Current Experimental Methods for Characterizing Protein–Protein Interactions

Mi Zhou,[a] Qing Li,[a] and Renxiao Wang*[a, b]
Protein molecules often interact with other partner protein molecules in order to execute their vital functions in living organisms. Characterization of protein–protein interactions thus plays a central role in understanding the molecular mechanisms of relevant protein molecules, elucidating the cellular processes and pathways relevant to health or disease for drug discovery, and charting large-scale interaction networks in systems biology research. A whole spectrum of methods, based on biophysical, biochemical, or genetic principles, have been developed to detect the time, space, and functional relevance of protein–protein interactions at various degrees of affinity and specificity. This article presents an overview of these experimental methods, outlining the principles, strengths and limitations, and recent developments of each type of method.

1. Introduction

Proteins are vital components of living organisms as they are the main components of the physiological metabolic pathways of cells. The field of proteomics aims at studying the expression, structure, and function of all proteins on the whole genome level. It has been estimated that human genome consists of 19000 genes coding for ~500 000 proteins.[1] Over 80% of all proteins do not exist in isolation, but rather interact with one another to form stable or transient complexes.[2] Protein–protein interactions (PPIs) thus play a central role in virtually all biological processes, such as metabolism, transport, structural organization, signal transduction, cell-cycle control, immune recognition, and gene transcription.[3] Completion of a genome sequencing project is normally accompanied with a vast growth of unannotated proteins. Identification of PPIs assists in functional assignment of novel proteins by situating them relative to other proteins in cellular pathways or functional classes,[4] derives a better understanding of the biological relevance of protein functions, and provides key information from which to elucidate the functioning of a living cell under both physiological and pathological conditions. Moreover, investigation of PPIs on a large scale, that is, so-called interactomes, is the foundation for systematic analysis of cellular networks. Cells can be envisioned as a complex web involving innumerable PPIs. Mapping PPI networks for model organisms and human to elucidate the organization of the proteome is one of the important goals for systems biology research, which paves the way to unravel the functional, logical and dynamical aspects of cellular systems.[5]

Notably, since aberrant PPIs contribute to the pathogenesis of numerous human diseases, PPIs are considered to be an emerging class of drug targets for therapeutic intervention.

Target-based drug discovery has slowly shifted from a protein-centric view, largely focusing on some validated therapeutic targets such as G-protein-coupled receptors (GPCRs), nuclear hormone receptors, ion channels and enzymes, towards a more holistic, pathway-centric view.[6] Modulating PPIs implicated in cancer, virology, cardiovascular, or immunology, such as Bcl-2/Bcl-xL–Bax/Bak,[7] p53–MDM2,[8] CCR5–HIV-1 gp120,[9] fibrinogen–GPIIb/IIIa,[10] CD80–CD28,[11] offers the opportunity to develop novel drugs for the treatment of the corresponding disease. In fact, about a dozen of PPI modulators have entered clinical trials or even successfully been developed into marketed drugs.[12–14]

Generally, PPIs can be classified into two main categories based on the contact interface: domain–domain and domain–motif interactions. Domain–domain interaction involves the binding of two globular domains that creates a large contact interface (~2000 Å²), with relatively strong affinities in the low nanomolar or even picomolar range. For domain–motif interaction, a short linear motif (up to 20–30 amino acid residues) on a protein molecule binds to its interacting partner with affinities in the low- to mid-micromolar range, forming a much smaller contact interface (~300–500 Å²).[15] A recent analysis estimated that the human proteome contains over 35 000 domains and 100 000 motifs.[16] This estimation provides a rough idea of the size of the protein–protein “interactome”.

A variety of experimental methods, based on either biophysical, biochemical, or genetic principles, have been developed to aid studies on PPIs. Each type of method has its own strengths and limitations in terms of sensitivity and specificity. Thus, they lay emphasis on different aspects of PPI analysis, for example, to identify binding partners of protein of interest, to generate structure details of protein complexes, to determine kinetic and thermodynamic constants of PPIs, to visualize and generate structure details of protein complexes, and to determine the time, space, and functional relevance of protein–protein interactions at various degrees of affinity and specificity. This article presents an overview of these experimental methods, outlining the principles, strengths and limitations, and recent developments of each type of method.

2. Biophysical Methods

A variety of biophysical methods can be used for the detection of PPIs, and the most prevalent are summarized in Table 1 and discussed in detail below.
2.1. Fluorescence polarization (FP)

The theory of fluorescence polarization (FP) was first described in 1926,[17] and it is based on the observation of the molecular movement of fluorophores in solution. When a fluorophore is excited by polarized light, it emits light with unequal intensities along different axes of polarization. The degree of polarization is inversely related to molecular rotation of the fluorophore, which is largely dependent on molecular mass. Typically, larger masses show slower rotation.[18] With adequate experimental design, an FP assay can be used to measure binding and dissociation between two molecules if one of the binding molecules is relatively small and labeled with a fluorophore (Figure 1 A). Commonly used fluorophores for an FP assay include fluorescein, rhodamine, BODIPY (boron-dipyrromethene), and Cy5 (cyanine-5) dyes, or their derivatives. Complex formation leads to an increase in FP signal (in millipolarization units, mP), which can be measured by a microplate reader. A binding curve can be generated by plotting mP values against multiple concentrations of the labeled molecule. Then, the dissociation constant (Kd) between two binding molecules can be determined from the curve.

FP assays have been widely used in PPI studies, for example to estimate the binding between initiation factor 3 (IF3) and 30S ribosomal subunits,[19] initiation factor 2 (IF2) and 30S ribosomal subunits,[20] gene 2.5 protein and DNA polymerase of bacteriophage T7,[21] and others. As PPIs are often governed by a small number of “binding epitopes”, for example, hot spots, utilizing short truncated motifs containing such epitopes mimics the interactions between two complete protein molecules and thus facilitates experimental measurements. Such an FP assay also provides an opportunity for the detection of inhibitors targeting the PPI.[22] For instance, an FP competition assay has been developed to identify inhibitors targeting the MDM2–p53 peptide interaction,[23] to screen small molecules that inhibit the ZipA–FtsZ peptide interaction,[24] to screen natural product extracts that disrupt the binding of antiapoptotic Bcl-2 family members to Bim BH3 peptides.[25] The advantages of an FP assay are its low cost, simple mix-and-read format without wash steps, and the high-throughput screening (HTS) capacity of the assay when carried out in a multiwell plate (96/384/1,536). However, like other fluorescence-based assays, it also suffers from some drawbacks, such as interference from auto-fluorescence, quenching, and light scattering.[26]

2.2. Surface plasmon resonance (SPR)

Surface plasmon resonance (SPR) is a quantum electromagnetic phenomenon that occurs at a metal-dielectric interface. When a photon of lights hits the metal surface (typically a gold film) at a certain angle of incidence, the light energy is transferred to collective excitations of free electrons, called surface plasmons, and generates electromagnetic waves that propagate along the interface. Resonant transfer of light energy to surface plasmons leads to a sharp attenuation in the reflected light. The angle of incidence at which minimum reflection occurs is defined as the SPR angle. A slight change near the interface (e.g., a change in refractive index in the vicinity of the metal surface) leads to a change in the SPR angle.[27,28] A biosensor system based on SPR has become one of the most popular sensing technology for measuring molecular interactions. In an SPR experiment, the bait protein is first immobilized onto the sensor chip, which is then perfused with buffer solution containing analytes. If interaction occurs, the accumulation of analytes on the sensor surface induces an increase in the refractive index at the sensor surface, and the SPR angle is shifted (Figure 1 B). During dissociation, the inverse effect is observed. The RU (termed response or resonance unit) is recorded to describe this signal change, where one RU is equivalent to a critical angle shift of 10-3 degree.[29]
Since the first commercial SPR-based analytical instrument launched by Biacore AB in 1990, other competing SPR biosensor systems have been developed by manufacturers such as Affinity Sensors, IBIS Technologies, Texas Instruments, Aviv, and BioTul AG. As the interface between the sensor surface and the interacting proteins to be studied is a vital component of this surface-sensitive sensor system, several robust and reproducible sensor chips from Biacore are commercially available, such as the CMS chip (carboxymethylated dextran covalently attached to a gold surface) for immobilization via amine, thiol, or carboxyl groups of proteins, the NTA chip (carboxymethylated dextran pre-immobilized with nitroliotriacetic acid) for immobilization of His-tagged proteins, SA chip (carboxymethylated dextran pre-immobilized with streptavidin) for immobilization of biotinylated proteins, and the CM5 chip (carboxymethylated dextran covalently attached to a gold surface) for immobilization via amine, thiol, aldehyde, or carboxyl groups of proteins, the NTA chip (carboxymethylated dextran pre-immobilized with nitroliotriacetic acid) for immobilization of His-tagged proteins, SA chip (carboxymethylated dextran pre-immobilized with streptavidin) for immobilization of biotinylated proteins.

