhibitor,  $\beta$ -lactamase inhibitor protein (Reichmann *et al.*, 2005), the RNase BARNASE and its inhibitor BARSTAR (Vaughan *et al.*, 1999; Reichmann *et al.*, 2005), and interferon IFN- $\alpha$ 2 and its receptor IFNA-R1 (Reichmann *et al.*, 2005, 2007). The interaction surface between TEM1  $\beta$ -lactamase and its inhibitor consists of five modules, where each module comprises a cluster of interacting residues (Reichmann *et al.*, 2005). Complex formation is very fast, and the complex is highly stable with a  $K_d$  in the femtomolar range. An important question is thus whether the molecular interactions that contribute to formation and stability of the complex are made up from an additive set of atomic interactions, or whether the interaction is made from more complex networks of cooperative interactions (Reichmann *et al.*, 2007). Surprisingly, individual bonds at the complex interaction surface can be altered or obliterated by mutation without major effects on the formation of the overall complex and with only small effects on binding affinity. This has been shown for a variety of proteins, e.g. the binding of IFN- $\alpha$ 2 to the IFNA-R1 receptor (Reichmann *et al.*, 2005, 2007). A careful analysis introducing single mutations alone or in combinations demonstrated cooperativity as well as intercluster additivity between the interface modules (Reichmann *et al.*, 2005). In contrast, the weak interaction between the bacterial signaling components CheA and CheY with an affinity of 2  $\mu$ M is characterized by fewer contacts between the two proteins, which are organized into only a single cluster. Thus in addition to knowledge of the thermodynamic and kinetic properties of a given interaction pair, structural information is important for

understanding the evolution of protein interaction networks. Systematic analysis of protein interactions combined with structural information may ultimately help to develop methods that will allow more accurate predictions of interactions and their properties (Shoemaker and Panchenko, 2007).

#### *Interactomics: the importance of systematic generation of protein interaction maps*

From the discussion above, it is obvious that a comprehensive map of both low- and high-affinity PPI among soluble and membrane proteins in the cell would be an invaluable asset for the understanding of biological processes and molecular mechanisms at the systems biology level. Such a map needs to include both binary protein interactions as well as larger complexes (Figure 1). Knowledge of the protein interaction network is a crucial prerequisite for understanding most cellular functions; especially the regulatory and signaling networks. Primary goals of the post-genomic era are: (i) the assignment of functions to each of the genes encoded in a given genome and (ii) their integration into metabolic and regulatory networks. While transcriptomics and proteomics are progressing rapidly, collection of other essential information for building these network maps – the mapping of protein interactions (interactome or associome), the profiling of intermediates (Meyer *et al.*, 2007), ions (Baxter *et al.*, 2007), and metabolic flux (Wiechert *et al.*, 2007) – will be a major focus of research in the coming years.

Network analysis requires methods amenable to high throughput (HT), such as yeast two-hybrid (Y2H) assays and affinity purification–mass spectrometry (AP-MS) for performing systematic screens (Table 1; Miernyk and Thelen, 2008; this issue, pp. 597–609). Benefits of HT analyses are

**Table 1** Methods for analyzing protein–protein interactions

	<i>In vitro</i>	<i>In vivo</i>
Affinity purification	AP-MS	AP-MS
Genetic test systems		Y2H <sup>b</sup> , mbSUS <sup>a, b</sup> , CytoTrap <sup>TM</sup> <sup>a</sup>
Fluorescence	FRET BRET	FRET <sup>a</sup> Split-GFP <sup>a</sup>
Plasmon resonance	Quantitative SPR (Boozer <i>et al.</i> , 2006)	
Crystal structure	Structure of complex	-
Calorimetry	Quantitative analysis of protein interactions	
AFM	Detection and quantitative analysis of protein interactions	
NMR	Quantitative analysis of large complexes (Sprangers and Kay, 2007)	STINT-NMR (Burz <i>et al.</i> , 2006)
Protein arrays	Identification and analysis of selectivity of protein interactions (Korf and Wiemann, 2005)	

AP-MS, affinity purification–mass spectrometry; Y2H, yeast two-hybrid; mbSUS, mating-based split ubiquitin system; FRET, fluorescence resonance energy transfer; BRET, bioluminescence resonance energy transfer; AFM, atomic force microscopy; SPR, surface plasmon resonance; STINT-NMR, mapping structural interactions using in-cell NMR spectroscopy.

<sup>a</sup>Suitable for membrane proteins.

<sup>b</sup>Amenable to high throughput.

that a single lab or small consortium with extensive experience in one method can carry out a whole or sub-genome screen and generate a complete dataset collected under comparable conditions in which the complement of all tested proteins serves as a multiparallel internal control, thus reducing the number of potential artifacts. A potential drawback of HT analysis may be that typically only a limited number of replica tests are performed (a single run of a matrix of 30,000 × 30,000 proteins covering the genome of a higher eukaryote adds up to close to a billion individual binary tests). Another drawback is that protein interactions are typically scored in an all-or-nothing scheme. Yeast two-hybrid analyses and screens, for example, often just score auxotrophy versus prototrophy using a binary code (Miller *et al.*, 2005; Uetz *et al.*, 2000). Furthermore, in the Y2H as well as other protein tagging approaches, test proteins are often overexpressed, thus modifying the relative concentrations of potential interaction partners from the *in vivo* state. Moreover, the use of heterologous systems can eliminate competing activities that exist in the native system and can also introduce novel competitors. Analysis of interactions in extracts, as typically performed in AP-MS experiments, may bring together proteins from different compartments in 'non-crowded' environments that do not reflect the *in vivo* situation. Therefore, overexpression and the elimination of competing interaction partners or the co-expression of proteins residing in different cellular compartments can lead to the detection of interactions that will not occur *in vivo*. Interactions detected in such screens are therefore designated 'potential interactions'. For cases in which a protein is found to interact with several or many other proteins, as may be expected for scaffolding proteins, orthogonal assays are required to determine the relevance of an interaction *in vivo*.

Taken together, HT screens are necessary to obtain an overview of the potential interactome, but extensive follow-up is required to unambiguously identify false positive and negative results and determine those interactions that are relevant for cellular function.

#### *The special case of membrane proteins*

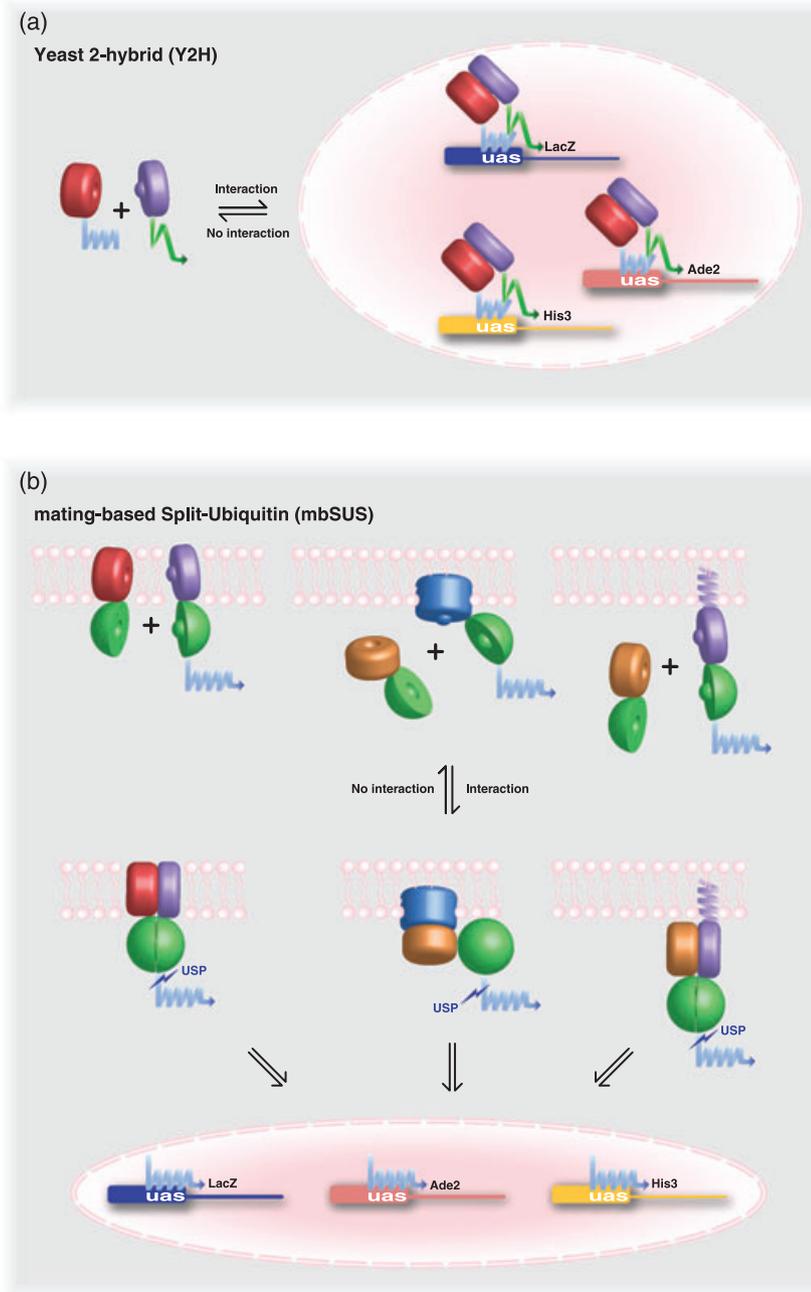
Membrane proteins play crucial roles in many biological processes. They control cell permeability (influx and efflux) for a myriad of compounds and are responsible for sensing chemical and physical stimuli from the environment (nutrients, hormones, pH, pathogens, etc.) to allow the organism to acclimate to changing conditions and to coordinate transport and metabolism. Despite the importance of membrane proteins (which represent 20–30% of the *Arabidopsis* proteome; Schwacke *et al.*, 2003) in carrying out functions such as transport, vesicular trafficking, energization, homeostasis, or signaling, little is known about their interactions with each other or with other proteins (Ludewig

*et al.*, 2003; Obrdlik *et al.*, 2004; Reinders *et al.*, 2002a). Yeast two-hybrid assays are depleted of membrane protein interactions (Xia *et al.*, 2006), because in the classical Y2H system the activation domain of a transcription factor, when fused to a membrane protein, will be retained at the membrane, and thus rendered unavailable for reconstitution of a functional transcription factor in the nucleus (Figure 2a). Moreover, membrane proteins are often toxic when expressed in *E. coli*, leading to under-representation of membrane protein open reading frames (ORFs) in cDNA expression libraries (Frommer and Ninnemann, 1995). Biochemical assays require optimization of solubilization in detergents (Kalipatnapu and Chattopadhyay, 2005) and subsequent reconstitution in lipid bilayers. Therefore, alternative methods, such as the split ubiquitin system (Obrdlik *et al.*, 2004) and advanced biochemical methods are required to provide maps of interactions between membrane proteins as well as the interface between membrane and soluble proteins.

#### **High-throughput interactomics tools**

To systematically analyze protein complexes at a sub- or full-genome level, several methods have been adapted for HT screens: Y2H systems, the mating-based split-ubiquitin system (mbSUS), and affinity purification of protein complexes followed by identification of proteins by mass spectroscopy (AP-MS; Miernyk and Thelen, 2008; this issue, pp. 597–609). Yeast two-hybrid and AP-MS methods have successfully been used to determine significant parts of protein interaction networks in *Saccharomyces cerevisiae* (Gavin *et al.*, 2002; Ho *et al.*, 2002; Ito *et al.*, 2000; Krogan *et al.*, 2006; Miller *et al.*, 2005; Uetz *et al.*, 2000), *Caenorhabditis elegans* (Li *et al.*, 2004; Walhout *et al.*, 2000), *Drosophila melanogaster* (Formstecher *et al.*, 2005; Giot *et al.*, 2003; Stanyon *et al.*, 2004), bacteria (Bartel *et al.*, 1996; Rain *et al.*, 2001), *Homo sapiens* (Rual *et al.*, 2005; Stelzl *et al.*, 2005), and *Arabidopsis thaliana* (de Folter *et al.*, 2005; Popescu *et al.*, 2007). As mentioned before, AP-MS methods detect the presence of primary or secondary interactions within a complex, whereas two-hybrid systems measure direct binary interactions (Figure 1).

Most of the detection systems are based on the reconstitution of a function of the two halves of a split protein. The canonical Y2H system consists of two components: the DNA-binding domain (DBD) from a transcription factor (generally Gal4 or LexA) fused to protein 'X', and the transcription activation domain (TAD; generally Gal4 or B42) fused to protein 'Y' (Figure 2a). When both chimeric proteins are co-expressed and localized to the nucleus and if protein 'X' interacts with protein 'Y', they reconstitute a functional transcription factor that activates transcription of marker genes in the nucleus. Since the first Y2H system developed by Fields and Song (1989), several modifications have been made to improve the quality of the data, including



**Figure 2.** Comparison of the mating-based split ubiquitin (mbSUS) and yeast two-hybrid (Y2H) systems.

(a) Classical Y2H system: a transcription factor is split into the activation (green arrow) and DNA-binding domain (blue zigzag line), which are fused to proteins 'X' and 'Y'. Interaction between the two partners will lead to activation of transcription of several reporter genes (*Ade2*, *His3*, and *LacZ*).

(b) The mbSUS: when protein 'X' (red or brown sphere) and protein 'Y' (purple or blue sphere) interact, a functional ubiquitin protein is reconstituted from the two domains (Nub, green hollow sphere; and Cub-PLV, green sphere with protrusion and the artificial transcription factor PLV as blue DNA binding and activation zigzag arrow). The protein fused to the Cub-PLV chimera must either be an integral transmembrane protein (purple sphere, top left), a peripheral membrane protein (blue, center), or attached to the membrane, e.g. by a lipid anchor (purple sphere with anchor, top left), otherwise Y-Cub-PLV can enter the nucleus and create a report in the absence of an interaction partner. In all three cases, interaction will lead to reconstitution of a 'functional' ubiquitin that will be recognized by endogenous ubiquitin-specific proteases (UBPs), leading to release of the artificial transcription factor PLV (blue zigzag arrow). The transcription factor, which contains a nuclear localization signal, will enter the nucleus, bind to operators (uas) to activate the transcription of several reporter genes (*Ade2*, *His3*, and *LacZ*).

the insertion of upstream activation sequences (UAS) into different promoters of the marker genes, use of low copy plasmids, implementation of multiple markers (*URA3*, *HIS3*, *ADE2*, *lacZ*, *GFP*), and use of negative selection of *de novo* autoactivators, for example the counter-selectable reporter CYH2 (for comprehensive reviews on Y2H see Vidalain *et al.* (2004) and Vidal (2005)).

Alternatively, protein complexes can be purified and analyzed by AP-MS complementing large-scale PPI datasets obtained by Y2H (Cusick *et al.*, 2005). Gavin *et al.* (2006) analyzed PPI in *S. cerevisiae* using a tandem affinity

purification (TAP) tag consisting of a calmodulin-binding domain, a protease cleavage site (TEV), and a protein A tag fused to 5500 ORFs. All C-terminal fusions were introduced into yeast by homologous recombination in order to express the tagged proteins under their native promoter in the native chromosomal environment. The tagged protein is then isolated by affinity purification along with its interacting partners, and their identities are determined by mass spectroscopy. As pointed out above, proteins from different compartments may associate in the extract, leading to false positives. Obviously, no ideal HT tool exists at present,

making extensive follow-up necessary to demonstrate the *in vivo* existence and relevance of interactions detected by any of the methods described here.

*The mating-based split-ubiquitin system (mbSUS) for membrane protein interactions*

To circumvent the problems associated with the analysis of membrane proteins using the classical Y2H, the mbSUS was developed (Figure 2b; Miller *et al.*, 2005; Obrdlik *et al.*, 2004). The split-ubiquitin system is similar to the classical Y2H as it uses yeast as a heterologous system and has a similar read-out, but it allows the detection of interactions of membrane proteins. The interaction must occur at the cytosolic face of any of the yeast membranes, including the nuclear envelope, endoplasmic reticulum (ER), Golgi, vacuole, mitochondria, and plasma membrane.

