1

Exploring Protein–Protein Interactions: Concepts, Methods, and Implications

Mi Zhou and Renxiao Wang

Fudan University, School of Pharmacy, Department of Medicinal Chemistry, 826 Zhangheng Road, Shanghai 201203, People's Republic of China

1.1 General Concepts of Protein–Protein Interactions

Proteins are the fundamental machinery that drive the vast majority of cellular processes. These versatile biomolecules rarely perform in isolation; instead, up to 80% of proteins engage in intricate interactions with one another, forming dynamic networks that underpin the functional complexity of living organisms [1]. This chapter aims to provide a comprehensive overview of protein–protein interactions (PPIs). We will begin by elucidating the fundamental concepts of PPIs, covering their definition, structural characteristics, and pivotal roles in both physiological and pathological processes. Building upon this foundation, we delve into the diverse spectrum of current methods employed in PPI studies, showcasing both experimental and computational approaches, along with their varied applications. The chapter culminates in a discussion of the profound implications of PPI research, illuminating its potential in advancing medical understanding, revolutionizing drug discovery, and catalyzing technological innovations across various fields.

1.1.1 Definition of Protein–Protein Interactions

PPIs refer to physical contacts between two or more proteins that occur within a defined biological context. First, these interactions are intentional, not arising from random encounters, but rather precisely orchestrated by specific biomolecular forces and mechanisms. Second, they should be nongeneric, distinct from basic cellular processes like protein synthesis or degradation. Also of note, the formation and regulation of PPIs are highly organized in time and space, governed by a complex interplay of factors such as cell type, cell cycle phase, developmental stage, protein modifications, absence or presence of cofactors and binding partners, and environmental conditions [2, 3].

Targeting Protein–Protein Interactions for Drug Discovery, First Edition. Edited by Jian Zhang.

1.1.2 Structural Properties of Protein–Protein Interactions

The intricate architecture of proteins is critical in determining the specificity, strength, and functional outcomes of their interactions. By unraveling these structural characteristics, researchers can gain deeper insights into the molecular mechanisms underlying PPIs and develop strategies to manipulate them for therapeutic purposes.

Proteins are sophisticated macromolecules comprising an array of structural and functional units that work in concert to perform diverse biological tasks. Among these, domains and short linear motifs (SLiMs) serve as two main classes of functional modules that mediate PPIs. Domains, typically spanning 50–200 residues, are independently folding structural units within proteins that often harbor distinct biological activities. SLiMs, on the other hand, are compact, recurring functional peptides consisting of 3–10 residues, primarily found within intrinsically disordered regions (IDRs) [4]. Indeed, many PPIs can be categorized into domain–domain interactions (DDIs) or domain–motif interactions (DMIs). DDIs usually underpin the formation of stable and long-lasting complexes, while DMIs are associated with transient and low-affinity interactions [5].

The surface regions where the direct physical interactions between two or more proteins occur are termed interfaces. They are highly diverse in terms of size, shape, and chemical properties, determining the specific recognition and binding between proteins engaged in different biological processes. According to statistics, one-sided size of an interface typically ranges from 200 to 2800 $Å^2$, with the majority falling between 200 and 1200 $Å^2$ [6]. Although considered relatively flat, interfaces possess a complex topography characterized by cavities, grooves, and protruding regions. Complementary geometric features ensure that proteins bind to each other in the correct orientation and with high specificity. Moreover, interfaces encompass a variety of chemical interactions, including hydrogen bonds, hydrophobic interactions, electrostatic forces, salt bridges, and disulfide bonds, which collectively account for the specificity and stability of PPIs [7].

Within the interfaces, there exist clusters of residues known as "hot spots," which make disproportionately large contributions to the binding affinity. Single-point mutations of these residues to alanine may cause a substantial increase in the binding free energy ($\Delta\Delta G \geq 2$ kcal/mol), highlighting their critical roles in stabilizing the PPIs [8]. Hot spots have a distinctive amino acid composition, enriched with tryptophan, arginine, and tyrosine, due to the unique physicochemical properties of these residues like bulky side chains, the propensity to form hydrophobic surfaces, and the capacity to engage in hydrogen bonding [9]. With regard to spatial organization, hot spots cluster within tightly packed regions rather than being randomly distributed, facilitating the removal of water molecules upon binding [10]. Besides, they are surrounded by moderately conserved and energetically less important residues, which form an O-ring to further occlude bulk solvent from the hot spots [9].

Proteins are not static, rigid structures; rather, they are dynamic entities capable of adopting multiple conformations. This inherent flexibility is critical for their diverse functions, particularly in PPIs. The traditional "lock and key" model, which

implies a preexisting perfect fit between interacting proteins, fails to capture the dynamic nature of PPIs [11]. Subsequently, more precise descriptions have been put forward. The "induced fit" model suggests that upon initial contact, proteins undergo conformational changes to achieve an optimal fit with their binding partners [12]. Meanwhile, the "conformational selection" model proposes that proteins exist in an equilibrium of various conformations, with binding events selecting and stabilizing the most favorable conformation [13].

1.1.3 Diverse Types of Protein–Protein Interactions

PPIs manifest in a striking diversity of forms, each tailored to perform a specific biological role. In the following sections, we will explore several prominent types of PPIs, varying in structural properties, molecular recognition mechanisms, and functional outcomes.

1.1.3.1 Enzyme–Substrate Interactions

Enzymes are biological catalysts that bind to specific substrates through their active sites and facilitate the conversion of substrates into product molecules. Key examples that illustrate the diverse roles of enzyme–substrate interactions include: (i) protein phosphorylation and dephosphorylation, where kinases and phosphatases add and remove phosphate groups to substrates, respectively, act as molecular switches in signal transduction [14]; (ii) proteolytic cleavage, an irreversible process mediated by proteases that catalyze the hydrolysis of peptide bonds in target proteins, precisely governs protein maturation, activation, stability, and localization [15]; (iii) histone modifications, such as acetylation and methylation, are carried out by enzymes like histone acetyltransferases (HATs) and histone methyltransferases (HMTs), contributing to epigenetic regulation of gene expression [16].

1.1.3.2 Receptor–Ligand Interactions

Cells possess an intricate communication network to sense and respond to environment cues and stimuli. This network is built upon cell surface receptors that bind external signaling molecules termed ligands, initiating downstream signaling cascades within the cell and manipulating various physiological processes like growth, development, immune response, and metabolism [17]. G protein-coupled receptors (GPCRs) constitute the largest family of these receptors, characterized by a unique architecture comprising seven transmembrane helices joined by intracellular and extracellular loops [18]. Chemokines, a class of small secreted proteins, exemplify protein ligands that bind to a specific subfamily of GPCRs called chemokine receptors. By sequentially binding to the N-terminal region and extracellular loops of their receptors, chemokines induce receptor conformational rearrangement and activation that direct immune cell migration and positioning during inflammation [19].

1.1.3.3 Antigen–Antibody Interactions

Antibodies are glycoproteins produced by B lymphocytes to recognize and bind to foreign substances known as antigens, which can be present on the surface of

invading pathogens like viruses and bacteria. This interaction launches an immune response to neutralize and eliminate pathogens, thus protecting the body from infection and disease [20]. An antibody is composed of two identical heavy chains and two identical light chains connected by disulfide bonds. Antigen recognition is achieved by the complementarity-determining regions (CDRs) located at the N-terminus of both heavy and light chains that precisely match antigen's epitope [21]. Beyond their pivotal role in the immune system, the exquisite specificity of antigen–antibody interactions has also been exploited for diagnostic and therapeutic applications.

1.1.3.4 Chaperone–Client Interactions

Chaperones function as protein quality control machinery, engaging with client proteins to assist in their de novo folding, subcellular translocation, and recovery from stress-induced misfolding and aggregation. Many chaperones belong to the heat shock protein (HSP) families, including HSP60, HSP70, HSP90, and HSP100 [22]. Chaperones from different families exhibit structural and functional diversity, reflected in their distinct binding patterns to clients. Some utilize ATP binding to allosterically regulate their conformation, facilitating efficient client binding and release. Others assemble into oligomeric complexes, creating an enclosed environment for client encapsulation [23]. Structural biology studies have revealed that, in some cases, chaperones bind their clients in interconverting conformational ensembles that are locally highly dynamic. This binding behavior allows clients to undergo folding while bound to chaperones and enables chaperones to accommodate a broad range of clients with varying folding properties and limited shape complementarity [24, 25].

1.1.3.5 Scaffold Interactions

Scaffold proteins are molecular organizers that bring together two or more signaling molecules like kinases and receptors into higher-order complexes to modulate their activities, thereby enabling the efficient spatial–temporal coordination of cellular signaling events [26]. This remarkable organizational capability of scaffolds is attributed to their unique modular architecture. Typically, they encompass discrete interaction domains capable of recognizing and binding to SLiMs (e.g. phosphotyrosine- and proline-containing sequences) on target proteins, along with IDRs that confer flexibility and functional versatility [27]. In addition to simply tethering targets in proximity, scaffolds can exert allosteric regulation on the targets as well. In this way, they can handle both the linear input–output signal transmission and complicated feedback loops within distinct pathways and even promote signal integration and crosstalk among various pathways [28, 29].

1.2 Functional Significance of Protein–Protein Interactions

Far from being static or isolated events, PPIs are dynamically woven into complex networks that enable proteins to communicate, coordinate, and cooperate in

carrying out the essential functions of life. Meanwhile, disruptions or aberrations in PPIs can lead to a cascade of dysfunction, contributing to the development of various diseases.

1.2.1 Cellular Signal Transduction

Signal transduction is the process by which extracellular signals are converted into cellular responses. It is organized into three distinct stages: reception, transduction, and response, each involving PPIs to ensure high fidelity of signal transmission (Figure 1.1).