SPR biosensor technology has various technical advantages, which has been proven useful for verifying PPIs. Small amounts of sample are needed (micro to sub-micrograms) without labeling requirements. It is capable of characterizing binding reaction in real time, determining kinetic and thermodynamic constants of biomolecular interactions, and allowing the analysis of interactions with a wide range of molecular weights, affinities and binding rate. A potential limitation is its sensitivity to interfering effects such as nonspecific interaction between the sensor surface and analyte. Additionally, the requirement for immobilization might disturb the interactions under measurement. The SPR technique has traditionally been used in low-throughput studies as standard SPR biosensor only contains three to four flow cells on a single sensor chip. Nowadays, SPR imaging (SPRI) is extended towards HTS. By utilizing protein array technology and a charge-coupled device (CCD) camera to make images based on the intensity of reflected light, SPRI has been developed to simultaneously process hundreds or thousands of samples. Furthermore, using the SPR technique coupled with mass spectrometry (MS) enables the discovery of novel PPIs.

2.3. Nuclear magnetic resonance (NMR)

Nuclear magnetic resonance (NMR) is a physical phenomenon whereby magnetically active nuclei in a strong magnetic field absorb electromagnetic radiation at distinct frequencies that are determined by the chemical environment of each nucleus. The most prominent applications of NMR spectroscopy are the characterization and determination of the structures of biological macromolecules in solution, including generating structural details of protein interactions at an atomic resolution. Various NMR parameters providing angular or distance data are primary sources of conformational information, including (1) chemical shift perturbation, which is the chemical shift (resonance frequency) change of a nucleus in the protein induced after addition of an interacting partner. In particular, heteronuclear single-quantum coherence (HSQC) NMR experiments are widely used to delineate protein–protein interfa-
The nuclear Overhauser effect (NOE) is the NMR equivalent of fluorescence resonance energy transfer (FRET), discussed later in Section 3.1, referring to the transfer of nuclear spin polarization from one nuclear spin population to another via dipolar cross-relaxation, the intensity of which is proportional to the inverse-sixth power of the distance between two nuclei. Many precise NOE-derived distance constraints between two proteins with tight interaction can be used to determine the full 3D structure of the complex. Residual dipolar couplings (RDCs), reporting on relative orientations among internuclear vectors irrespective of their distance separations, can provide orientation restraints for structural characterization of protein complexes. Other NMR data includes cross-saturation, pseudo-contact shift, J coupling, and so on. A combination of these NMR-derived restraints can help to obtain a more refined structure of a given PPI with a very wide range of dissociation constants (from $10^{-6}$ to $10^{-2}$ M).

As the conventional application of NMR spectroscopy is often oriented to soluble proteins, solid-state NMR spectroscopy has evolved to study membrane proteins involved in PPIs in liquid crystalline lipid bilayers as a more native-like membrane mimetic environment. Another attractive progress in NMR technology is the in-cell NMR spectroscopy, which allows the exploration of PPIs under physiological conditions in living cells.

### 2.4. Circular dichroism (CD)

Circular dichroism (CD) spectroscopy measures differences in absorption of left and right circularly polarized light that arise due to structural asymmetry, and it is often reported in units of ellipticity. A CD spectrum of a protein in the far-UV region (190–250 nm) arises from the amide groups on the protein backbone and is sensitive to the secondary structure of the protein including alpha-helix, beta-sheet, and random coil structures. While a CD spectrum in the near-UV (250–320 nm) region arises from aromatic amino acids and disulfide bonds, and is sensitive to certain aspects of the tertiary structure of the protein. Thus, changes in the CD spectrum can provide information about the changes in protein conformation.

Various methods have been used to analyze protein conformational changes accompanying PPIs by studying the CD spectrum.
spectra, such as CONTIN: a program that estimates the changes in secondary structure of the catalytic subunit of cAMP-dependent protein kinase when bound by Kemptide,\textsuperscript{[51]} MLR: a program designed to elucidate the mechanism and structural requirements of the interaction of tropomodulin with tropomyosin,\textsuperscript{[52]} SELCON, CDNN, K2D, and so on.\textsuperscript{[53]}

As CD spectroscopy is a quantitative technique and the change in CD upon binding is directly proportional to the amount of complex formed, the association or dissociation constant of the complex can also be determined by using CD by direct titration of one protein by another.\textsuperscript{[52, 54]} An enhanced version, termed synchrotron radiation circular dichroism (SRCD), extends the utility and applications of conventional CD by using a synchrotron ring instead of a xenon arc lamp as the light source. This version enables collection of data in the vacuum UV region (<190 nm), detection of spectra with improved signal-to-noise ratios, and measurements for PPIs involving either induced-fit or rigid-body mechanisms.\textsuperscript{[55, 56]} Besides information related to conformational changes, analysis of CD spectra obtained at different temperatures can provide thermodynamic parameters (such as binding constants) of interacting protein molecules.\textsuperscript{[57]}

Generally, CD spectroscopy is a quick method that does not require large amounts of proteins or extensive data processing, and it can be conducted under varying conditions, including solvent, temperature, and pH. However, as a spectroscopic method for characterization of protein structures, CD gives less specific information than NMR spectroscopy.

2.5. Static and dynamic light scattering (SLS/DLS)

Light scattering refers to a physical process involving the interaction of light with a small particle or molecule. When an incident photon encounters a molecule, it perturbs the electron orbits within the molecule and creates an induced dipole moment, which becomes a source of electromagnetic radiation with energy radiated in all directions. The evaluation of the scattered light with regard to its intensity and wavelength often yields valuable information about the scattering molecules, such as molar mass, binding affinity, and absolute stoichiometry of complex interactions. This makes light scattering a versatile tool for characterization of macromolecules and their interactions in solution.\textsuperscript{[58]} In contrast to most methods for characterization, light scattering is an absolute measurement without requiring outside calibration standards.

The most widely used light scattering techniques are static light scattering (SLS) and dynamic light scattering (DLS). In SLS, also known as classical light scattering or multiple angle light scattering, time-averaged scattered light intensity is collected at multiple different angles, with weight average molar mass, size and shape of molecules, and second virial coefficient reflecting PPIs calculated from this data. In DLS, also known as quasi-elastic light scattering, the fluctuations in the intensity of scattered light at a certain angle as the effect of the Brownian motion of the scattering molecules is detected, yielding information about translational diffusion coefficient and hydrodynamic radius of molecules.\textsuperscript{[59]} Combining a light scattering detection technique and a separation system, such as size-exclusion chromatography, provides a good method for the semi-quantitative characterization of PPIs with high sensitivity and high resolution in short run times.\textsuperscript{[58, 60]} As interaction studies should be performed under true equilibrium, a series of batch SLS or DLS measurements over varying compositions referred to as composition-gradient SLS (CG-SLS) or DLS (CG-DLS) have been used to determine equilibrium properties including association stoichiometry and binding affinity of interacting macromolecules in solution.\textsuperscript{[61, 62]}

2.6. Analytical ultracentrifugation (AUC)

Analytical ultracentrifugation (AUC) is a classical method for detecting the concentration distribution of the macromolecule in solution in real-time under centrifugal force, and providing first-principle hydrodynamic and thermodynamic information about the size, shape, molar mass, association energy, association stoichiometry and thermodynamic nonideality of the molecules.\textsuperscript{[63]} Two types of AUC experiments exist, sedimentation velocity (SV) and sedimentation equilibrium (SE). An SV experiment is usually conducted at a very high rotor speed, leading to the depletion of macromolecules from the meniscus region at the air/solution interface and the formation of a concentration boundary that moves toward the bottom of the centrifuge cell as a function of time.\textsuperscript{[64]} It has been employed to determine the relative energetics of the homo- and hetero-dimerization of BirA by using SEDFIT, SC-ISOTHERM, and SEDANAL\textsuperscript{[65]} and to reveal the significance of hydrolysable ATP in the complex formation between cpn60 and cpn10.\textsuperscript{[66]}

In contrast, an SE experiment is conducted at lower rotor speeds, when the centrifugal and back-diffusive forces come to thermodynamic equilibrium. It yields a steady-state concentration distribution that is a function of buoyant solute mass, mass distribution, and mass action parameters for interacting systems, such as the equilibrium association constant or equivalent dissociation constant.\textsuperscript{[67]} SE has been applied to evaluate the $K_d$ value for the CD2–CD48 interaction,\textsuperscript{[68]} and to study the stoichiometry and affinity of the SCF–Kit interaction.\textsuperscript{[69]} A variant of the SE technique called “tracer SE”, in which one of the solute components is labeled, facilitates the study of both self- and hetero-associations over an unprecedentedly broad range of concentration with increased resolution.\textsuperscript{[70]} The disadvantages of SE include long duration for the completion of the experiment.