The system solves the problem of a classical Y2H that a transcription factor will be non-functional when fused to a membrane protein but still makes use of reconstitution of two halves of a split protein, here ubiquitin (split-ubiquitin system, SUS). The concept of SUS relies on the release of a transcription factor from the membrane if two membrane proteins interact. The SUS uses a ubiquitin split into two halves: The N-terminal domain of ubiquitin (Nub) can reconstitute a functional ubiquitin when co-expressed with its other C-terminal half (Cub) (Johnsson and Varshavsky, 1994). Nub mutants such as NubG (containing mutation Ile13Gly) with reduced affinity to Cub reconstitute the full-length ubiquitin only when brought into its vicinity via interaction of the two fusion partners. In SUS, protein 'X' is fused to the NubG and protein 'Y' is fused to the Cub fused to an artificial transcription factor composed of a tag (IgG-binding domains of *Staphylococcus aureus* protein A), LexA DNA-binding domain, and the activation domain of VP16 (PLV) (Stagljar *et al.*, 1998). When 'X' interacts with 'Y', the Nub and Cub moieties are brought together and a functional ubiquitin molecule can be reconstituted, triggering action of endogenous ubiquitin-specific proteases, thus cleaving the reconstituted ubiquitin from their fused membrane proteins and releasing the transcription factor PLV into the cytosol. The transcription factor diffuses into the nucleus where it activates transcription of marker genes (Figure 2b). The SUS system was further improved to make it amenable for HT by using a mating approach to bring together bait and prey in one cell (mbSUS; Miller *et al.*, 2005; Obrdlik *et al.*, 2004).

*Practical considerations for mbSUS analysis*

The first critical step for mbSUS analysis is cloning of the membrane protein ORF. As mentioned above, membrane proteins are often toxic when expressed in *E. coli*, and the first step for the generation of Nub/Cub fusions is the cloning in *E. coli* (Frommer and Ninnemann, 1995). There are two

ways around this problem: (i) cloning PCR products directly in yeast by using *in vivo* recombination (Miller *et al.*, 2005; Obrdlik *et al.*, 2004) or (ii) using secure *E. coli* vectors as typically used in the Gateway™ technology (Invitrogen, <http://www.invitrogen.com/>). *In vivo* cloning in yeast is faster since it eliminates an intermediate cloning step in *E. coli* because yeast is directly co-transformed with the PCR product (having homologous overhang) and the linearized vector. However, it requires additional steps to verify the sequence of the PCR-derived inserts in both the NubG and Cub-PLV vectors. The Gateway™ technology takes advantage of the commercially available Gateway™ entry vectors which carry multiple *rrnB* sequences acting as transcriptional termination signals upstream of the insertion site, therefore reducing read-through expression and thus toxicity in *E. coli*. Alternatively, usage of an *E. coli* strain reducing the plasmid copy number may also reduce potential toxicity (e.g. CopyCutter from Epicentre Biotechnologies, <http://www.epibio.com/>).

The currently available vectors for mbSUS (available from the Arabidopsis Biological Resource Center (ABRC); <http://www.biosci.ohio-state.edu/~plantbio/Facilities/abrc/index.html>) offer the possibility of cloning the gene of interest into either low- or high-copy plasmids (X-NubG or NubG-X) or a low-copy plasmid (Y-Cub-PLV). Therefore, the expression level (in yeast) of fused proteins can be manipulated via choice of the copy number of the plasmid. The use of a methionine-regulated promoter further expands the control over the expression level of Y-Cub-PLV. An HT pilot screen (Obrdlik *et al.*, 2004) indicated that the low-copy vector provides more stringent conditions; while the use of high-copy plasmids establishes less stringent conditions but offers higher sensitivity (Grefen *et al.*, 2007).

An important feature to consider before cloning ORFs into the mbSUS vectors is the topology of the membrane protein. In order to produce a read-out in the nucleus, the Cub-PLV and NubG fusions both must be present in the cytosol. Thus fusions have to be made accordingly, and at present, the system does not allow for analysis of proteins in which both N- and C-termini are located inside an organelle or outside the cell. For large-scale screens, structural information and prediction tools may be used to evaluate the potential topologies, as has been done systematically for Arabidopsis and rice (*Oryza sativa*) membrane proteins (<http://aramem.non.botanik.uni-koeln.de>). The detector domains must be fused to either a cytosolic N- or C-terminus; suitable vectors for N- or C-terminal fusions are available (<http://www.associomics.org/>). Obviously, Cub-PLV must be fused to an integral or membrane-associated protein (Figure 2b); a soluble protein fused to the Cub-PLV moiety would diffuse into the nucleus without having to interact with the NubG fused protein and activate the transcription of the markers leading to a false-positive read-out. Since membrane protein predictions are not sufficiently accurate, especially when

only a single hydrophobic domain is present or when a protein contains a hydrophobic leader peptide (Xia *et al.*, 2006), it is necessary to control for activation of the reporters in the absence of NubG. It is possible to artificially add a transmembrane domain or a membrane anchor to also allow the use of soluble proteins as Cub-PLV fusions. In contrast, both membrane or soluble proteins can be used as NubG fusions, enabling us to test not only the complement of membrane protein/membrane protein interactions, but also to target the interface between the membrane and the cytosol, which includes most of the interactions important for the initial steps of signaling cascades.

#### Potential pitfalls

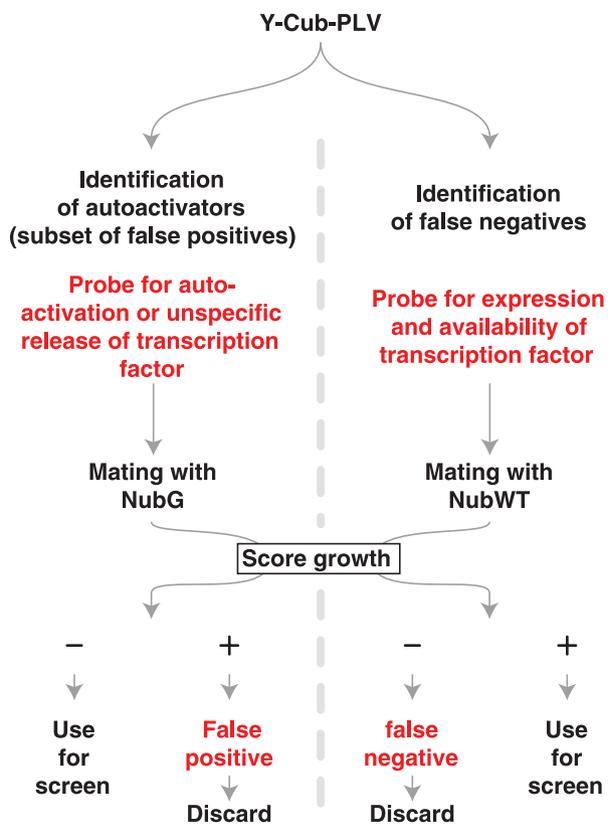
A major source of false-positive output of the Y2H system is caused by activation of transcription of the reporter genes by the protein fused to the DBD independently of the protein fused to TAD (Rual *et al.*, 2005). A second source of false-positives in HT datasets is *de novo* autoactivators, which may represent up to 10% of the baits (protein fused to the DBD). *De novo* autoactivators emerge during the course of the screen by spontaneous mutations (Rual *et al.*, 2005). Strategies for reducing artifacts include verification that several or all reporters score positive, by counter-selection of CYH2-containing vectors on cycloheximide (currently not implemented in the mbSUS), or by addition of 3-amino-1,2,4-triazole (3-AT) to increase the stringency of HIS3 selection (3-AT is a competitive inhibitor of imidazole-glycerol-phosphate dehydratase, the *HIS3* gene product). Furthermore, the reliability of the data can be increased significantly by retesting the original ORF-fusion clones to exclude mutations during the selection phase.

In the case of the mbSUS, Y-Cub-PLV fusions comprise a functional transcription factor; thus if the fusion protein is not prevented from diffusion into the nucleus via attachment to a membrane, transcription of the reporter genes will be activated in the absence of X-NubG. For example proteins that contain a hydrophobic core but no membrane domain may bioinformatically be classified as membrane proteins by mistake, thus mbSUS analysis will yield a positive read-out in the absence of NubG. In addition, false positives may arise from proteolysis of the fusion and release of the PLV transcription factor by unknown processes (e.g. the quality control mechanisms in the ER). Strategies to eliminate both types of false positives include testing for reporter activity in the absence of NubG or mating of each Y-Cub-PLV fusion with soluble free NubG. False negatives may arise from low abundance of the Y-Cub-PLV due to low expression or low stability of the fusion protein, or to a lack of accessibility of the PLV to ubiquitin protease cleavage. In contrast to the I<sup>13</sup>G mutant (NubG), the wild-type N-terminal ubiquitin domain (NubWT) can readily interact with the C-terminal ubiquitin domain. Thus co-expression of the Y-Cub-PLV

fusion with NubWT may be used to test for Y-Cub-PLV expression and PLV accessibility without the need for fused proteins to interact (Figure 3). In certain cases, fine tuning of the expression of the Cub-PLV fusion can be achieved by using the methionine-repressible promoter and titration using different methionine concentrations in the medium and/or using 3-aminotriazole, or using other Nub affinity mutants (Raquet *et al.*, 2001) will allow optimization of the selection conditions for autoactivators or for clones showing low expression.

#### Strategies for HT analysis

The large number of assays that need to be performed in order to determine the whole complement of potential protein interactions require HT technologies. For Y2H, one approach is to use standard liquid handling robotics (Miller *et al.*, 2005). The number of individual assays to be



**Figure 3.** Scheme for controls to determine false-positive and false-negative read-out from mating-based split ubiquitin system (mbSUS) analyses.

Protein 'Y' (which is putatively inserted or anchored in one of the cellular membranes) is fused to Cub-PLV. In a first step, some of the false-positive data can be eliminated by testing for reporter activity (HIS3 prototrophy, red chromophore formation for ADE2 or lacZ activity) either in the absence of NubG or by mating with yeast cells expressing soluble NubG (without a fusion partner). In a second step, functional expression of the Cub-PLV fusion as well as accessibility of Cub-PLV for interaction is tested by mating with cells expressing the soluble wild-type version of Nub (NubWT), which interacts with Cub-PLV fusions in the absence of an interaction partner for protein 'Y'.

performed can be reduced by 2D and 3D pooling strategies (Jin *et al.*, 2006; Rual *et al.*, 2005; Zhong *et al.*, 2003). One potential pooling strategy for mbSUS screens would be to create subpools of the Cub library, and to perform a first set of screens of Cub subpools interacting with individual NubG proteins, followed by a round of deconvolution for each pool yielding a positive read-out. This requires testing each NubG fusion against the individual members of the pools it was found to interact with. For mbSUS screens, the optimal pool size was estimated to be about five.

An alternative pooling strategy was developed for screening the human genome interactome (Rual *et al.*, 2005). In this approach, each bait is mated to individual pools of 188 TAD-ORFs in a microplate and positive colonies were identified from each pool. Subsequently, the positive colonies were retested and sequenced to identify the interacting prey (TAD-ORF). Potential interactors were then again retested from original clones (Rual *et al.*, 2005). The choice of the pooling strategy depends on the number of interactions detected in a matrix as well on the promiscuity of protein interactions in the collection and thus has to be designed on the basis of pre-screens for each individual system.

A disadvantage of many of the large-scale Y2H screens is that the output is not quantitative but rather a visual binary score of prototrophy. The information content of Y2H screens could be improved by determining growth curves quantitatively using a fluorescent marker such as GFP. While the reporter output may not necessarily correlate with the affinity of the underlying interaction, the quantitative data may help to reduce artifacts and improve standardization over multiple assays performed over the data collection period. Titration of the promoters may provide additional insights into properties of the interactions.

### Fluorescence spectroscopy and imaging technologies for analysis of protein–protein interactions

Neither *in vitro* tests nor Y2H provide data on the interaction of proteins in their native environment. New imaging technologies, coupled with the development of genetically encoded fluorescent proteins (FPs) and the increasing capability of software for image acquisition and analysis, have enabled *in vivo* studies of protein functions and processes. Genetically encoded FPs are at the core of a variety of approaches to probe PPI in living cells (Table 2). The most

**Table 2** Imaging-based methods for detecting protein–protein interactions by fluorescence resonance energy transfer (FRET)

Method	Specific potential	Important points
Filter-based fluorescence intensity ratio imaging	Simple system allowing for fluorescence bleed-through corrections Image captures (sensitized emission) is rapid and very suited for time-lapse or 3D	Need to capture image from a sample and two reference images (donor alone and acceptor alone) Stoichiometry of donor and acceptor is difficult to establish in live cell imaging Pixel shifts caused by filter change If sensitized emission is captured: correction for bleed-through need to be done either by following capturing emission from controls or by spectral un-mixing Measures are equipment-specific, determination of FRET requires calibration (Vogel <i>et al.</i> , 2006)
Ratio imaging by spectral unmixing	Requires either a confocal microscope with spectrophotometric capacity or multiple bandpass filters	Algorithm for spectral unmixing may not be implemented appropriately when FRET occurs in the sample
Acceptor photobleaching	Simple system; can be performed on a wide-field system Best when combined with an independent second method	Long bleaching time in wide-field microscopy can induce phototoxicity and reduce cell viability Does not correct for fluorescence bleed-through Destructive, non-dynamic Physical properties of fluorescent proteins may compromise FRET measurements
Fluorescence lifetime imaging (FLIM)	Considered the gold standard for FRET analysis of protein interactions Independent of fluorophore stoichiometry Can determine dynamic interactions	Both the time-domain and the frequency-domain mode require specialized equipment Frequency domain: assumption is that the donor has a single exponential decay which is not necessarily the case with biological samples (i.e. cyan fluorescent protein lifetime fits better a double exponential) Time domain: requires long exposures which can cause photodamage, thus typically requires fixation which may lead to artifacts
Anisotropy and anisotropy decay	Can measure homo-FRET and thus permits multiplexing	

popular methods are (i) co-localization of two labeled proteins, (ii) FRET measurements where protein 'X' is fused to a donor FP while protein 'Y' is fused to an acceptor FP, and (iii) protein-fragment complementation assays (PCA) consisting of a split protein that reconstitutes a function upon interaction of protein 'X' and protein 'Y' fused to the different moieties of the split protein (Figure 4).

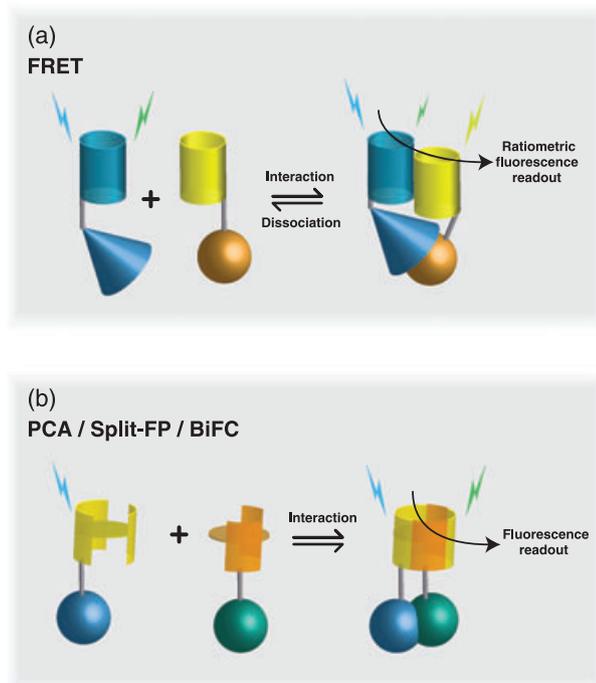
#### Fluorescence resonance energy transfer (FRET)

Fluorescence resonance energy transfer refers to a quantum mechanical effect between a given pair of fluorophores, i.e. a fluorescent donor and an acceptor, where, upon excitation of the donor, energy is transferred from the donor to the acceptor in a non-radiative manner via dipole-dipole coupling (resonance) (Förster, 1948; Jares-Erijman and Jovin, 2003). As a result of FRET between a donor and acceptor a

portion of the energy absorbed by the donor is emitted in a spectral window that is characteristic of the acceptor. Fluorescence resonance energy transfer is characterized by the efficiency of the energy transfer,  $E$ , which is defined as the fraction of the photons absorbed by the donor and transferred to the acceptor (Figure 4a).  $E$  is a function of the inverse sixth power of the distance ( $r$ ) between the two fluorophores [ $E = R_0^6 / (R_0^6 + r^6)$ ]. The distance at which energy transfer is 50% is known as the Förster distance ( $R_0$ ) and is a unique property of a given FRET pair.  $R_0$  depends on the extent of spectral overlap (overlap integral) between donor emission and acceptor absorption ( $>30\%$ ;  $J(\lambda)$ ), the quantum yield of the donor ( $Q_D$ ), the refractive index ( $n$ ) of the medium, and the relative orientation of the dipole moment ( $\kappa$ ) of the donor and acceptor:  $R_0 = 9.78 \times 10^3 [(\kappa^2 n^{-4} Q_D J(\lambda))]^{1/6}$  (Jares-Erijman and Jovin, 2003; Lakowicz, 2006). Because of its exquisite dependence on molecular distance, FRET has been described as a molecular ruler (Stryer, 1978), which operates in the range of 1–10 nm; a distance relevant for most molecules engaged in complex formation or conformational changes. Although the contribution of the dipole orientation compromises FRET as an accurate measure of molecular distance, FRET is capable of resolving molecular interactions and conformations with a spatial resolution exceeding the inherent diffraction limit of conventional optical microscopy (Jares-Erijman and Jovin, 2003).