The reception stage commences with receptor–ligand interactions. To be exact, a variety of extracellular signaling molecules, ranging from neurotransmitters and hormones to proteins like chemokines, thrombin, and growth factors, are recognized by their corresponding receptors embedded in the membrane, such as

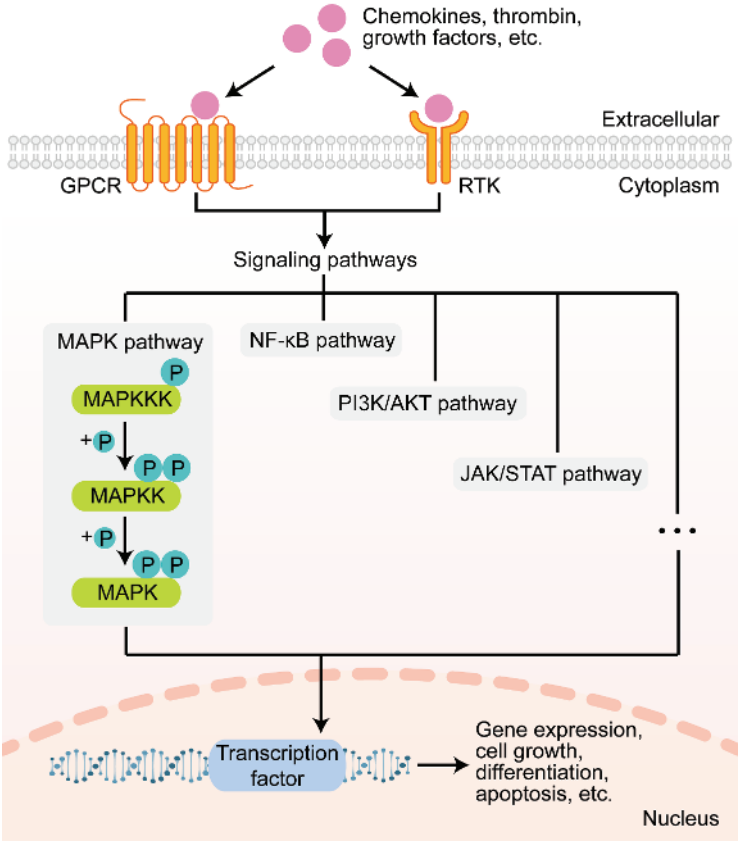

Figure 1.1 A graphic representation of the signal transduction process. Extracellular signals, upon binding to their receptors, are transmitted through intracellular signaling pathways, culminating in a series of cellular responses. (Source: Mi Zhou and Renxiao Wang.)

GPCRs and receptor tyrosine kinases (RTKs) [17]. Upon ligand binding, receptors generally undergo conformational changes, activating associated proteins or triggering homodimerization and autophosphorylation to initiate downstream signaling pathways [30, 31].

During the transduction stage, the signal is propagated from the cell surface to the intracellular targets through an elaborate network of signaling cascades. These cascades are characterized by a series of posttranslational modifications, predominantly phosphorylation events driven by kinase–substrate interactions. A classic example is the mitogen-activated protein kinase (MAPK) pathway, where a cascade of three kinases, namely MAPK kinase kinase (MAPKKK), MAPK kinase (MAPKK), and MAPK, phosphorylate and activate each other in a sequential manner [32]. A similar phosphorylation-dependent activation mechanism is also observed in other crucial pathways, including nuclear factor kappa B (NF-κB) pathway, phosphoinositide 3-kinase (PI3K)/protein kinase B (AKT) pathway [33], and Janus kinase (JAK)/signal transducer and activator of transcription (STAT) pathway [34]. While some interactions serve to relay and amplify the signal, others involving phosphatases and inhibitory proteins lead to signal decay and termination [35]. Moreover, adaptor, scaffold, and docking proteins further fine-tune signal transduction by manipulating the assembly, function, and localization of relevant proteins [36].

Finally, the signal cascades converge on the regulation of their target proteins (e.g. transcription factors, enzymes, cytoskeletal proteins, and apoptotic proteins), eliciting the desired cellular responses such as changes in gene expression, cell growth, differentiation, and apoptosis.

1.2.2 Regulation of Gene Expression

Gene expression in eukaryotes encompasses the processes of transcription, RNA processing, and translation, wherein PPIs play a critical regulatory role, contributing to the accurate transfer of genetic information from DNA to messenger RNA (mRNA) to protein.

At the onset of transcription activation, transcriptional activators bind to specific DNA sequences and then recruit coactivators to improve transcription efficiency. Some coactivators like HMTs and HATs perform histone modifications and chromatin remodeling, rendering the DNA more accessible to the transcriptional machinery [37]. Others, such as the mediator complex, bridge RNA polymerase II (Pol II) and general transcription factors onto core promoters, resulting in the assembly of a pre-initiation complex (PIC) that directs the initiation of transcription [38] (Figure 1.2a). On the other hand, the interplay between transcriptional repressors, co-repressors, and histone-modifying enzymes, like histone deacetylases (HDACs), can suppress the transcription process [39].

Once transcribed, the precursor mRNA (pre-mRNA) undergoes several processing steps to become the mature mRNA. These include 5′-end capping, splicing, and 3′-end cleavage and polyadenylation, each orchestrated by specialized molecular machinery (Figure 1.2b). Capping involves the sequential recruitment of capping

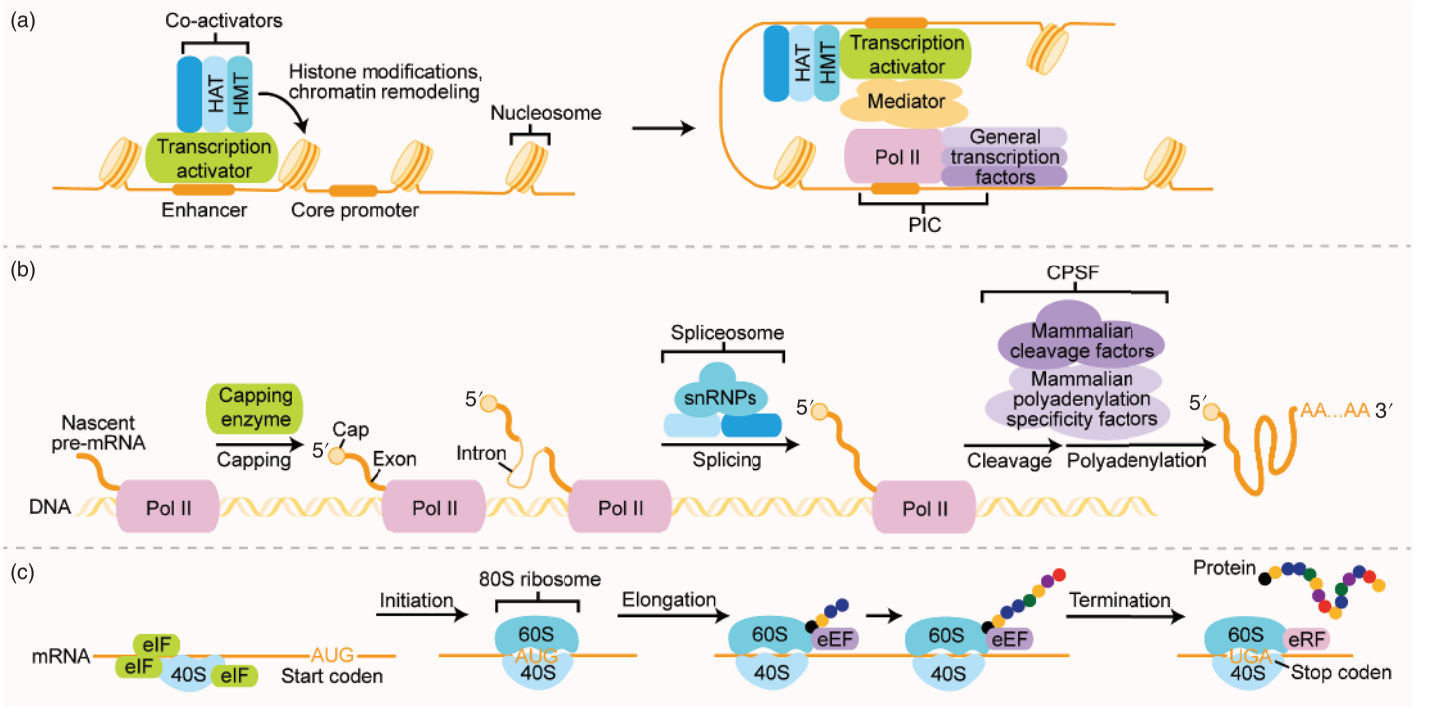

Figure 1.2 Simplified schematic of the gene expression processes in eukaryotes, including (a) transcription, (b) RNA processing, and (c) translation. (Source: Mi Zhou and Renxiao Wang.)

enzymes onto the Pol II, where they catalyze modifications to the 5′ end of the nascent pre-mRNA [40]. Next, splicing is executed by a highly dynamic and massive ribonucleoprotein complex termed the spliceosome, which is assembled by ordered binding and release of small nuclear ribonucleoproteins (snRNPs) and numerous other splicing factors [41]. Finally, 3′-end processing is performed by the cleavage and poly-adenylation specificity factor (CPSF), a large multi-subunit complex that utilizes endonuclease, poly(A) polymerase, and accessory proteins to cleave the pre-mRNA and add a poly(A) tail [42].

During translation initiation, a suite of eukaryotic translation initiation factors (eIFs) mediate the assembly of the 40S and 60S ribosomal subunits at the mRNA start codon, culminating in the formation of the 80S ribosome ready for protein synthesis [43]. Elongation ensues, with eukaryotic elongation factors (eEFs) assisting the ribosome in moving along the mRNA and adding amino acids to the growing polypeptide chain [44]. At last, upon encountering a stop codon, the ribosome triggers the recruitment of eukaryotic release factors (eRFs), promoting the release of the nascent peptide and marking the termination of translation [45] (Figure 1.2c).