Generally speaking, AUC is a powerful method for studying macromolecules over a wide range of concentrations and in a wide variety of solvents. The results produced by this method are absolute and independent of standards for comparison as it relies on the principal property of mass and the fundamental law of gravitation.

2.7. Calorimetric methods

Calorimetry relies on the fact that all chemical reactions are accompanied by an endothermic or exothermic change in energy. In PPIs, the change in conformation of interacting pro-
teins from the free to the bound state produces a plethora of bond rearrangements, including hydrogen bonds, van der Waals interactions, and hydrophobic interactions formed or broken, which result in an overall change in the Gibbs energy of the system.\(^7\) Calorimetric methods can directly determine the thermodynamic parameters as a probe for an interaction by detecting the heat energy changes. Currently, two super-sensitive calorimetric techniques with commercially available instruments, that is, isothermal titration calorimetry (ITC) and differential scanning calorimetry (DSC), are widely applied.

ITC directly quantifies the heat change in a reaction at constant temperature or heat capacity at multiple temperatures; it is easy to perform with commercially available high-sensitivity instruments, such as MicroCal VP-ITC and iTC200. In a typical ITC experiment, one interacting protein placed in the injection syringe is titrated in small aliquots into another in a sample cell, which is temperature controlled and coupled to a reference cell via a sensitive thermopile/thermocouple circuit. The heat either released (an exothermic reaction) or absorbed (an endothermic reaction) at equilibrium conditions over the course of titration can be measured directly with sufficient sensitivity (in the order of several hundred nanojoules)\(^7\) (Figure 1C). ITC does not suffer from constraints of molecular size, shape or chemical constitution. Neither is there any need for immobilization or modification of proteins.\(^7\) However, extreme care must be taken in the use of a buffer in an ITC experiment, that is, both interacting proteins in the sample cell as well as the syringe should be dialyzed in an identical buffer to minimize the artifacts generated by mismatched buffer components.\(^7\) The main parameters extracted from ITC experiment include association constant (\(K_a\)), enthalpy change (\(\Delta H\)), reaction stoichiometry (\(n\)), heat capacity change (\(\Delta C_p\)), free energy change (\(\Delta G\)) and entropy change (\(\Delta S\)).\(^7\) These data help us better understand both the affinities and mechanisms of PPIs, for example, estimating the hydrophobic effect of a residue at protein–protein interface by calculating changes in solvent portion of the binding entropy,\(^7\) determining exact contributions to binding including solvation, protonation, and structural rearrangements,\(^7\) validating oligomeric state of a protein during binding events according to stoichiometry data.\(^7\)

DSC measures the heat effects of temperature-induced transitions of proteins in dilute solutions. Most modern DSC instruments, such as nano-DSC from Calorimetry Science Corp., have two cells: a sample cell containing interacting proteins and a reference cell containing only solvent. By applying a controlled temperature program heating up to 100 °C or cooling down to 0 °C at a constant rate, the power required to maintain identical temperature between two cells is converted into molar heat capacity and recorded as a function of temperature.\(^8\) From a DSC thermogram, the transition temperature (\(T_m\)), enthalpy change (\(\Delta H\)) and heat capacity change (\(\Delta C_p\)) can be obtained. As the thermal stability of proteins will be affected by complex formation, DSC experiments can be used to study PPIs by observing shifts in \(T_m\) values.\(^8\)

### 2.8. Microscale thermophoresis (MST)

Microscale thermophoresis (MST) emerged as a new powerful technique for quantitative analysis of PPIs in free solution, which has been applied widely in academia and industry. This technique detects the directed motion of molecules in a microscopic temperature gradient, an effect termed thermophoresis. This effect is highly sensitive to binding-induced changes in various molecular properties, such as size, charge, conformation or hydration shell.\(^9\) In a typical MST experiment, one of the interacting proteins is fluorescently labeled. It is normally done by using \(N\)-hydroxysuccinimide (NHS) dye to react with a primary amine group on the labeled protein, or a recombinant protein fused to a fluorescent protein such as green fluorescent protein (GFP). Alternatively, MST can be performed in a label-free manner, in which the intrinsic UV-fluorescent signal of the aromatic amino acid residues is detected. A MST instrument uses an infrared (IR) laser to heat a defined sample volume inside a capillary and produces a temperature gradient spanning 2–6 K. The thermophoretic movement of the fluorescent protein is measured by detecting the fluorescence emission of the heat spot within the sample before, during, and after the IR laser is switched on. Before heating, the initial fluorescence is recorded at a constant level. Once the IR laser is switched on, the increase of temperature leads to an abrupt decrease in fluorescence intensity, so called temperature jump (\(T\)-jump). After that, the fluorescence decreases slowly (thermophoresis) until reaching a steady state. When the IR laser is switched off, the fluorescence recovery (inverse \(T\)-jump and back-diffusion) is observed (Figure 1D). During this procedure, both \(T\)-jump and the thermophoresis process can be affected by a binding event due to changes in size, charge or hydration shell. Hence, MST can be applied to determine dissociation constants of PPIs via titration, in which a constant concentration of fluorescent protein and varying concentrations of its nonfluorescent binding partner are applied.\(^8\)

MST has several advantages over other in vitro biophysical methods, including easy setting, fast measurement (binding affinity is determined in 10 min), and low sample consumption (a few microliters at nanomolar concentrations). It allows quantitative measurement of high-affinity interactions with dissociation constant in the picomolar range. In addition, it can be performed in any buffer or even in complex bioliquids like blood serum or cell lysate.\(^8\)

### 3. Biochemical Methods

#### 3.1. Fluorescence and bioluminescence resonance energy transfer (FRET and BRET)

Fluorescence resonance energy transfer (FRET) refers to the nonradiative (dipole–dipole) energy transfer from a donor fluorophore (fused to the “bait”) upon excitation to an acceptor fluorophore (fused to the “prey”).\(^9\) In PPI studies, intermolecular FRET only occurs when the bait and the prey are engaged in complex formation that brings two fluorophores in very close proximity (1–10 nm) (Figure 2A). Furthermore, a donor–
Figure 2. Illustration of some biochemical methods for PPI detection. A) Fluorescence resonance energy transfer (FRET) between a CFP-fused bait and a YFP-fused prey, where CFP and YFP act as the donor and the acceptor fluorophore, respectively. If the bait and the prey do not interact, excitation of CFP results in light emission (475 nm) by CFP only. When bait–prey interaction occurs, CFP and YFP are brought into proximity, leading to energy transfer from CFP to nearby YFP. Then, YFP emission can be detected at 528 nm. B) Bioluminescence resonance energy transfer (BRET) between a Rluc-fused bait and a EYFP-fused prey, where Rluc and EYFP act as the donor luciferase and the acceptor fluorophore, respectively. If bait and prey do not interact, only the blue-emitting spectrum (480 nm) of Rluc can be detected. When bait–prey interaction occurs, bioluminescence energy generated by Rluc in the presence of its substrate (coelenterazine) is transferred to EYFP, which in turn generates light emission at 530 nm. C) AlphaScreen. Upon excitation at 680 nm, a photosensitizing agent on the donor bead converts ambient oxygen to singlet oxygen. When bait–prey interaction brings the donor and the acceptor bead into close proximity, energy is transferred from singlet oxygen to thioxene derivatives on the acceptor bead, resulting in light emission at 520–620 nm. D) Protein-fragment complementation assay (PCA). Complementary C-terminal and N-terminal fragments of GFP are fused to the bait and the prey, respectively. If the bait and the prey do not interact, the fragments remain unstructured and lack functional activity. When the bait–prey interaction occurs, two fragments are brought together to form a full-length GFP with the native structure and function. E) Pull-down assay. In a GST pull-down assay, the bait is expressed as a GST-fused protein, and it is incubated with cell or tissue lysate containing the prey. The GST–bait–prey complex is captured by glutathione-coupled beads and eluted for analysis by SDS-PAGE. F) Tandem affinity purification (TAP). In this type of assay, the TAP tag (including CBP, a TEV cleavage site, and ProtA) is fused to the bait, and then the whole construction is introduced into the host cell. The TAP–bait–prey complex is isolated from cell lysate via affinity column containing IgG beads. After washing, TEV protease is added to release the CBP–bait–prey complex, which is then immobilized on calmodulin-coated beads. After washing, the native complex is eluted by chelating calcium using EGTA. G) Co-immunoprecipitation (Co-IP). The bait–prey complex is captured by glutathione-coupled beads and eluted for analysis by SDS-PAGE. H) ELISA. In the case of direct ELISA, the bait is immobilized on a microwell plate, and then cell lysate containing the prey is applied to the wells. After incubation and washing away unbound proteins, the bait–prey complex is detected by an HRP-conjugated antibody against the prey. Finally, a substance of the HRP enzyme is added, which in turn produces detectable fluorescence- or absorbance-based signal via the enzymatic reaction. I) Proximity ligation assay (PLA). The bait and prey are recognized by their respective PLA probes consisting of antibodies and single-stranded DNA oligonucleotides, where the free oligonucleotide ends are brought into proximity. An oligonucleotide connector is added to hybridize to the complementary oligonucleotides on the PLA probes. A ligase seals the nick between two PLA probes and forms a new DNA strand that can be amplified by PCR.
acceptor FRET pair requires an overlap between the emission spectrum of the donor and the excitation spectrum of the acceptor to obtain efficient energy transfer. The widely used FRET pairs come from variants of GFP for example, cyan fluorescent protein (CFP)–yellow fluorescent protein (YFP), blue fluorescent protein (BFP)–GFP [92], CyPet–Ypet [93], MiCy–mKO [94], cerulean fluorescent protein (CrFP)–YFP [95] and mVenuses–mStrawberry [96]. Such large fluorophore can also be replaced by a small, membrane-permeable fluorescein derivative termed FlAsH, which targets to a short tetracysteine motif (e.g., CCPGGCC) introduced to the protein sequence without disturbing the protein function [97].