#### The advantages of FRET over co-localization

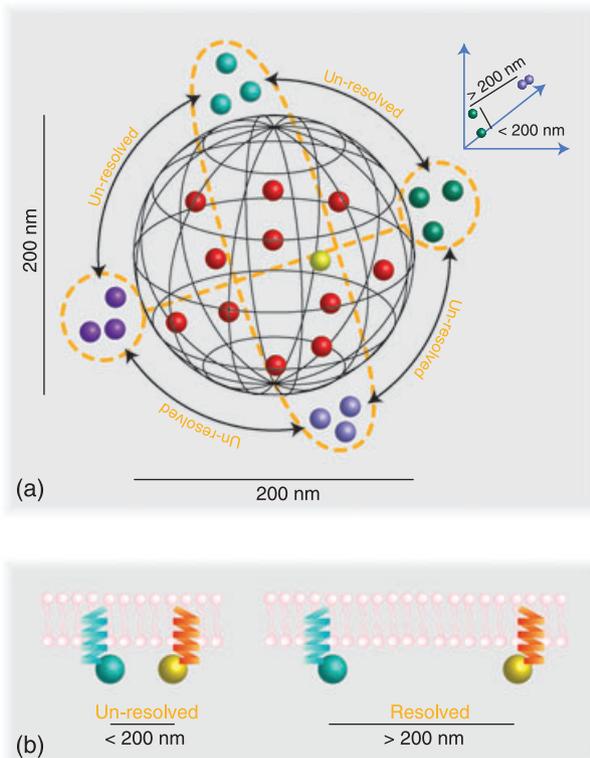
Resolution is defined as the smallest distance between two points within an image, which can be separated and still be distinguished. Resolution depends on the wavelength of the light imaged and on the numerical aperture (NA) of the objective ( $1.22\lambda/2NA$  for wide-field epifluorescence). For a high-NA objective, the resolution is thus about 200 nm at best. If we assume a sphere with a diameter of 200 nm (corresponding to a volume of  $4.2 \times 10^{-3}$  fl; Figure 5a), it could contain up to 140,000 densely packed GFP molecules [volume of a single GFP molecule is about  $30 \times 10^{-9}$  fl; a very similar discussion is presented by Vogel *et al.* (2006)]. This simple calculation demonstrates that there are many ways that two proteins could be contained within a co-localized volume without physically interacting. Fluorescence resonance energy transfer increases the apparent resolution by restricting the volume occupied by two interrogated fluorophores. The FRET volume as calculated by Vogel *et al.* (2006) is  $4 \times 10^{-6}$  fl, which is much smaller than the optical resolution volume but still large enough space to contain about 100 GFP molecules. Thus, co-localization even by FRET may be one of several pieces of circumstantial evidence for a PPI, but on its own is insufficient to conclude that two proteins are in a complex. To illustrate this, two proteins may reside in the same vesicle, one in the lumen and the other on the surface, so they will appear



**Figure 4.** Comparison of fluorescence resonance energy transfer (FRET) and protein fragment complementation assay (PCA) [split fluorescent protein (FP) or bimolecular fluorescence complementation (BiFC) methods].

(a) FRET: when two proteins 'X' (blue cone) and 'Y' (orange ball) are in sufficiently close vicinity (2–8 nm), e.g. in the case of an interaction between the two fusion partners (direct or indirect), resonance energy transfer will occur between the donor fluorophore (blue cylinder) and the acceptor fluorophore (yellow cylinder).

(b) In the PCA (also named split-FP or BiFC), a fluorescent protein, e.g. GFP or Venus, is split into two halves (yellow and orange half-cylinders). When expressed separately, the split-FP halves do not form a functional fluorophore. However, when the fused proteins 'X' and 'Y' interact, a FP is reconstituted creating a stable and quasi-irreversible complex generating a functional fluorophore. Irreversibility is advantageous for sensitivity, but increases the possibility of artifacts, especially when the fused proteins are overexpressed.



**Figure 5.** The optical path resolution is critical for the interpretation of imaging-based interaction analyses.

(a) Objects (e.g. GFP molecules) are shown in a three-dimensional space at a scale corresponding to the maximal optical resolution of 200 nm for conventional fluorescence microscopy. Since the volume of a pixel is significantly larger than the volume of a single GFP molecule, co-localization within a pixel cannot be used as a proof of a protein–protein interaction (Vogel *et al.*, 2006). Thus the lower the resolution, the lower the confidence will be. At the highest resolution, none of the objects (red and yellow) within the sphere (black meshed structure) can be resolved, whereas objects at opposite sides of the sphere are resolved. Due to the form of the point-spread function, objects in different z-planes will appear to be co-localized. Apparently, the higher the expression, the higher the possibility of false-positive data.

(b) A similar limitation applies to membrane proteins; however, here proteins are limited to a two-dimensional space (Fung and Stryer, 1978). In this case the reduced degrees of freedom and diffusion in a plane increase the chance of random collisions and produce a positive read-out for proteins that normally do not interact. This also applies to split fluorescent protein approaches.

co-localized, but they do not interact. Or, one protein may be evenly distributed in the plasma membrane, whereas the other is highly localized to puncta, where it physically interacts with the evenly distributed protein. In both cases, additional evidence for interaction/non-interaction is required.

There is a 'critical concentration' of the acceptor at which chance diffusion alone will place an acceptor within an  $R_0$  distance of a donor. For a cyan fluorescent protein (CFP)/yellow fluorescent protein (YFP) FRET pair, with an  $R_0$  of 5.4 nm, the 'critical concentration' of YFP is about 2.8  $\mu\text{M}$  and this corresponds to only about 6.7 YFP molecules in the FRET volume. Due to the increased chance of random collisions in a two-dimensional space, this is an even more

important consideration when we consider interactions that occur in a restricted plane such as a membrane (Figure 5b) or in a restricted volume inside an organelle. These simple calculations demonstrate that there is a need to carefully evaluate positive FRET results to exclude the possibility that FRET is due to random collisions. Energy transfer efficiency can be estimated fairly easily (see below) and can be calibrated (Koushik *et al.*, 2006; Thaler *et al.*, 2005; Vogel *et al.*, 2006). The observed values should be compared with a theoretical value for the donor–acceptor pair to evaluate the potential of random collisions. In addition, independent methods may have to be applied to verify interactions suggested by positive FRET results. The need for appropriate controls also applies to negative FRET results. Low transfer efficiency may be caused by the absence of a molecular interaction, by a stoichiometry of donor-to-acceptor other than 1:1 (Vogel *et al.*, 2006), or by excitation of the acceptor in the donor excitation channel (bleed-through).

#### Methods for FRET determination and analysis of FRET changes

Most methods evaluate energy transfer efficiency as the relative fluorescence intensity of the donor FP in the presence or absence of the acceptor FP. The most popular methods employed are: (i) filter-based FRET (ratio-imaging/sensitized emission), (ii) spectral imaging, (iii) acceptor photobleaching, (iv) lifetime measurements (fluorescence lifetime microscopy, FLIM), and (v) a combination of the above (Table 2). The various measurement modes were recently reviewed by Jares-Erijman and Jovin (2003, 2006).

Filter-based FRET acquires fluorescence intensity of the donor (excitation- and emission-specific filters), acceptor (excitation- and emission-specific filters) and acceptor-sensitized emission (excitation of the donor and capture of acceptor emission) by using either two-filter or three-filter configurations. Filter-based FRET is probably the most problematic method if FRET is not intramolecular (when the donor is not physically linked to the acceptor, a case that applies to most PPI studies) because it requires acquisition and registration of multiple images, correction for spectral bleed-through as well as knowledge of the stoichiometry of donor and acceptor. In many studies, only the sensitized emission (emission of the acceptor after excitation of the donor) is measured. However, when the donor is not physically linked to the acceptor, this method neither considers bleed-through (donor emission passed through the acceptor emission filter and direct excitation of acceptor by donor excitation filter) nor the concentration of the FPs. A better technique involves acquisition of multiple images from different samples expressing donor alone, the acceptor alone or both the donor and the acceptor: (i) donor excitation–donor emission, (ii) acceptor excitation–acceptor emission, (iii) donor excitation–acceptor emission. This

normalization provides for stringent correction for bleed-through fluorescence (Berney and Danuser, 2003). It also allows for estimation of donor/acceptor stoichiometry and for the presence of FRET and non-FRET signals in each acquired image (Berney and Danuser, 2003; Gordon *et al.*, 1998). The method relies, however, on the assumption that the cellular concentration of the samples expressing donor alone is the same as the concentration of the donor in cells co-expressing donor and acceptor. If the same promoters are used for expression of donor and acceptor, competition for limiting factors may occur, leading to differences in donor expression levels. In live cell imaging, determination of donor levels in the presence or absence of acceptor may be difficult to achieve since it requires acquisition of three images from three different transformed lines; and expression levels will probably vary in their individual transformants.

An alternative to filter-based FRET measurements is spectral imaging followed by linear unmixing (Zimmermann *et al.*, 2003). Emission spectra for each pixel are acquired (e.g. using a confocal microscope equipped with spectral sensor or using a slit-scanning spectral system such as the SpectralDV) and deconvoluted by spectral unmixing to obtain the 'pure' emission for each fluorophore. This system has the additional advantage that it can correct for autofluorescence. Although this method corrects for spectral bleed-through and autofluorescence, depending on the algorithm used, it may underestimate the contribution of the donor and overestimate the acceptor when FRET occurs (Thaler *et al.*, 2005). Vogel and colleagues have implemented a method in which the emission spectra are captured at two different excitation wavelengths in order to calculate the contribution of energy transfer (Thaler *et al.*, 2005). Moreover, they implemented a FRET calibration system that can be applied to normalize data and make them comparable between different imaging systems (Vogel *et al.*, 2006).

Since in live cell imaging the proportion of the two fluorophores cannot always be established reliably, an alternative method for measuring FRET is to determine the emission intensity of the donor before and after acceptor photobleaching (also named acceptor photobleaching or donor dequenching). This method requires the donor to be relatively photostable while the acceptor is photolabile. Assuming that the donor is not affected by the light used to bleach the acceptor, emission from the donor increases after photobleaching of the acceptor when FRET occurs. With patterned illumination it is possible to photobleach a defined area of the sample making it possible to measure emission of the donor both with and without an acceptor in a single image. Acceptor photobleaching method is sensitive to incomplete bleaching of the acceptor; if the acceptor is bleached to only 30% of its original intensity, this can create an error of up to 50% in quantifying FRET (Berney and Danuser, 2003). When using a wide-field microscope with a

conventional mercury arc light source, bleaching times can be as long as 20 min; thus diffusion of FPs might be a significant problem and 100% photobleaching might be difficult to achieve. Furthermore, one has to consider phototoxicity and possible loss of cell viability. Intense laser light may reduce exposure times but present a similar if not greater hazard for phototoxicity. Photobleaching can be performed on fixed cells to avoid diffusion of the FPs from unbleached areas; however, fixatives such as formaldehyde have been shown to differentially quench FP fluorescence and fixation may cause artifacts due to cross-linking of proteins during the fixation process (Chen *et al.*, 2006).

An alternative to measuring fluorescence intensity to estimate FRET efficiency is fluorescence lifetime imaging (FLIM). This has the advantage of being largely independent of fluorophore concentrations as it measures the relaxation time of an excited fluorophore after a short pulse of excitation light rather than the number of photons emitted (Biskup *et al.*, 2007; Wouters, 2006). Fluorescence lifetime imaging is also independent of light scattering and refraction in different regions of the specimen (Bastiaens and Squire, 1999). The fluorescence lifetime corresponds to the average time a fluorophore remains in the excited state following excitation and shows a lifetime characteristic for a given FP (Lakowicz, 2006). Fluorescence decays exponentially. Fluorescence resonance energy transfer withdraws energy from the donor and thus leads to a reduction of its lifetime. Two methods are used for measuring fluorescence lifetime: time- and frequency-domain FLIM. Both methods are limited by the small number of photons recorded and thus may require high-intensity excitation light. Time-domain measurements use a pulsed laser for excitation and time-resolved image acquisition to quantify donor lifetime directly. The fluorophore is excited by femtosecond light pulses and the time at which photon arrives after each pulse is measured, yielding a histogram of decay times (Lakowicz, 2006; van Munster and Gadella, 2005). Advanced methods that distinguish FRET from other parameter changes such as quenching employ spectrally resolved fluorescence lifetime measurements with use of streak cameras or time-correlated single-photon counting (Biskup *et al.*, 2007). The frequency-domain method measures lifetimes indirectly using excitation of the sample/probe by continuous light with sinusoidally modulated intensity coupled with sinusoidally modulated detection. Lifetimes are then calculated as a function of phase and amplitude changes of the signal (Lakowicz, 2006). Time-domain FLIM has been successfully used in plant cells to demonstrate the interaction between the two receptor-like kinases BRI1 and BAK1 (SERK3, Russinova *et al.*, 2004), the AAA-ATPase CDC48A and the receptor-like kinase SERK1 (Aker *et al.*, 2006), and the AvrA10-dependent interaction between the transcription factor WRKY2 and the MLA10 receptor (Shen *et al.*, 2007). Typically, integration times for exposure for

these analyses were in the range of 60–120 sec. Frequency-domain FLIM has also been successfully used to analyze the interaction of G-protein subunits in plants (Adjobo-Hermans *et al.*, 2006). A potential drawback of frequency-domain FLIM is that decay is measured indirectly. This may be problematic, since typically only a fraction of the donor associates with an acceptor, leading to overlaid decay components for donors alone and donors in the FRET vicinity to an acceptor (Biskup *et al.*, 2007). These multicomponent decays can be readily resolved with time-domain analysis, while for frequency-domain analysis the decays need to be differentiated using acquisition at multiple frequencies of the modulation of the excitation light (Redford and Clegg, 2005). Typically, time- and frequency-domain methods are considered to be similar with respect to the detection of a given number of photons detected and are comparable over several orders of magnitude (Gratton *et al.*, 2003; Philip and Carlsson, 2003). Two factors need to be considered when choosing between the two systems: the need for speed of acquisition in the case of dynamic interactions versus the sensitivity required at low fluorophore concentrations. Time-domain FLIM has been reported to have a higher signal-to-noise ratio for dim samples but required extended integration, while frequency-domain FLIM may be advantageous when rapid image acquisition of brighter samples is required to study dynamic processes (Gratton *et al.*, 2003; Philip and Carlsson, 2003). Taken together, the recent development of new hardware for FLIM detection provides opportunities to localize and characterize PPIs efficiently.

#### Choice of fluorophores

The choice of the optimal FRET pair lies in the different physical properties of the FPs. Ideal FPs will have a high quantum yield, a high extinction coefficient, a large Stokes shift, good photostability, low sensitivity to the cellular environment (ionic interactions, pH), and, for FLIM, a suitable lifetime that can be measured with the available equipment. Furthermore, these features should be similar for both the donor and acceptor fluorophores, and the excitation and emission spectra of the FRET pair should be separated as far as possible. For recent reviews on FPs and their properties, see Shaner *et al.* (2004, 2005) and Dixit *et al.* (2006).

