Protein Interaction Network Biology in Neuroscience

Avik Basu, Peter EA Ash, Benjamin Wolozin, and Andrew Emili*

Mapping the intricate networks of cellular proteins in the human brain has the potential to address unsolved questions in molecular neuroscience, including the molecular basis of cognition, synaptic plasticity, long-term potentiation, learning, and memory. Perturbations to the protein–protein interaction networks (PPIN) present in neurons, glia, and other cell-types have been linked to multifactorial neurological disorders. Yet while knowledge of brain PPINs is steadily improving, the complexity and dynamic nature of the heterogeneous central nervous system in normal and disease contexts poses a formidable experimental challenge. In this review, the recent applications of functional proteomics and systems biology approaches to study PPINs central to normal neuronal function, during neurodevelopment, and in neurodegenerative disorders are summarized. How systematic PPIN analysis offers a unique mechanistic framework to explore intra- and inter-cellular functional modules governing neuronal activity and brain function is also discussed. Finally, future technological advancements needed to address outstanding questions facing neuroscience are outlined.

Introduction

At the molecular level, core brain functions involve physical interactions among a diverse array of cell surface and intracellular proteins, which form dynamic spatiotemporal-regulated functional modules underlying and linking the different cell-types of the central nervous system (CNS).¹,² For instance, synaptic transmission depends on transient and stable protein–protein interactions (PPI) among the hundreds of components that form the presynaptic and post-synaptic compartments. Neurotransmission depends on physical associations among a network of scaffold proteins, calcium sensors, and SNARE (soluble N-ethylmaleimide–sensitive factor attachment protein receptor) proteins at the pre-synapse; while multiprotein complexes formed from different scaffold proteins, notably Synapse Associated Protein 90/Post Synaptic Density protein 95-associated proteins, provide a receptive structural and functional framework at the post-synaptic region.³

More generally, PPI mediate or regulate virtually every aspect of neuronal cell behavior, including cell-to-cell communication, neurogenesis, synaptogenesis, and even cell death by autophagy or apoptosis.⁴ It is also increasingly recognized that the coherent cellular organization of these PPINs (PPIN) play a central role in terms of accurate information processing, cognition, memory, and reflex behaviors that govern neural communication. Conversely, PPIN dysregulation has been shown to cause protein aggregation, neuronal stress and dysfunction, and cell death associated with various neurodegenerative diseases (ND)⁵,⁶ and neurodevelopmental disorders (NDD).⁷ Detailed knowledge of this “interactome” architecture in both healthy and diseased brain contexts therefore offers the potential to advance mechanistic understanding of normal neuronal function as well as the causal basis of neurological disorders.

To achieve these goals, experimental methods to study PPIN in the CNS have been devised to advance neuroscience and personalized medicine. Mass spectrometry (MS)-based methods, in particular, have emerged for mapping PPIN in model systems ranging from nematodes, fruit fly and mice, to human cells or tissue. These have led to unbiased surveys of the specialized multi-protein architecture of synapses and sub-synaptic structures, such as the synaptic cleft⁸ and both the excitatory and inhibitory post-synaptic density.⁹ Yet the discovery of pathophysiological relevant PPINs is complicated due to the complex and heterogeneous nature of the human brain.

In this review, we explore state-of-the-art proteomic techniques and computational tools available to neuroscientists interested in PPIN. These approaches complement traditional genetic, genomic, and cell biology strategies, such as genome-wide
association studies (GWAS), imaging, and transcriptomic profiling, to illuminate the molecular mechanisms underlying key brain functions and the multifactorial basis of neurological disorders. We also survey its recent applications to NDs such as Alzheimer’s Disease (AD), to NDDs like Angelman syndrome (AS), and other specific pathological contexts. We conclude by defining major current challenges and promising future directions that should enable a better mechanistic understanding of important neuronal processes and diseases that remain unclear to this day.

Methods to Analyze Multi-Protein Complexes

Historically, molecular profiling studies of the CNS typically involve measuring differences in the relative abundance of mRNAs or proteins, without understanding their underlying physical or functional connectivity. By revealing inter-relationships, the study of PPIN aims to provide a more informative mechanistic context for interpreting standard genetic or proteomic datasets. For example, by overlaying gene expression or protein abundance data on a PPIN as a graphical model, one can illuminate functionally coherent parts of the network that form during development or that are altered in a particular disease state.

A plethora of experimental and allied computational techniques are available to explore the network biology relevant to neuroscience, but we focus on approaches coupling liquid chromatography to mass spectrometry (LC/MS) as a readout (Figure 1). As summarized in Table 1, these methods have respective strengths and weaknesses (e.g., ease of deployment, scale-up, capability, sensitivity versus specificity).[10] What is striking is how the toolkit has evolved to fit the modern research needs of neurobiologists.