1.2.3 Immune Response

The immune response is built upon a foundation of intricate cell–cell communication, with PPIs serving as the key messengers. These interactions, occurring between and within immune and nonimmune cells, are indispensable for recognizing pathogens, activating signal cascades, and mounting an effective defense against invading threats.

1.2.3.1 Immune Cell Migration

Migration is crucial for immune cells to patrol for foreign antigens and traffic to inflammation or infection sites. This process often involves transmigration across epithelial barriers, relying on adhesion factors-governed PPIs. The initial capture and rolling of immune cells along the endothelial surface are mediated by interactions between selectins (e.g. P-selectin) on endothelial cells and their ligands (e.g. P-selectin glycoprotein ligand-1 [PSGL-1]) on immune cells. Film adhesion is subsequently established through interactions between integrins (e.g. lymphocyte function-associated antigen-1 [LFA-1]) on immune cells and their ligands (e.g. intercellular adhesion molecule-1 [ICAM-1]) on endothelial cells [46]. Finally, platelet endothelial cell adhesion molecule-1 (PECAM-1) molecules on both endothelial and immune cells engage in homophilic interactions, promoting transendothelial migration [47].

1.2.3.2 T-Cell Antigen Recognition and Activation

When a foreign antigen invades the body, the T-cell receptor (TCR) on the T-cell surface specifically recognizes and binds to an antigen peptide presented by major histocompatibility complex (MHC) molecules on the surface of antigen-presenting cells (APCs). In parallel, co-receptors, such as cluster of differentiation 4 (CD4) and CD8, on the T-cell surface also engage with the MHC molecules, reinforcing

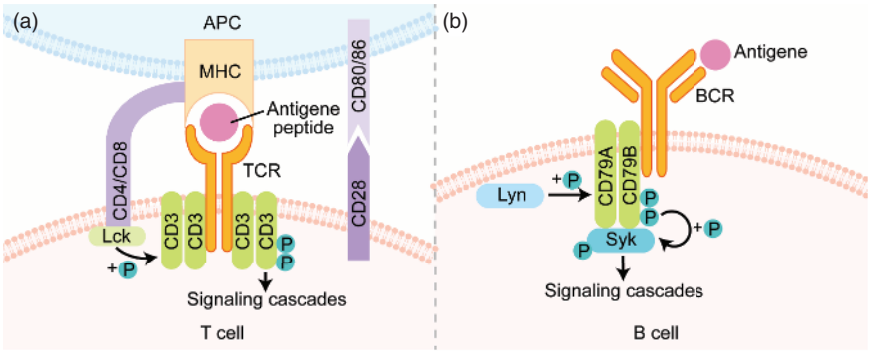

Figure 1.3 Illustration of TCR signaling (a) and BCR signaling (b) during antigen recognition. (Source: Mi Zhou and Renxiao Wang.)

the stability of the TCR-peptide-MHC complex. Co-receptor engagement brings the lymphocyte-specific protein tyrosine kinase (Lck) into proximity with the TCR-CD3 complex. Lck then induces the phosphorylation of CD3 molecules, initiating downstream signaling cascades that ultimately lead to T-cell activation, proliferation, and differentiation. Additionally, PPIs between co-stimulatory molecules, such as CD28 on T cells and CD80/86 on APCs, enhance TCR-induced signals and prevent T-cell anergy [48] (Figure 1.3a).

1.2.3.3 B-Cell Antigen Recognition and Activation

Upon engagement with its cognate antigen, the B-cell receptor (BCR) undergoes clustering on the plasma membrane. This event triggers the phosphorylation of CD79A and CD79B signaling subunits on BCR by Src family tyrosine kinases, primarily Lck/Yes-related novel protein tyrosine kinase (Lyn). Phosphorylated CD79A and CD79B then serve as docking sites for the recruitment and activation of spleen tyrosine kinase (Syk). Once activated, Syk propagates signals to downstream signaling cascades [49, 50] (Figure 1.3b). Similarly, PPIs involving co-receptors and co-stimulatory molecules can augment BCR-induced signals [51]. Ultimately, these signaling events can drive B-cell differentiation into plasma cells, specialized in secreting antibodies that target specific antigens and neutralize or clear them from the body.

1.2.4 Protein Degradation Pathway

To maintain protein homeostasis, cells employ sophisticated quality control mechanisms that promptly identify and eliminate aberrant or unwanted proteins. Two major protein degradation systems, the ubiquitin-proteasome system (UPS) and the autophagy-lysosomal pathway, are essential for this vital task.

UPS selectively targets substrate proteins, tagging them with poly-ubiquitin chains for degradation by the proteasome. The ubiquitination process is carried out by the sequential action of three key enzymes: E1 ubiquitin-activating enzymes, E2-ubiquitin-conjugating enzymes, and E3 ubiquitin ligases. E1 initiates the process

by activating ubiquitin, forming a thioester bond between ubiquitin's C-terminus and its own catalytic cysteine. After that, E1 recruits E2 and transfers ubiquitin to E2's catalytic cysteine. Finally, E3 recognizes and binds to the substrate along with the E2-ubiquitin complex, facilitating the transfer of ubiquitin from E2 onto the substrate [52]. The degradation process is executed by the 26S proteasome, a large multi-subunit protease complex consisting of two main components: 20S core particle (CP) and 19S regulatory particles (RPs). Each component is made up of multiple distinct subunits, with dedicated chaperones assisting in stabilizing intermediate structures and ensuring the proper incorporation of individual subunits. Once assembled, the 20S CP and 19S RPs unite to form the functional 26S proteasome, poised for the degradation of poly-ubiquitylated substrates [53].

The autophagy-lysosomal pathway focuses on the sequestration of cargo (e.g. damaged organelles and protein aggregates) within autophagosomes, which then fuse with lysosomes for cargo degradation and recycling. This process is mainly driven by a set of autophagy-related (ATG) proteins and their interactions. The unc-51-like kinase 1 (ULK1) complex initiates phagophore nucleation, followed by the recruitment of the class III phosphatidylinositol 3-kinase (PI3K) complex that promotes autophagosome formation and elongation. Two ubiquitin-like conjugation systems, namely the ATG12 and microtubule-associated protein 1 light chain 3 (LC3)/GABA type A receptor-associated protein (GABARAP) systems, further contribute to autophagosome maturation [54]. Cargo recognition and delivery are mediated by cargo receptors like sequestosome 1 (p62/SQSTM1), which interact with LC3 on the autophagosome membrane, thereby guiding the cargo for degradation within autophagolysosomal compartment [55].

1.2.5 Disease Mechanisms

PPIs are fundamental for maintaining cellular homeostasis. However, their dysregulation, whether caused by genetic anomalies or the introduction of foreign proteins, plays a significant role in the development and progression of various diseases.

1.2.5.1 Cancer

Neoplastic progression entails the acquisition of a panel of functional capabilities that enable tumor growth and metastasis, termed "hallmarks of cancer" [56]. These hallmarks are inextricably linked to dysregulated PPIs. For instance, aberrant activation of oncoproteins, such as rat sarcoma virus (RAS), induces sustained proliferative signaling by engaging downstream effectors [57]. Negative regulation of the tumor suppressor p53 by mouse double minute 2 homolog (MDM2) is one mechanism whereby cancer cells evade growth suppression [58]. Abnormal expression of anti-apoptotic B-cell lymphoma-2 (BCL-2) family proteins, through sequestering pro-apoptotic counterparts, often contributes to cell death resistance [59]. Hypoxic tumor microenvironment stimulates the production of pro-angiogenic factors like vascular endothelial growth factor (VEGF), which activate their receptors on endothelial cells to induce angiogenesis [60]. Various cancer cells overexpress

programmed cell death ligand 1 (PD-L1) and exploit the PD-L1/programmed cell death protein 1 (PD-1) pathway to escape immune destruction [61].

1.2.5.2 Neurodegenerative Diseases

Alzheimer's disease (AD) is characterized by two hallmark pathologies, amyloid-β (Aβ) plaques and neurofibrillary tangles. The former arises from the sequential proteolytic cleavage of amyloid precursor protein (APP) by secretases, generating Aβ peptides that accumulate and deposit into amyloid plaques [62]. The latter comprises insoluble aggregates of hyperphosphorylated tau protein, a consequence of an imbalanced phosphorylation and dephosphorylation [63]. Parkinson's disease (PD) is marked by the pathological aggregation of α-synuclein (α-syn) into insoluble fibrillar structures called Lewy bodies. While genetic and environmental factors contribute to α-syn aggregation, its engagements with various other proteins, including those implicated in neurodegenerative diseases (e.g. tau) and regulators of actin cytoskeleton and neurotransmission, can accelerate this process [64]. Likewise, Huntington's disease (HD) is driven by the self-aggregation of mutant huntingtin (mHTT) protein. Abnormal PPIs between mHTT and other proteins disrupt crucial cellular processes such as axonal transport, mitochondrial fission, and gene transcription, resulting in neuronal and mitochondrial dysfunction [65].

1.2.5.3 Infectious Disease

The pathogenesis of infections caused by viruses, bacteria, fungi, or parasites hinges critically on a complex interplay between pathogen proteins and host proteins. These PPIs underpin every stage of infections, including:

(i) Invasion: Viruses gain entry into the host by utilizing their surface proteins to recognize specific host cell receptors. A notable example is the binding of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike (S) protein to the human angiotensin-converting enzyme 2 (ACE2) receptor, which mediates the viral attachment and entry into the respiratory epithelial cells [66]. Bacteria and fungi, on the other hand, rely on surface-exposed adhesins that interact with corresponding host cell receptors or extracellular matrix proteins, enabling their attachment and colonization [67, 68].

(ii) Immune Evasion: Pathogens employ various strategies to escape or suppress host's immune defenses. Some, like human cytomegalovirus (HCMV), encode specific proteins that target and interfere with MHC molecules, disrupting antigen presentation [69]. Others, such as *Streptococcus pyogenes*, secrete immunoglobulin-degrading enzymes to cleave antibodies, rendering them ineffective [70].