Another criterion for the choice of fluorophores for FRET is the extent of overlap between the emission spectrum of the donor and the excitation spectrum of the acceptor. On the one hand it is advantageous to obtain a FRET pair with a large spectral overlap, since the Förster equation says that transfer efficiency depends on the overlap of donor emission and acceptor excitation. On the other hand, depending on the Stokes shift of the two fluorophores, a large spectral overlap between donor emission and acceptor excitation can lead to contamination of the acceptor-sensitized

emission with the donor emission (bleed-through). For example, the CFP/YFP FRET pair commonly used for genetically encoded sensors displays a large spectral overlap; however, because the Stokes shift for YFP is small, significant bleed-through may be observed. Alternatively, a fluorescence filter set has to be used that is shifted towards longer wavelengths, reducing the fluorescence intensity since a smaller fraction of the YFP emission spectrum can be collected. The GFP-S<sup>65</sup>T/YFP pair, which has a larger overlap integral compared with CFP/YFP, yields a higher energy transfer efficiency. However, due to bleed-through, FRET cannot be measured reliably using filters but has to be acquired by determining the emission spectra followed by linear spectral unmixing (Zimmermann *et al.*, 2002). When the Stokes shift is small and there is significant overlap between the excitation and emission spectra of a single fluorophore, as in the case of YFP, homotransfer can occur. Homotransfer is defined as energy transfer between two identical molecules (Lakowicz, 2006). In summary, the ideal pair will have a large spectral overlap integral regarding donor emission and acceptor excitation, while excitation of the acceptor is minimal to permit maximal FRET and emission from the acceptor should be minimal in the donor channel.

Among the different available FPs, the CFP/YFP pair is the most popular for FRET measurements (Rizzo *et al.*, 2006). However, blue fluorescent protein (BFP)/GFP (Heim and Griesbeck, 2004; Mitra *et al.*, 1996), CyPet/Ypet (Nguyen and Daugherty, 2005), MiCy/mKO (Karasawa *et al.*, 2004), T-Sapphire/mOrange (Shaner *et al.*, 2004), cerulean/YFP (Aker *et al.*, 2006, 2007), and mVenus (Venus-A<sup>206</sup>K)/mStrawberry (Adjobo-Hermans *et al.*, 2006) have also been used. Blue fluorescent protein was replaced fairly rapidly by CFP because of CFP's higher quantum yield and enhanced photostability. Furthermore, the excitation wavelength of BFP (380 nm) is more cytotoxic than CFP (434 nm) and autofluorescence of plant material below 420 nm is more likely to be a problem. Although CFP is a better donor than BFP it is still dimly fluorescent compared with YFP. Another disadvantage of CFP for lifetime FRET measurements is that its decay kinetic fits a double exponential, complicating interpretation of lifetime results (Rizzo *et al.*, 2006). The standard CFP and YFP are known to form weak dimers; thus it is recommended to use monomeric variants to exclude artifacts [EYFP-A<sup>206</sup>K (mYFP) and ECFP-A<sup>206</sup>K (mCFP), Zacharias *et al.*, 2002; ]. Alternatives to CFP are Cerulean (Rizzo *et al.*, 2004), Azurite (Mena *et al.*, 2006) and mTFP1 (Teal fluorescent protein, Ai *et al.*, 2006) all of which are brighter (higher quantum yield and higher extinction coefficient) and have single-component decay kinetics. Cerulean is more susceptible to photobleaching than CFP, while mTFP1 appears to be as photostable as EGFP, making it the most stable cyan FP. mTFP1 has the additional advantage of having its excitation maximum at 462 nm and can thus be

excited with lasers commonly installed on confocal microscopes. Similarly, mCitrine (mYFP Q<sup>69M</sup>) and Venus are better alternatives to YFP since they are less sensitive to the ionic conditions, including changes in pH or chloride within the physiologically relevant range (Griesbeck *et al.*, 2001; Nagai *et al.*, 2002).

#### *Practical considerations for FRET measurements*

The proper use of FRET efficiency measurements to characterize molecular interactions requires that correction be made for: (i) bleed-through fluorescence (excitation of the acceptor fluorophore through the donor excitation filter and donor emission signal through the acceptor emission filter) and (ii) stoichiometry of donor and acceptor fluorophores. The simplest case for FRET measurements is if the fluorophores are covalently coupled, as is the case with genetically encoded FRET sensors (Lalonde *et al.*, 2005). In this case it may not even be necessary to correct for bleed-through (Deuschle *et al.*, 2006; Fehr *et al.*, 2005; Gu *et al.*, 2006; Okumoto *et al.*, 2005). Whenever the relative levels of the FRET partners are not equimolar, FRET measurements are problematic. Correction techniques for the relative levels of donor and acceptor fluorophores have been developed (Berney and Danuser, 2003; Gordon *et al.*, 1998). The rule described above for co-localization, i.e. that the concentration of donor and/or acceptor should not reach the 'critical concentration', e.g. about 2.8  $\mu\text{M}$  for soluble CFP/YFP FRET pairs, and a significantly lower value for the 'critical concentration' for membrane proteins also applies to all FRET measurements. Besides the proper choice of FP pair for FRET, the excitation, dichroic, and emission filters need to be chosen carefully to maximize excitation and emission and minimize bleed-through (Shaw, 2006).

Traditionally, FRET has been used as a spectroscopic ruler to measure the distance between two sites on a protein, such as intrinsic tryptophan acting as donor and an acceptor dye covalently bound to the protein (Stryer, 1978). The widespread availability of spectrofluorometers and fluorescence microscopes together with the advent of genetically encoded fluorophores has made FRET measurements comparatively accessible and affordable. Energy transfer efficiency can be estimated fairly easily and be compared with a theoretical value for the donor-acceptor pair to evaluate the potential for random collision. In most cases, rather than measuring the actual FRET efficiency a proxy is determined, in the simplest case the ratio of the peak emission of the two fluorophores (Vogel *et al.*, 2006). A number of factors affect the apparent energy transfer, such as the relative dipole orientation (if it is not random;  $\kappa^2 \neq 2/3$ ), the MOLAR ratio of the partners (if different from 1:1), protein concentrations above the critical level, and technical problems such as bleed-through. While the theoretical achievable transfer efficiency may be 50%, the measured efficiency could be

10%; thus extensive additional information is needed before conclusions on the molecular interaction can be drawn. Considering photophysical properties of FPs and the possibility of homotransfer (if donor concentration reaches the 'critical concentration'), it is good practice to verify that donor quenching upon FRET also results in an increased fluorescence intensity of the acceptor or in a decrease in the lifetime of the donor (Subramaniam *et al.*, 2003), or to obtain actual values of FRET for different FRET pairs by using calibration systems as developed by Vogel's group (Koushik *et al.*, 2006; Thaler *et al.*, 2005). The different methods for measuring FRET efficiency have been compared using different set-ups and microscope/electronic configurations (Gordon *et al.*, 1998; Koushik *et al.*, 2006; Pelet *et al.*, 2006; Rizzo *et al.*, 2006). Taken together, PPI analysis using fluorophore-based assays requires careful control of a number of parameters to exclude artifacts.

#### *Bioluminescence resonance energy transfer (BRET)*

Bioluminescence resonance energy transfer (BRET) is similar to FRET in that energy transfer occurs between a donor and acceptor (Subramaniam *et al.*, 2006; Xu *et al.*, 1999). The method is suitable for the analysis of protein interactions in extracts and is suitable for imaging at least at lower resolution (Xu *et al.*, 2007). The major difference of BRET lies in the RET donor, which in the case of BRET makes use of luciferase which catalyses the oxidation of luciferin to emit light. The energy of the reaction can be transferred by RET to an acceptor (e.g. GFP or YFP) if luciferase and the fluorophore are within a radius of 50 Å. The most frequently used BRET pairs are coelenterazine/GFP (or YFP) or the DeepBlueC™/UV-GFP (though DeepBlueC does not appear to work in plants; Subramaniam *et al.*, 2004). Because this reaction occurs in the dark, it does not require excitation light, hence there is no risk of photodamage, no acceptor photobleaching, no fluorescence bleed-through, and, due to the lack of excitation, no problems caused by autofluorescence of the sample. Despite the apparently higher sensitivity of BRET over FRET, the emission generated is limited in intensity, requiring long integration times; thus while BRET can be visualized at the tissue and cellular levels with a sensitive camera (e.g. modified electron bombardment CCD) it cannot be used for analyzing dynamic interactions since exposure times are prohibitive (Xu *et al.*, 2007).

#### *Fluorescence correlation spectroscopy (FCS)*

Fluorescence correlation spectroscopy (FCS) measures fluctuations in fluorescence intensity caused by the diffusion or conformational changes of fluorescently labeled molecules in a small interrogated volume, typically created by a confocal microscope (Lakowicz, 2006). Fluorescence correlation spectroscopy can be used to measure several prop-

erties of a labeled molecule including the number of molecules in the interrogated volume, their diffusion rate, flow rate, aggregate formation, and rotational dynamics (with polarized light) (Schwille *et al.*, 1999). In a typical application, as a diffusing fluorophore moves into the interrogated volume a burst of photons begins due to multiple cycles of excitation and emission, and ends when the fluorophore leaves the interrogated volume. The duration of bursts is correlated with the diffusion rate. Although a confocal microscope is used to create the excitation volume (interrogated volume) in which the single molecule is observed, this method is not an imaging technique that provides spatial information in a living cell; it is rather used to study molecular interactions *in vitro* and *in vivo* (Goedhart *et al.*, 2000; Hink *et al.*, 2003; Köhler *et al.*, 2000).

The diffusion rate depends on the size of the molecule and its interaction with other molecules; this dependence makes FCS a valuable method for measuring a wide range of binding interactions such as PPI. Because the diffusion time scales with the cubic root of molecular mass, interaction of binding partners must result in a significant increase in mass to be detected by FCS. In fact, dimerization, or simple doubling of mass, is difficult to detect as it causes an increase in the diffusion coefficient of only 26% (Bacia *et al.*, 2006; Lakowicz, 2006; Meseth *et al.*, 1999). Thus, the smaller of the binding partners should always be the labeled component to maximize the increase in mass upon interaction and binding. An alternative methodology, which can be applied to interactions that result in small changes in mass, is fluorescence cross-correlation spectroscopy (FCCS; Schwille *et al.*, 1997). In FCCS interacting partners are labeled with different fluorophores and the intensity fluctuations of the two species are cross-correlated. If the two molecules interact, their intensities will tend to fluctuate together. It should be noted that FCCS does not measure the physical interaction of two partners but calculates the probability of having both partners in the same restricted volume at the same time.

#### *Practical considerations for FCS measurements*

Fluorescence correlation spectroscopy measurements are affected by light scattering, autofluorescence, photobleaching of the fluorophores, the detection limit of the microscope, and photodamage of living material (Schwille *et al.*, 1999). As is the case with all fluorescence microscopy, the power of the excitation energy must also be adjusted carefully to find the best compromise between signal detection on the one hand and photobleaching/phototoxicity on the other. Light scattering at the cell wall can be a special problem when analyzing plant cells. Two-photon fluorescence excitation has several advantages for *in vivo* analysis in plants: (i) a smaller interrogated volume reduces both light scattering and phototoxicity and (ii) longer excitation

wavelengths can reduce phototoxicity and improve depth penetration (Schwille *et al.*, 1999).

The amplitude of fluctuations in fluorescence intensity is inversely proportional to the number of molecules measured at the same time, meaning that high concentrations will diminish the effect of fluctuations and result instead in the measurement of an average intensity (Müller *et al.*, 2003). For plant cells, Hink *et al.* (2003) used concentrations below 5  $\mu\text{M}$ , while Schwille *et al.* (1997) recommend the use of concentration ranges of 1–100 nM. As a consequence of the low concentration of the proteins used, interactions can only be detected if they have sufficiently high affinity for FCS/FCCS. It is important to note that FCS/FCCS may be limited by interactions where diffusion is relatively rapid. For example, proteins in the membrane may diffuse slowly and not transit the interrogation volume frequently enough to provide reliable fluctuations over experimentally practical integration periods. In live cells, cytoplasmic streaming (which is particularly rigorous in plant cells) may violate assumptions of diffusion being the chief determinant of fluorophore motion.

The choice of fluorescent proteins for FCS applications in plant cells entails many of the same consideration that are true for FRET; increasing fluorophore detectability by using enhanced versions of GFP (Schwille *et al.*, 1999), CFP and YFP (Hink *et al.*, 2003), and S65TmGFP4 (Köhler *et al.*, 1997), and avoiding fluorescent protein variants that have a tendency to oligomerize on their own. However, by contrast to FRET, in FCCS it is important that the two fluorophores have non-overlapping spectra to avoid energy transfer and bleed-through fluorescence. In fact, if energy transfer (or photobleaching) occurs, the apparent diffusion coefficient will decrease. Other important considerations, such as proper alignment of the observation volumes for each wavelength for FCCS are outlined by Schwille *et al.* (1997, 1999).

#### *Protein fragment complementation assays (PCA)*

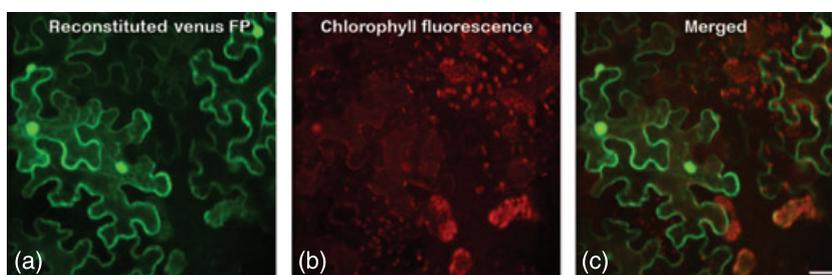
The two-hybrid systems described above are based on the concept that two separately expressed domains of a split protein cannot complement each other except if the local concentration is increased (see above discussion on 'critical concentration'). In the case of the classical Y2H system, a transcription factor is split genetically into its DNA-binding and transcription activation domains. The two domains do not appear to be able to reconstitute a functional transcription factor without the help of an interaction. The inability to reconstitute is most probably due to the absence of a stable interaction interface in the absence of the covalent linkage of the two subunits (see discussion above on 'structural basis of protein interaction') as is the case for split-GFP, split-ubiquitin, or split  $\beta$ -galactosidase. The two domains even do

not reassemble when a nuclear localization signal is added as in the case of Y2H vectors. It is also possible that the halves cannot fold by themselves, although the observation of diffractable crystals suggests that the activation domain can fold on its own (Chattopadhyaya and Pal, 2004). In contrast, the two halves of the much simpler ubiquitin do reconstitute a 'functional' ubiquitin on their own that is recognized by the ubiquitin-specific proteases (Johnsson and Varshavsky, 1994). Only by introducing mutations can the affinity (or the folding) of the two halves towards each other be reduced to make them suitable for interaction screens. It has been known for many years that when proteins are split into two polypeptide chains and these peptides are co-expressed in a cell, a significant portion can reconstitute to form a functional protein and this reconstitution from fragments has been used to study protein folding (de Prat-Gay, 1996). This phenomenon has also been used to study protein structure, folding, function, and evolution. One of the classical examples is  $\alpha$ -complementation of  $\beta$ -galactosidase in *E. coli* (Galarneau *et al.*, 2002; Rossi *et al.*, 1997; Spotts *et al.*, 2002). Other examples include the extracellular yeast invertase (Schonberger *et al.*, 1996), as well as lactose and sucrose transporters (Bibi and Kaback, 1992; Reinders *et al.*, 2002b; Wrubel *et al.*, 1994),  $\beta$ -lactamase (Galarneau *et al.*, 2002), luciferase (Paulmurugan *et al.*, 2002; Remy and Michnick, 2006), dihydrofolate reductase (DHFR; Pelletier *et al.*, 1998, 1999), TEV protease (Wehr *et al.*, 2006), and FP variants (Ghosh *et al.*, 2000; Hu *et al.*, 2002). Some complementation assays require the presence of an exogenous substrate (e.g.  $\beta$ -lactamase, luciferase,  $\beta$ -galactosidase, DHFR). The molecular mechanism(s) for the reconstitution of functional proteins from unfolded domains is largely unknown, and multiple folding pathways may exist. One may assume that the interfaces formed between the domains are sufficient to reconstitute a stable protein, or to induce folding of the two separate polypeptides post-translationally.

Similar to the cases described above, FPs can also be split and reconstituted, a process that may be driven in part by

$\beta$ -strand addition, a feature of a number of PPIs (Ghosh *et al.*, 2000; Wrubel *et al.*, 1994). However, reconstitution of split FPs is quasi-irreversible with a half-life time estimated at 10 years and cannot be used for dynamic studies (Magliery *et al.*, 2005). N-GFP and C-GFP (split at 157–158) do not seem to reassociate when combined at concentrations of 100  $\mu$ M in solution or when co-expressed in bacteria (Ghosh *et al.*, 2000). The reconstitution assays (BiFC or split-GFP) can be used to study a variety of processes such as protein folding (Cabantous *et al.*, 2005), and similar to other two-hybrid systems as a tool to detect protein interactions with subcellular resolution (Ghosh *et al.*, 2000). It has been proposed that the reconstitution from FP fragments has an advantage over FRET interactions studies by having a higher dynamic range, with practically no fluorescence in the absence of an interaction to high levels of fluorescence after fusion to interacting proteins. The absence of fluorescence of the non-interacting domains has been ascribed to the inability of the two separate domains to fold by themselves (Magliery *et al.*, 2005).

