We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,300

Open access books available

172,000

International authors and editors

190M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Unbiased Identification of Extracellular Protein–Protein Interactions for Drug Target and Biologic Drug Discovery

Shengya Cao and Nadia Martinez-Martin

Abstract

Technological improvements in unbiased screening have accelerated drug target discovery. In particular, membrane-embedded and secreted proteins have gained attention because of their ability to orchestrate intercellular communication. Dysregulation of their extracellular protein–protein interactions (ePPIs) underlies the initiation and progression of many human diseases. Practically, ePPIs are also accessible for modulation by therapeutics since they operate outside of the plasma membrane. Therefore, it is unsurprising that while these proteins make up about 30% of human genes, they encompass the majority of drug targets approved by the FDA. Even so, most secreted and membrane proteins remain uncharacterized in terms of binding partners and cellular functions. To address this, a number of approaches have been developed to overcome challenges associated with membrane protein biology and ePPI discovery. This chapter will cover recent advances that use high-throughput methods to move towards the generation of a comprehensive network of ePPIs in humans for future targeted drug discovery.

Keywords: drug discovery, high-throughput screening, extracellular protein–protein interactions, unbiased target discovery, receptors, membrane proteins, secreted proteins

1. Introduction: targeting ePPIs to address disease burden

The World Health Organization estimates that over 70% of deaths in 2016 worldwide were due to non-communicable diseases like cardiovascular disease (CVD) and cancer. This number is expected to grow to over 80% by 2060 [1]. Even if these diseases arise from environmental damage, the disease states usually depend on altered cellular communication, driven at the molecular level by altered ePPIs. For example, interactions between immune cells and arterial walls through adhesion proteins can initiate positive feedback loops which drive atherosclerotic plaque formation in CVD [2]. Similarly, while genetic mutations are the root cause of cancer, aberrant cell–cell interactions allow cancer cells to evade the immune system [3], migrate [4], siphon nutrients [5] and ignore signals to stop growing [6]. ePPIs also contribute to communicable diseases, which, highlighted by the

(ongoing as of this writing) COVID-19 pandemic, can rapidly increase human deaths with the introduction of a novel pathogen. As with many pathogens, the virus underlying the pandemic, SARS-CoV-2, exploits host cell-surface receptors to enter cells to replicate and spread [7, 8].

Because ePPIs are often central to the initiation and progression of diseases, they offer opportunities for molecular intervention using drugs. Greater understanding of the ePPIs underlying diseases allows them to be effectively targeted and manipulated to reverse disease phenotypes. For example, for CVD, several efforts to target different cytokines are showing promise in stemming the progression of atherosclerosis [9]. The development of cancer immunotherapies in the last decade has revolutionized cancer treatment. These treatments block ePPIs between immune checkpoint proteins such as CTLA-4 or PD-L1 and their binding partners to reinvigorate the body's defenses [10]. Even with SARS-CoV-2, an antibody cocktail (REGN-COV2) that blocks the ePPI between the virus spike protein and receptors on the cell, has been shown to stop viral entry and has gained emergency authorization for use in COVID-19 patients [11]. These examples highlight that identifying and targeting ePPIs can have strong therapeutic benefits in a variety of known and emerging diseases that make up a significant portion of human disease burden worldwide (**Figure 1**).

Despite the importance of ePPIs for both understanding and treating disease, our understanding of this field remains limited, especially compared to other classes of protein–protein interactions (PPIs). A main reason for this disparity is that common techniques for general PPI discovery are not well suited for ePPIs. Interactions between individual secreted or membrane proteins are typically weak, making them difficult to capture. Membrane proteins are biochemically recalcitrant and tend to misfold or aggregate outside of a native membrane context making them incompatible with many readouts designed for soluble proteins. Extracellular proteins also pick up many complex and heterogeneous post-translational modifications on their journey out of the cell, including specific disulfide bonds designed for the non-reducing extracellular environment. Since these can play roles in ePPIs but are not well characterized, they can be missed by common non-native expression

Figure 1. Examples of therapeutically relevant ePPIs. (A) The tumor microenvironment consists of a complex mix of cell types that communicate through ePPIs. One example is the expression of immune checkpoint proteins such as PD-L1 on cancer cells, which inhibits cytotoxic T-cell function, allowing the cancer cells to evade the immune system. Drugs targeting these ePPIs are the foundation for the cancer immunotherapies, which have provided significant benefits for cancer patients. Many other ePPIs in this space are under active investigation. (B) SAR-CoV-2 uses its spike protein to co-op the ACE2 receptor for viral entry into host cells and initiate viral replication and infection. Strategies for blocking this interaction are being explored to address the COVID-19 pandemic.

systems [12, 13]. Altogether, these biochemical features make most available technologies suboptimal and as a result, ePPIs are remarkably underrepresented in current databases.

Because of the difficulties with ePPI discovery, many new approaches have been developed to specifically identify human ePPIs that play roles in homeostasis and disease. While past low-throughput methods and focused studies have provided fundamental insights into specific receptors and pathways, the rapid explosion in sequencing, mass spectrometry (MS), targeted mutagenesis and high-throughput screening techniques has made the exhaustive identification of ePPIs a realistic goal. Here, we will address how new techniques deal with unique challenges associated with ePPIs and highlight the progress towards to the elucidation of a comprehensive network map of all human ePPIs.

2. Methods for detecting ePPIs

From biophysical approaches to *in vivo* studies, a number of methods have been developed or are being improved that have the potential to enable unbiased ePPIs discovery. The majority of methods can be categorized into a few broad technological concepts: biochemical fractionation, affinity purification, protein-fragment complementation, proximity labeling, direct protein interaction detection and computational modeling. As is the case for other disciplines, deciphering the complexities of extracellular interactions requires a multipronged approach. Since the different approaches provide different types of information, these methodologies are complementary. Especially as these categories have matured, many new techniques bridge the different concepts to balance the various benefits and shortcomings and push for increased throughput. The specific method-of-choice will depend on the expertise, equipment and overall resources available in each laboratory (Table 1).

Approach	Direct Interaction	Weak Interaction Detection	Correct Modification	Automation Preferred	False +	False –
Biochemical Fractionation	Variable	+	+++	No	++	+++
Affinity Purification	+	+	+++	Variable	++	+++
Protein-fragment complementation	++	++	++	Yes	++	++
Proximity labeling	+	+++	++	No	+++	+
Direct interaction screens	+++	++ – +++	+ – +++	Yes	+	+++

Table 1.
Comparison of approaches for unbiased detecting ePPIs.

2.1 Biochemical fractionation

2.1.1 Concept description

Biochemical fractionation is the splitting of a complex lysate, typically cell or tissue extract, into simpler mixtures to identify the simplest solution that retains a

certain biochemical property. The measurable property could be complex cellular activities, such as the stimulation of cell migration, or simple ones such as binding to a target protein (**Figure 2**).