(iii) Replication: Pathogens often hijack host cellular machinery for their replication needs. For example, expression of the human immunodeficiency virus (HIV) genome is regulated by its trans-activator of transcription (TAT) protein, which assembles with host's positive transcription elongation factor b (P-TEFb) onto the viral promoter, thus enhancing the production of viral transcripts [71].

1.3 Methods for Analyzing Protein–Protein Interactions

The field of PPI research has undergone a transformative evolution with the advent of diverse and sophisticated methodologies, ranging from traditional biochemical techniques to high-throughput technologies and advanced computational algorithms. Together, these complementary strategies form a powerful toolkit for researches to uncover insights into PPIs, from atomic-level details to intricate systemic networks.

1.3.1 Experimental Methods

A diverse array of experimental methods, grounded in either biophysical, biochemical, or genetic principles, have been established to identify and characterize PPIs across various biological contexts and scales. These methods offer multifaceted capabilities, including:

1.3.1.1 Structure Determination

Elucidating the 3D structure of PPIs may provide valuable insights into the geometry and physicochemical characteristics of the binding interfaces and even shed light on potential therapeutic interventions [72]. The primary techniques currently employed for high-resolution protein structure determination are X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and cryogenic electron microscopy (cryo-EM). X-ray crystallography offers detailed information on atom positions and chemical bonds by analyzing X-ray diffraction patterns generated from protein crystals [73]. As the workhorse of structural biology, it accounts for over 80% of experimental structures deposited in the Protein Data Bank (PDB), with the median resolution plateauing at ~2 Å as of 1990 [74]. Inherent limitations of crystallography are the requirement to obtain well-diffracting crystals and the occurrence of crystal packing artifacts. NMR spectroscopy, eliminating the crystallization process, enables structural studies of proteins in solution with the aid of structural restraints acquired from different NMR experiments. Given its strength in uncovering protein dynamics and flexibility over a range of timescales, this technique is well-suited for investigating intrinsically disordered proteins, proteins in transient states, and weak interactions [75, 76]. Whereas drawbacks, such as the necessity of isotope labeling and size limit (typically <50 kDa), should be taken into account. Driven by technological breakthroughs over the past decade, cryo-EM has experienced a "resolution revolution" and is gaining widespread popularity [77]. It relies on direct visualization of proteins embedded in a thin layer of vitreous ice, thus unraveling protein structures in close-to-physiological state with a routine resolution of 2.5–4.5 Å [78]. Compared with other techniques, it is more tolerant to sample quantity and purity. Although high-resolution reconstruction of small-size proteins (<100 kDa) remains challenging for cryo-EM, novel strategies have been applied to break through the barrier [79]. In conclusion, each technique has its own pros and cons, rendering them complementary tools to decipher the intricate 3D structures of proteins and protein complexes [80, 81].

1.3.1.2 Affinity, Kinetics, and Thermodynamics Measurement

Quantitative in vitro analysis of PPIs in terms of affinity, kinetics, and thermodynamics is essential for exploring the underlying mechanisms of molecular recognition. Binding affinity, defined as the strength of the interaction between biomolecules that bind reversibly, is typically quantified through the equilibrium dissociation constant (K_d) [82]. A variety of experimental methods have been established to obtain the K_d values, including fluorescence polarization (FP), surface plasmon resonance (SPR), biolayer interferometry (BLI), isothermal titration calorimetry (ITC), differential scanning calorimetry (DSC), microscale thermophoresis (MST), fluorescence and bioluminescence resonance energy transfer (FRET and BRET), AlphaScreen, protein-fragment complementation assay (PCA), enzyme-linked immunosorbent assays (ELISA), and so forth [83, 84]. Of the methods mentioned earlier, SPR and BLI can also yield kinetic parameters, specifically the association rate constant (k_{on}) and dissociation rate constant (k_{off}) [85]. Both technology platforms enable real-time and label-free monitoring of binding events and have a wide range of applications, such as investigating the assembly and activation of signaling complexes [86], antibody evaluation and selection [87], and protein engineering [88]. On the other hand, calorimetric methods like ITC and DSC can furnish extra thermodynamic insight. In the case of ITC, a comprehensive thermodynamic profile, including Gibbs free energy change (ΔG), enthalpy change (ΔH), entropy change (ΔS), and heat capacity change (ΔC_p), can be acquired from a single experiment [89]. Combining these data with structural information can further interpret the driving forces for the binding [90]. As a final note, prior to choosing the appropriate methods, factors such as the need for immobilization or labeling, affinity range, sample and time consumption, throughput performance, and the potential risks of false positives and false negatives should be carefully considered. Besides, integrating and validating results from multiple methods is crucial to ensure data accuracy and reliability.

1.3.1.3 Large-Scale PPI Network Mapping

Systematic mapping of PPI network within a cell or an organism, termed the "interactome," is an ongoing endeavor aimed at deciphering previously uncharacterized proteins, global proteome organization and function, and genotype–phenotype relationships. For nearly three decades, proteome-wide interactomes of yeast, human, and other model organisms have been delineated and progressively expanded using diverse high-throughput experimental techniques [91, 92]. One powerful method to identify binary PPIs is yeast two-hybrid (Y2H) screening. In the original Y2H, PPIs are detected inside the yeast nucleus based on the functional reconstitution of a transcription factor. Thereafter alternative versions have broadened its applicability to membrane proteins, proteins with posttranslational modifications, and PPIs in mammalian cells [93]. In 2020, an ORFeome resource covering 17,408 protein-coding genes was established, and over 150 million protein pairs were screened using Y2H, generating the largest human binary protein interactome consisting of ~53,000 high-quality PPIs [94]. In contrast, co-complex mapping is typically carried out via affinity purification coupled to mass spectrometry (AP-MS). In AP-MS, a bait protein and its binding partners are isolated from

cell lysate using bait-specific antibody and then subjected to mass spectrometric analysis for identification of the purified components. AP-MS excels at detecting PPIs under near physiological conditions but falls short in capturing weak or transient interactions [95]. It was recently employed in the BioPlex project for the identification of dual human interactomes that encompass 118,162 and 70,966 PPIs in 293T and HCT116 cells, respectively [92]. Despite significant experimental efforts, current human interactomes remain incomplete and present high levels of noise [96]. Methodological advances, integration of various data sources, and incorporation of context-specific (e.g. cell-, tissue-, and disease-specific) information may enhance both interactome coverage and quality, providing researchers with a more comprehensive and dynamic perspective.

1.3.2 Computational Methods

Despite remarkable efforts made by experimental techniques, they often struggle to capture transient and weak interactions or those involving intrinsically disordered proteins. Besides, they are time-consuming and labor-intensive in general. Computational methods offer a complementary and efficient alternative, harnessing vast datasets and predictive models to infer and analyze PPIs.

1.3.2.1 Sequence-Based Methods

The abundance of protein sequence data in public databases has propelled the development of strategies for PPI prediction solely based on sequence information. Herein are some widely adopted approaches: (i) Ortholog-based methods are premised on the evolutionary conservation of PPIs across species, utilizing known PPIs in one species to identify orthologous interactions ("interologs") in another species [97]. (ii) Domain-based methods are built on the concept that PPIs are primarily mediated by domains. By integrating experimentally determined PPI datasets with domain information sourced from databases (e.g. Pfam [98]), DDIs can be inferred and serve as the basis for predicting novel PPIs [99]. (iii) Co-evolution analysis assumes that amino acid substitutions in one protein are often accompanied by compensatory substitutions in its interacting partner to maintain functional integrity. Multiple sequence alignments are employed to quantify the co-evolutionary relationships between residue pairs within proteins of interact, with highly co-evolving pairs presumed to be in physical or functional contact, suggesting a potential PPI [100]. (iv) Machine learning-based methods have gained considerable prominence, progressing from traditional algorithms like support vector machines (SVMs) and random forests (RFs) to advanced deep learning models such as deep neural networks (DNNs) and recurrent neural networks (RNNs). These models extract sequence-based features (e.g. amino acid composition, physicochemical properties, and evolutionary information) from large data sets, capture the underlying patterns, and make accurate predictions for novel protein pairs [101, 102]. To date, protein sequence data remain the primary source for computational PPI prediction. Ongoing improvements are driven by the growing availability of high-quality PPI data sets, the introduction of innovative

sequence encoding schemes, as well as the integration with diverse biological data such as structure information [103].

1.3.2.2 Structure-Based Methods

Structural characterization of PPIs remains a challenge due to their inherent complexity and technological limitations. In fact, fewer than 5% of human PPIs have been structurally resolved through experimental techniques or homology modeling [104]. Molecular docking has emerged as a rapid and cost-effective alternative to bridge this gap by predicting 3D structures of protein complexes. During the docking process, unbound protein structures or structural models are used as input, generating numerous putative binding modes through sampling. The most plausible ones are then selected by means of scoring and ranking [105]. A wide variety of sophisticated docking programs have been developed, each employing different sampling algorithms and scoring functions, including ZDOCK [106], HADDOCK [107], ClusPro [108], and RossettaDock [109]. Continuous efforts have been made to improve prediction accuracy and reliability, such as accounting for backbone and side-chain flexibility, incorporating experimental data like interface details and distance restraints, and integrating deep learning techniques [110, 111]. Molecular dynamics (MD) simulation is another powerful computational technique widely employed to investigate the structural properties of proteins and their complexes. It applies Newtonian mechanics to study the motion of atoms and molecules within a system at nanosecond to microsecond timescales. Analysis of the output trajectories may provide atomic-level insights into the dynamic behavior of protein molecules [112]. In the realm of PPI studies, MD simulation is utilized to assess complex stability, analyze residue flexibility, estimate interaction strengths, and explore the impact of various conditions (e.g. environmental factors, solvents, and mutations) on PPIs [113]. Popular tools for MD simulation include AMBER [114], GROMACS [115], CHARMM [116], and NAMD [117]. Future developments are focused on extending the simulation timescales, improving the accuracy of force fields, and integrating artificial intelligence techniques for enhanced prediction and analysis capabilities.