With the development of optimized FRET pairs, FRET has been applied to in vitro binding assay for screening of PPI inhibitors [98,99]. Combined with microscopy imaging technologies, it can also be used to visualize and quantify PPIs in real time in living cells, with high spatiotemporal resolution [100–102]. As FRET microscopy relies on the ability to collect fluorescent signals from protein interactions, it suffers from some drawbacks, including auto-fluorescence, detector noise, optical noise, photobleaching, and spectral bleedthrough (partial overlap of donor and acceptor emission wavelengths). [103] Therefore, several methods have been used to correct for these problems and calculate FRET efficiency, each with its own strengths and weaknesses [104], including sensitized emission [105,106] acceptor photobleaching [107–109] fluorescence life-time imaging microscopy (FLIM) [110–112] spectral imaging [113,114] and fluorescence polarization imaging [115,116]. FRET experiments can also be performed using a flow cytometer. This type of experiment allows the analysis of a large number of cells in a short time, producing statistically more robust results than the microscope technique. However, it cannot be used to analyze intracellular structures. [108,117,118] Applicability of FRET with a laser scanning cytometer (LSC) combines attributes of both microscopy and flow cytometry, and yields information on a pixel-by-pixel and cell-by-cell basis [119].

Time-resolved FRET (TR-FRET) is a special form of FRET in which long-life emission fluorophores (lanthanides such as europium and terbium) are used as donors. Commonly used donor–acceptor pairs includes europium chelate or cryptate–allophycocyanin [120,121], europium chelate–Cy5 [122], europium chelate–Dy647 [123], terbium cryptate–d2 [124] and terbium–AF488 [125]. As carriers of lanthanides, chelates and cryptate not only absorb light and transfer energy to lanthanides, but also protect lanthanides from quenching by water molecules [126]. Labeling the protein of interest is usually carried out by using primary antibodies or antibodies against epitopes, such as glutathione S-transferase (GST), to which the donor or acceptor fluorophores are conjugated. Since the donor species used in TR-FRET has longer emission lifetime (micro- to milliseconds) as compared with conventional fluorophores (1–50 ns), it permits a temporal delay between donor excitation and detection of acceptor emission. This property enables TR-FRET to circumvent limitations of conventional FRET, such as cross-talk, photobleaching, and bleedthrough. It also decreases interference from short-life background fluorescence, such as autofluorescence from library compounds and endogenous fluorescent proteins in biological fluids or serum [126–128]. Besides quantifying association and dissociation events through time-gated fluorescence intensity measurements, TR-FRET can also be performed in a simple mix-and-read format for HTS of PPI inhibitors [121–123].

Bioluminescence resonance energy transfer (BRET) is analogous to FRET except that the energy donor is a bioluminescent protein rather than a fluorescent protein. Upon oxidation of its substrate, the luminescent donor protein, typically Renilla luciferase (Rluc), emits light to excite the fluorescent acceptor protein in close proximity. [129] As BRET does not require extrinsic excitation by a light source, thereby avoiding the technical problems associated with FRET, such as auto-fluorescence, photobleaching, simultaneous excitation of both donor and acceptor fluorophores [130]. In the original version of BRET, nowadays called BRET1, enhanced YFP (EYFP) is used as the acceptor and coelenterazine h is used as the substrate for Rluc [131] (Figure 2B). Thereafter, several other versions of BRET with different substrates and energy donor–acceptor pairs have been developed, for example, BRET2, with GFP (a blue-shifted variant of GFP) as the acceptor and a modified coelenterazine, DeepBlueC, as the substrate for Rluc, leading to increased spectral separation between donor and acceptor emission [132]. Extended BRET (eBRET) enables the investigation of PPIs in real time over extended timescales by using a protected form of coelenterazine h, termed EnduRen [133]. BRET3 exhibits several fold improvement in light intensity by using RLuc8 and mOrange as the donor and acceptor, respectively [134]. NanoBRET, with a coelenterazine derivative (furimazine) as the donor and a chloroalkane derivative of nochloro TOM (NCT) dye as the acceptor, provides improved spectral resolution and greater light intensity [135]. Sometimes microscopic imaging for BRET is limited due to its low level of light emission from bioluminescent reactions. With improvements in sensitive CCD cameras and advanced microscopes, new derivations of BRET technology become suitable for capturing luminescent light and progress to towards BRET imaging in single cells, tissues, and live animals [136,137,138].

3.2. AlphaScreen and AlphaLISA

Both AlphaScreen and AlphaLISA are bead-based proximity assays used to detect biomolecular interactions, where the acronym “Alpha” stands for “amplified luminescent proximity homogeneous assay”. The key components of these assays are donor and acceptor beads (250–350 nm in diameter) coated with different functional groups served as a photosensitizer and chemiluminescent unit, respectively. Donor and acceptor beads are brought into proximity by the interactions between the binding proteins immobilized on them. Upon excitation at 680 nm, phthalocyanine coated on the donor beads converts ambient oxygen to a more excited singlet state. Each donor bead can generate approximately 60 000 singlet oxygen molecules, resulting in exceptionally high signal amplification. Within its 4 µs half-life, singlet oxygen travels over a constrainted distance (< 200 nm) and triggers a cascade of energy transfer to thiocyanate derivatives within the acceptor beads, ultimate-
ly generating a luminescent signal. In the case of AlphaScreen, dyes within the acceptor beads are thioxene, anthracene, and rubrene, which emit light at 520–620 nm; while in the case of AlphaLiSA, anthracene and rubrene are replaced by europium chelate, which has a much narrower emission wavelength bandwidth centered around 615 nm (Figure 2C).[138, 139]

Alpha-based assays provide a versatile and sensitive technique for characterizing PPIs as well as their modulators in an HTS format.[140–146] They have the following technical advantages: (1) broad energy transfer distance (200 nm) compared with FRET (~10 nm), which allows a greater range of analyte that can be studied, such as full-length proteins and immuno-complexes; (2) detection of interactions with a wide range of affinities ($K_d$ values ranging from picomolar to millimolar); (3) ease of use. Many types of beads coupled to secondary antibodies, streptavidin, or to glutathione or nickel chelate for capture of expression tags are available.[139] However, as each type of Alpha beads has a characteristic bead capacity, there exists a “hook point”, where an excess of the target protein oversaturates the donor or acceptor beads and inhibits their association, resulting in a progressive signal decrease. Hence, performing a cross-titration assay beforehand to select an appropriate protein concentration below the hook point is essential for the success of an Alpha-based assay.[147]

3.3. Protein-fragment complementation assay (PCA)

In this type of assay, two proteins of interest (the bait and the prey) are fused to a complementary N- or C-terminal fragment of a reporter protein that is rationally dissected. If bait and prey interact, reporter protein fragments are brought into proximity leading to the folding and reconstitution of the activity of the reporter protein (Figure 2D).[148] Commonly used reporter proteins include dihydrofolate reductase (DHFR),[149] beta-lactamase,[150] beta-galactosidase,[151] yeast cytosine deaminase (yCD),[152] ubiquitin,[153] GFP and it variants,[154] luciferase,[155] and so on. They provide detectable effects, such as cell survival on selective medium, the change of colorimetric signals, or the appearance of fluorescence or luminescence, for desired applications:

(1) Split-DHFR PCA is based on the theory that DHFR can catalyze the reduction of dihydrofolate into tetrahydrofolate, which is crucial for cell growth and proliferation. This simple survival selection assay is widely used as a systematic approach for large-scale analysis, for example, to monitor signal transduction network in eukaryotes,[156] to map the yeast protein interactome,[157] or to study the global organization of yeast PPI networks in response to environmental perturbations.[158]

(2) In split-ubiquitin PCA, if bait–prey interaction occurs, re-assembly of ubiquitin recruits ubiquitin-specific protease that cleaves off a reporter protein attached to the C terminus of ubiquitin. Such reporter proteins for detecting the reassociation of ubiquitin fragments include haemagglutinin (HA)-tagged DHFR,[153] LexA-VP16,[159] RUra3p,[160] and Rgpt2,[161] whose release results in a shift in Western blot by using anti-HA antibody, LacZ and His3 reporter genes activation, cell resistance to 5-fluoroorotic acid, and cell resistance to hypoxanthine/aminopterin/thymidine and sensitivity to 6-thioguanine, respectively. This method has been employed in PPI analysis involving transcription factors and membrane proteins, which cannot be studied in conventional yeast two-hybrid system (see Section 4.2).[162, 163]

(3) PCA based on fluorescent proteins, also called bimolecular fluorescence complementation (BiFC), offers the advantage of trapping weak protein complexes, due to its irreversible chromophores formation, and enables direct visualization of PPIs at high spatial resolution in any subcellular compartment in a wide variety of cell types and organisms. Recently, a tripartite split-GFP association system based on the association of three fragments of GFP, two short peptides each tagged to bait and prey, and a third large detector fragment, was established. This method enables the minimization of aggregation and folding interference problem typically encountered in a traditional split-GFP assay with bulky and poorly folded fragments of GFP, and thus enables PPIs to be monitored both in vivo and in vitro.[164] However, the irreversibility of BiFC prevents dynamic analysis of complex dissociation and partner exchange. Furthermore, spontaneous association of fluorescent proteins independent of an interaction between bait and prey can cause false-positive results.[165]

(4) The split-luciferase assay utilizes luciferase as a reporter protein to catalyze the oxidation of its substrate luciferin and produce bioluminescence. Compared with irreversible fluorescence readout PCA, its reversibility allows the detection of kinetic and equilibrium aspects of the formation and disruption of PPIs[156] and application in drug discovery for the HTS of chemicals that target a specific PPI in living cells.[167] Such an assay has the advantage of a high signal-to-noise ratio due to low cellular background luminescence, however, the inefficiency of the chemical reaction required to create excited-state luminescent substrates compared with the excitation of fluorescent proteins represents a limitation.[168]

3.4. Affinity chromatography

Affinity chromatography is a biochemical separation technique for purification of a specific molecule from complex mixtures, based on highly specific interactions between two molecules, for example, enzyme and substrate, receptor and ligand, antibody and antigen. It greatly enhances the speed and efficiency of protein purification and also provides the technology platform to perform PPI studies, such as pull-down, tandem affinity purification (TAP), and antibody-based methods including co-immunoprecipitation (Co-IP) and enzyme-linked immunosorbent assay (ELISA).

In the original pull-down assay devised by Kaelin et al.,[169] the bait protein fused to GST was expressed in bacteria and immobilized on glutathione-coupled beads to capture interacting proteins from a cell lysate. After washing away nonspecific proteins, bound proteins were eluted and identified by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 2E). This method is useful for confirming suspected interactions or discovering unknown interactions. Other fusion tags such as polyhistidine,[170] HaloTag (a modified dehaloge-
or biotin.\textsuperscript{174} K\textsubscript{m} can also be used to express the bait protein captured by Ni-NTA agarose, HaloLink resin, or streptavidin magnetic beads, respectively. To decrease the amount of protein needed for a single experiment, Schulte et al. developed SiNgled Bead Affinity Detection (SINBAD) as a miniaturized pull-down method. Using affinity-tagged bait and color-coded prey with quantum dots, binding events can be directly visualized on the surface of a single affinity chromatography bead by fluorescence microscopy, without elution and gel separation of the isolated protein complexes.\textsuperscript{173} Another bead-based pull-down assay termed Bead Halo requires affinity-tagged bait and fluorescently tagged prey, with a positive interaction visible as a fluorescent halo around the beads. It can detect specific low-affinity interactions (beyond the \( K_{D} \approx 5\ \mu M \) range), which can be removed during wash steps in original pull-down assay.\textsuperscript{174}

The TAP method is a two-step affinity purification strategy that employs a dual-affinity TAP tag fused to the bait protein. In the original system, the TAP tag consists of two IgG-binding units of protein A of Staphylococcus aureus (ProtA) and a calmodulin-binding peptide (CBP) separated by a TEV protease cleavage site. In the first step, the tagged protein along with its interacting partners are immobilized on IgG-containing resin by the ProtA fraction, and then eluted by TEV protease cleavage. In the second step, the eluent is incubated with calmodulin-coated beads to remove contaminants and remaining TEV protease, and then finally eluted by using chelating agents (Figure 2F).\textsuperscript{175} These two affinity purification steps of TAP significantly decrease nonspecific background compared with single-step purification.

Coupled to MS for further protein identification, TAP has been successfully applied to large-scale analysis of protein complexes in lower organisms\textsuperscript{176,178} and to map smaller interactomes and signaling pathways in mammalian cells.\textsuperscript{179–182} A broad variety of TAP tags are available, such as Myc, FLAG, polyhistidine, HA, streptavidin-binding peptide (SBP), Strep II tag, and Streptococcus protein G (ProtG), to improve protein yields and recovery in mammalian or insect cells other than yeast.\textsuperscript{183}

Co-IP is a popular technique to identify PPIs and has been used in thousands of experiments. It shares the fundamental principle of the antigen–antibody reaction, utilizing a bait-specific antibody to precipitate the bait protein (i.e., the antigen) as well as its binding partners in the sample solution, followed by MS or Western blot analysis (Figure 2G).\textsuperscript{184} Both monoclonal and polyclonal antibodies can be used in co-IP and have their own advantages in different applications. Monoclonal antibodies recognize a single epitope on antigen with highly specificity to decrease background noise and cross-reactivity. In contrast, polyclonal antibodies recognize multiple epitopes, providing more robust detection and giving stronger signals. Considerable amounts of protein extract, typically 100–1000 µg, are required for a standard co-IP experiment, which might be unsuitable for cell culture assays conducted in a 96-well microtiter plate format or biopsies of tumor samples.

A microsphere-based co-IP assay (µco-IP) was developed that employs only 5–10 µg of protein extract.\textsuperscript{185} As co-IP is performed with whole cell or tissue lysate under nondenaturating conditions, it reflects a physiological association of endogenous proteins without any artificial effects, such as overexpression or modification of the protein of interest. However, it is not as sensitive as other affinity chromatography method, since the concentration of the antigen is much lower in its natural conditions. One solution involves the use of a cross-linking strategy in which cells or tissues are treated with chemical or photoactive cross-linkers to join interacting proteins covalently.\textsuperscript{186}

ELISA is a plate-based assay designed for detecting the presence of an antigen or an antibody in sample solution. It has been widely applied as a diagnostic tool in medical sciences, a quality control in various industries, as well as a quantitative method for assessing PPIs at physiological conditions and screening for PPI inhibitors.\textsuperscript{187–189} To measure a PPI in the direct ELISA format, the protein of interest is noncovalently coated to a polystyrene microwell plate through passive adsorption. After washing away excess free protein, the remaining available binding surfaces of the wells are blocked with nonreacting protein such as bovine serum albumin (BSA). Then, the sample solution containing the binding partner protein is added and incubated in the wells to interact with the immobilized protein. After another washing step, a primary antibody against the binding partner protein with an attached enzyme, for example horseradish peroxidase (HRP) or alkaline phosphatase (AP), is applied as a detection antibody. Finally, a substrate for the enzyme is added, which elicits a chromogenic or fluorescent signal upon enzymatic reaction. The amount of the binding partner protein bound to the immobilized protein is quantified through absorbance- or fluorescence-based readout.\textsuperscript{190} (Figure 2H). ELISA can also be performed in an indirect format in which the binding partner protein is detected by a primary antibody and an enzyme-linked secondary antibody, or in a sandwich format in which the protein of interest is coated to the plate through a pre-coated “capture” antibody. ELISA exhibits superior sensitivity due to signal amplification through the enzymatic oxidation or hydrolysis of substrates to produce enhanced color or fluorescence.\textsuperscript{191} However, the multiple incubation and washing steps needed in an ELISA are time consuming and can disrupt weak interactions (e.g., \( K_{D} \approx 1\ \mu M \)).\textsuperscript{192}