2.1.2 Concept pros

Biochemical fractionation does not require any knowledge of the components and can be an unbiased technique. It is a versatile concept since any biochemical property can be studied, from *in vivo* tissue level responses to molecular PPIs. Some degree of fractionation is easily combined with other techniques to reduce the starting complexity and to improve data interpretability.

2.1.3 Concept cons

The results from biochemical fractionations are dependent on the particular purification steps used and can be highly variable. Due to the multiple purification steps, this approach can also be labor and time intensive. The different purification steps can inactivate proteins by inducing misfolding or removing key co-factors. This is especially true for ePPIs that involve membrane proteins, which can lose activity if extracted from membranes [12, 13].

2.1.4 Specific applications

Biochemical fractionations played a role historically in identifying some of the first extracellular signaling proteins like cytokines using activities such as macrophage migration or bacteria killing [14]. In the 21st century, this approach has identified stable soluble PPIs proteome-wide by fractionating cell lysates down to the level of co-eluting protein complexes and identifying them using MS [15]. While the specific purification steps used for soluble proteins are unlikely to be applicable to ePPIs, alternative centrifugation-based fractionation successfully recovered biochemically active membranes from crude fruit fly extracts [16]. Direct application of affinity purification from crude extracts without enriching for synaptic membranes did not recover known ePPIs [16]. However, using biochemical fractionation to enrich for synaptic components was necessary for the identification of key proteins in synapse formation using an affinity purification approach (described in the next section) [17].

Figure 2.

Biochemical fractionation can be used to reduce the complexity of a mixture while maintaining the desired activity. While traditionally performed in series, fractionation can also be done in parallel with modern purification techniques.

2.2 Affinity purification

2.2.1 Concept description

Affinity purification involves isolating a target-of-interest in non-denaturing conditions to enable co-isolation of any binding partners that are stably attached. The most common implementation is immunoprecipitation, where an antibody, generally attached to a solid substrate like a bead, plate or column, is used to specifically recognize the target-of-interest. Any factors that are not stably bound to the protein of interest are washed away by flushing the solid substrate with buffer. Proteins that survive the washes are identified (**Figure 3**).

2.2.2 Concept pros

Affinity purification allows for the direct isolation of a target-of-interest from complex mixtures. It is versatile and can be combined with many other approaches.

2.2.3 Concept cons

To isolate the target-of-interest, there needs to be a reagent, like an antibody, that will specifically and tightly bind the target. Since such reagents are not always readily available, the target may need to be tagged and introduced exogenously, which can affect target behavior. Affinity purification for unbiased identification of binding partners requires either large tagged libraries or access to MS. Importantly, to be identified, binding partners need to survive cell lysis and washes. This has limited the applicability of this otherwise widely-utilized approach in the study of ePPIs. Cell lysis and target extractions typically require membrane solubilization, which can disrupt membrane protein-dependent interactions. In addition, affinity purification workflows often miss detection of low affinity interactions, which are typical of ePPIs.

2.2.4 Specific applications

The most widely used version of this concept is affinity purification followed by binding partner identification using mass spectrometry (AP/MS) (**Table 2**). MS allows for the unbiased identification of interactors in their endogenous form in virtually any cell type or tissue. Traditionally, AP/MS studies were mostly restricted to one or a few targets-of-interest. However, recent technological advances, dominated

Figure 3. Affinity purification isolates a protein and any stably interacting protein from a complex solution such as a cell or tissue lysate. Shown here is the most common implementation where an antibody is used to isolate the target protein, followed by unbiased identification of binding proteins using MS. the antibody is attached to a solid substrate to allow for physical manipulations of the target protein.

Approach	Readout
AP-MS	Mass spectrometry
LUMIER	Luciferase activity
XL-MS	Mass spectrometry

Table 2.
Affinity purification techniques covered in this section. These approaches differ primarily in their readout method.

by the BioPlex project, have driven the development of a systematic pipeline that enables high-throughput AP/MS. Such efforts have already resulted in an interaction network with nearly 120,000 interactions identified for over 14,000 proteins in HEK293 cells [18].

A second and more recently developed method is the luminescence-based mammalian interactome mapping (LUMIER). In this approach, a library of epitope-tagged (specifically FLAG-tagged) constructs are co-transfected with a target-of-interest fused to Renilla luciferase. An anti-FLAG antibody is then used to pull down the tagged protein and binding is assayed by reading out luciferase activity of the immunoprecipitate [19]. Though in principle this approach offers increased sensitivity, this technique requires that both the target-of-interest and the library are tagged and expressed using artificial constructs. Thus, the applicability of this method for unbiased screening greatly depends on the accessibility to large libraries of tagged constructs.

While these approaches excel at identifying soluble interactions, ePPIs struggle to survive the processing steps and are noticeably underrepresented in both the LUMIER and even the much more comprehensive BioPlex dataset. One way to address some of the challenges associated to ePPIs is to combine affinity purification with cross-linking, turning transient ePPIs into permanent covalent linkages. By using cross-linking in combination with mass spectrometry (XL-MS), ePPIs can be identified in an unbiased manner. While cross-linking stabilizes weak interactions, XL-MS still has associated challenges like the presence of unproductive cross-links or combinatorial database search space. To overcome these, newly developed cross-linkers used for XL-MS can include affinity tags or MS cleavable moieties [20]. However, these increase cross-linker size and the chances of cross-linking nearby, non-interacting proteins. Existing cross-linking reagents also primarily target reactive amines, limiting the number of protein interactions that can be captured. Cross-linking protein complexes also tends make them less soluble [21]. This is worsened by the fact that ePPIs often involve membrane proteins which already present solubility challenges that complicates down-stream processing. Since extracellular proteins are often heterogeneously post-translationally modified [22], they can be challenging to identify in mass databases for MS experiment. Overall, XL-MS represents one of the few techniques that does not have a bias against ePPIs over soluble PPIs.

While most cross-linking approaches select for general features of proteins like reactive amine groups, alternative strategies have been developed that specifically target cell-surface receptors using trifunctional cross-linkers. These methods typically have one moiety that covalently attaches the cross-linker to the target protein-of-interest, a soluble protein that ranges from peptides, to antibodies or even complex entities such as viral particles. A second moiety links to glycosylated receptor proteins bound to or near the target, and a third moiety enables purification. Three molecules and associated workflows have been described: TRICEPS [23], followed by ASB [24] and HATRIC [25]. These approaches have the potential

to enable unbiased study of targets of diverse nature in physiologically relevant settings such as cells expressing endogenous receptors, thus offering an attractive option for ePPI discovery.

2.3 Protein-fragment complementation

2.3.1 Concept description

Protein-fragment complementation refers to methods where a protein with reporter activity is split in two and fused to two proteins being tested for binding. Since the two halves do not interact with each other on their own, reporter activity is only recovered when the halves are fused to interacting proteins. The archetypal example of this approach is the yeast two-hybrid system (Y2H) which uses two halves of a transcriptional factor that can drive the expression of a report gene (**Figure 4**).