1.3.2.3 Network-Based Methods

The identification of unmapped interactions within a PPI network can be regarded as the task of predicting missing links between nodes in a graph representation, where nodes denote proteins and links represent interactions between them. Such network-based prediction approaches can be broadly categorized as follows [118]: (i) Similarity-based methods leverage local properties of nodes to establish a similarity score function for link prediction. A classic example is the common neighbor method, which assumes that two proteins sharing a greater number of common neighbors exhibit higher similarity and are thus more likely to interact [119]. (ii) Probabilistic methods aim to identify model parameters that best explain the inherent structure of a network, such as community structure. Taking the stochastic block model as an example, each protein is assigned to multiple communities, reflecting its participation in various biological processes. This model operates on the principle that proteins within the same community are more

prone to interact with each other compared to proteins from different communities [120]. (iii) Factorization-based methods constitute a class of graph embedding techniques that factorize the network's adjacency matrix to build low-dimensional representations for each node. Geometric Laplacian Eigenmap Embedding (GLEE) is one such method that exploits the simplex geometry of the Laplacian matrix to capture the most basic structural property of the graph, demonstrating strong competence in predicting unobserved interactions in PPI networks [121]. (iv) Machine learning-based methods make predictions by learning patterns from training data. As a case in point, a conditional generative adversarial network (cGAN) can be conditioned on the raw topological features of the network, enabling the generator to produce new probable links [122].

1.4 Implications of the Basic Research on Protein–Protein Interactions

Given their essential role in biological functions, deciphering the intricate PPI network not only holds paramount significance in the realm of molecular biology but also has far-reaching implications in medical advances, drug discovery, and biotechnology innovations.

1.4.1 Advancing Disease Understanding and Diagnosis

PPIs play pivotal roles in cellular functions, from normal physiological processes to the development and progression of various diseases. In cancer, genomic alterations give rise to oncogenic PPIs, driving uncontrolled cell proliferation and metastasis [123]. Neurodegenerative diseases often involve mutations, posttranslational modifications, or misfolding of aggregation-prone proteins, resulting in aberrant PPIs that exacerbate cytotoxicity [124]. Infectious diseases heavily rely on host–pathogen PPIs to facilitate pathogen entry, replication, and dissemination within the host organism [125]. Consequently, identifying and characterizing these interactions is essential for unraveling the molecular basis of diseases. The COVID-19 pandemic serves as a compelling case in point. Investigation of the interaction between the SARS-CoV-2 spike protein and the human ACE2 receptor advanced our knowledge of the virus's entry mechanism into host cells [126]. Further affinity measurement of different SARS-CoV-2 variants toward ACE2 helped to assess the virus's infectivity, transmissibility, and potential threat to public health, offering invaluable information for epidemiological studies [127].

Current molecular diagnostics mainly focus on identifying genetic alterations and individual protein biomarkers. However, emerging evidence suggests that altered PPI patterns are strongly linked to disease progression and clinical outcomes, holding significant potential as novel diagnostic and prognostic biomarkers [128]. One notable example lies in the application of proximity ligation assay (PLA) to detect protein complexes formed between epidermal growth factor receptor (EGFR) and growth factor receptor-bound protein 2 (GRB2) in non-small cell lung cancer

(NSCLC) patient specimens. This approach allows for *in situ* assessment of EGFR signaling activity and has proven to be a superior predictor of clinical response to targeted therapies compared to conventional strategies based on EGFR mutations or protein abundance [129]. As our understanding of disease-related PPIs deepens and advanced methods for PPI profiling in clinical specimens emerge, PPI-based diagnostics, alongside genomic and proteomic analyses, will become increasingly powerful tools for precision and personalized medicine.

1.4.2 Driving Target-Based Drug Discovery

Given their central role in numerous disease pathways, PPIs have emerged as attractive targets for therapeutic intervention. However, modulating these interactions presents challenges due to the complex and dynamic nature of PPI interfaces, which are often characterized by extensive, shallow surfaces with limited deep cavities. PPIs lack endogenous small-molecule ligands that act as a starting point for drug discovery [130]. Despite these obstacles, substantial efforts have been made to overcome these barriers. Structural biology techniques can pinpoint hot spots at PPI interfaces, serving as a basis for the structure-based design of PPI modulators. Computational methods, such as virtual screening and molecular docking, can predict the binding affinity and mode of molecules at binding sites, accelerating lead discovery. By integrating these approaches, researchers can efficiently identify and optimize PPI modulators, opening a new avenue for drug discovery.

Over the past decade, the field of PPI modulators has witnessed remarkable progress, with quite a few candidates being approved for medical use or entering clinical trials for the treatment of various diseases, particularly solid tumors [131]. Small molecules constitute the most abundant class of PPI modulators, offering advantages such as oral bioavailability and membrane permeability. Classic PPI targets include interactions between antiapoptotic BCL-2 family proteins and their proapoptotic counterparts, X-linked inhibitor of apoptosis (XIAP) and caspase-9, and MDM2-p53 [132]. A significant milestone was achieved in 2016 with the approval of venetoclax, a BCL-2 inhibitor, for the treatment of chronic lymphocytic leukemia (CLL), making a breakthrough in the development of PPI-modulating therapies [133]. The second category comprises antibodies, which are well-suited for covering large PPI interfaces but typically restricted to extracellular targets. Their primary focus lies in manipulating PPIs involving immune checkpoint molecules, such as PD-1/PD-L1 or CD40/CD40 ligand (CD40L) interaction, thereby exerting profound influences on immune responses [134, 135]. Peptides represent the third category of PPI modulators, exhibiting high selectivity but hampered by a short half-life. To address this, chemical modifications like cyclization are frequently employed to improve their pharmacokinetic properties [136].

1.4.3 Fostering Innovations in Biotechnology

Biotechnology often harnesses the power of protein interactions to develop innovative solutions and products. This is vividly illustrated in synthetic biology, where

naturally sourced or computationally designed PPIs serve as building blocks for constructing sophisticated genetic and protein circuits, endowing living cells with novel functionalities [137]. Generally, PPIs can be engineered to function as toggle switches and logic gates within synthetic genetic circuits, offering precise control over gene expression [138, 139]. Furthermore, by incorporating PPIs that govern posttranscriptional modifications like phosphorylation, synthetic protein circuits can rapidly respond to environmental changes, enabling dynamic modulation of protein interactions, localization, and degradation [140]. These programmable circuits hold immense potential for numerous applications. For instance, researchers created a chemically disruptable heterodimer based on the interaction between BCL-X_L (B-cell lymphoma-extra large) and BH3 (BCL-2 homology 3)-mimic peptide. Incorporating this PPI into the design of chimeric antigen receptor (CAR) led to a controllable CAR-T cell therapy, whose activity could be modulated by timed administration of a small-molecule BCL-X_L inhibitor [141].

Protein assemblies found in nature, such as actin fibers, microtubules, and viral capsids, showcase the remarkable role of PPIs in organizing complex architectures. Inspired by this, scientists have employed rational design to fabricate protein assemblies with diverse superstructures like nanowires, nanotubes, 2D and 3D lattices, and cage structures [142]. This nanotechnology paves the way for the development of advanced biomaterials with tailored properties and functionalities. As an example, coiled-coils are ubiquitous protein oligomerization motifs that have the inherent ability to wrap each other into supercoiled helices. These structures can be engineered to undergo higher-order assembly into fibrous network-based hydrogels [143]. Compared to inorganic and polymeric hydrogels, such protein-based hydrogels exhibit higher biocompatibility and better flexibility, rendering them ideal candidates for tissue engineering scaffolds and drug delivery carriers. Another example comes from virus-like particles (VLPs), which are self-assembled nanostructures made up of viral structural proteins, mimicking the architecture and symmetry of native viruses but lacking viral genetic material. Demonstrating high immunogenicity alongside an excellent safety profile, VLPs are gaining popularity in vaccine development [144].

Immunosensors are analytical devices that leverage the specificity of antigen–antibody interactions to detect and quantify target molecules in complex samples, with the resulting immunoreaction converted into an electrochemical, optical, or other detectable signal [145]. Enabled by artificially produced polyclonal or monoclonal antibody, these biosensors offer detection of a wide array of analytes, encompassing tumor biomarkers (cancer antigen 15-3, carcinoembryonic antigen, and α-fetoprotein) [146], pathogens responsible for infectious diseases (e.g. SARS-CoV-2, *Streptococcus pneumoniae*, and *Legionella pneumophila*) [147], foodborne pathogenic microorganisms (e.g. *Salmonella*, *Escherichia coli*, and *Listeria monocytogenes*) [148], food allergens (e.g. ovalbumin in eggs, β-lactoglobulin in milk, and Ara h 1 in peanuts) [149], and pathogens in crops and water resources [145]. The simplicity, rapidity, and sensitivity of immunosensors have led to their widespread adoption across human health, food and agricultural safety, environmental monitoring, and other fields.

1.5 Conclusions and Perspectives

In the post-genomic era, where the focus has shifted from studying individual genes to exploring the complex interplay of biomolecules, the investigation of PPIs stands as a critical frontier. At the heart of virtually every biological function, from cellular signaling and gene expression to immune response and metabolic pathways, lies an intricate network of PPIs. By elucidating these interactions, researchers can gain valuable insights into how cells respond to environmental stimuli, communicate with one another, and maintain homeostasis. The significance of PPI research is further underscored by the increasing number of studies that have identified disease-associated PPIs. This knowledge has opened new avenues for therapeutic intervention, leading to the development of novel drugs that target specific PPIs involved in disease progression.