Mapping Interactions via Affinity Pulldown

Due to their relatively straightforward application, immunoprecipitation and affinity pull-down techniques are widely used in neuroscience to explore bimolecular interacting partners of select protein targets.[11,12] By combining this concept into a high-throughput (HTP) pipeline, systematic immunoprecipitation coupled to MS (IP-MS) and affinity-tagging/purification coupled to MS (AP-MS) workflows have emerged as the basis of global screening platforms for surveying “interactomes” on a large-scale and relatively unbiased manner (Figure 2A).[13,14] These methods often succeed at identifying stably associated interactors, such as the components of multi-subunit macromolecular complexes. While the power of AP-MS method has been demonstrated for ND/NDD (see below), they can miss potentially important transient interactions, for example at the synapse that mediates neurotransmission. Potential artifacts can also result from ectopic overexpression of proteoforms in a non-physiological context, and also through interactions occurring as a result of the harsh extraction procedure, so that resulting PPIN data does not necessarily accurately reflect the in situ interactome.

Global Biochemical Profiling of PPIN

The burgeoning integration of chemical crosslinking with MS (CX-MS) Figure 2B offers a synergistic technique to map binary
interaction interfaces of potential interest to neurobiologists.[15] The advent of innovative crosslinker chemistries, such as various spacer arm lengths, expanded side-chain reactivities, and MS cleavability (which facilitates data interpretation), provide valuable distance constraints for structural modeling of multi-protein complexes with near-atomic resolution.[16] Yet due to limitations in specificity, efficiency, recovery, and challenges with regards to data analysis, CF/CX-MS are underutilized for large-scale PPI studies in neuroscience.

One emerging approach to study PPIN in a near-native physiological context is the use of biochemical co-fractionation of native protein complexes combined with MS (CF-MS), Figure 2C which allows researchers to probe the endogenous macromolecules present in a complex mammalian tissue such as the brain.[17] Complex protein mixtures can be fractioned by the brain.[17] Complex protein mixtures can be fractioned by macromolecules present in a complex mammalian tissue such as the brain.[17] Complex protein mixtures can be fractioned by macromolecules present in a complex mammalian tissue such as the brain.[17] Complex protein mixtures can be fractioned by macromolecules present in a complex mammalian tissue such as the brain.[17] Complex protein mixtures can be fractioned by macromolecules present in a complex mammalian tissue such as the brain.[17]

### Table 1. Summary of strengths and limitations of experimental methods for mapping PPI.

| Approach | Key Features | Drawbacks | Method Reference | Neuro Study |
|----------|--------------|-----------|-----------------|-------------|
| IP-MS    | - Relatively easy to set up | - Miss transient/weak interactions | [14] | [11] |
|          | - Background subtraction with isotype controls and knock out model possible | - Dependent on good quality antibody for pulldown | | |
| AP-MS    | - Moderately high throughput | - Miss transient or weak interactions | [13,15] | [12,60] |
|          | - Protein tagging easier with commercial ORFs | - High false discovery rate due to background | | |
|          | - Choice of available epitopes, makes study design flexible | - Overexpression of bait might lead to artifact | | |
| Biotin Ligase based proximity assays (BioID and TurboID) | - Can detect transient and weak interactions | - Protein tagging required by genetic manipulation | [29] | [9,38] |
|          | - In vivo application possible | - Tagging with large enzyme can alter the function/localization of bait | | |
|          | - Study context-dependent interactions in model systems in time window (10 min to 24 h) | - Relatively low throughput | | |
|          | - Amicable to CRISPR mediated knock in and/or inducible promoter system for optimal experiment design | - Larger radius resulting in promiscuous labeling | | |
| Peroxidase based proximity ligation (HRP, APEX2) | - Applicable to studying transient interactions (<1 min) | - Overexpression and/or constitutive expression might lead to artifacts | [32] | [8,33] |
|          | - Limited labeling radius ideal for studying direct interactors | | | |
|          | - Compatible with electron microscopy | | | |
|          | - Protein tagging not required | | | |
|          | - Labels limited set of very close interactors | | | |
|          | - Background subtraction with isotype controls possible | | | |
|          | - Spatiotemporal manipulation possible with light activation | | | |
| Micro-mapping | - Not suitable for in vivo applications due to toxicity of peroxidase reagents | - Protein tagging required and large enzyme tag can alter the function/localization of bait | | |
|          | | - Relatively low throughput | | |
|          | | - Not suitable for in vivo applications due to toxicity of peroxidase reagents | | |
| CF-MS    | - Simultaneous global discovery of endogenous assemblies | - Overexpression and/or constitutive expression might lead to artifacts | [19] | [18] |
|          | - Good for identifying native stable complexes | | | |
|          | - Protein tagging or over-expression not required | | | |
| CX-MS    | - High-resolution information suitable for structural inference | | [17] | [16] |
|          | - Range of crosslinkers offers experimental flexibility | - High degree of nonspecific crosslinking needs optimization | | |