2.3.2 Concept pros

Protein-fragment complementation is usually performed using living cells allowing proteins to be maintained in relatively native conditions. Reporter activity often have an amplification step that allows for the sensitive detection of even weak interactions [26]. While protein-fragment complementation technically reads out proximity, reasonable linker lengths can select for small distances. The interaction has to persist long enough for the activity to be reconstituted, reducing false positives rates when compared to some other proximity-based techniques.

2.3.3 Concept cons

Protein-fragment complementation mandates the tagging of proteins with non-native sequences for the reporter readout. These tags can be substantial in size and affect the behavior of the proteins being tagged. Since the reporter activity depends only on the reporter portion being in close proximity, this approach does not guarantee a direct interaction. Most systems only test binary interactions by design since only the tagged proteins are being assayed.

2.3.4 Specific applications

Y2H has been used extensively to detect PPIs since its conception (**Table 3**). The use of yeast allows for low-cost high-throughput testing of interactions. This technique has now been used to detect interactions between 90% of human proteins [27]. However, Y2H is not suited for ePPI discovery. The expression of human proteins in yeast may result in non-native post-translational

Figure 4. General schematic for the protein-fragment complementation approach. Two proteins suspected of interaction are each tagged with half of a split reporter protein. The reporter activity is detected only if the split reporter is brought together by an interaction between the two tagged proteins.

Approach	Split reporter system
Yeast two-hybrid	Split Gal4 transcription factor driving reporter expression
MYTH/MaMTH	Split ubiquitin restricting localization of a transcription factor

Table 3.
Specific protein-fragment complementation approaches covered in this section. These differ primarily in the specifics of the split reporter activity.

modifications relevant for function, but more importantly, Y2H actively selects against ePPIs because the interaction must occur in the nucleus to drive transcriptional readout.

To complement the classic Y2H approaches and overcome the pitfalls related to ePPIs, several systems specifically targeting membrane proteins have been developed. The membrane yeast two-hybrid (MYTH) [28] and its mammalian counterpart, mammalian membrane two-hybrid (MaMTH) [29] require that at least one protein being tested is anchored to the plasma membrane. Both of these approaches use a split ubiquitin system where one of the two halves is fused to a membrane protein and a transcription factor. Tethering the transcription factor to the membrane protein keeps it out of the nucleus, preventing reporter expression. When the membrane protein interacts with a protein containing the second half of ubiquitin, a cleavage event occurs, releasing the transcription factor to translocate to the nucleus and initiate reporter expression. In combination with targeted libraries, this approach has been used for the high-throughput detection of interactions between receptor tyrosine kinases and phosphatases [30]. However, since these techniques rely on the endogenous ubiquitin machinery for cleavage, they mandate that both binding partners be expressed in the same cells, limiting applicability of these techniques for detection of in-trans interactors.

2.4 Proximity labeling

2.4.1 Concept description

Proximity labeling techniques identify possible PPIs by covalently modifying proteins that are in close proximity, typically within a few nanometers. In most cases, the label includes an affinity tag like biotin which allows the labeled proteins to be purified and identified using MS (**Figure 5**).

Figure 5.
General schematic of a proximity labeling experiment. A target-of-interest is fused to an enzyme that labels nearby proteins with an affinity tag. That tag can then be used to isolate the proteins for identification.

2.4.2 Concept pros

The different proximity proteomics methods have represented some of the most significant advances in the field of PPI detection, and in particular membrane protein interaction discovery. From the initial development of BioID and its further iterations in BioID2 and TurboID, as well as the more recently developed MicroMap, these techniques have substantially increased the sensitivity for detection of a range of interactions, including weak, transient interactions by translating them into permanent covalent linkages. In addition, proximity proteomics approaches are generally applicable to complex physiological systems and cellular models of interest, and can bypass over-expression of proteins-of-interest and laborious libraries. Furthermore, these approaches also offer the advantage of temporal control, though currently this is typically on the tens of minutes time-scale.

2.4.3 Concept cons

While proximity labeling typically does not require any special equipment, the unbiased identification of proximal proteins requires access to MS. Since these approaches fundamentally readout proximity, the PPI is inferred. Especially if used in complex physiological contexts, the possibility of identifying neighboring but not directly interacting proteins means that this approach has the greatest challenge when it comes to data interpretability. Additionally, the most popular proximity proteomics methods, BioID and APEX, require expression of the protein-of-interest fused to a bulky tag, followed by over-expression of the fusion protein. Although generally applicable to physiologically relevant systems, experimental conditions may require optimization to ensure that the overall behavior of the target-of-interest is not altered by tagging or over-expression.

Currently, these techniques have been applied primarily to detect interactions between proteins on the same cells. In many cases, the utility for in-trans interactions remains to be demonstrated.

2.4.4 Specific applications

The different proximity labeling techniques vary based on what enzymes or chemistries are used to accomplish the labeling. The field of proximity proteomics has been predominantly driven by the development of enzyme-catalyzed proximity labeling. These typically used a promiscuous biotin ligase (from BioID to the much faster TurboID) or a peroxidase (usually APEX or horse radish peroxidase (HRP)) to create a highly reactive biotin species that can only diffuse a short distance before reacting with nearby proteins or water [31]. When a target protein is tagged with one of these enzymes and substrate added, proteins in its vicinity are biotinylated, allowing them to be isolated and identified using MS. The particular techniques differ slightly in their tradeoffs. The peroxidases tend to be more broadly reactive and requiring less labeling time. However, they require the addition of a biotin conjugate and hydrogen peroxide, both of which could be toxic to cells. While the biotin ligases do not have this problem, they are on average slower, though the more recently engineered TurboID can achieve efficient labeling within minutes (**Table 4**) [31].

This concept has been incorporated into specific techniques for identifying ePPIs like selective proteomic proximity labeling assay using tyramine (SPPLAT). SPPLAT uses an HRP-conjugated antibody recognizing a cell surface protein.

Approach	Substrate/Cross-linker Moieties
Biotin ligases (BioID, TurboID)	Biotin
Peroxidases (APEX, HRP)	Different biotin conjugates
SPPLAT	Tyramide biotin conjugates
EMARS	Aryl azide biotin conjugates
MicroMap	Diazirine biotin conjugates

Table 4. *Specific approaches mentioned in this section for using proximity labeling and cross-linking. These techniques differ in the labeling enzymes that they use (for labeling) or the chemistries of the substrate.*

Since the antibody cannot diffuse across the plasma membrane, it specifically targets ePPIs without any tagging of proteins, and thus enabling studies in unmodified cellular settings [32]. Another approach is enzyme-mediated activation of radical sources (EMARS) which also uses an HRP-conjugated antibody. However, EMARS uses a biotin fused to an aryl azide group giving it a large labeling radius of 200–300 nm, making it more suitable for characterizing entire microdomains rather than ePPIs [33]. Though the use of antibodies has advantages, genetically tagging a protein with HRP can allow these types of techniques to be performed in the physiological context in an organism. For example, the use of a CD2-HRP fusion protein with a membrane-impermeable biotin-phenol allowed the identification of cell-type specific neural protein cross talk in the fly brain [34].