An implicit condition of these systems is that the two halves do not reconstitute by themselves and that only the interaction of protein 'X' with protein 'Y' (where X and Y are fused to the split domains of an FP, respectively) triggers the reconstitution of an FP. It is known, however, that the large N-terminal fragment of split-GFP can pre-form a chromophore under certain conditions (Demidov and Broude, 2006). Moreover, certain fragment combinations can spontaneously reconstitute in the absence of attenuators (Cabantous *et al.*, 2005). Moreover, when expressed from the strong CaMV35S promoter using the *Nicotiana benthamiana* transient infiltration system, the soluble Venus halves yield significant fluorescence levels even in the absence of a fusion partner (Figure 6; SL and WBF, unpublished). The reconstituted FP shows a similar subcellular localization to natively expressed GFP with localization to the nucleus. A similar autoassembly of the two FP halves or assembly of the N-terminal fusion with a protein-of-interest and the free C-terminal domain has been observed in plant cytosol and



**Figure 6.** Autoreconstitution of split-Venus in plants.

Split-halves of Venus (nVenus: amino acids 1–155; cVenus: amino acids 155–238), when co-expressed, reconstitute a functional fluorescent protein in the absence of fusion partners (SL and WF, unpublished results). Tobacco (*Nicotiana benthamiana*) leaves were infiltrated with an *Agrobacterium tumefaciens* suspension (OD 0.1) of each construct (35S-nVenus-Nos and 35S-cVenus-Nos) and imaged 60 h post-infiltration using a Nipkow spinning disk confocal microscope (for method see Deuschle *et al.*, 2006): (a) Venus channel; (b) chlorophyll fluorescence; (c) merge. Bar represents 20  $\mu$ m. No fluorescent signal was obtained when the halves were expressed alone.

even when the halves were targeted to the ER (Cabantous *et al.*, 2005; Walter *et al.*, 2004; Zamyatnin *et al.*, 2006). These data suggest that in the plant systems used, the 'critical concentration' (see discussion in sections on co-localization and FRET) is reached such that the two halves can reconstitute a fluorescent protein in the absence of fusion partners. As a consequence it is only possible to observe differences in the rate of formation of the reconstituted and correctly folded split FP. Since the dissociation constant of complex formation is negligible, reconstitution is driven only by the *on rate* (association constant). Thus any factor that enhances formation of a functional fluorophore will affect the time when fluorescence becomes detectable after infiltration/transfection (Table 3). Factors that contribute to changes in the association rate include association of fusion partners, effects of the fusion on protein steady-state levels of the chimeras, and effects of the fusions on folding. Their steady-state levels, the expression level of the chimera, protein fragment turnover, accessibility of the fragments for reassembly, changes in free diffusion, and effects of the fusion on the (pre-) folding of GFP halves will affect the on rate. In other words, the split-FP system measures a number of parameters such as binding of fusion partners, protein folding, protein turnover, and accessibility for interaction, etc. Given the existence of these multiple factors that could contribute to the rate of reconstitution, from first principles, the split-FP system does not measure protein interactions alone but measures all of the parameters mentioned above. Additional experiments are necessary to differentiate which of the factors is responsible for increased reconstitution rates, with the binding affinity of the fusion partners being only one of several possibilities. The tests used for distinguishing between these possibilities include the use of mutations in the fusion partners, which affect formation of the fluorophore. However, these mutations may affect any of the other parameters as well, such as protein stability. It is important to note here also that at least high-affinity complexes are typically not affected by single-point mutations (Reichmann *et al.*, 2007). Lowering the expression level as well as novel split-FP systems in which the dissociation rate is increased may constitute means to improve the suitability for interaction, folding, and stability assays.

**Table 3** Parameters measured with the split fluorescent protein (FP) systems

Contribution of fusion partners to FP association rate
Contribution of fusion partners to FP maturation
Contribution of fusion partners to (pre-)folding of FP halves
Contribution of fusion partners to available number of FP halves
Contribution of fusion partners to stability of FP halves
Contribution of FP halves to stability of fusion partners
Steric accessibility of FPs in fusions for reconstitution

### Quantitative analysis of protein interactions

A variety of methods have been developed to detect and quantify PPI both *in vitro* and *in vivo* (Table 1), including surface plasmon resonance (SPR; BIAcore™, McDonnell, 2001; Schuck, 1997b), isothermal titration calorimetry (ITC; Kumaran and Jez, 2007; Perozzo *et al.*, 2004; ), and analytical ultracentrifugation (reviewed in Lakey and Raggett, 1998). Modern plasmon resonance instrumentation enables the characterization of interactions at least in moderate throughput. Quantitative information on protein interactions in the pico- to nanomolar affinity range can also be obtained by mass spectrometry, such as matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF), and novel quantitative approaches are being developed in this area, e.g. intensity fading (Krogan *et al.*, 2006; Yanes *et al.*, 2007). The free energies of protein–protein association as determined using electrospray ionization mass spectrometry (ESI-MS) correlate accurately with values obtained by solution enzyme assays (Krishnaswamy *et al.*, 2006). The sensitivity of the mass spectrometry approaches can be increased to detect weak and transient interactions using chemical cross-linking approaches (Vasilescu and Figeys, 2006). Another method that can be applied for the detection of PPIs is atomic force microscopy (AFM; Hinterdorfer and Dufrene, 2006) which measures the specific interaction forces involved in protein interactions (Lin *et al.*, 2005), and which has been used to quantify the dissociation kinetics of protein complexes, for example for the SNARE complex (Yersin *et al.*, 2003).

#### Surface plasmon resonance (SPR)

Probably the simplest method for the analysis of thermodynamic and kinetic parameters of protein interactions *in vitro* is surface plasmon resonance (SPR). Surface plasmon resonance measures the change in refractive index of a solvent near a surface (typically a gold film) that occurs during complex formation or dissociation (McDonnell, 2001; Schuck, 1997a). One partner (the bait) is bound to the surface of a chip that is coated with the gold foil, in the case of proteins via affinity tags such as nickel-chelation of a His-tagged protein (Rich and Myszka, 2000). Thus a prerequisite for application of SPR is that sufficient protein can be obtained in a heterologous expression system. The chip is mounted onto the instrument which measures SPR using an evanescent wave in real time and the chip is perfused with a buffer using a microfluidic device. Then a solution containing the prey ligand is added. If the prey binds to the bait, the SPR signal changes, producing a new equilibrium due to association of bait and prey. After this new steady state has been reached, the chip is perfused with a buffer solution, the ligand dissociates and a new equilibrium will be established.

From this reversible reaction, the association and dissociation as well as the binding constants can be determined. Modern instruments (e.g. BLAcore™) provide for the analysis of interactions between multiple baits and prey with medium throughput. Obviously, this method is best suited for the analysis of interactions of soluble proteins since it is difficult to bind a membrane protein, e.g. a transporter, in its native conformation to the chip. As pointed out above, knowledge of the thermodynamic and kinetic parameters would be a highly valuable asset for our understanding of protein interactions, and in order to obtain insights into the detection range regarding affinities of the various HT methods.

### Databases for protein interactions

Physical protein interaction data are available for most model organisms. However, the number of PPI data generated from experimental approaches varies widely depending on the organism. Compared with animal systems, very few experimentally derived interaction data for plants are available today. Several repositories for PPI data exist, including IntAct (Hermjakob *et al.*, 2004b), bioGRID (Stark *et al.*, 2006), BIND (Gilbert, 2005), DIP (Xenarios *et al.*, 2002), KEGG (Franca-Koh *et al.*, 2006), MINT (Chatr-Aryamontri *et al.*, 2007), and MIPS (Pagel *et al.*, 2005). Both hand-curated data from the literature and bulk data from HT screens are available in these databases. In addition, TAIR has some Arabidopsis protein interaction data curated from the literature (<http://www.arabidopsis.org/portals/proteome/proteinInteract.jsp>). In addition, new databases such as the database for Kinetic Data of Bio-molecular Interactions (KDBI) provide kinetic data of PPI derived from literature curation. The KDBI contains information about binding or reaction events, participating molecules (name, synonyms, molecular formula, classification, SWISS-PROT AC or CAS number), binding or reaction equation, and kinetic data (Ji *et al.*, 2003). The proteomics community has developed and adapted a standard for protein interaction data in XML called Proteomics Standards Initiative Molecular Interaction (PSI-MI; Hermjakob *et al.*, 2004a). Several open-source, free software applications for visualizing and analyzing protein interaction data exist. Some of the popular ones include Cytoscape (Shannon *et al.*, 2003) and Osprey (Breitkreutz *et al.*, 2003).

### Construction and analysis of protein interaction networks

The PPI data can be utilized largely in two ways. One can start with a protein or complex of interest and determine which proteins physically interact. Alternatively, one can analyze the entire network of the interacting proteins of a system to learn about any high-level organizing principles of complexes and interacting proteins. Here, we briefly

describe the methods and findings of the study of PPI networks from this global perspective.

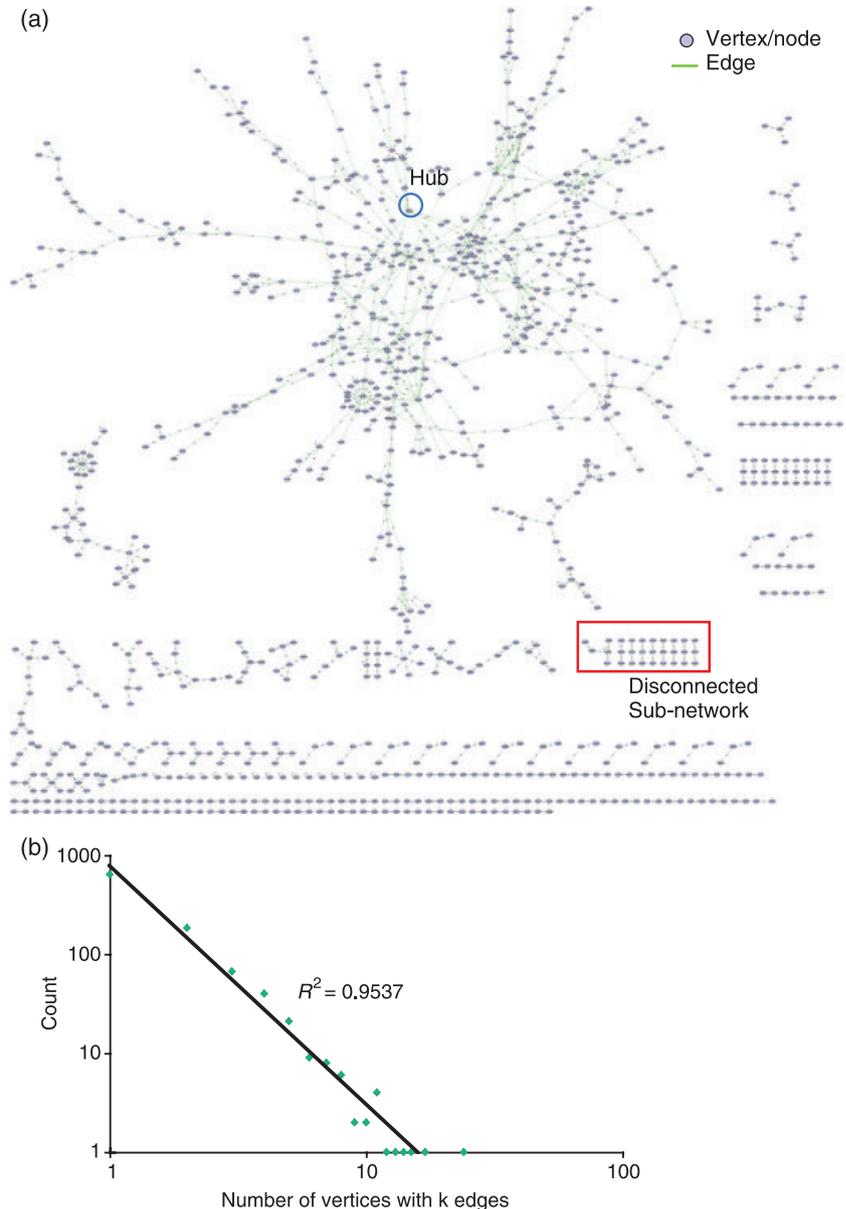
Genome-wide PPI networks are currently available for several model organisms such as *S. cerevisiae*, *C. elegans*, and *D. melanogaster* (Hermjakob *et al.*, 2004b). Although these networks represent only a fraction of the complete interactomes, investigation of these networks is a first step towards a systems-biology understanding of cells and organisms. In *S. cerevisiae*, physical interactions between proteins have been identified with large-scale HT experiments using the Y2H method as well as direct purification of complexes using AP-MS analyses (Ito *et al.*, 2001; Uetz *et al.*, 2000). The number of PPI data for *S. cerevisiae* has been increasing at a rapid rate, currently totaling 18 272 interactions among 4920 proteins in DIP database (as of 6 September 2007). Similarly, a large fraction of PPI data exist for *C. elegans* (Hermjakob *et al.*, 2004b; Li *et al.*, 2004; Xenarios *et al.*, 2002) and *D. melanogaster* (Giot *et al.*, 2003; Hermjakob *et al.*, 2004b; Xenarios *et al.*, 2002). At present, experimentally derived plant protein interactions are much fewer than animals. For *A. thaliana*, there are about 1800 interactions among 1000 proteins that are curated from literature available from TAIR, IntAct, and BIND. Large-scale HT protein interaction projects had not been published for plants at the time of this writing, though there are at least two projects under way to generate large-scale interaction data (<http://www.associomics.org/>, J. Ecker, Salk Institute for Biological Studies, La Jolle, CA, USA, personal communication). Meanwhile, attempts have been made to extrapolate potential interaction data for Arabidopsis from interacting orthologs in *S. cerevisiae*, *C. elegans*, *D. melanogaster*, and *H. sapiens* (Donaldson *et al.*, 2003; Geisler-Lee *et al.*, 2007). A total of 1159 high-confidence, 5913 medium-confidence, and 12907 low-confidence interactions were predicted for 3617 proteins (Geisler-Lee *et al.*, 2007). For *O. sativa*, 8902 interactions for 1879 proteins have been predicted (Yu *et al.*, 2005).

When a substantial portion of an interactome is available, it becomes feasible to study the interactome using graph topological/theoretical analysis methods to obtain insights into PPI properties from a systems view. A PPI network is typically represented as a graph  $G = (V, E)$  (Harary, 1969) where each vertex (a node representing a protein) in  $V$  (total number of proteins in the network) is connected to an interacting protein by a line (called an edge in graph analysis) in  $E$  (total number of interactions in a network) (Figure 7a). Such protein interactions may be derived from individual experimental datasets, public databases, or with the help of PPI prediction tools. Graph visualization and analysis tools such as Graphviz (Gansner and North, 2000) and Cytoscape (Shannon *et al.*, 2003) can be used to draw PPI networks in two or even three dimensions. Some graph visualization tools, including Pajek (Batagelj and Mrvar, 1998) and LGL (Adai *et al.*, 2004), apply a force-directed

**Figure 7.** Protein–protein interaction (PPI) network.

(a) Scale-free topology of *Saccharomyces cerevisiae* interaction network from 1004 proteins and 948 interactions (<http://depts.washington.edu/sfields>; Uetz *et al.*, 2000). Drawn with Cytoscape (Shannon *et al.*, 2003).

(b) Power-law degree distribution shows that the PPI network is a scale-free network. The  $R^2$  value of 0.9637 indicates that the linear trendline fits the data very well.



graphical representation (representing vertices as physical objects that exert forces on each other such that vertices that are connected by an edge attract to ensure that they are placed closely and pairs of graph vertices repulse from other pairs to ensure that non-related vertices are placed at larger distances. The resulting graph layout is an energy-minimal state of the force system.) guided by a minimal spanning tree (the portion of the network that represents a tree where all the vertices are connected fully and the number of edges is less than or equal to that of every other spanning tree) of the network and are useful in visualizing and exploring large biological networks.