Technological advancements have dramatically accelerated the exploration of PPIs, with the number of human binary PPIs identified through large-scale screens increasing from approximately 2700 in 2005 to 14,000 in 2014 and reaching 53,000 in 2020 [94]. Even so, the available PPI data represent only a fraction of the estimated human interactome, which is believed to contain 650,000 PPIs [150]. While experimental techniques have made significant strides, challenges remain in terms of throughput, sensitivity, and the ability to capture transient or weak interactions. Leveraging advanced computational techniques, such as molecular docking, machine learning algorithms, and network analysis, can complement experimental findings and expand relevant data resources. Moreover, the emergence of cutting-edge technologies, like single-cell analysis, holds promise for elucidating cell-to-cell variability and the spatiotemporal heterogeneity of PPIs [151]. As the field of PPI research continues to evolve, our understanding of PPIs will be profoundly enhanced, leading to groundbreaking discoveries and advancements that benefit humanity.

References

1 Berggard, T., Linse, S., and James, P. (2007). Methods for the detection and analysis of protein-protein interactions. *Proteomics* 7: 2833–2842.

2 De Las Rivas, J. and Fontanillo, C. (2010). Protein-protein interactions essentials: key concepts to building and analyzing interactome networks. *PLoS Computational Biology* 6: e1000807.

3 Ngounou Wetie, A.G., Sokolowska, I., Woods, A.G. et al. (2014). Protein-protein interactions: switch from classical methods to proteomics and bioinformatics-based approaches. *Cellular and Molecular Life Sciences* 71: 205–228.

4 Van Roey, K., Uyar, B., Weatheritt, R.J. et al. (2014). Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation. *Chemical Reviews* 114: 6733–6778.

5 Seo, M.H. and Kim, P.M. (2018). The present and the future of motif-mediated protein-protein interactions. *Current Opinion in Structural Biology* 50: 162–170.

6 Yan, C., Wu, F., Jernigan, R.L. et al. (2008). Characterization of protein-protein interfaces. *The Protein Journal* 27: 59–70.

7 Keskin, O., Tuncbag, N., and Gursoy, A. (2016). Predicting protein-protein interactions from the molecular to the proteome level. *Chemical Reviews* 116: 4884–4909.

8 Moreira, I.S., Fernandes, P.A., and Ramos, M.J. (2007). Hot spots – a review of the protein-protein interface determinant amino-acid residues. *Proteins* 68: 803–812.

9 Bogan, A.A. and Thorn, K.S. (1998). Anatomy of hot spots in protein interfaces. *Journal of Molecular Biology* 280: 1–9.

10 Keskin, O., Ma, B., and Nussinov, R. (2005). Hot regions in protein-protein interactions: the organization and contribution of structurally conserved hot spot residues. *Journal of Molecular Biology* 345: 1281–1294.

11 Fischer, E. (1894). Influence of configuration on the action of enzymes. *Berichte der Deutschen Chemischen Gesellschaft* 27: 2985–2993.

12 Koshland, D.E. (1958). Application of a theory of enzyme specificity to protein synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 44: 98–104.

13 Ma, B., Kumar, S., Tsai, C.J. et al. (1999). Folding funnels and binding mechanisms. *Protein Engineering* 12: 713–720.

14 Seok, S.H. (2021). Structural insights into protein regulation by phosphorylation and substrate recognition of protein kinases/phosphatases. *Life (Basel)* 11: 957.

15 Klein, T., Eckhard, U., Dufour, A. et al. (2018). Proteolytic cleavage-mechanisms, function, and "Omic" approaches for a near-ubiquitous post-translational modification. *Chemical Reviews* 118: 1137–1168.

16 Liu, R., Wu, J., Guo, H. et al. (2023). Post-translational modifications of histones: mechanisms, biological functions, and therapeutic targets. *MedComm (2020)* 4: e292.

17 Ullo, M.F. and Case, L.B. (2023). How cells sense and integrate information from different sources. *WIREs Mechanisms of Disease* 15: e1604.

18 Cheng, L., Xia, F., Li, Z. et al. (2023). Structure, function and drug discovery of GPCR signaling. *Molecular Biomedicine* 4: 46.

19 Bhusal, R.P., Foster, S.R., and Stone, M.J. (2020). Structural basis of chemokine and receptor interactions: key regulators of leukocyte recruitment in inflammatory responses. *Protein Science* 29: 420–432.

20 Kapingidza, A.B., Kowal, K., and Chruszcz, M. (2020). Antigen-antibody complexes. *Sub-Cellular Biochemistry* 94: 465–497.

21 Sela-Culang, I., Kunik, V., and Ofran, Y. (2013). The structural basis of antibody-antigen recognition. *Frontiers in Immunology* 4: 302.

22 Saibil, H. (2013). Chaperone machines for protein folding, unfolding and disaggregation. *Nature Reviews. Molecular Cell Biology* 14: 630–642.

23 Kim, Y.E., Hipp, M.S., Bracher, A. et al. (2013). Molecular chaperone functions in protein folding and proteostasis. *Annual Review of Biochemistry* 82: 323–355.

24 Hiller, S. and Burmann, B.M. (2018). Chaperone-client complexes: a dynamic liaison. *Journal of Magnetic Resonance* 289: 142–155.

25 Hiller, S. (2019). Chaperone-bound clients: the importance of being dynamic. *Trends in Biochemical Sciences* 44: 517–527.

26 Zeke, A., Lukacs, M., Lim, W.A. et al. (2009). Scaffolds: interaction platforms for cellular signalling circuits. *Trends in Cell Biology* 19: 364–374.

27 Cortese, M.S., Uversky, V.N., and Dunker, A.K. (2008). Intrinsic disorder in scaffold proteins: getting more from less. *Progress in Biophysics and Molecular Biology* 98: 85–106.

28 Good, M.C., Zalatan, J.G., and Lim, W.A. (2011). Scaffold proteins: hubs for controlling the flow of cellular information. *Science* 332: 680–686.

29 Pan, C.Q., Sudol, M., Sheetz, M. et al. (2012). Modularity and functional plasticity of scaffold proteins as p(l)acemakers in cell signaling. *Cellular Signalling* 24: 2143–2165.

30 Jiang, H., Galtes, D., Wang, J. et al. (2022). G protein-coupled receptor signaling: transducers and effectors. *American Journal of Physiology. Cell Physiology* 323: C731–C748.

31 Paul, M.D. and Hristova, K. (2019). The RTK interactome: overview and perspective on RTK Heterointeractions. *Chemical Reviews* 119: 5881–5921.

32 Yoon, S. and Seger, R. (2006). The extracellular signal-regulated kinase: multiple substrates regulate diverse cellular functions. *Growth Factors* 24: 21–44.

33 He, Y., Sun, M.M., Zhang, G.G. et al. (2021). Targeting PI3K/Akt signal transduction for cancer therapy. *Signal Transduction and Targeted Therapy* 6: 425.

34 Hu, X., Li, J., Fu, M. et al. (2021). The JAK/STAT signaling pathway: from bench to clinic. *Signal Transduction and Targeted Therapy* 6: 402.

35 Lemmon, M.A., Freed, D.M., Schlessinger, J. et al. (2016). The dark side of cell signaling: positive roles for negative regulators. *Cell* 164: 1172–1184.

36 Buday, L. and Tompa, P. (2010). Functional classification of scaffold proteins and related molecules. *The FEBS Journal* 277: 4348–4355.

37 Talukdar, P.D. and Chatterji, U. (2023). Transcriptional co-activators: emerging roles in signaling pathways and potential therapeutic targets for diseases. *Signal Transduction and Targeted Therapy* 8: 427.

38 Soutourina, J. (2018). Transcription regulation by the Mediator complex. *Nature Reviews. Molecular Cell Biology* 19: 262–274.

39 Payankaulam, S., Li, L.M., and Arnosti, D.N. (2010). Transcriptional repression: conserved and evolved features. *Current Biology* 20: R764–R771.

40 Garg, G., Dienemann, C., Farnung, L. et al. (2023). Structural insights into human co-transcriptional capping. *Molecular Cell* 83: 2464–2477.e2465.

41 Kastner, B., Will, C.L., Stark, H. et al. (2019). Structural insights into nuclear pre-mRNA splicing in higher eukaryotes. *Cold Spring Harbor Perspectives in Biology* 11: a032417.

42 Boreikaite, V. and Passmore, L.A. (2023). 3′-End processing of eukaryotic mRNA: machinery, regulation, and impact on gene expression. *Annual Review of Biochemistry* 92: 199–225.

43 Brito Querido, J., Diaz-Lopez, I., and Ramakrishnan, V. (2024). The molecular basis of translation initiation and its regulation in eukaryotes. *Nature Reviews. Molecular Cell Biology* 25: 168–186.

44 Xu, B., Liu, L., and Song, G. (2021). Functions and regulation of translation elongation factors. *Frontiers in Molecular Biosciences* 8: 816398.

45 Hellen, C.U.T. (2018). Translation termination and ribosome recycling in eukaryotes. *Cold Spring Harbor Perspectives in Biology* 10: a032656.

46 Cappenberg, A., Kardell, M., and Zarbock, A. (2022). Selectin-mediated signaling-shedding light on the regulation of integrin activity in neutrophils. *Cells* 11: 1310.

47 Privratsky, J.R. and Newman, P.J. (2014). PECAM-1: regulator of endothelial junctional integrity. *Cell and Tissue Research* 355: 607–619.

48 Shah, K., Al-Haidari, A., Sun, J. et al. (2021). T cell receptor (TCR) signaling in health and disease. *Signal Transduction and Targeted Therapy* 6: 412.

49 Efremov, D.G., Turkalj, S., and Laurenti, L. (2020). Mechanisms of B cell receptor activation and responses to B cell receptor inhibitors in B cell malignancies. *Cancers (Basel)* 12: 1396.

50 Tkachenko, A., Kupcova, K., and Havranek, O. (2023). B-cell receptor signaling and beyond: the role of Igalpha (CD79a)/Igbeta (CD79b) in normal and malignant B cells. *International Journal of Molecular Sciences* 25: 10.