The newest addition to the proximity labeling family is MicroMap, which uses entirely orthogonal chemistry to the existing techniques. MicroMap uses an antibody to detect the target-of-interest, which is then recognized by a secondary antibody conjugated to a photocatalyst. The photocatalyst absorbs blue light to catalyze the activation of a biotin conjugate molecule in its vicinity. This approach uses a more reactive chemical moiety than the biotin ligase or peroxide approaches, which allow for an even smaller radius of labeling and thus, is more likely to detect direct PPIs. Using MicroMap, the authors proposed a new set of putative binders partners for key immune receptors such as PD-L1 [35].

2.5 Direct protein–protein interaction screens

2.5.1 Concept description

Direct interaction screens encompass a wide variety of techniques that have several features in common. First, there is a query protein that is the target-of-interest. Second, the query protein is tested for binding to a library containing possible binding partners presented as recombinant proteins or receptors expressed on cells. Third, a positive signal in the screen directly reads out an interaction between the query and a given binding partner in the library, using detection methods that vary depending on the approach. Major distinguishing factors between the various direct PPI-screening techniques include: the level of multimerization of the target protein (from monomers to oligomeric proteins), the form of the library of binding partners being screened (protein-based vs. cell-based formats), as well as the degree of purification required (purified protein vs. conditioned media) (Figure 6).

2.5.2 Concept pros

This approach typically allows for the opportunity to control most aspects of the screen such as protein concentration and buffer conditions. This approach is also

Figure 6.
General schematic for direction protein interactions screens. The library of proteins can be directly immobilized on a solid substrate or be expressed on cells. Alternative multimerization methods have been developed to present the query protein of interest, reviewed in the text.

generally amenable to scale-up. Therefore many modern libraries have high coverage of at least specific protein families. The simplicity and the fact that the readout reflects direct PPIs also generally leads to straight-forward data analysis.

2.5.3 Concept cons

Many of these approaches use purified proteins and may require the use of ectodomains rather than the full-length protein. While the ectodomain is sufficient for binding in many instances, this requirement makes it difficult to identify ePPIs that use multiple ectodomains or transmembrane domains for binding; a behavior documented for the family of seven-transmembrane-domain-containing G protein-coupled receptors (GPCRs). These approaches also typically require that the target-of-interest be screened against a library, which needs to be comprehensive for truly unbiased identification if ePPIs. The generation and maintenance of a large library, either as recombinant proteins or plasmids for expression on cells, can be costly and may require access to automation.

2.5.4 Specific applications

While published work tends to take advantage of specific combinations of the type of target presentation and the type of library, significant mixing and matching is possible due to the similarity in the overall conceptual framework. Therefore, we will talk about the major types of target presentation and library separately and mention any incompatibilities. Also, since there are a large diversity of library formats, we divided the formats into protein-based libraries and cell-based libraries, though the same target presentation strategies can be used for both (Table 5).

Target Presentation	Protein-based Library	Cell-based Library
Dimer (Fc tag)	Protein microarray (on slide)	cDNA libraries (on slide)
Pentamer (COMP tag)	Purified protein (in plate)	cDNA libraries (on plate)
Beads-based Multimer (variable)	Purified protein (SPR chip)	CRISPRa gRNA libraries
	Purified protein (Magneto-sensor chip)	Knock-down or Knock-out libraries
	Conditioned media protein (in plate)	

Table 5.
Lists of target presentations and protein and cell-based library formats covered in this section.

2.5.4.1 Target presentation methods

To be able to directly assay interactions, targets-of-interests are typically presented as recombinant protein for this type of approach. While secreted proteins are soluble and can be directly screened, membrane proteins tend to misfold and aggregate if they are extracted from membranes because of their hydrophobic transmembrane domains [12, 13]. Since ectodomains are usually the portion of transmembrane proteins available for direct ePPIs, typically only ectodomains are used for direct ePPI screening.

A number of multimerization approaches that increase query protein avidity and therefore facilitate detection of transient interactions have been developed. In particular, there are three dominant strategies: dimerization induced by fusing ectodomains to the constant Fc region of antibodies [36], pentamerization induced by fusing ectodomains to the rat cartilage oligomeric matrix protein (COMP) [37] and higher order multimerization using small beads with high protein-binding capacity, usually in the form of protein A-coated or streptavidin-coated beads.

While increased multimerization is a major factor for increasing sensitivity with target presentation, the readout method to measure target binding also varies. Using Fc-tagged dimers allows the detection of the target using a variety of secondary antibodies or protein A/G that bind to Fc regions with high affinity [36]. However, using enzymatic readouts can add a high degree of signal amplification that allows for increased sensitivity. Therefore, the approach used to generate the largest ePPI networks to date uses pentamerization combined with an enzymatic β -lactamase colorimetric assay [37–40]. As for bead-based approaches, the specific readout can be magnetic, fluorescent or chemiluminescent depending on the specific screening method used [41, 42].

Lastly, some of the recent high-throughput technologies use conditioned media enriched for the target-of-interest rather than purified proteins. Using conditioned media involves direct capture of secreted protein or protein ectodomains in the absence of protein purification, thus minimizing potential inactivation of the proteins due to purification steps. The use of conditioned media can also save time and resources, helping to make the approaches more accessible to different laboratory and more amenable to scaling up [37–40].

2.5.4.2 Protein-based library formats

Different protein-based library formats can allow for different levels of throughput and information collected about the binding interactions. The most common and high-throughput approaches are generally qualitative, detecting whether the interaction is present, but not providing quantitative information such as kinetic parameters. One example of this type of library is the protein microarray, which for ePPIs, contains different purified secreted proteins or ectodomains directly spotted on slides. Only small amounts of each protein are used, allowing for the dense tiling of thousands of proteins per slide. The compact format allows slides to be covered with a small volume of fluorescently-labeled target protein, rinsed and imaged using microscopy [43]. While this is a convenient format, the construction of the protein microarrays is often costly because it requires all of the proteins to be purified.

Another type of library for qualitative ePPI identification uses plate-based screening formats. The use of plates allows for the easy addition of proteins and controlled washes without the need for specialized microfluidics. While purified proteins can be used, plate-based formats allow for the direct capture of secreted tagged proteins from conditioned media. Capture of biotinylated proteins using streptavidin-coated plates [37] or Fc-tagged proteins using Protein A-coated plates [38–40] followed by washing allows for the effective purification of library proteins

in wells while adding sensitivity by multimerizing (in the case of multivalent binding of streptavidin to biotin) or capturing already multimerized proteins. This approach also allows the use of enzymatic liquid phase readouts: β -lactamase-based colorimetric assays or luciferase-based luminescence assays which provides an additional degree of signal amplification. The plate-based approach also gives one value per well, allowing for simple data analysis and the greatest interpretability.