As a result of the analysis of existing PPI networks, it has been shown that PPI networks from large-scale, systematic experiments have similar scale-free (most pro-

teins interact with just a few partners and a small number of proteins interact with many partners) and small-world topology (the shape of the network is significantly different from random graph models) (Guelzim *et al.*, 2002; Milo *et al.*, 2002; Yeager-Lotem *et al.*, 2004). Common topological and biological features of these PPI networks lead to numerous biological hypotheses, which will be described briefly in this section.

Small-world networks (Watts and Strogatz, 1998) were discovered according to their clustering coefficient (a measure that weighs the cohesiveness of a graph) and their mean-shortest path length (average length of the shortest paths between two vertices). Small-world networks, as compared with random graphs with the same scale (same number of vertices in a graph), are characterized by cluster-

ing coefficients that are significantly higher, and mean shortest-path length much lower than those in random graphs. By definition, small-world networks have a high portion of cliques (a graph in which every vertex is connected to every other vertex in the graph) and sub-graphs (a graph whose vertex and edge sets are subsets of those in the larger graph  $G$ ) that are a few edges short of being cliques. This follows from the requirement of a high clustering coefficient. Secondly, the majority of pairs of vertices will be connected by at least one short path. This follows from the requirement that the mean shortest path length be small. If the average length of the shortest paths for every two vertices is small, then by chance any two vertices will be connected by a short path. Here we assume the shortest path follows normal distribution and the standard deviation is not high. Barabasi and Albert hypothesize that the prevalence of small-world networks in biological systems may reflect an evolutionary advantage of such architecture, since small-world networks are more robust to perturbations than other networks (Barabasi and Albert, 1999). In this case, this feature would provide an advantage to biological systems that are subject to damage by mutation or viral infection.

In addition to being a of a small-world type, biological networks, including PPI networks, tend to be scale-free (Guelzim *et al.*, 2002; Milo *et al.*, 2002; Yeger-Lotem *et al.*, 2004). In scale-free networks, the number of connections per protein follows a power-law distribution (Figure 7b) such that most vertices have only a few connections, while a few number of vertices are highly connected (often called 'hubs'). These hubs can shape the way in which the network operates because the hubs and their direct neighbors occupy the majority of the network. Since hubs interconnect multiple sub-networks, it is not surprising that deletion of such hubs is often lethal, which may be expected because the loss of a centrally connected component probably affects multiple cellular processes (Jeong *et al.*, 2001).

These global features of the PPI network suggest general principles of the architecture of the network, but do not help in elucidating what makes one PPI network different from another. To do this, it is important to analyze features of the network at a more local level. Apart from the global topological characteristics, complex PPI networks are very different from each other in several respects such as size, shape, and connectivity, but they all share striking local properties: the presence of many small dense sub-networks, called network modules (also called clusters, Alon, 2003). A network module is defined as a set of vertices that have strong interactions and a common biological function (e.g. a ribosome, proteosome, or photosystem I). A module has boundary vertices (vertices interacting with other vertices outside a module) that control the input/output interactions (interactions between vertices inside a module and vertices outside a module) with the rest of the network.

A module also has internal vertices (vertices that do not interact with any vertex outside a module) that do not significantly interact with vertices outside the module. Modules in PPI networks of different sizes have been found using different methods, included the highly connected sub-graphs (HCS; Hartuv *et al.*, 2000) and detection of molecular complexes (MCODE; Bader and Hogue, 2003). Modular biological networks may have an advantage over non-modular network by having the capacity to be more readily reconfigured to adapt to new conditions. The modules in PPI networks also make the networks more robust to perturbation by isolating the perturbation to the module that is affected (Gerhart and Kirschner, 1997; Lipson *et al.*, 2002). Recent analysis on experimentally derived PPI networks showed that with increasing number of proteins in a PPI network, the number of vertices in individual modules increases while the number of identified modules decreases (Przulj *et al.*, 2004).

Another useful local property of networks is the existence of network motifs, which are significantly recurring local topological patterns (Milo *et al.*, 2002). If modules represent functional units, network motifs could be said to represent structural units of a network. While relatively less widely studied than the global topological features, network motifs can lead to a better understanding of various classes of complex networks, as some network motifs may be particular to specific classes of networks. For example, certain triad and tetrad motifs are found to appear commonly in gene transcription networks of *S. cerevisiae* and *E. coli* but not in any other kinds of networks (Milo *et al.*, 2002). In addition, network motifs can also unravel the basic structural elements that underlie the hierarchical and modular architecture of complex natural networks such as PPI networks. It is interesting to note that similarity in network motif topology does not necessarily stem from duplication. Evolution, by constant tinkering, appears to converge on these network motifs in different non-homologous systems, presumably because they are optimally suited to carry out key functions (Wagner, 2003). Network motifs can be detected by frequent sub-graph mining algorithms such as mfinder (Kashtan *et al.*, 2004), hSigGram, vSigGram (Kuramochi and Karypis, 2004), and FPF (Schreiber and Schwobbermeyer, 2004), which compare the patterns found in the target network with those found in suitably randomized networks. Once a dictionary of network motifs and their functions is established, one could envision researchers detecting network motifs in new networks just as protein domains are currently used in annotating new protein sequences (Alon, 2003).

#### Assessing the accuracy of high-throughput PPI data

As described in the section 'Practical considerations for mbSUS analysis', it is important to minimize the number of

'false positives' and 'false negatives' in the PPI data and to evaluate the reliability of the data obtained. It is very important to obtain experimental values for the reliability of the data, e.g. by carrying out independent replicates and independent confirmation by alternative methods, to enable bioinformatic analysis of the reliability of a network. Bioinformatic methods for assessing the reliability of each candidate PPI as well as the network have been developed. Two common methods for detecting false positives are based on data integration and topology.

Data integration-based methods assess the accuracy of PPI data by taking advantage of the difference in results derived from different approaches. For example, Bader *et al.* (2004) developed a quantitative method to compute confidence values for PPIs with a logistic regression approach (a statistical regression model for binomially distributed response/dependent variables). A training set is generated by comparing networks from Y2H and AP-MS experiments. Pairs of proteins that are directly connected to each other in an AP-MS network and are less than two edges apart in the Y2H network were selected as positive examples, and proteins that are directly connected in one network but far apart in the second network were selected as negative examples. These training sets were used to build a hyperplane (a high-dimensional generalization of the concept of a line in Euclidean plane geometry and a plane in three-dimensional Euclidean geometry) that maximally separates the high-confidence pairs from the rest using a logistic regression model. The logistic regression model is then used to predict confidence scores for pair-wise interactions in the full dataset. The high-confidence interactions in Bader's experiments show high agreement with manually curated database annotations. This approach relies on the use of comprehensive datasets, which at present are available only for a few organisms. A major drawback of this approach is that the two methods detect different types of protein complexes: while Y2H methods determine binary PPIs, AP-MS experiments identify complexes, in which two proteins do not have to interact directly. Moreover, it has to be taken into account that the PPIs detected with the two methods may actually differ biologically, for example with respect to their ability to detect PPIs differing in kinetics and binding affinities (see the description of properties of PPIs).

Topology-based methods model the expected topological characteristics of true PPI networks, and then devise mathematical measures to assess the reliability of candidate interactions. For example, Saito and colleagues developed a series of computational measures called interaction generalities (IG1 and IG2) (Saito *et al.*, 2002, 2003) to assess the reliability of PPI. The IG1 measure is based on the idea that proteins that appear to have many interacting partners, which in turn have no further interactions, were likely to be false positives (Saito *et al.*, 2002). This is a reasonable model for Y2H data, as 'sticky' proteins or proteins that accumulate

beyond the 'critical' concentration in Y2H assays may interact with proteins non-specifically or autoactivate the reporter without interacting with another protein. These proteins appear to interact with a large number of random proteins in the experimental data. IG1 is a local measure, which does not consider the topological properties of the protein interaction network beyond the candidate protein pair. As such, it only addresses the 'sticky protein' error but does not correct other types of experimental error that could also lead to false positives such as PPIs detected between overexpressed proteins. The IG2 measure (Saito *et al.*, 2003) was developed to incorporate topological properties of interactions beyond the candidate interacting pairs by considering the five possible topological relationships of a third protein, C, with a candidate interacting pair (A, B), which increases the statistical power of determining reliability. The IG2 measure uses the weighted sum of the five topological components with respect to C. The weights were assigned a priori by performing a principal components analysis (a technique used to reduce multidimensional datasets to lower dimensions by keeping lower-order principal components and ignoring higher-order ones: such low-order components often contain the 'most important' aspects of the data) on the entire interaction network. Experimental results demonstrated that IG2 performed better as measured by more coherence in gene ontology (GO) functional, processing, and component annotations of the interacting proteins than IG1 (Saito *et al.*, 2003).

## Conclusions

Because of the importance of protein interactions for biology, a wide spectrum of advanced methods have been developed to collect the complete interactomes of organisms from all kingdoms. The generation of interactome maps for plants is lagging behind animal systems, and systematic large-scale network analysis remains a major necessity for systems biology. Independent of the organism of interest, membrane proteins are typically depleted from genome-wide analyses for technical reasons. Thus, data on both membrane protein/membrane protein interactions as well as the interface between membrane proteins and the soluble interactome are largely missing despite their importance for advancing our understanding of the communication of cells with their environment. The mbSUS and AP-MS methods tailored for membrane protein interactions will be a means to fill this gap. All of the HT analyses will require extensive follow-up to verify the existence and relevance of the data. In addition to verification, two areas that will require specific attention in the near future will be analysis of localized interactions within a cell and determination of the structural, thermodynamic, and kinetic properties of these interactions. Localized reactions are critical, since many signaling processes occur in small regions of a

cell, as exemplified by localized signaling processes in chemotaxis of *Dictyostelium* (Franca-Koh *et al.*, 2006). Tools for systematic determination of structures and for analysis of the biochemical properties of the interactions are available, including SPR, NMR, and AFM. More sophisticated statistical analyses to address the accuracy of the HT PPI and bioinformatic mining and analysis to relate the PPI information to biological context and derive new principles underlying these processes will be required. Ultimately integration of PPI data with other datasets from transcriptomics, metabolomics including fluxomics (Wiechert *et al.*, 2007), phosphoproteomics (Nühse *et al.*, 2004), and comparative genomics will enable the quantitative modeling of biological processes in plant cells to facilitate our understanding of the interplay between molecular machineries and biological processes in biological systems.

### Acknowledgements

The authors want to thank C. Biskup (University of Jena, Germany) for extremely fruitful discussion on FLIM. We are grateful to NSF for grant support for the large-scale membrane protein interactome project in Arabidopsis (<http://www.associomics.org/>) under the NSF 2010 program (MCB-0618402).

### References

- A dai, A.T., Date, S.V., Wieland, S. and Marcotte, E.M. (2004) LGL: creating a map of protein function with an algorithm for visualizing very large biological networks. *J. Mol. Biol.*, **340**, 179–190.
- Adjobo-Hermans, M.J., Goedhart, J. and Gadella, T.W., Jr (2006) Plant G protein heterotrimers require dual lipidation motifs of G $\alpha$  and G $\gamma$  and do not dissociate upon activation. *J. Cell Sci.*, **119**, 5087–5097.
- Ai, H.W., Henderson, J.N., Remington, S.J. and Campbell, R.E. (2006) Directed evolution of a monomeric, bright and photostable version of *Clavularia* cyan fluorescent protein: structural characterization and applications in fluorescence imaging. *Biochem. J.* **400**, 531–540.
- Aker, J., Borst, J.W., Karlova, R. and de Vries, S. (2006) The *Arabidopsis thaliana* AAA protein CDC48A interacts *in vivo* with the somatic embryogenesis receptor-like kinase 1 receptor at the plasma membrane. *J. Struct. Biol.* **156**, 62–71.
- Aker, J., Hesselink, R., Engel, R., Karlova, R. *et al.* (2007) *In vivo* hexamerization and characterization of the Arabidopsis AAA ATPase CDC48A-complex using FRET-FLIM and FCS. *Plant Physiol.*, **145**, 339–350.
- Alberts, B. (1998) The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell*, **92**, 291–294.
- Alon, U. (2003) Biological networks: the tinkerer as an engineer. *Science*, **301**, 1866–1867.
- Bacia, K., Kim, S.A. and Schwille, P. (2006) Fluorescence cross-correlation spectroscopy in living cells. *Nat. Methods*, **3**, 83–89.
- Bader, G.D. and Hogue, C.W. (2003) An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*, **4**, 2.
- Bader, J.S., Chaudhuri, A., Rothberg, J.M. and Chant, J. (2004) Gaining confidence in high-throughput protein interaction networks. *Nat. Biotechnol.*, **22**, 78–85.
- Barabasi, A.L. and Albert, R. (1999) Emergence of scaling in random networks. *Science*, **286**, 509–512.
- Bartel, P.L., Roecklein, J.A., SenGupta, D. and Fields, S. (1996) A protein linkage map of *Escherichia coli* bacteriophage T7. *Nat. Genet.* **12**, 72–77.
- Bastiaens, P.I.H. and Squire, A. (1999) Fluorescence lifetime imaging microscopy: spatial resolution of biochemical processes in the cell. *Trends Cell Biol.* **9**, 48–52.
- Batagelj, V. and Mrvar, A. (1998) Pajek – program for large network analysis. *Connections*, **21**, 47–57.
- Baxter, I., Ouzzani, M., Orcun, S., Kennedy, B. *et al.* (2007) Purdue Ionomics Information Management System. An integrated functional genomics platform. *Plant Physiol.* **143**, 600–611.
- Berney, C. and Danuser, G. (2003) FRET or no FRET: a quantitative comparison. *Biophys. J.* **84**, 3992–4010.
- Bibi, E. and Kaback, H.R. (1992) Functional complementation of internal deletion mutants in the lactose permease of *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **89**, 1524–1528.
- Biskup, C., Zimmer, T., Kelbauskas, L., Hoffmann, B. *et al.* (2007) Multi-dimensional fluorescence lifetime and FRET measurements. *Microsc. Res. Tech.* **70**, 442–451.
- Booser, C., Kim, G., Cong, S.X., Guan, H.W. *et al.* (2006) Looking towards label-free biomolecular interaction analysis in a high-throughput format: a review of new surface plasmon resonance technologies. *Curr. Opin. Biotechnol.* **17**, 400–405.
- Breitkreutz, B.J., Stark, C. and Tyers, M. (2003) Osprey: a network visualization system. *Genome Biol.* **4**, R22.
- Burz, D.S., Dutta, K., Cowburn, D. and Shekhtman, A. (2006) Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR). *Nat. Methods*, **3**, 91–93.
- Cabantous, S., Terwilliger, T.C. and Waldo, G.S. (2005) Protein tagging and detection with engineered self-assembling fragments of green fluorescent protein. *Nat. Biotechnol.* **23**, 102–107.
- Chatr-Aryamontri, A., Ceol, A., Montecchi Palazzi, L., Nardelli, G. *et al.* (2007) MINT: the Molecular INteraction database. *Nucleic Acids Res.* **35**, D572–D574.
- Chattopadhyaya, R. and Pal, A. (2004) Improved model of a LexA repressor dimer bound to recA operator. *J. Biomol. Struct. Dyn.* **21**, 681–689.
- Chen, H., Puhl, H.L., Koushik, S.V., Vogel, S.S. *et al.* (2006) Measurement of FRET efficiency and ratio of donor to acceptor concentration in living cells. *Biophys. J.* **91**, L39–L41.
- Cusick, M.E., Klitgord, N., Vidal, M. and Hill, D.E. (2005) Interactome: gateway into systems biology. *Hum. Mol. Genet.* **14**, R171–R181.
- de Folter, S., Immink, R.G.H., Kieffer, M., Parenicova, L. *et al.* (2005) Comprehensive interaction map of the Arabidopsis MADS box transcription factors. *Plant Cell*, **17**, 1424–1433.
- de Prat-Gay, G. (1996) Association of complementary fragments and the elucidation of protein folding pathways. *Protein Eng.* **9**, 843–847.
- Demidov, V. and Brode, N. (2006) Profluorescent protein fragments for fast bimolecular fluorescence complementation *in vitro*. *Nat. Protoc.* **1**, 714–719.
- Deuschle, K., Chaudhuri, B., Okumoto, S., Lager, I. *et al.* (2006) Rapid metabolism of glucose detected with FRET glucose nanosensors in epidermal cells and intact roots of Arabidopsis RNA-silencing mutants. *Plant Cell*, **18**, 2314–2325.
- Dixit, R., Cyr, R. and Gilroy, S. (2006) Using intrinsically fluorescent proteins for plant cell imaging. *Plant J.* **45**, 599–615.
- Donaldson, I., Martin, J., de Bruijn, B., Wolting, C. *et al.* (2003) PreBIND and Textomy - mining the biomedical literature for protein-protein interactions using a support vector machine. *BMC Bioinformatics*, **4**, 11.