51 Susa, K.J., Seegar, T.C., Blacklow, S.C. et al. (2020). A dynamic interaction between CD19 and the tetraspanin CD81 controls B cell co-receptor trafficking. *eLife* 9: e52337.

52 Toma-Fukai, S. and Shimizu, T. (2021). Structural diversity of ubiquitin E3 ligase. *Molecules* 26: e52337.

53 Kim, Y., Kim, E.K., Chey, Y. et al. (2023). Targeted protein degradation: principles and applications of the proteasome. *Cells* 12: e2300083.

54 Siva Sankar, D. and Dengjel, J. (2021). Protein complexes and neighborhoods driving autophagy. *Autophagy* 17: 2689–2705.

55 Johansen, T. and Lamark, T. (2020). Selective autophagy: ATG8 family proteins, LIR motifs and cargo receptors. *Journal of Molecular Biology* 432: 80–103.

56 Hanahan, D. and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144: 646–674.

57 Tatli, O. and Dinler Doganay, G. (2021). Recent developments in targeting RAS downstream effectors for RAS-driven cancer therapy. *Molecules* 26: 7561.

58 Liu, Y., Su, Z., Tavana, O. et al. (2024). Understanding the complexity of p53 in a new era of tumor suppression. *Cancer Cell* 42: 946–967.

59 Qian, S., Wei, Z., Yang, W. et al. (2022). The role of BCL-2 family proteins in regulating apoptosis and cancer therapy. *Frontiers in Oncology* 12: 985363.

60 Liu, Z.L., Chen, H.H., Zheng, L.L. et al. (2023). Angiogenic signaling pathways and anti-angiogenic therapy for cancer. *Signal Transduction and Targeted Therapy* 8: 198.

61 Han, Y., Liu, D., and Li, L. (2020). PD-1/PD-L1 pathway: current researches in cancer. *American Journal of Cancer Research* 10: 727–742.

62 Zhou, R., Yang, G., and Shi, Y. (2020). Macromolecular complex in recognition and proteolysis of amyloid precursor protein in Alzheimer's disease. *Current Opinion in Structural Biology* 61: 1–8.

63 Muralidar, S., Ambi, S.V., Sekaran, S. et al. (2020). Role of tau protein in Alzheimer's disease: the prime pathological player. *International Journal of Biological Macromolecules* 163: 1599–1617.

64 Wang, J., Dai, L., Chen, S. et al. (2024). Protein-protein interactions regulating alpha-synuclein pathology. *Trends in Neurosciences* 47: 209–226.

65 Wanker, E.E., Ast, A., Schindler, F. et al. (2019). The pathobiology of perturbed mutant huntingtin protein-protein interactions in Huntington's disease. *Journal of Neurochemistry* 151: 507–519.

66 Jackson, C.B., Farzan, M., Chen, B. et al. (2022). Mechanisms of SARS-CoV-2 entry into cells. *Nature Reviews. Molecular Cell Biology* 23: 3–20.

67 Kline, K.A., Fälker, S., Dahlberg, S. et al. (2009). Bacterial adhesins in host-microbe interactions. *Cell Host & Microbe* 5: 580–592.

68 Kumari, A., Tripathi, A.H., Gautam, P. et al. (2021). Adhesins in the virulence of opportunistic fungal pathogens of human. *Mycology* 12: 296–324.

69 Noriega, V., Redmann, V., Gardner, T. et al. (2012). Diverse immune evasion strategies by human cytomegalovirus. *Immunologic Research* 54: 140–151.

70 Brouwer, S., Rivera-Hernandez, T., Curren, B.F. et al. (2023). Pathogenesis, epidemiology and control of Group A Streptococcus infection. *Nature Reviews. Microbiology* 21: 431–447.

71 Clark, E., Nava, B., and Caputi, M. (2017). Tat is a multifunctional viral protein that modulates cellular gene expression and functions. *Oncotarget* 8: 27569–27581.

72 Winter, C., Henschel, A., Tuukkanen, A. et al. (2012). Protein interactions in 3D: from interface evolution to drug discovery. *Journal of Structural Biology* 179: 347–358.

73 Brito, J.A. and Archer, M. (2020). Chapter 10 - Structural biology techniques: X-ray crystallography, cryo-electron microscopy, and small-angle X-ray scattering. In: *Practical Approaches to Biological Inorganic Chemistry*, 2e (ed. R.R. Crichton and R.O. Louro), 375–416. Elsevier.

74 Burley, S.K., Berman, H.M., Duarte, J.M. et al. (2022). Protein data bank: a comprehensive review of 3D structure holdings and worldwide utilization by researchers, educators, and students. *Biomolecules* 12: 1425.

75 Alderson, T.R. and Kay, L.E. (2021). NMR spectroscopy captures the essential role of dynamics in regulating biomolecular function. *Cell* 184: 577–595.

76 Hu, Y., Cheng, K., He, L. et al. (2021). NMR-based methods for protein analysis. *Analytical Chemistry* 93: 1866–1879.

77 Kuhlbrandt, W. (2014). Biochemistry. The resolution revolution. *Science* 343: 1443–1444.

78 Fukuda, Y., Stapleton, K., and Kato, T. (2023). Progress in spatial resolution of structural analysis by cryo-EM. *Microscopy (Oxf)* 72: 135–143.

79 Wentinck, K., Gogou, C., and Meijer, D.H. (2022). Putting on molecular weight: enabling cryo-EM structure determination of sub-100-kDa proteins. *Current Research in Structural Biology* 4: 332–337.

80 Wang, H.W. and Wang, J.W. (2017). How cryo-electron microscopy and X-ray crystallography complement each other. *Protein Science* 26: 32–39.

81 Geraets, J.A., Pothula, K.R., and Schroder, G.F. (2020). Integrating cryo-EM and NMR data. *Current Opinion in Structural Biology* 61: 173–181.

82 Kastritis, P.L. and Bonvin, A.M. (2013). On the binding affinity of macromolecular interactions: daring to ask why proteins interact. *Journal of the Royal Society, Interface* 10: 20120835.

83 Zhou, M., Li, Q., and Wang, R. (2016). Current experimental methods for characterizing protein-protein interactions. *ChemMedChem* 11: 738–756.

84 Poluri, K.M., Gulati, K., and Sarkar, S. (2021). Experimental methods for determination of protein–protein interactions. In: *Protein-Protein Interactions: Principles and Techniques: Volume I* (ed. K.M. Poluri, K. Gulati, and S. Sarkar), 197–264. Singapore: Springer.

85 Murali, S., Rustandi, R.R., Zheng, X. et al. (2022). Applications of surface plasmon resonance and biolayer interferometry for virus-ligand binding. *Viruses* 14: 717.

86 Zhou, L., Todorovic, V., Kakavas, S. et al. (2018). Quantitative ligand and receptor binding studies reveal the mechanism of interleukin-36 (IL-36) pathway activation. *The Journal of Biological Chemistry* 293: 403–411.

87 Ylera, F., Harth, S., Waldherr, D. et al. (2013). Off-rate screening for selection of high-affinity anti-drug antibodies. *Analytical Biochemistry* 441: 208–213.

88 Nastle, L., Deuschle, F.C., Morath, V. et al. (2023). FerryCalin: an engineered lipocalin protein directed against the transferrin receptor with potential for brain drug delivery. *Chembiochem* 24: e202200795.

89 Roselin, L.S., Lin, M.S., Lin, P.H. et al. (2010). Recent trends and some applications of isothermal titration calorimetry in biotechnology. *Biotechnology Journal* 5: 85–98.

90 Brookes, J.C., Gray, E.R., Loynachan, C.N. et al. (2023). Thermodynamic analysis of an entropically driven, high-affinity nanobody-HIV p24 interaction. *Biophysical Journal* 122: 279–289.

91 Bludau, I. and Aebersold, R. (2020). Proteomic and interactomic insights into the molecular basis of cell functional diversity. *Nature Reviews. Molecular Cell Biology* 21: 327–340.

92 Huttlin, E.L., Bruckner, R.J., Navarrete-Perea, J. et al. (2021). Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. *Cell* 184: 3022–3040.e3028.

93 Stynen, B., Tournu, H., Tavernier, J. et al. (2012). Diversity in genetic in vivo methods for protein-protein interaction studies: from the yeast two-hybrid system to the mammalian split-luciferase system. *Microbiology and Molecular Biology Reviews* 76: 331–382.

94 Luck, K., Kim, D.K., Lambourne, L. et al. (2020). A reference map of the human binary protein interactome. *Nature* 580: 402–408.

95 Low, T.Y., Syafruddin, S.E., Mohtar, M.A. et al. (2021). Recent progress in mass spectrometry-based strategies for elucidating protein-protein interactions. *Cellular and Molecular Life Sciences* 78: 5325–5339.

96 Sharifi Tabar, M., Parsania, C., Chen, H. et al. (2022). Illuminating the dark protein-protein interactome. *Cell Reports Methods* 2: 100275.

97 Matthews, L.R., Vaglio, P., Reboul, J. et al. (2001). Identification of potential interaction networks using sequence-based searches for conserved protein-protein interactions or "interologs". *Genome Research* 11: 2120–2126.

98 Mistry, J., Chuguransky, S., Williams, L. et al. (2021). Pfam: the protein families database in 2021. *Nucleic Acids Research* 49: D412–D419.

99 Deng, M., Mehta, S., Sun, F. et al. (2002). Inferring domain-domain interactions from protein-protein interactions. *Genome Research* 12: 1540–1548.

100 Burger, L. and van Nimwegen, E. (2008). Accurate prediction of protein-protein interactions from sequence alignments using a Bayesian method. *Molecular Systems Biology* 4: 165.

101 Tang, T., Zhang, X., Liu, Y. et al. (2023). Machine learning on protein-protein interaction prediction: models, challenges and trends. *Briefings in Bioinformatics* 24: bbad076.