While the plate-based approach is generally the most scalable options, other techniques trade some scale for quantitative information on ePPIs. In particular, microfluidics, automation and miniaturization has pushed label-free biophysical techniques to be more high-throughput. For example, the combination of microfluidics and either surface plasmon resonance (SPR) or magneto-nanosensors has increase the scale enough to study all combinatorial interactions between a small number of proteins, making it especially adept at addressing complex cross-talk between small interaction networks [44, 45]. While SPR is the gold standard technique for biophysical characterization of protein interactions and calculation of kinetic parameters, the magneto-nanosensor platform provides higher degrees of sensitivity, and therefore, requiring less material to detect weak ePPIs. However, it requires the use of magnetic nanoparticles conjugated to the target-of-interest. The nanoparticles are are flowed over patches of library proteins printed on magneto-nanosensors that detect a change in electrical resistance if a nanoparticle is nearby [45]. Another technique that also provides similar information is bilayer interferometry (BLI) which translates protein binding into a light interference signal. While typically less sensitive than SPR and the magneto-nanosensor platform, BLI excels in its ease-of-use. BLI uses small, disposable sensors that can be coupled to targets-of-interest, typically through the capture of tag like Fc-tags or biotin. The sensors are then simply dipped into wells containing the potential binding partners in solution. With advances in automated and miniaturized BLI setups, it can be used to screen for interactions in high-throughput, provided that libraries of recombinant proteins are available. This technology helped identify the PVR-TIGIT interaction [46] which is mechanistic foundations of the anti-TIGIT immunotherapy [47].

2.5.4.3 Cell-based library formats

Even though protein-based libraries have many advantages such as storability and easy data interpretation, they can often fail to detect ePPIs because of biochemical challenges associated with membrane proteins. Many membrane proteins lose activity when truncated into soluble ectodomains or extracted from membranes. In addition, the complex cellular membrane environment can provide important protein and non-protein co-factors, orient and cluster membrane proteins and assist in high-order complex formation. Therefore, hard to purify receptors are often screened against cDNA libraries expressing membrane proteins directly on cells. This is especially true for important drug targets like GPCRs and ion channels [48] which have multiple transmembrane domains and typically small extracellular regions.

To screen for interactions using cell-based formats, libraries are used to either induce loss-of-function (lack of binding) or gain-of-function (increased binding). In the loss-of-function approach, possible binders of a target-of-interest are knocked down or knocked out either randomly using chemical mutagens or transposons like gene trap [49] or in a targeted manner with siRNA, zinc-finger nuclease, transcription activator-like effector nuclease (TALEN) or clustered regularly interspaced short palindromic repeat (CRISPR) libraries [50]. When the target-of-interest is incubated with the cells, the target should not interact if the interaction partner has been depleted. However, it may be difficult to identify cells that bind

the target-of-interest. In addition, the interaction must be simple enough that knocking down one interaction partner causes a detectable decrease in binding.

To avoid these limitations, an alternative approach is to overexpress receptors that may participate in ePPIs. This is most commonly done using cDNA libraries. The DNA libraries are spotted on slides [51] or plated into wells [52], with each spot or well containing a vector encoding for a different protein. Cells are then added and transfected to induce them to overexpress the protein and display them on the plasma membrane. If the cells are expressing the receptor for the target-of-interest, this can be detected by increased target binding to the surface of the cell. This approach has been successfully utilized to deorphanize secreted factors [53], interactions between immune receptors [54], or identify glycan-dependent recognition of specific ligands [55]. However, the generation and management of cDNA libraries that have significant coverage of membrane proteins can be expensive and not accessible to many investigators. In addition, selective expression of the myriad of possible receptors isoforms that may participate in ligand binding makes truly comprehensive cDNA libraries infeasible. One way to address isoform-specific expression while facilitating library management is to use CRISPR activation (CRISPRa). In this implementation, CRISPR-Cas9 fused to transcriptional activation domains is coupled to guide RNAs selectively targeting cell-surface genes to overexpress receptor proteins. A high coverage CRISPRa guide library targeting most cell surface proteins has been recently utilized to identify novel receptor-ligand interactions [56].

2.6 Computational models

2.6.1 Concept description

Computational models cover a large range of concepts that attempt to predict PPIs based on existing knowledge of the biochemistry of protein binding and features of proteins, such as the sequence, conserved residues or structural features.

2.6.2 Concept pros

Computational models can offer relatively less resource-consuming and faster alternatives to experimental research. They allow for the theoretical exploration of PPIs without regard for experimental challenges related to expression of proteins or development of workflows or platforms. Modern machine learning approaches may also identify unintuitive features that are the most predictive for interactions such as unappreciated modifications. They can also draw from larger pools of information, taking into account protein expression patterns, genetic variations and dysregulation in disease.

2.6.3 Concept cons

Computational modeling approaches to identify PPIs, not to mention ePPIs, are still in their infancy, with overall low rates of accuracy. Many are based on our existing knowledge of experimentally determined interactions, which may have biases and is incomplete. Approaches that attempt to model binding interfaces are too computationally expensive to be high-throughput even when experimentally determined protein structures exist [57].

2.6.4 Specific applications

Since computational approaches remain immature for human ePPIs, we will mostly highlight the different computational resources and a few different

Approach	Details
STRING	Database of PPIs predictions based on public data
BioGRID	Database of curated public PPI data
BioPlex	Database of human PPI identified by AP/MS
Coevolution	PPI prediction based on coevolution of residues
PICTree	Functional clustering using hidden Markov models
FpClass	PPI prediction using machine learning

Table 6.
List of computational resources and approaches covered in this section.

approaches rather than try to describe a list of the major algorithms. However, this is a rapidly developing area mirroring the explosion in available experimental datasets, including data from all of the approaches mentioned so far as well as expression data for cell types, tissue and now single cells identifying which proteins are at the same places at the same times (**Table 6**) [58, 59].

The increased availability of comprehensive databases for PPIs, and more recently ePPIs, have fueled diverse computational approaches. Efforts like STRING [60] and BioGRID [61] which collects and curates public data on PPIs, are often drawn on for model development and are also important resources for individual researchers looking for the next interaction to drug. There are also many databases that document the progress of specific approaches like BioPlex which contains thousands of human interactions identified by AP/MS [18] as well as the Research Collaboratory for Structural Bioinformatics Protein Data Bank which captures many structures showing the molecular details of PPIs [62]. However, even here ePPIs have posed a challenge because we do not have a definitive list of all proteins that reach the cell surface in various tissues and cell types, though ongoing efforts are trying to experimentally answer that question [59, 63–65]. Recently, the human surfaceome was estimated using a machine learning model to predict the cell surface localization of almost 3000 proteins [66].

Modeling approaches that actually attempt to predict ePPIs range in terms of the types of information they try to account for. While not yet applied to human ePPIs, the use of residue-residue coevolution in combination with structure modeling successfully predicted many ePPIs in bacteria [67]. Another approach, PICTree, focused on the structurally related immunoglobulin superfamily (IgSF) of proteins, using knowledge of family members with known binding partners and sequence conservation to predict new interactions using [68]. Lastly, some approaches use broad information sets about a gold standard set of interactions. For example, FpClass trains their model on everything amino acid makeup to post-translational modifications to expression patterns [69]. However, this still resulted in an estimated false discovery rate of 60%, which shows that while modeling can assist in hypothesis generation, there is more work to be done before modeling would take the place of experimental approaches.