- Fehr, M., Okumoto, S., Deuschle, K., Lager, I. *et al.* (2005) Development and use of fluorescent nanosensors for metabolite imaging in living cells. *Biochem. Soc. Trans.* **33**, 287–290.
- Fields, S. and Song, O.K. (1989) A novel genetic system to detect protein–protein interactions. *Nature*, **340**, 245–246.
- Formstecher, E., Aresta, S., Colluar, V., Hamburger, A. *et al.* (2005) Protein interaction mapping: a *Drosophila* case study. *Genome Res.* **15**, 376–384.
- Förster, T. (1948) Intermolecular energy migration and fluorescence. *Ann. Phys.* **2**, 55–75.
- Franca-Koh, J., Kamimura, Y. and Devreotes, P. (2006) Navigating signaling networks: chemotaxis in *Dictyostelium discoideum*. *Curr. Opin. Genet. Dev.* **16**, 333–338.
- Frommer, W. and Ninnemann, O. (1995) Heterologous expression of genes in bacterial, fungal, animal, and plant cells. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **46**, 419–444.
- Fung, B.K.K. and Stryer, L. (1978) Surface density determination in membranes by fluorescence energy-transfer. *Biochemistry*, **17**, 5241–5248.
- Galarneau, A., Primeau, M., Trudeau, L.E. and Michnick, S.W. (2002)  $\beta$ -Lactamase protein fragment complementation assays as *in vivo* and *in vitro* sensors of protein–protein interactions. *Nat. Biotechnol.* **20**, 619–622.
- Gansner, E.R. and North, S.C. (2000) An open graph visualization system and its applications to software engineering. *Software Pract. Exp.* **30**, 1203–1233.
- Gavin, A.C., Bosche, M., Krause, R., Grandi, P. *et al.* (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*, **415**, 141–147.
- Gavin, A.C., Aloy, P., Grandi, P., Krause, R. *et al.* (2006) Proteome survey reveals modularity of the yeast cell machinery. *Nature*, **440**, 631–636.
- Geisler-Lee, J., O'Toole, N., Ammar, R., Provart, N.J. *et al.* (2007) A predicted interactome for arabidopsis. *Plant Physiol.*, **145**, 317–329.
- Gerhart, J. and Kirschner, M. (1997) *Cells, Embryos, and Evolution: Toward a Cellular and Developmental Understanding of Phenotypic Variation and Evolutionary Adaptability*. Oxford: Blackwell Science.
- Ghosh, I., Hamilton, A.D. and Regan, L. (2000) Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein. *J. Am. Chem. Soc.* **122**, 5658–5659.
- Gilbert, D. (2005) Biomolecular interaction network database. *Brief Bioinformatics*, **6**, 194–198.
- Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, A. *et al.* (2003) A protein interaction map of *Drosophila melanogaster*. *Science*, **302**, 1727–1736.
- Goedhart, J., Hink, M.A., Visser, A., Bisseling, T. *et al.* (2000) *In vivo* fluorescence correlation microscopy (FCM) reveals accumulation and immobilization of Nod factors in root hair cell walls. *Plant J.* **21**, 109–119.
- Gordon, G.W., Berry, G., Liang, X.H., Levine, B. *et al.* (1998) Quantitative fluorescence resonance energy transfer measurements using fluorescence microscopy. *Biophys. J.* **74**, 2702–2713.
- Gratton, E., Breusegem, S., Sutin, J., Ruan, Q. *et al.* (2003) Fluorescence lifetime imaging for the two-photon microscope: time-domain and frequency-domain methods. *J. Biomed. Opt.* **8**, 381–390.
- Grefen, C., Lalonde, S. and Obrdlik, P. (2007) Split-ubiquitin system for identifying protein–protein interactions in membrane and full-length proteins. *Curr. Protoc. Neurosci.* **Unit 5.27**, 5.27.1–5.27.40.
- Griesbeck, O., Baird, G.S., Campbell, R.E., Zacharias, D.A. *et al.* (2001) Reducing the environmental sensitivity of yellow fluorescent protein. Mechanism and applications. *J. Biol. Chem.* **276**, 29188–29194.
- Grigoriev, A. (2003) On the number of protein–protein interactions in the yeast proteome. *Nucleic Acids Res.* **31**, 4157–4161.
- Gu, H., Lalonde, S., Okumoto, S., Looger, L.L. *et al.* (2006) A novel analytical method for *in vivo* phosphate tracking. *FEBS Lett.* **580**, 5885–5893.
- Guelzim, N., Bottani, S., Bourgine, P. and Kepes, F. (2002) Topological and causal structure of the yeast transcriptional regulatory network. *Nat. Genet.* **31**, 60–63.
- Harary, F. (1969) *Graph Theory*. Boston: Addison-Wesley.
- Hartuv, E., Schmitt, A.O., Lange, J., Meier-Ewert, S. *et al.* (2000) An algorithm for clustering cDNA fingerprints. *Genomics*, **66**, 249–256.
- Heim, N. and Griesbeck, O. (2004) Genetically encoded indicators of cellular calcium dynamics based on troponin C and green fluorescent protein. *J. Biol. Chem.* **279**, 14280–14286.
- Hermjakob, H., Montecchi-Palazzi, L., Bader, G., Wojcik, R. *et al.* (2004a) The HUPO PSI's molecular interaction format – a community standard for the representation of protein interaction data. *Nat. Biotechnol.* **22**, 177–183.
- Hermjakob, H., Montecchi-Palazzi, L., Lewington, C., Mudali, S. *et al.* (2004b) IntAct: an open source molecular interaction database. *Nucleic Acids Res.* **32**, D452–D455.
- Hink, M.A., Borst, J.W. and Visser, A. (2003) Fluorescence correlation spectroscopy of GFP fusion proteins in living plant cells. *Method Enzymol.* **361**, 93–112.
- Hinterdorfer, P. and Dufrene, Y.F. (2006) Detection and localization of single molecular recognition events using atomic force microscopy. *Nat. Methods*, **3**, 347–355.
- Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D. *et al.* (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*, **415**, 180–183.
- Hu, C.-D., Chinenov, Y. and Kerppola, T.K. (2002) Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. *Mol. Cell*, **9**, 789–798.
- Huang, X.Y., Holden, H.M. and Raushel, F.M. (2001) Channeling of substrates and intermediates in enzyme-catalyzed reactions. *Annu. Rev. Biochem.* **70**, 149–180.
- Islam, M., Wallin, R., Wynn, R., Conway, M. *et al.* (2007) A novel branched-chain amino acid metabolite: Protein–protein interactions in a supramolecular complex. *J. Biol. Chem.*, **282**, 11893–11903.
- Ito, T., Tashiro, K., Muta, S., Ozawa, R. *et al.* (2000) Toward a protein–protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. *Proc. Natl Acad. Sci. USA*, **97**, 1143–1147.
- Ito, T., Chiba, T., Ozawa, R., Yoshida, M. *et al.* (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl Acad. Sci. USA*, **98**, 4569–4574.
- Jares-Erijman, E.A. and Jovin, T.M. (2003) FRET imaging. *Nat. Biotechnol.* **21**, 1387–1395.
- Jares-Erijman, E.A. and Jovin, T.M. (2006) Imaging molecular interactions in living cells by FRET microscopy. *Curr. Opin. Chem. Biol.* **10**, 409–416.
- Jeong, H., Mason, S.P., Barabasi, A.L. and Oltvai, Z.N. (2001) Lethality and centrality in protein networks. *Nature*, **411**, 41–42.
- Ji, Z.L., Chen, X., Zhen, C.J., Yao, L.X. *et al.* (2003) KDBI: Kinetic Data of Bio-molecular Interactions database. *Nucleic Acids Res.*, **31**, 255–257.

- Jin, F.L., Hazbun, T., Michaud, G.A., Salcius, M. *et al.* (2006) A pooling-deconvolution strategy for biological network elucidation. *Nat. Methods*, **3**, 183–189.
- Johnsson, N. and Varshavsky, A. (1994) Split Ubiquitin as a sensor of protein interactions *in vivo*. *Proc. Natl Acad. Sci. USA*, **91**, 10340–10344.
- Kalipatnapu, S. and Chattopadhyay, A. (2005) Membrane protein solubilization: recent advances and challenges in solubilization of serotonin<sub>1A</sub> receptors. *IUBMB Life*, **57**, 505–512.
- Karasawa, S., Araki, T., Nagai, T., Mizuno, H. *et al.* (2004) Cyan-emitting and orange-emitting fluorescent proteins as a donor/acceptor pair for fluorescence resonance energy transfer. *Biochem. J.* **381**, 307–312.
- Kashtan, N., Itzkovitz, S., Milo, R. and Alon, U. (2004) Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. *Bioinformatics*, **20**, 1746–1758.
- Kerrien, S., Alam-Faruque, Y., Aranda, B., Bancarz, I. *et al.* (2007) IntAct – open source resource for molecular interaction data. *Nucleic Acids Res.* **35**, D561–565.
- Köhler, R.H., Cao, J., Zipfel, W.R., Webb, W.W. *et al.* (1997) Exchange of protein molecules through connections between higher plant plastids. *Science*, **276**, 2039–2042.
- Köhler, R.H., Schwille, P., Webb, W.W. and Hanson, M.R. (2000) Active protein transport through plastid tubules: velocity quantified by fluorescence correlation spectroscopy. *J. Cell Sci.*, **113**, 3921–3930.
- Korf, U. and Wiemann, S. (2005) Protein microarrays as a discovery tool for studying protein–protein interactions. *Expert Rev. Proteomics*, **2**, 13–26.
- Koushik, S.V., Chen, H., Thaler, C., Puhl, H.L. *et al.* (2006) Cerulean, Venus, and Venus(Y67C) FRET reference standards. *Biophys. J.* **91**, L99–L101.
- Kozer, N., Kuttner, Y.Y., Haran, G. and Schreiber, G. (2007) Protein-protein association in polymer solutions: from dilute to semidilute to concentrated. *Biophys. J.* **92**, 2139–2149.
- Krishnaswamy, S.R., Williams, E.R. and Kirsch, J.F. (2006) Free energies of protein–protein association determined by electrospray ionization mass spectrometry correlate accurately with values obtained by solution methods. *Protein Sci.* **15**, 1465–1475.
- Krogan, N.J., Cagney, G., Yu, H.Y., Zhong, G.Q. *et al.* (2006) Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature*, **440**, 637–643.
- Kumaran, S. and Jez, J.M. (2007) Thermodynamics of the interaction between O-acetylserine sulfhydrylase and the C-terminus of serine acetyltransferase. *Biochemistry*, **46**, 5586–5594.
- Kuramochi, M. and Karypis, G. (2004) *Finding Frequent Patterns in a Large Sparse Graph*. Lake Buena Vista, Florida: SIAM International Conference on Data Mining.
- Lakey, J.H. and Raggett, E.M. (1998) Measuring protein–protein interactions. *Curr. Opin. Struct. Biol.* **8**, 119–123.
- Lakowicz, J.R. (2006) *Principles of Fluorescence Spectroscopy*, 3rd edn. New York: Springer.
- Lalonde, S., Ehrhardt, D.W. and Frommer, W.B. (2005) Shining light on signaling and metabolic networks by genetically encoded biosensors. *Curr. Opin. Plant Biol.* **8**, 574–581.
- Li, S.M., Armstrong, C.M., Bertin, N., Ge, H. *et al.* (2004) A map of the interactome network of the metazoan *C. elegans*. *Science*, **303**, 540–543.
- Lin, S., Chen, J., Huang, L. and Lin, H. (2005) Measurements of the forces in protein interactions with atomic force microscopy. *Curr. Proteomics*, **2**, 55.
- Lipson, H., Pollack, J.B. and Suh, N.P. (2002) On the origin of modular variation. *Evolution*, **56**, 1549–1556.
- Long, S.B., Campbell, E.B. and MacKinnon, R. (2005) Crystal structure of a mammalian voltage-dependent Shaker family K<sup>+</sup> channel. *Science*, **309**, 897–903.
- Loqué, D., Lalonde, S., Looger, L., von Wirén, N. *et al.* (2007) Trans-activation and cooperativity in an ammonium transporter complex. *Nature*, **446**, 195–198.
- Ludewig, U., Wilken, S., Wu, B.H., Jost, W. *et al.* (2003) Homo- and hetero-oligomerization of ammonium transporter-1 NH<sub>4</sub><sup>+</sup> uniporters. *J. Biol. Chem.* **278**, 45603–45610.
- MacKinnon, R. (1991) Determination of the subunit stoichiometry of a voltage-activated potassium channel. *Nature*, **350**, 232–235.
- Magliery, T.J., Wilson, C.G.M., Pan, W.L., Mishler, D. *et al.* (2005) Detecting protein–protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. *J. Am. Chem. Soc.*, **127**, 146–157.
- McDonnell, J.M. (2001) Surface plasmon resonance: towards an understanding of the mechanisms of biological molecular recognition. *Curr. Opin. Chem. Biol.*, **5**, 572–577.
- Mena, M.A., Treynor, T.P., Mayo, S.L. and Daugherty, P.S. (2006) Blue fluorescent proteins with enhanced brightness and photostability from a structurally targeted library. *Nat. Biotechnol.*, **24**, 1569–1571.
- Meseth, U., Wohland, T., Rigler, R. and Vogel, H. (1999) Resolution of fluorescence correlation measurements. *Biophys. J.* **76**, 1619–1631.
- Meyer, R., Steinfath, M., Lisec, J., Becher, M. *et al.* (2007) The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA*, **104**, 4759–4764.
- Miernyk, J.A. and Thelen, J.J. (2008) Biochemical approaches for discovering protein–protein interactions. *Plant J.* **53**, 597–609.
- Miller, J.P., Lo, R.S., Ben-Hur, A., Desmarais, C. *et al.* (2005) Large-scale identification of yeast integral membrane protein interactions. *Proc. Natl Acad. Sci. USA*, **102**, 12123–12128.
- Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N. *et al.* (2002) Network motifs: simple building blocks of complex networks. *Science*, **298**, 824–827.
- Mitra, R.D., Silva, C.M. and Youvan, D.C. (1996) Fluorescence resonance energy transfer between blue-emitting and red-shifted excitation derivatives of the green fluorescent protein. *Gene*, **173**, 13–17.
- Mitra, R.M., Gleason, C.A., Edwards, A., Hadfield, J. *et al.* (2004) A Ca<sup>2+</sup> /calmodulin-dependent protein kinase required for symbiotic nodule development: Gene identification by transcript-based cloning. *Proc. Natl Acad. Sci. USA*, **101**, 4701–4705.
- Müller, J.D., Chen, Y. and Gratton, E. (2003) Fluorescence correlation spectroscopy. *Methods Enzymol.* **361**, 69–92.
- Nagai, T., Ibata, K., Park, E.S., Kubota, M. *et al.* (2002) A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **20**, 8790.
- Nguyen, A. and Daugherty, P. (2005) Evolutionary optimization of fluorescent proteins for intracellular FRET. *Nat. Biotechnol.* **23**, 355–360.
- Nooren, I.M. and Thornton, J.M. (2003) Diversity of protein–protein interactions. *EMBO J.* **22**, 3486–3492.
- Nühse, T.S., Stensballe, A., Jensen, O.N. and Peck, S.C. (2004) Phosphoproteomics of the arabidopsis plasma membrane and a new phosphorylation site database. *Plant Cell*, **16**, 2394–2405.
- Obrdlik, P., El-Bakkoury, M., Hamacher, T., Cappellaro, C. *et al.* (2004) K<sup>+</sup> channel interactions detected by a genetic system optimized for systematic studies of membrane protein interactions. *Proc. Natl Acad. Sci. USA*, **101**, 12242–12247.
- Okumoto, S., Looger, L.L., Micheva, K.D., Reimer, R.J. *et al.* (2005) Detection of glutamate release from neurons by genetically