Proximity ligation assay (PLA) is a novel technique developed for analyzing an endogenous protein via an amplifiable DNA reporter with high specificity and sensitivity.\textsuperscript{193} It allows detection, quantification, and localization of PPIs in cells and tissues. In the latter case, a pair of proximity probes, consisting of antibodies or aptamers to which short single-stranded DNA oligonucleotides have been attached, are designed to bind to their respective target proteins. Upon interaction between target proteins, an added oligonucleotide connector can hybridize to both oligonucleotides of the proximity probes. New DNA strands are formed by joining oligonucleotide ends in ligation reaction and can be further amplified by real-time polymerase chain reaction (PCR) for highly sensitive detection.\textsuperscript{194–196} (Figure 2I). Alternatively, rolling circle amplification (RCA) can be employed to generate a localized amplification product that is further hybridized with fluorescence-la-
beled oligonucleotides. Hence, interactions of endogenous proteins, served as specific fluorescent dots, can be visualized and counted in situ with confocal microscopy. This proximity ligation in situ assay (P-LISA) can also be performed in combination with flow cytometry, providing a statistical approach of analyzing PPIs on a single-cell basis.

P-LISA has advantages over other fluorescence-based methods, such as FRET, BRET, and BiFC. It allows detection of endogenous proteins in cellular context without being affected by overexpression. Besides, it provides powerful signal amplification via RCA because the protein complex becomes labelled by several hundred copies of the fluorophore. However, it should be noted that PLA only detects the proximity of proteins within a certain distance, typically up to a few tens of nanometers, rather than physical molecular interactions.

3.5. Cross-linking

Cross-linking is a process that links two or more of the proteins in a complex through covalent bonds by using a cross-linker, which is defined as a chemical reagent containing two or more reactive groups connected by a spacer or linker region. This technique has been applied to characterize low-affinity PPIs that are difficult to detect by other methods or to stabilize transient PPIs in a dynamic process both in vitro and in vivo. For a specific PPI discovery and characterization, an appropriate cross-linker should be selected empirically. Cross-linkers are varied with respect to the chemical reactivity, spacer arm length, water solubility, cell membrane permeability, cleavability, and so on; most cross-linkers are commercially available from Pierce Chemicals (Rockford, IL, USA). For instance, the most frequently used cross-linker is the homo-bifunctional cross-linker containing two NHS esters, which target primary amines in proteins with high reaction efficiency under physiological conditions (pH 7–8). Hetero-bifunctional cross-linkers usually consist of one amino-reactive group and one thiol-reactive group such as maleimide, pyridyldisulfide, and iodoacetyl. Photo-reactive cross-linkers with photo-reactive sites such as aryl azide, diazirine, and benzophenone, can be induced by irradiation with UV-light to attach binding proteins. Membrane-permeable cross-linkers such as formaldehyde enable to detect PPIs in their native environment. Moreover, isotope-labeled cross-linkers, cleavable cross-linkers, fluorogenic cross-linkers, and trifunctional cross-linkers containing a biotin group have been employed to facilitate identification and affinity purification of cross-linker containing species from complicated reaction products.

In fact, cross-linking is always applied in combination with other downstream methods for further analysis of the cross-linked proteins, involving SDS-PAGE to separate cross-linked from non-cross-linked proteins, TAP or immunoaffinity chromatography for affinity-based purification of cross-linked products, MS techniques to identify the interacting partners. A considerable weakness of this method is the high risk of detecting nonspecific interactions, as it joins proteins in close vicinity that might not be in direct contact. Additionally, identification of the cross-linked proteins can be quite cumbersome due to low abundance of the cross-linked species and data complexity of tandem mass spectrometry (MS–MS) spectra.

4. Genetic Methods

4.1. Phage display

Phage display was initially described by Smith in 1985, who reported that foreign DNA can be inserted into filamentous phage gene III to create a fusion protein displayed on the surface of phage particles. Since then, this technique has been extensively used in multiple applications, including isolating monoclonal antibodies, mapping epitope of the antigen involved in antibody binding, identifying substrates of various proteases, developing modulators of both active and allosteric sites of an enzyme, finding receptor agonists and antagonists, as well as searching for interacting partners of a protein.

In the case of PPI studies, by using M13 filamentous phage display, a foreign DNA sequence encoding prey protein or peptide is inserted into the pIII or pVIII gene encoding minor or major coat protein, respectively, creating fusions of prey with coat protein displayed on phage surface in either a high- or low-valency format, resulting in a direct linkage between the genotype and the phenotype. Subsequently, pools of phages exposing different prey proteins or peptides are incubated with a bait protein, which is immobilized on a solid surface. Specific interacting phages get captured whereas unspecific phages are washed off. Binding preys are then eluted, amplified by reinfecion of Escherichia coli cells, enriched more than 1000-fold over ordinary phages by successive rounds of a process called “panning”, and finally identified by sequencing the phage DNA (Figure 3A). Therefore, the great advantages of phage display originate in the incorporation of the protein and genetic information into a single phage particle, which allows fast determination of selected phages, as well as multiplex selection and amplification steps, which make it more sensitive in identifying unknown binding proteins compared with other bioassays.

The preferred systems that have been used for phage display include the filamentous M13 phage and the lytic T7 phage. M13 phage display provides easy manipulation in library construction but encounters misfolding problems as the displayed proteins are secreted through the periplasmic space of the E. coli membrane. While T7 phage display, with fusion proteins displayed on the C terminus of capsid protein 10B, does not suffer from such issues as the displayed protein released from the host cell through lysis. A high-quality, large-diversity phage display library constructed from designed synthetic oligonucleotides, cDNA, or open reading frames (ORFs), is an essential requirement for PPI screening. For example, Karkkainen et al. designed codon-optimized genes encoding for 296 human SH3 domains and cloned them into a M13-derived phagemid vector for objective identification of SH3 domains that bind to target proteins of interest. Lin et al. identified a kaposin A-interacting protein by using a phage display human brain cDNA library. In general, phage display with cDNA libraries is inefficient and uncommon, as a cDNA library
A transcriptional activator in yeast GAL4 contains two distinct domains, a DNA-binding domain (DBD) and an activation domain (AD), which independently are nonfunctional as a transcription factor, but reconstitute a functional GAL4 transcription factor when they are in close proximity.[219] Inspired by this finding, Fields and Song first described the Y2H method in 1989.[220] They fused two yeast proteins bait and prey to the DBD and AD of GAL4, respectively. As bait and prey were known to interact, co-expression of these two hybrid proteins in yeast reconstituted proximity of the GAL4 domains, leading to transcription of a GAL1-lacZ reporter gene that could be analyzed by β-galactosidase activity measurement (Figure 3B).

Y2H is a powerful technique for identifying and analyzing PPIs in vivo. However, as Y2H requires the translocation of the interacting proteins into the nucleus, it’s not suitable for membrane proteins or proteins localized in other subcellular compartments. Other unsuited proteins include transcriptional activators, which produce spontaneous activation of the transcription of the reporter gene independent of bait–prey interaction. Moreover, Y2H cannot detect PPIs involving post-translational modifications that do not take place in yeast.[221]

To circumvent the limitations of Y2H and extend its applicability, a number of variant two-hybrid systems have been developed:

1. The repressed transactivator (RTA) system was established for detection of interactions with transcriptional activator proteins.[222] An auto-activating bait protein is fused to DBD of GAL4, whereas prey is fused to the repression domain (RD) of a transcription repressor TUP1. Therefore, bait–prey interaction causes repression of a reporter gene. This method has been applied to screen for various transcriptional activators, such as VP16,[222] c-Myc,[222–224] Mitf,[225] and androgen receptor.[226–228]

2. RNA polymerase III (Pol III) system can be used to assess the interactions between RNA polymerase II regulatory proteins and their partners, which cannot be analyzed by the classic, RNA polymerase II-dependent Y2H. In the RNA Pol III system, protein prey is fused to τ138p, which can activate RNA polymerase III-mediated transcription.[229]

3. Sos recruitment system (SRS) is based on the Ras signaling pathway instead of transcriptional regulation in classic Y2H. In SRS, bait is fused to mammalian Sos, and prey is fused to a membrane localization signal. Interaction between bait and prey results in Sos membrane localization and Ras activation, allowing a Cdc25-2 temperature sensitive yeast strain grow at the restrictive temperature.[230] Since SRS does not depend on a transcriptional readout, it can study interactions involving transcriptional activators or repressors. Furthermore, it is suitable for identifying proteins that do not target well to the nucleus or undergo post-translational modifications in the cytoplasm.