3. Summary

Extracellular protein–protein interactions are an important set of possible drug targets. They are commonly dysregulated in disease and can be targeted to alter disease phenotypes. Practically, ePPIs are exposed on the cell surface, making them easier to access using therapeutic approaches. However, because of challenges associated with ePPI biochemistry, most membrane proteins and secreted factors do

not have identified interactions. Elucidating the extracellular interaction networks in humans as well as their dysregulation during disease will be key to understand basic biology and fuel new or improved drug development efforts. To tackle this daunting challenge, researchers have applied genetic, chemical, biochemical and computational approaches to come up with an ever-growing list of ePPIs. Here we have reviewed the progress made in the last decade in technologies suitable for the study of ePPIs. In particular, we discuss those approaches that can be applied to the high throughput screening of ePPIs in an unbiased fashion.

4. Future perspectives

As costs are continually falling on readouts like sequencing and mass spectrometry, and as throughput increases with better automation and computational analysis, the future looks bright in the field of ePPI identification. More and more techniques will cross over the categories that we have laid out, finding middle points that balance the various tradeoffs of ease, interpretability, and physiological relevance.

One exciting development that 2020 brought was the release of two large-scale efforts using ePPI-optimized pentamer-based direct interaction screening approaches. These efforts each systematically tested hundreds of thousands of pairwise interactions, focusing on the IgSF of single-pass transmembrane proteins, the largest family of secreted and membrane-expressed proteins in the human genome [39, 40]. These large interaction networks identified hundreds of new interactions and present the most extensive ePPI network maps to date.

Once the interactions are found, we need to be able to manipulate them in humans to cure diseases. While not the topic of this chapter, several exciting developments on the drug development front holds much promise for targeting ePPIs. New highly selective inhibitors that recognize the transmembrane domains of protein, such as the isoform-selective inhibitor of the Nav1.7 channel, can provide novel classes of chemical inhibitors of transmembrane proteins to disrupt ePPIs [70]. While cytokines often offer desirable ways to manipulate many immune functions, they are often like playing with fire because of their many disparate effects. However, with improvements in protein design, completely artificial cytokine mimics can now be made which can be highly selective for activities that are desired and counter selected for activities that are not [71].

One major challenge that lies ahead is to not to just identify ePPIs but to identify disease relevant human ePPIs. Along these lines, a recently published map of the IgSF highlighted the power of big data integration, showing that the combination of clinical data with a focus on the protein pair participating in ePPIs gave greater predictive value than each of the proteins alone [39], suggesting that targeting specific ePPIs may be more beneficial than targeting an individual protein. Another challenge is the reliance on animal models. Plasma membrane and secreted factors are some of the least conserved of all proteins [72], having to evolve to adapt to our unique physiology. As more complex human ePPI networks are discovered, it will be a challenge to understand their impacts at the organismal level. Whether it be organoid systems or better functional assays, with the rapid growth in ePPI identification technologies, soon we'll have to find high-throughput ways to ask, what do they do?

5. Executive summary

- Extracellular protein–protein interactions (ePPIs) make for good drug targets because they control many biological processes and are accessible to therapeutic agents.

- An estimated third of all human genes encode for proteins that may be involved in ePPIs, necessitating high-throughput approaches for unbiased discovery.
- General PPI detection techniques often fail to overcome challenges posed by extracellular proteins, leading to the development of ePPI-specific approaches.
- EPPI-specific technologies address some of these challenges directly, such as using multimerization to strengthen characteristically weak interactions or assaying interactions on cells to avoid difficult membrane protein purifications.
- Techniques typically balance several tradeoffs, mainly: control and interpretability versus physiological relevance, rates of false positive versus false negative results, and scale and coverage versus time and expense.
- Specific techniques fall into broad categories: biochemical fractionation, affinity purification, protein–fragment complementation, proximity labeling, direct interaction and computational modeling that can be synergistic for ePPI discovery.
- New ePPIs are still being discovered with the aid of new techniques, suggesting that many remain to be found. The methodologies discussed in this chapter should set the bases for identification and characterization of novel ePPIs in humans and other model organisms.

Acknowledgements

We thank Genentech Reviewers for critically reading the manuscript. Figures were created with BioRender.com. **Figure 1A** was adapted from “Tumor Microenvironment” and **Figure 1B** was adapted from “Proposed Therapeutic Treatments for COVID-19 Targeting Viral Entry Mechanism”, by BioRender.com (2020). Retrieved from <https://app.biorender.com/biorender-templates>

Conflict of interest

Both authors are Genentech employees and own shares in the Roche/Genentech group.

IntechOpen

IntechOpen

Author details

Shengya Cao and Nadia Martinez-Martin*
Genentech, South San Francisco, CA, USA

*Address all correspondence to: nadia.martinez@almirall.com

IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

References

- [1] WHO|Projections of mortality and causes of death, 2016 to 2060. [cited 2020 Nov 18]. Available from: https://www.who.int/healthinfo/global_burden_disease/projections/en/
- [2] Libby P, Ridker PM, Maseri A. Inflammation and atherosclerosis. *Circulation*. 2002;105(9):1135-1143.
- [3] Vinay DS, Ryan EP, Pawelec G, Talib WH, Stagg J, Elkord E, et al. Immune evasion in cancer: Mechanistic basis and therapeutic strategies. *Semin Cancer Biol*. 2015;35:S185–S198.
- [4] Mayor R, Carmona-Fontaine C. Keeping in touch with contact inhibition of locomotion. Vol. 20, *Trends in Cell Biology*. 2010. p. 319-28.
- [5] Carmeliet P, Jain RK. Angiogenesis in cancer and other diseases. *Nature*. 2000 Sep;407(6801):249-257.
- [6] Abercrombie M. Contact inhibition and malignancy. *Nature*. 1979 Sep;281(5729):259-262.
- [7] Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*. 2020;181(2):271-280.e8.
- [8] Shang J, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, et al. Cell entry mechanisms of SARS-CoV-2. *Proc Natl Acad Sci U S A*. 2020;117(21).
- [9] Marchio P, Guerra-Ojeda S, Vila JM, Aldasoro M, Victor VM, Mauricio MD. Targeting early atherosclerosis: A focus on oxidative stress and inflammation. *Oxid Med Cell Longev*. 2019;2019(Ldl).
- [10] Hegde PS, Chen DS. Top 10 Challenges in Cancer Immunotherapy. *Immunity*. 2020;52(1):17-35.
- [11] Hansen J, Baum A, Pascal KE, Russo V, Giordano S, Wloga E, et al. Studies in humanized mice and convalescent humans yield a SARS-CoV-2 antibody cocktail. *Science* (80-). 2020;369(6506):1010-4.
- [12] Wright GJ, Martin S, Bushell KM, Söllner C. High-throughput identification of transient extracellular protein interactions. *Biochem Soc Trans*. 2010;38(4):919-922.
- [13] Martinez-Martin N. Technologies for proteome-wide discovery of extracellular host-pathogen interactions. *J Immunol Res*. 2017;2017.
- [14] Oppenheim JJ. Cytokines: Past, present, and future. *Int J Hematol*. 2001;74(1):3-8.
- [15] Havugimana PC, Hart GT, Nepusz T, Yang H, Turinsky AL, Li Z, et al. A census of human soluble protein complexes. *Cell*. 2012;150(5):1068-1081.
- [16] Depner H, Lützkendorf J, Babkir HA, Sigrist SJ, Holt MG. Differential centrifugation-based biochemical fractionation of the *Drosophila* adult CNS. *Nat Protoc*. 2014;9(12):2796-2808.
- [17] Oswald D, Khorramshahi O, Gupta VK, Banovic D, Depner H, Fouquet W, et al. Cooperation of Syd-1 with Neurexin synchronizes pre- with postsynaptic assembly. *Nat Neurosci*. 2012;15(9):1219-1226.
- [18] Huttlin EL, Bruckner RJ, Navarrete-Perea J, Cannon JR, Baltier K, Gebreab F, et al. Dual Proteome-scale Networks Reveal Cell-specific Remodeling of the Human Interactome. *bioRxiv*. 2020;2020.01.19.905109. Available from: <https://www.biorxiv.org/content/biorxiv/early/2020/01/19/2020.01.19.905109.full.pdf>