- encoded surface-displayed FRET nanosensors. *Proc. Natl Acad. Sci. USA*, **102**, 8740–8745.
- Pagel, P., Kovac, S., Oesterheld, M., Brauner, B. *et al.* (2005) The MIPS mammalian protein–protein interaction database. *Bioinformatics*, **21**, 832–834.
- Paulmurugan, R., Umezawa, Y. and Gambhir, S.S. (2002) Non-invasive imaging of protein–protein interactions in living subjects by using reporter protein complementation and reconstitution strategies. *Proc. Natl Acad. Sci. USA*, **99**, 15608–15613.
- Pelet, S., Previte, M.J.R. and So, P.T.C. (2006) Comparing the quantification of Förster resonance energy transfer measurement accuracies based on intensity, spectral, and lifetime imaging. *J. Biomed. Opt.*, **11**, 34017.
- Pelletier, J.N., Campbell-Valois, F.X. and Michnick, S.W. (1998) Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *Proc. Natl Acad. Sci. USA*, **95**, 12141–12146.
- Pelletier, J.N., Arndt, K.M., Pluckthun, A. and Michnick, S.W. (1999) An *in vivo* library-versus-library selection of optimized protein–protein interactions. *Nat. Biotechnol.*, **17**, 683–690.
- Pellicena, P. and Kuriyan, J. (2006) Protein–protein interactions in the allosteric regulation of protein kinases. *Curr. Opin. Struct. Biol.*, **16**, 702–709.
- Perozzo, R., Folkers, G. and Scapozza, L. (2004) Thermodynamics of protein–ligand interactions: history, presence, and future aspects. *J. Recept. Signal Transduct.*, **24**, 1–52.
- Philip, J. and Carlsson, K. (2003) Theoretical investigation of the signal-to-noise ratio in fluorescence lifetime imaging. *J. Opt. Soc. Am. A*, **20**, 368–379.
- Popescu, S., Popescu, G., Bachan, S., Zhang, Z. *et al.* (2007) Differential binding of calmodulin-related proteins to their targets revealed through high-density *Arabidopsis* protein microarrays. *Proc. Natl Acad. Sci. USA*, **104**, 4730–4735.
- Przulj, N., Wigle, D.A. and Jurisica, I. (2004) Functional topology in a network of protein interactions. *Bioinformatics*, **20**, 340–348.
- Rain, J.C., Selig, L., De Reuse, H., Battaglia, V. *et al.* (2001) The protein–protein interaction map of *Helicobacter pylori*. *Nature*, **409**, 211–215.
- Raquet, X., Eckert, J.H., Muller, S. and Johnsson, N. (2001) Detection of altered protein conformations in living cells. *J. Mol. Biol.*, **305**, 927–938.
- Redford, G.I. and Clegg, R.M. (2005) Polar plot representation for frequency-domain analysis of fluorescence lifetimes. *J. Fluoresc.*, **15**, 805–815.
- Reichmann, D., Rahat, O., Albeck, S., Meged, R. *et al.* (2005) The modular architecture of protein–protein binding interfaces. *Proc. Natl Acad. Sci. USA*, **102**, 57–62.
- Reichmann, D., Rahat, O., Cohen, M., Neuvirth, H. *et al.* (2007) The molecular architecture of protein–protein binding sites. *Curr. Opin. Struct. Biol.*, **17**, 67–76.
- Reinders, A., Schulze, W., Kühn, C., Barker, L. *et al.* (2002a) Protein–protein interactions between sucrose transporters of different affinities co-localized in the same enucleate sieve element. *Plant Cell*, **14**, 1567–1577.
- Reinders, A., Schulze, W., Thaminy, S., Stagljar, I. *et al.* (2002b) Intra- and intermolecular interactions in sucrose transporters at the plasma membrane detected by split-ubiquitin system and functional assays. *Structure*, **10**, 763–772.
- Remy, I. and Michnick, S.W. (2006) A highly sensitive protein–protein interaction assay based on *Gussia* luciferase. *Nat. Methods*, **3**, 977–979.
- Rich, R.L. and Myszka, D.G. (2000) Advances in surface plasmon resonance biosensor analysis. *Curr. Opin. Biotechnol.*, **11**, 54–61.
- Rizzo, M.A., Springer, G.H., Granada, B. and Piston, D.W. (2004) An improved cyan fluorescent protein variant useful for FRET. *Nat. Biotechnol.*, **22**, 445–449.
- Rizzo, M.A., Springer, G., Segawa, K., Zipfel, W.R. *et al.* (2006) Optimization of pairings and detection conditions for measurement of FRET between cyan and yellow fluorescent proteins. *Microsc. Microanal.*, **12**, 238–254.
- Rosenberg, O.S., Deindl, S., Comolli, L.R., Hoelz, A. *et al.* (2006) Oligomerization states of the association domain and the holoenzyme of  $Ca^{2+}$ /CaM kinase II. *FEBS Lett.*, **273**, 682–694.
- Rossi, F., Charlton, C.A. and Blau, H.M. (1997) Monitoring protein–protein interactions in intact eukaryotic cells by beta-galactosidase complementation. *Proc. Natl Acad. Sci. USA*, **94**, 8405–8410.
- Rual, J.F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T. *et al.* (2005) Towards a proteome-scale map of the human protein–protein interaction network. *Nature*, **437**, 1173–1178.
- Russinova, E., Borst, J.W., Kwakitaal, M., Cano-Delgado, A. *et al.* (2004) Heterodimerization and endocytosis of *Arabidopsis* brassinosteroid receptors BRI1 and AtSERK3 (BAK1). *Plant Cell*, **16**, 3216–3229.
- Saito, R., Suzuki, H. and Hayashizaki, Y. (2002) Interaction generality, a measurement to assess the reliability of a protein–protein interaction. *Nucleic Acids Res.*, **30**, 1163–1168.
- Saito, R., Suzuki, H. and Hayashizaki, Y. (2003) Construction of reliable protein–protein interaction networks with a new interaction generality measure. *Bioinformatics*, **19**, 756–763.
- Schevitz, R.W., Otwinowski, Z., Joachimski, A., Lawson, C.L. *et al.* (1985) The 3-dimensional structure of Trp repressor. *Nature*, **317**, 782–786.
- Schonberger, O., Knox, C., Bibi, E. and Pines, O. (1996) Split invertase polypeptides form functional complexes in the yeast periplasm *in vivo*. *Proc. Natl Acad. Sci. USA*, **93**, 9612–9617.
- Schreiber, F. and Schwobbermeyer, H. (2004) *Towards Motif Detection in Networks: Frequency Concepts and Flexible Search*. Camerino, Italy: NETTAB'04.
- Schuck, P. (1997a) Use of surface plasmon resonance to probe the equilibrium and dynamic aspects of interactions between biological macromolecules. *Annu. Rev. Biophys. Biomol. Struct.*, **26**, 541–566.
- Schuck, P. (1997b) Reliable determination of binding affinity and kinetics using surface plasmon resonance biosensors. *Curr. Opin. Biotechnol.*, **8**, 498–502.
- Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K. *et al.* (2003) ARAMEMNON, a novel database for *Arabidopsis* integral membrane proteins. *Plant Physiol.*, **131**, 16–26.
- Schwille, P., Meyer-Almes, F.J. and Rigler, R. (1997) Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution. *Biophys. J.*, **72**, 1878–1886.
- Schwille, P., Haupts, U., Maiti, S. and Webb, W.W. (1999) Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. *Biophys. J.*, **77**, 2251–2265.
- Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N.G. *et al.* (2004) Improved monomeric red, orange and yellow fluorescent proteins derived from *Drosophila* sp red fluorescent protein. *Nat. Biotechnol.*, **22**, 1567–1572.
- Shaner, N.C., Steinbach, P.A. and Tsien, R.Y. (2005) A guide to choosing fluorescent proteins. *Nat. Methods*, **2**, 905–909.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S. *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, **13**, 2498–2504.
- Shaw, S.L. (2006) Imaging the live plant cell. *Plant J.*, **45**, 573–598.

- Shen, K., Teruel, M.N., Connor, J.H., Shenolikar, S. *et al.* (2000) Molecular memory by reversible translocation of calcium/calmodulin-dependent protein kinase II. *Nat. Neurosci.* **3**, 881–886.
- Shen, Q.H., Saijo, Y., Mauch, S., Biskup, C. *et al.* (2007) Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses. *Science*, **315**, 1098–1103.
- Shoemaker, B. and Panchenko, A. (2007) Deciphering protein–protein interactions. Part I. Experimental techniques and databases. *PLoS Comput. Biol.* **3**, e42.
- Spotts, J.M., Dolmetsch, R.E. and Greenberg, M.E. (2002) Time-lapse imaging of a dynamic phosphorylation dependent protein–protein interaction in mammalian cells. *Proc. Natl Acad. Sci. USA*, **99**, 15142–15147.
- Sprangers, R. and Kay, L.E. (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature*, **445**, 618–622.
- Stagljar, I., Korostensky, C., Johnsson, N. and te Heesen, S. (1998) A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins *in vivo*. *Proc. Natl Acad. Sci. USA*, **95**, 5187–5192.
- Stanyon, C.A., Liu, G.Z., Mangiola, B.A., Patel, N. *et al.* (2004) A Drosophila protein–interaction map centered on cell-cycle regulators. *Genome Biol.* **5**, R96.
- Stark, C., Breitkreutz, B.J., Reguly, T., Boucher, L. *et al.* (2006) BioGRID: a general repository of interaction datasets. *Nucleic Acids Res.* **34**, D535–D539.
- Stelzl, U., Worm, U., Lalowski, M., Haenig, C. *et al.* (2005) A human protein–protein interaction network: a resource for annotating the proteome. *Cell*, **122**, 957–968.
- Stryer, L. (1978) Fluorescence energy transfer as a spectroscopic ruler. *Annu. Rev. Biochem.* **47**, 819–846.
- Subramaniam, V., Hanley, Q.S., Clayton, A.H.A. and Jovin, T.M. (2003) Photophysics of green and red fluorescent proteins: Implications for quantitative microscopy. In *Methods Enzymol.* **360**, 178–201.
- Subramanian, C., Kim, B.H., Lyssenko, N.N., Xu, X. *et al.* (2004) The Arabidopsis repressor of light signaling, COP1, is regulated by nuclear exclusion: Mutational analysis by bioluminescence resonance energy transfer. *Proc. Natl Acad. Sci. USA*, **101**, 6798–6802.
- Subramanian, C., Woo, J., Cai, X., Xu, X.D. *et al.* (2006) A suite of tools and application notes for *in vivo* protein interaction assays using bioluminescence resonance energy transfer (BRET). *Plant J.* **48**, 138–152.
- Thaler, C., Koushik, S.V., Blank, P.S. and Vogel, S.S. (2005) Quantitative multiphoton spectral imaging and its use for measuring resonance energy transfer. *Biophys. J.* **89**, 2736–2749.
- Trakselis, M.A., Alley, S.C. and Ishmael, F.T. (2005) Identification and mapping of protein–protein interactions by a combination of cross-linking, cleavage, and proteomics. *Bioconjug. Chem.* **16**, 741–750.
- Uetz, P., Giot, L., Cagney, G., Mansfield, T.A. *et al.* (2000) A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature*, **403**, 623–627.
- van Munster, E.B. and Gadella, T.W.J. (2005) Fluorescence lifetime imaging microscopy (FLIM). *Adv. Biochem. Engin./Biotechnol.* **95**, 143–175.
- Vasilescu, J. and Figey, D. (2006) Mapping protein–protein interactions by mass spectrometry. *Curr. Opin. Biotechnol.* **17**, 394–399.
- Vaughan, C.K., Buckle, A.M. and Fersht, A.R. (1999) Structural response to mutation at a protein–protein interface. *J. Mol. Biol.* **286**, 1487–1506.
- Vidal, M. (2005) Interactome modeling. *FEBS Lett.* **579**, 1834–1838.
- Vidalain, P.O., Boxem, M., Ge, H., Li, S.M. *et al.* (2004) Increasing specificity in high-throughput yeast two-hybrid experiments. *Methods*, **32**, 363–370.
- Vogel, S.S., Thaler, C. and Koushik, S.V. (2006) Fanciful FRET. *Sci. STKE*, **2006**, re2.
- Wagner, A. (2003) Does selection mold molecular networks? *Sci. STKE*, **2003**, pe41.
- Walhout, A.J.M., Sordella, R., Lu, X.W., Hartley, J.L. *et al.* (2000) Protein interaction mapping in *C. elegans* using proteins involved in vulval development. *Science*, **287**, 116–122.
- Walter, M., Chaban, C., Schutze, K., Batistic, O. *et al.* (2004) Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. *Plant J.* **40**, 428–438.
- Wang, T. and Montell, C. (2007) Phototransduction and retinal degeneration in Drosophila. *Eur. J. Physiol.* **454**, 821–847.
- Watts, D.J. and Strogatz, S.H. (1998) Collective dynamics of ‘small-world’ networks. *Nature*, **393**, 440–442.
- Wehr, M.C., Laage, R., Bolz, U., Fischer, T.M. *et al.* (2006) Monitoring regulated protein–protein interactions using split TEV. *Nat. Methods*, **3**, 985–993.
- Wei, N. and Deng, X.W. (2003) The COP9 signalosome. *Annu. Rev. Cell. Dev. Biol.* **19**, 261–286.
- Wiechert, W., Schweissgut, O., Takanaga, H. and Frommer, W.B. (2007) Fluxomics: mass spectrometry versus quantitative imaging. *Curr. Opin. Plant Biol.* **10**, 323–330.
- Wouters, F.S. (2006) The physics and biology of fluorescence microscopy in the life sciences. *Contemp. Phys.* **47**, 239–255.
- Wrubel, W., Stochaj, U. and Ehring, R. (1994) Construction and *in vivo* analysis of new split lactose permeases. *FEBS Lett.* **349**, 433–438.
- Xenarios, I., Salwinski, L., Duan, X.Q.J., Higney, P. *et al.* (2002) DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions. *Nucleic Acids Res.* **30**, 303–305.
- Xia, Y., Lu, L.J. and Gerstein, M. (2006) Integrated prediction of the helical membrane protein interactome in yeast. *J. Mol. Biol.* **357**, 339–349.
- Xu, X.D., Soutto, M., Xie, Q., Servick, S. *et al.* (2007) Imaging protein interactions with bioluminescence resonance energy transfer (BRET) in plant and mammalian cells and tissues. *Proc. Natl Acad. Sci. USA*, **104**, 10264–10269.
- Xu, Y., Piston, D.W. and Hirschie Johnson, C. (1999) A bioluminescence resonance energy transfer (BRET) system: application to interacting circadian clock proteins. *Proc. Natl Acad. Sci. USA*, **96**, 151–156.
- Yanes, O., Villanueva, J., Querol, E. and Aviles, F.X. (2007) Detection of non-covalent protein interactions by ‘intensity fading’ MALDI-TOF mass spectrometry: applications to proteases and protease inhibitors. *Nat. Protoc.* **2**, 119–130.
- Yeger-Lotem, E., Sattath, S., Kashtan, N., Itzkovitz, S. *et al.* (2004) Network motifs in integrated cellular networks of transcription-regulation and protein–protein interaction. *Proc. Natl Acad. Sci. USA*, **101**, 5934–5939.
- Yersin, A., Hirling, H., Steiner, P., Magnin, S. *et al.* (2003) Interactions between synaptic vesicle fusion proteins explored by atomic force microscopy. *Proc. Natl Acad. Sci. USA*, **100**, 8736–8741.
- Yu, J., Wang, J., Lin, W., Li, S. *et al.* (2005) The Genomes of *Oryza sativa*: a history of duplications. *PLoS Biol.* **3**, e38.
- Zacharias, D.A., Violin, J.D., Newton, A.C. and Tsien, R.Y. (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science*, **296**, 913–916.

- Zamyatnin, A.A., Solovyev, A.G., Bozhkov, P.V., Valkonen, J.P.T. et al.** (2006) Assessment of the integral membrane protein topology in living cells. *Plant J.* **46**, 145–154.
- Zhong, J.H., Zhang, H.M., Stanyon, C.A., Tromp, G. et al.** (2003) A strategy for constructing large protein interaction maps using the yeast two-hybrid system: regulated expression arrays and two-phase mating. *Genome Res.* **13**, 2691–2699.
- Zimmermann, T., Rietdorf, J., Girod, A., Georget, V. et al.** (2002) Spectral imaging and linear un-mixing enables improved FRET efficiency with a novel GFP2-YFP FRET pair. *FEBS Lett.* **531**, 245–249.
- Zimmermann, T., Rietdorf, J. and Pepperkok, R.** (2003) Spectral imaging and its applications in live cell microscopy. *FEBS Lett.* **546**, 87–92.