4. An improved version of SRS is the Ras recruitment system (RRS), in which bait is fused to constitutively active mammalian Ras instead of Sos.[231]
(5) The G protein fusion system allows the interaction between an integral membrane protein bait and a soluble prey to be studied. By this method, the prey fused to the G protein γ-subunit is localized at the membrane upon interaction with bait, leading to the disruption of G protein signaling and downstream transcriptional events.\(^{[232]}\)

(6) SCINEX-P system is developed to analyze interactions between extracellular or transmembrane proteins in the endoplasmic reticulum (ER), in which bait and prey are fused to truncated Ire1p. Bait–prey interaction causes dimerization of Ire1p, which activates unfolded protein response signaling, leading to Hac1p-dependent activation of reporter genes.\(^{[233]}\)

(7) Reverse two-hybrid system is an upside-down version of Y2H, in which disruption of the bait–prey interaction decreases the expression of a counter-selectable reporter gene, such as URA3.\(^{[234]}\) CVH2\(^{[235]}\) and GAL1,\(^{[236]}\) and allows cell growth under selective conditions. This method facilitates the detection of dissociation of PPIs, and the selection of small-molecule inhibitors against PPIs that could be developed as therapeutic agents.\(^{[247]}\)

(8) Bacterial two-hybrid system using E. coli as a host organism offers some advantages over Y2H, including ease of use, lack of cellular compartmentalization, rapid library screening due to high transformation efficiency, and fast growth rate observed with E. coli, and increased sensitivity due to the absence of endogenous proteins that compete for the interactions.\(^{[146]}\)

A number of bacterial two-hybrid systems have been studied, which are based on recruitment of RNA polymerase to activate a lacZ reporter gene,\(^{[238]}\) complementation of signaling enzymes such as adenylate cyclase,\(^{[239]}\) dimerization of LexA derivatives to reconstitute a functional repressor,\(^{[240]}\) and reconstitution of polyhydroxyalkanoates synthesis regulatory protein PhaR.\(^{[241]}\)

(9) Mammalian two-hybrid system allows for the studies of interactions of proteins in native physiological context with proper folding and post-translational modifications. One unique system termed MAPPIT (mammalian protein–protein interaction trap) is founded on type I cytokine signal transduction, in which bait–prey interaction restores JAK–STAT signaling and finally induces transcription of a reporter gene coupled to a STAT-dependent promoter.\(^{[242]}\) This approach has been well validated for large-scale interactomics studies and high-throughput drug screening.\(^{[243]}\)

(10) Fluorescent two-hybrid (F2H) assay takes advantage of cell lines harboring a stable integration of a lac operator array. In such an assay, a fluorescently labeled bait is fused to a lac repressor and then immobilized. Interaction between the bait and a differently labeled prey leads to co-localization of fluorescent signals. This technique enables visualization and analysis of PPIs from different subcellular compartments including nucleus, cytoplasm, and mitochondria in living cells in real time.\(^{[244–246]}\)

4.3. Protein microarray

Protein microarray technology is a high-throughput method that allows simultaneous analysis of thousands of proteins arrayed at high density on a solid surface (Figure 3C). Among various forms of protein microarrays, functional protein microarray is constructed by immobilizing large number of full-length functional proteins or protein domains onto the chips, which is used to study biochemical properties and activities of an entire proteome in a single experiment including assessing PPIs.\(^{[247]}\) A major breakthrough came from a study by Zhu et al.\(^{[248]}\) They cloned 5800 yeast ORFs, expressed and purified their corresponding proteins with His\(^{6}\) fusion tag, spotted these proteins onto nickel-coated slides, and finally generated the first proteome microarray containing 80% yeast proteins. After incubating the microarrays with biotinylated calmodulin and Cy3-labeled streptavidin, calmodulin-interacting proteins were identified by fluorescence signal detection. Since then, proteome microarrays have been successfully employed for identification of protein interaction partners, that is, to discover calmodulin-binding proteins with microarray containing 1133 proteins in Arabidopsis thaliana.\(^{[249]}\) to investigate novel links of protein kinases to distinct cellular pathways using yeast proteome microarray composed of ~4200 GST fusion yeast proteins,\(^{[250]}\) to profile the BAG3 interactome by using a human proteome microarray comprising ~17000 unique full-length human ORFs.\(^{[251]}\)

The core technologies and corresponding challenges in functional protein microarray include:

(1) Construction of a comprehensive expression clone library. Although proteome microarrays have been fabricated from the proteomes of yeast,\(^{[248,250]}\) E. coli,\(^{[252]}\) coronavirus,\(^{[253]}\) Arabidopsis thaliana,\(^{[249]}\) and human,\(^{[251,254]}\) expression-ready ORF collections for other species are highly desired.

(2) High-throughput protein production, including protein expression and purification. Expression in a homologous system is ideal for analysis of yeast and E. coli proteins, because these proteins are in their natural environment and thus fold properly.\(^{[248,252]}\) While for most eukaryotic proteins, expression in a heterologous system such as in E. coli can cause solubility problems due to improper folding and the loss of native post-translational modification. For protein purification, affinity tags such as His\(^{6}\), GST, FLAG, maltose binding protein, are fused to either the C or N terminus of these proteins, which can facilitate purification by using specific resins. As the traditional cloning—expressing—purification—printing approach is time consuming and labor intensive, a nucleic acid programmable protein array (NAPPA) was developed. It spots expression plasmids on the chip and uses a mammalian cell-free expression system to allow the proteins to be expressed and immobilized in situ through affinity tags without requiring purification.\(^{[255]}\)

(3) Protein immobilization. As proteins have a 3D structure that is critical to their function, it is necessary to develop effective protein immobilization methods that do not cause activity loss or incorrect orientation of the capture proteins. To this end, various methods have been developed, such as noncovalent absorption,\(^{[256]}\) covalent binding by chemical bonding,\(^{[257,258]}\) and affinity capture.\(^{[248]}\)

(4) Detection method. Readout systems based on a label-dependent strategy, such as fluorescence, radioactivity, chemi-
luminescence, enzymatic signal amplification, or label-free strategy such as surface plasmon resonance, mass spectrometry, atomic force microscopy, can be used to detect signals within each microspot. [10]

5. Summary & Outlook

The human interactome is estimated to contain ~170,000 binary PPI pairs, the majority of which remains to be explored. [260] Revealing the temporal and spatial relationships between interacting proteins will certainly help to establish a better understanding of the molecular mechanism of biological processes as well as the identification of a wider range of potential targets for drug discovery. Biophysical methods, such as AUC and ITC, are often based on long-standing physical principles. They have started to show good performance in studying the hydrodynamic and thermodynamic aspects of PPIs owing to dramatic improvements in instrumentation. Biochemical methods, such as affinity chromatography, are also frequently applied together with MS techniques to identify the binding partners of the protein of interest. During the past decade, genomics revolution has assisted in the construction of genome-scale cDNA and ORF collections, which facilitates PPI identification and PPI network mapping on a large scale with high-throughput genetic methods, such as Y2H, phage display, and protein microarray. It should be kept in mind that all methods have their own limitations, and these can lead to false positives and negatives in detection. Therefore, validation of PPIs using methods based on different techniques is recommended to produce reliable results. For example, potential PPIs discovered by genetic methods can be further confirmed by biochemical assays. Finally, it is important to note that, since experimental measurements have generated a wealth of data, there is an increasing demand for publicly accessible databases that collect and record such data in some standard formats. More powerful bioinformatic tools are also much needed to convert raw data into useful knowledge.

Acknowledgements

The authors are grateful for the financial support from the National Science Foundation of China (grant 81430083, 81172984, 21072213), the Ministry of Science and Technology of China (863 High-Tech grant 2012AA020308), and the Macao Special Administrative Region Science and Technology Development Fund (grant 055/2013/A2).