- [19] Barrios-Rodiles M, Brown KR, Ozdamar B, Bose R, Liu Z, Donovan RS, et al. High-throughput mapping of a dynamic signaling network in mammalian cells. *Science* (80-). 2005;307(5715):1621-5.
- [20] Chavez JD, Mohr JP, Mathay M, Zhong X, Keller A, Bruce JE. Systems structural biology measurements by in vivo cross-linking with mass spectrometry. *Nat Protoc*. 2019;14(8): 2318-2343.
- [21] Iacobucci C, Götze M, Sinz A. Cross-linking/mass spectrometry to get a closer view on protein interaction networks. *Curr Opin Biotechnol*. 2020;63:48-53.
- [22] Kolarich D, Jensen PH, Altmann F, Packer NH. Determination of site-specific glycan heterogeneity on glycoproteins. *Nat Protoc*. 2012;7(7): 1285-1298.
- [23] Frei AP, Jeon OY, Kilcher S, Moest H, Henning LM, Jost C, et al. Direct identification of ligand-receptor interactions on living cells and tissues. *Nat Biotechnol*. 2012;30(10):997-1001.
- [24] Tremblay TL, Hill JJ. Biotin-Transfer from a trifunctional crosslinker for identification of cell surface receptors of soluble protein ligands. *Sci Rep*. 2017;7(December 2016):1-9.
- [25] Sobotzki N, Schaefroth MA, Rudnicka A, Koetemann A, Marty F, Goetze S, et al. HATRIC-based identification of receptors for orphan ligands. *Nat Commun*. 2018;9(1):1-8.
- [26] Estojak J, Brent R, Golemis EA. Correlation of two-hybrid affinity data with in vitro measurements. *Mol Cell Biol*. 1995;15(10):5820-5829.
- [27] Luck K, Kim DK, Lambourne L, Spirohn K, Begg BE, Bian W, et al. A reference map of the human binary protein interactome. *Nature*. 2020;580(7803):402-408.
- [28] Stagljar I, Korostensky C, Johnsson N, Te Heesen S. A genetic system based on split-ubiquitin for the analysis of interactions between membrane proteins in vivo. *Proc Natl Acad Sci U S A*. 1998;95(9):5187-5192.
- [29] Petschnigg J, Groisman B, Kotlyar M, Taipale M, Zheng Y, Kurat CF, et al. The mammalian-membrane two-hybrid assay (MaMTH) for probing membrane-protein interactions in human cells. *Nat Methods*. 2014;11(5):585-592.
- [30] Yao Z, Darowski K, St-Denis N, Wong V, Offensperger F, Villedieu A, et al. A Global Analysis of the Receptor Tyrosine Kinase-Protein Phosphatase Interactome. *Mol Cell*. 2017;65(2): 347-360.
- [31] Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, Carr SA, et al. Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol*. 2018;36(9):880-898.
- [32] Rees JS, Li XW, Perrett S, Lilley KS, Jackson AP. Selective proteomic proximity labeling assay using tyramide (SPPLAT): A quantitative method for the proteomic analysis of localized membrane-bound protein clusters. *Curr Protoc Protein Sci*. 2015;2015(April): 19.27.1-19.27.18.
- [33] Honke K, Kotani N. Identification of cell-surface molecular interactions under living conditions by using the enzyme-mediated activation of radical sources (EMARS) method. *Sensors (Switzerland)*. 2012;12(12): 16037-16045.
- [34] Li J, Han S, Li H, Udeshi ND, Svinkina T, Mani DR, et al. Cell-Surface Proteomic Profiling in the Fly Brain Uncovers Wiring Regulators. *Cell*. 2020;180(2):373-386.e15.