102 Casadio, R., Martelli, P.L., and Savojardo, C. (2022). Machine learning solutions for predicting protein–protein interactions. *WIREs Computational Molecular Science* 12: e1618.

103 Khatun, M.S., Shoombuatong, W., Hasan, M.M. et al. (2020). Evolution of sequence-based bioinformatics tools for protein-protein interaction prediction. *Current Genomics* 21: 454–463.

104 Burke, D.F., Bryant, P., Barrio-Hernandez, I. et al. (2023). Towards a structurally resolved human protein interaction network. *Nature Structural & Molecular Biology* 30: 216–225.

105 Huang, S.Y. (2014). Search strategies and evaluation in protein-protein docking: principles, advances and challenges. *Drug Discovery Today* 19: 1081–1096.

106 Pierce, B.G., Wiehe, K., Hwang, H. et al. (2014). ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics* 30: 1771–1773.

107 van Zundert, G.C.P., Rodrigues, J., Trellet, M. et al. (2016). The HADDOCK2.2 Web server: user-friendly integrative modeling of biomolecular complexes. *Journal of Molecular Biology* 428: 720–725.

108 Kozakov, D., Hall, D.R., Xia, B. et al. (2017). The ClusPro web server for protein-protein docking. *Nature Protocols* 12: 255–278.

109 Lyskov, S. and Gray, J.J. (2008). The RosettaDock server for local protein-protein docking. *Nucleic Acids Research* 36: W233–W238.

110 van Noort, C.W., Honorato, R.V., and Bonvin, A. (2021). Information-driven modeling of biomolecular complexes. *Current Opinion in Structural Biology* 70: 70–77.

111 Sunny, S. and Jayaraj, P.B. (2022). Protein-protein docking: past, present, and future. *The Protein Journal* 41: 1–26.

112 Sinha, S., Tam, B., and Wang, S.M. (2022). Applications of molecular dynamics simulation in protein study. *Membranes (Basel)* 12: 844.

113 Shukla, R. and Tripathi, T. (2020). Molecular dynamics simulation of protein and protein–ligand complexes. In: *Computer-Aided Drug Design* (ed. D.B. Singh), 133–161. Singapore: Springer.

114 Case, D.A., Aktulga, H.M., Belfon, K. et al. (2023). AmberTools. *Journal of Chemical Information and Modeling* 63: 6183–6191.

115 Pall, S., Zhmurov, A., Bauer, P. et al. (2020). Heterogeneous parallelization and acceleration of molecular dynamics simulations in GROMACS. *The Journal of Chemical Physics* 153: 134110.

116 Brooks, B.R., Brooks, C.L. 3rd, Mackerell, A.D. Jr., et al. (2009). CHARMM: the biomolecular simulation program. *Journal of Computational Chemistry* 30: 1545–1614.

117 Phillips, J.C., Hardy, D.J., Maia, J.D.C. et al. (2020). Scalable molecular dynamics on CPU and GPU architectures with NAMD. *The Journal of Chemical Physics* 153: 044130.

118 Wang, X.W., Madeddu, L., Spirohn, K. et al. (2023). Assessment of community efforts to advance network-based prediction of protein-protein interactions. *Nature Communications* 14: 1582.

119 Haque, M., Sarmah, R., and Bhattacharyya, D.K. (2018). A common neighbor based technique to detect protein complexes in PPI networks. *Journal, Genetic Engineering & Biotechnology* 16: 227–238.

120 Wang, X., Hu, P., and Hu, L. (2020). A novel stochastic block model for network-based prediction of protein-protein interactions. In: *Intelligent Computing Theories and Application*. Cham: Springer International Publishing.

121 Torres, L., Chan, K.S., Eliassi-Rad, T. et al. (2020). GLEE: Geometric Laplacian Eigenmap Embedding. *Journal of Complex Networks* 8: 1–17.

122 Balogh, O.M., Benczik, B., Horvath, A. et al. (2022). Efficient link prediction in the protein-protein interaction network using topological information in a generative adversarial network machine learning model. *BMC Bioinformatics* 23: 78.

123 Sharifi Tabar, M., Francis, H., Yeo, D. et al. (2022). Mapping oncogenic protein interactions for precision medicine. *International Journal of Cancer* 151: 7–19.

124 Calabrese, G., Molzahn, C., and Mayor, T. (2022). Protein interaction networks in neurodegenerative diseases: from physiological function to aggregation. *The Journal of Biological Chemistry* 298: 102062.

125 Poluri, K.M., Gulati, K., Tripathi, D.K. et al. (2023). Protein-protein interactions in host–pathogen interactions. In: *Protein-Protein Interactions: Pathophysiological and Therapeutic Aspects: Volume II* (ed. K.M. Poluri, K. Gulati, D.K. Tripathi, et al.), 207–264. Singapore: Springer Nature Singapore.

126 Zhou, P., Yang, X.L., Wang, X.G. et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579: 270–273.

127 Nurrohman, D.T. and Chiu, N.F. (2024). Unraveling the dynamics of SARS-CoV-2 mutations: insights from surface plasmon resonance biosensor kinetics. *Biosensors (Basel)* 14: 99.

128 Yoon, T.Y. and Lee, H.W. (2019). Shedding light on complexity of protein-protein interactions in cancer. *Current Opinion in Chemical Biology* 53: 75–81.

129 Smith, M.A., Hall, R., Fisher, K. et al. (2015). Annotation of human cancers with EGFR signaling-associated protein complexes using proximity ligation assays. *Science Signaling* 8: ra4.

130 Alzyoud, L., Bryce, R.A., Al Sorkhy, M. et al. (2022). Structure-based assessment and druggability classification of protein-protein interaction sites. *Scientific Reports* 12: 7975.

131 Lu, H., Zhou, Q., He, J. et al. (2020). Recent advances in the development of protein-protein interactions modulators: mechanisms and clinical trials. *Signal Transduction and Targeted Therapy* 5: 213.

132 Wu, D., Li, Y., Zheng, L. et al. (2023). Small molecules targeting protein-protein interactions for cancer therapy. *Acta Pharmaceutica Sinica B* 13: 4060–4088.

133 Eradat, H. (2019). Venetoclax for the treatment of chronic lymphocytic leukemia. *Current Hematologic Malignancy Reports* 14: 469–476.

134 Twomey, J.D. and Zhang, B. (2021). Cancer immunotherapy update: FDA-approved checkpoint inhibitors and companion diagnostics. *The AAPS Journal* 23: 39.

135 Tang, T., Cheng, X., Truong, B. et al. (2021). Molecular basis and therapeutic implications of CD40/CD40L immune checkpoint. *Pharmacology & Therapeutics* 219: 107709.

136 Philippe, G.J.B., Craik, D.J., and Henriques, S.T. (2021). Converting peptides into drugs targeting intracellular protein-protein interactions. *Drug Discovery Today* 26: 1521–1531.

137 Rosa, S., Bertaso, C., Pesaresi, P. et al. (2021). Synthetic protein circuits and devices based on reversible protein-protein interactions: an overview. *Life (Basel)* 11: 1171.

138 Shui, S., Gainza, P., Scheller, L. et al. (2021). A rational blueprint for the design of chemically-controlled protein switches. *Nature Communications* 12: 5754.

139 Moon, T.S., Lou, C., Tamsir, A. et al. (2012). Genetic programs constructed from layered logic gates in single cells. *Nature* 491: 249–253.

140 Gordley, R.M., Williams, R.E., Bashor, C.J. et al. (2016). Engineering dynamical control of cell fate switching using synthetic phospho-regulons. *Proceedings of the National Academy of Sciences of the United States of America* 113: 13528–13533.

141 Giordano-Attianese, G., Gainza, P., Gray-Gaillard, E. et al. (2020). A computationally designed chimeric antigen receptor provides a small-molecule safety switch for T-cell therapy. *Nature Biotechnology* 38: 426–432.

142 Hamley, I.W. (2019). Protein assemblies: nature-inspired and designed nanostructures. *Biomacromolecules* 20: 1829–1848.

143 Hill, L.K., Meleties, M., Katyal, P. et al. (2019). Thermoresponsive protein-engineered coiled-coil hydrogel for sustained small molecule release. *Biomacromolecules* 20: 3340–3351.

144 Nooraei, S., Bahrulolum, H., Hoseini, Z.S. et al. (2021). Virus-like particles: preparation, immunogenicity and their roles as nanovaccines and drug nanocarriers. *Journal of Nanobiotechnology* 19: 59.

145 Kumar, H., Dhalaria, R., Guleria, S. et al. (2024). Immunosensors in food, health, environment, and agriculture: a review. *Environmental Chemistry Letters* 22: 2573–2605.

146 Palakollu, V.N., Veera Manohara Reddy, Y., Shekh, M.I. et al. (2024). Electrochemical immunosensing of tumor markers. *Clinica Chimica Acta* 557: 117882.

147 Pavia, C.S. and Plummer, M.M. (2021). The evolution of rapid antigen detection systems and their application for COVID-19 and other serious respiratory infectious diseases. *Journal of Microbiology, Immunology, and Infection* 54: 776–786.

148 Zhan, K., Chen, L., Li, S. et al. (2024). A novel metal-organic framework based electrochemical immunosensor for the rapid detection of Salmonella typhimurium detection in milk. *Food Chemistry* 444: 138672.

149 Hosu, O., Selvolini, G., and Marrazza, G. (2018). Recent advances of immunosensors for detecting food allergens. *Current Opinion in Electrochemistry* 10: 149–156.

150 Stumpf, M.P., Thorne, T., de Silva, E. et al. (2008). Estimating the size of the human interactome. *Proceedings of the National Academy of Sciences of the United States of America* 105: 6959–6964.

151 Wen, Y., Xie, D., and Liu, Z. (2022). Advances in protein analysis in single live cells: principle, instrumentation and applications. *TrAC Trends in Analytical Chemistry* 152: 116619.