Keywords: biochemical methods · biophysical methods · genetic methods · protein–protein interactions

[1] I. Eskundia, D. Juan, J. M. Rodriguez, A. Frankish, M. Diekhans, J. Harrow, J. Vazquez, A. Valencia, M. L. Tress, Hum. Mol. Genet. 2014, 23, 5866–5878.
[2] T. Berggård, S. Linsé, P. James, Proteomics 2007, 7, 2833–2842.
[3] P. Braun, A. C. Gingras, Proteomics 2012, 12, 1478–1498.
[4] J. Pelletier, S. Sidhu, Curr. Opin. Biotechnol. 2001, 12, 340–347.
[5] M. Vidal, M. E. Cusick, A. L. Barabási, Cell 2011, 144, 986–998.
[6] H. Ruffner, A. Bauer, T. Bouwmeester, Drug Discovery Today 2007, 12, 709–716.
[7] C. Tse, A. R. Shoemaker, J. Adicettes, M. G. Anderson, J. Chen, S. Jin, E. F. Johnson, K. C. Marsh, M. J. Mitten, P. Nimmer, L. Roberts, S. K. Tahir, Y. Xiao, X. Yang, H. Zhang, S. Fesik, H. S. Rosenberg, S. W. Elmore, Cancer Res. 2008, 68, 3421–3428.
[8] B. T. Yu, L. Vassilev, Curr. Top. Microbiol. Immunol. 2011, 348, 151–172.
[9] P. Dörre, M. Westby, S. Dobbs, P. Griffin, B. Irvine, M. Macartney, J. Mori, G. Rickett, C. Smith-Burchell, C. Napier, R. Webster, D. Armour, D. Price, B. Stammgen, A. Wood, M. Moras, Antimicrob. Agents Chemother. 2005, 49, 4721–4732.
[10] K. Peerlincx, I. De Lepeleire, M. Goldberg, D. Farrell, J. Barrett, E. Hand, D. Panebianco, H. Deckmyn, J. Vermeylen, J. Arnout, Circulation 1993, 88, 1512–1517.
[11] K. Uvebrant, D. da Graça Thirge, A. Rosen, M. Akesson, H. Berg, B. Walpe, P. Bjork, J. Biomol. Screening 2007, 12, 464–472.
[12] P. Tompa, N. E. Davey, T. J. Gibson, M. M. Babu, Mol. Cell 2014, 55, 161–169.
[13] F. Perrin, J. Phys. Radium. 1926, 7, 390–401.
[14] Y. Du in Protein–Protein Interactions (Eds.: C. L. Meyerkord, H. Fu), Vol. 1278 of the series Molecular Biology (Series Ed.: J. M. Walker), Springer, New York, 2015, p. 529–544.
[15] J. W. J. W. Hershley, Biochemistry 1981, 20, 5859–5865.
[16] J. W. J. W. Hershley, J. Biol. Chem. 1982, 257, 1215–1220.
[17] Y. T. Kim, S. Tabor, J. E. Churchich, C. C. Richardson, J. Biol. Chem. 1992, 267, 15032–15040.
[18] W. A. Lea, A. Simeonov, Expert Opin. Drug Discovery 2011, 6, 17–32.
[19] S. M. Knight, N. Umezawa, H. S. Lee, S. H. Gellman, B. K. Kay, Anal. Biochem. 2002, 300, 230–236.
[20] C. H. Kenny, W. Ding, K. Kelleher, S. Benard, E. G. Dushin, A. G. Sutherland, L. Mosyak, R. Kroz, G. Elstadi, Anal. Biochem. 2003, 323, 224–233.
[21] C. A. Hassig, F. Y. Zeng, P. Kung, M. Kiankarnim, S. Kim, P. W. Diaz, D. Zhai, K. Welsh, S. Morshedian, Y. Su, B. O’Keefe, D. J. Newman, Y. Rusman, H. Kaur, C. E. Salomon, S. G. Brown, B. Baire, A. R. Michel, T. R. Hoye, S. Francis, G. I. Georg, M. A. Walters, D. B. Divilanska, G. P. Roth, A. E. Wright, J. C. Reed, J. Biomol. Screening 2014, 19, 1201–1211.
[22] J. C. Owicik, J. Biomol. Screening 2000, 5, 297–306.
[23] A. Madeira, E. Ohman, A. Nilsson, B. Sjogren, P. E. Andren, P. Svenningson, Nat. Protoc. 2009, 4, 1023–1037.
[24] C. Boozer, G. Kim, S. Song, H. Guan, T. Londergan, Curr. Opin. Biotechnol. 2006, 17, 400–405.
[25] H. H. Nguyen, J. Park, S. Kang, M. Kim, Sensors 2015, 15, 10481–10510.
[26] J. M. McDonnell, Curr. Opin. Chem. Biol. 2001, 5, 572–577.
[27] R. L. Rich, D. G. Myszka, Curr. Opin. Biotechnol. 2000, 11, 54–61.
[28] S. Oestenæicher, W. F. Blum, B. Schmidt, T. Braulke, B. Kubler, J. Biol. Chem. 2005, 280, 9994–10000.
[29] E. F. Vernon-Wilson, W. J. Kee, A. C. Willis, A. N. Barclay, D. L. Simmons, M. H. Brown, Eur. J. Immunol. 2000, 30, 2130–2137.
[30] A. Rogstam, S. Linse, A. Lindqvist, P. James, L. Wagner, T. Berggård, Biochem. J. 2007, 401, 353–363.
[31] J. Homola, Anal. Bioanal. Chem. 2003, 377, 528–539.
[32] J. S. Yuk, H. S. Kim, J. W. Jung, S. H. Jung, J. S. Lee, W. J. Kim, J. A. Han, Y. M. Kim, K. S. Ha, Biosens. Bioelectron. 2006, 21, 1521–1528.
[33] F. Remy-Martin, M. El Osta, G. Lucchi, R. Zeggari, T. Leblois, S. Bellon, P. Ducoroy, W. Boireau, Anal. Bioanal. Chem. 2012, 404, 423–432.
[34] P. M. Thompson, M. R. Beck, S. L. Campbell, in Protein–Protein Interactions (Eds.: C. L. Meyerkord, H. Fu), Vol. 1278 of the series Methods in Molecular Biology (Series Ed.: J. M. Walker), Springer, New York, 2015, p. 267–279.
[35] D. S. Garrett, Y. J. Seok, A. Peterkosky, G. M. Clore, A. M. Gronenborn, Biochemistry 1997, 36, 4393–4398.
[36] M. K. Greene, K. Maskos, S. J. Landry, Proc. Natl. Acad. Sci. USA 1998, 95, 6108–6113.
[37] D. S. Garrett, Y. J. Seok, A. Peterkosky, A. M. Gronenborn, G. M. Clore, Nat. Struct. Biol. 1999, 6, 166–173.
[252] C. S. Chen, E. Korobkova, H. Chen, J. Zhu, X. Jian, S. C. Tao, C. He, H. Zhu, Nat. Methods 2008, 5, 69–74.
[253] H. Zhu, S. Hu, G. Jona, X. Zhu, N. Kreiswirth, B. M. Willey, T. Mazzulli, G. Liu, Q. Song, P. Chen, M. Cameron, A. Tyler, J. Wang, J. Wen, W. Chen, S. Compton, M. Snyder, Proc. Natl. Acad. Sci. USA 2006, 103, 4011–4016.
[254] J. S. Jeong, L. Jiang, E. Albino, J. Marrero, H. S. Rho, J. Hu, S. Hu, C. Vera, D. Bayron-Poueymiroy, Z. A. Rivera-Pacheco, L. Ramos, C. Torres-Castro, J. Olan, J. Bonaventura, J. D. Boeke, W. Y. Yap, I. Pino, D. J. Eichinger, H. Zhu, S. Blackshaw, Mol. Cell. Proteomics 2012, 11, O111.016253.
[255] N. Ramachandran, E. Hainsworth, B. Bhullar, S. Eisenstein, B. Rosen, A. Y. Lau, J. C. Walter, J. LaBaer, Science 2004, 305, 86–90.
[256] J. Ziauddin, D. M. Sabatini, Nature 2001, 411, 107–110.
[257] G. MacBeath, S. L. Schreiber, Science 2000, 289, 1760–1763.
[258] Y. Lee, E. K. Lee, Y. W. Cho, T. Matsui, I. C. Kang, T. S. Kim, M. H. Han, Proteomics 2003, 3, 2289–2304.
[259] S. Hu, Z. Xie, J. Qian, S. Blackshaw, H. Zhu, Wiley Interdiscip. Rev. Syst. Biol. Med. 2011, 3, 255–268.
[260] A. Schoenrock, B. Samanfar, S. Pitre, M. Hooshyar, K. Jin, C. A. Phillips, H. Wang, S. Phanse, K. Omidi, Y. Gui, M. Alamgir, A. Wong, F. Barrenas, M. Babu, M. Benson, M. A. Langston, J. R. Green, I. Dehne, A. Golshani, BMC Bioinf. 2014, 15, 383.

Received: October 21, 2015
Revised: January 8, 2016
Published online on February 11, 2016