- [35] Geri JB, Oakley J V, Reyes-Robles T, Wang T, McCarver SJ, White CH, et al. Microenvironment mapping via Dexter energy transfer on immune cells. *Science*. 2020;367(6482):1091-1097.
- [36] Czajkowsky DM, Hu J, Shao Z, Pleass RJ. Fc-fusion proteins: New developments and future perspectives. *EMBO Mol Med*. 2012;4(10):1015-1028.
- [37] Bushell KM, Söllner C, Schuster-Boeckler B, Bateman A, Wright GJ. Large-scale screening for novel low-affinity extracellular protein interactions. *Genome Res*. 2008;18(4):622-630.
- [38] Martinez-martin N, Marcandalli J, Huang CS, Lanzavecchia A, Perez L, Martinez-martin N, et al. An Unbiased Screen for Human Cytomegalovirus Identifies Neuropilin-2 as a Central Viral Receptor Article An Unbiased Screen for Human Cytomegalovirus Identifies Neuropilin-2 as a Central Viral Receptor. *Cell*. 2018;174(5):1158-1171.e19.
- [39] Verschueren E, Husain B, Yuen K, Sun Y, Paduchuri S, Senbabaoglu Y, et al. The Immunoglobulin Superfamily Receptome Defines Cancer-Relevant Networks Associated with Clinical Outcome. *Cell*. 2020;182(2):329-344.e19.
- [40] Wojtowicz WM, Vielmetter J, Fernandes RA, Siepe DH, Eastman CL, Chisholm GB, et al. A Human IgSF Cell-Surface Interactome Reveals a Complex Network of Protein-Protein Interactions. *Cell*. 2020;182(4):1027-1043.e17.
- [41] Bosse R, Illy C, Chelsky D. Principles of AlphaScreen, Amplified Luminescent Proximity Homogenous Assay. *Appl Note*. :1-8.
- [42] Husain B, Ramani SR, Chiang E, Lehoux I, Paduchuri S, Arena TA, et al. A platform for extracellular interactome discovery identifies novel functional binding partners for the immune receptors B7-H3/CD276 and PVR/CD155. *Mol Cell Proteomics*. 2019;18(11):2310-2323.
- [43] Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, Bertone P, et al. Global analysis of protein activities using proteome chips. *Science* (80-). 2001;293(5537):2101-5.
- [44] Piliarik M, Vaisocherová H, Homola J. A new surface plasmon resonance sensor for high-throughput screening applications. *Biosens Bioelectron*. 2005;20(10 SPEC. ISS.):2104-10.
- [45] Lee JR, Bechstein DJB, Ooi CC, Patel A, Gaster RS, Ng E, et al. Magneto-nanosensor platform for probing low-affinity protein-protein interactions and identification of a low-affinity PD-L1/PD-L2 interaction. *Nat Commun*. 2016;7:1-9.
- [46] Yu X, Harden K, Gonzalez LC, Francesco M, Chiang E, Irving B, et al. The surface protein TIGIT suppresses T cell activation by promoting the generation of mature immunoregulatory dendritic cells. *Nat Immunol*. 2009;10(1):48-57.
- [47] Johnston RJ, Comps-Agrar L, Hackney J, Yu X, Huseni M, Yang Y, et al. The Immunoreceptor TIGIT Regulates Antitumor and Antiviral CD8+ T Cell Effector Function. *Cancer Cell*. 2014;26(6):923-937.
- [48] Santos R, Ursu O, Gaulton A, Bento AP, Donadi RS, Bologa CG, et al. A comprehensive map of molecular drug targets. *Nat Rev Drug Discov*. 2016;16(1):19-34.
- [49] Carette JE, Guimaraes CP, Varadarajan M, Park AS, Wuethrich I, Godarova A, et al. Haploid genetic screens in human cells identify host factors used by pathogens. *Science* (80-). 2009;326(5957):1231-5.

- [50] Fernández A, Josa S, Montoliu L. A history of genome editing in mammals. *Mamm Genome*. 2017;28(7-8):237-246.
- [51] Freeth J, Soden J. New Advances in Cell Microarray Technology to Expand Applications in Target Deconvolution and Off-Target Screening. *SLAS Discov*. 2020;25(2):223-230.
- [52] Wood L, Wright GJ. High-Content Imaging for Large-Scale Detection of Low-Affinity Extracellular Protein Interactions. *SLAS Discov*. 2019;24(10):987-999.
- [53] Mullican SE, Lin-Schmidt X, Chin CN, Chavez JA, Furman JL, Armstrong AA, et al. GFRAL is the receptor for GDF15 and the ligand promotes weight loss in mice and nonhuman primates. *Nat Med*. 2017;23(10):1150-1157.
- [54] Bhatt RS, Berjis A, Konge JC, Mahoney KM, Klee AN, Freeman SS, et al. KIR3DL3 is an inhibitory receptor for HHLA2 that mediates an alternative immunoinhibitory pathway to PD1. *Cancer Immunol Res*. 2020;canimm.0315.2020.
- [55] Sun Y, Senger K, Baginski TK, Mazloom A, Chinn Y, Pantua H, et al. Evolutionarily conserved paired immunoglobulin-like receptor α (PILR α) domain mediates its interaction with diverse sialylated ligands. *J Biol Chem*. 2012;287(19):15837-15850.
- [56] Chong ZS, Ohnishi S, Yusa K, Wright GJ. Pooled extracellular receptor-ligand interaction screening using CRISPR activation. *Genome Biol*. 2018;19(1):1-16.
- [57] Macalino SJY, Basith S, Clavio NAB, Chang H, Kang S, Choi S. Evolution of in silico strategies for protein-protein interaction drug discovery. *Molecules*. 2018;23(8).
- [58] Petryszak R, Keays M, Tang YA, Fonseca NA, Barrera E, Burdett T, et al. Expression Atlas update - An integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Res*. 2016;44(D1):D746-D752.
- [59] Thul PJ, Lindskog C. The human protein atlas: A spatial map of the human proteome. *Protein Sci*. 2018;27(1):233-244.
- [60] Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47(D1):D607-D613.
- [61] Oughtred R, Stark C, Breitkreutz BJ, Rust J, Boucher L, Chang C, et al. The BioGRID interaction database: 2019 update. *Nucleic Acids Res*. 2019;47(D1):D529-D541.
- [62] Burley SK, Berman HM, Bhikadiya C, Bi C, Chen L, Di Costanzo L, et al. RCSB Protein Data Bank: Biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res*. 2019;47(D1):D464-D474.
- [63] Bausch-Fluck D, Hofmann A, Bock T, Frei AP, Cerciello F, Jacobs A, et al. A mass spectrometric-derived cell surface protein atlas. *PLoS One*. 2015;10(4):1-22.
- [64] Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, et al. The Human Cell Atlas. *Elife*. 2017 Dec 5;6:1-31.
- [65] Roberts B, Haupt A, Tucker A, Grancharova T, Arakaki J, Fuqua MA, et al. Systematic gene tagging using CRISPR/Cas9 in human stem cells to

illuminate cell organization. *Mol Biol Cell*. 2017;28(21):2854-2874.

[66] Bausch-Fluck D, Goldmann U, Müller S, van Oostrum M, Müller M, Schubert OT, et al. The in silico human surfaceome. *Proc Natl Acad Sci*. 2018;201808790.

[67] Cong Q, Anishchenko I, Ovchinnikov S, Baker D. Protein interaction networks revealed by proteome coevolution. *Science* (80-). 2019;365(6449):185-9.

[68] Yap EH, Rosche T, Almo S, Fiser A. Functional clustering of immunoglobulin superfamily proteins with protein-protein interaction information calibrated hidden markov model sequence profiles. *J Mol Biol*. 2014;426(4):945-961.

[69] Braun P, Tasan M, Dreze M, Barrios-Rodiles M, Lemmens I, Yu H, et al. An experimentally derived confidence score for binary protein-protein interactions. *Nat Methods*. 2009;6(1):91-97.

[70] Ahuja S, Mukund S, Deng L, Khakh K, Chang E, Ho H, et al. Structural basis of Nav1.7 inhibition by an isoform-selective small-molecule antagonist. *Science* (80-). 2015;350(6267).

[71] Silva DA, Yu S, Ulge UY, Spangler JB, Jude KM, Labão-Almeida C, et al. De novo design of potent and selective mimics of IL-2 and IL-15. *Nature*. 2019;565(7738):186-191.

[72] Ramilowski JA, Goldberg T, Harshbarger J, Kloppman E, Lizio M, Satagopam VP, et al. A draft network of ligand-receptor-mediated multicellular signalling in human. *Nat Commun*. 2015;6.