Proximity Ligation to Capture Interactions In Vivo

Detecting weak or transient interactions that are induced in a precise spatiotemporal manner is an important challenge. The
advant of protein proximity labeling (PL) techniques that use enzyme fusions to covalently couple biotin to physically-adjacent proteins (within a roughly 10 nm radius in the cellular milieu)[27] has opened an entirely new way of looking at the dynamics of protein associations in a living cell context. PL involves expressing target protein fusions to engineered enzymes such as biotin ligases (BioID[28] and TurboID/mini TurboID[29]) or peroxidases (horse radish peroxidase (HRP),[30] ascorbate peroxidase (APEX)[31]) Figure 3. Interacting proteins, biotinylated in vivo, are enriched either at the protein (e.g., using streptavidin-conjugated beads) or peptide level (e.g., using anti-biotin antibody beads)[32] and subsequently identified by MS.

HRP and APEX/APEX2 rapidly oxidize exogenous biotin-phenol reagent, using H2O2 as a co-substrate, into highly reactive phenoxyl radical (<1 ms half-life) that tags immediate neighboring proteins (Figure 3A). HRP shows much higher catalytic activity in the oxidative environment of the secretory system and at the cell surface and is therefore ideal for applications targeting CNS receptors at the plasma membrane[33] and synaptic cleft.[8] APEX offers rapid, inducible labeling, producing a limited reactive radius better suited to studying dynamic interactions, such as during synaptic transmission or pathogenic infections, but requires treating cells with toxic H2O2 to initiate labeling. The low permeability of biotin-phenol limits cytosolic labeling in both cases.

On the other hand, biotin ligase-based approaches utilize highly permeable biotin and intra-cellular ATP for labeling (Figure 3B). Earlier variants show low activity; BioID, derived from mutant Escherichia coli biotin ligase (BirA),[34] requires lengthy (∼24 h) labeling, TurboID-based enzymes,[29] optimized using directed evolution, have substantially higher activity, requiring ∼10 min labeling, which enables probing dynamic processes with much higher spatiotemporal resolution. In addition, they are active at lower temperatures, enabling PL in organisms such as flies, worms, yeast, and plants. A smaller miniTurbo variant with a lower affinity for biotin allows for less background and selective labeling.[29] Recent developments include split-TurboID and split-APEX, consisting of two inactive fragments that are reconstituted via specific PPI in vivo,[35] are especially useful to minimize artifacts in structure-function, localization, and to map organelle-restricted interactions such as at the synapse.

In general, biotin ligases are best suited for investigating the interaction partners of cytoplasmic-oriented proteins[9,36] whereas peroxidase-based methods allow the study of dynamic events involving membrane proteins like G-protein coupled receptor signaling.[37] To ensure success, it is critical to validate the bait-enzyme fusion is functional and properly localized. To overcome extreme overexpression artifacts, endogenous tagging by CRISPR-Cas9-mediated genetic knock-in, or by adeno-associated virus or lentivirus delivery, is preferred over transient transfection. Although virus-mediated delivery is more popular due to ease of application, CRISPR-Cas9-based endogenous tagging systems are in gaining popularity and have been used both in cell culture[38] as well as in animal models.[39] The user must optimize the amount of starting material (tissue amount, DNA load, cell count, replicates) and time-course to maximize labeling of high-confidence interactors with less background.

A recently described technique, termed Photo-PL (PPL), known colloquially as MicroMapping (Figure 3C), combines the selectivity of antibodies with photocatalyzed reagents to enable labeling of the immediate molecular neighborhood of cell surface targets.[40] In this approach, researchers tag a light-activated photocatalyst to an antibody targeting a specific protein of interest or
Figure 3. Workflows of experimental methods used for detection of PPIN by proximity ligation methods A) APEX, B) BioID, and C) Micromapping. All of the methods involve tagging of neighboring or interacting proteins with biotin and subsequent enrichment by streptavidin pull down.

epitope. After engaging with its target, the photocatalyst tagged antibody captures energy from a suitable blue light source and transfers it to nearby compounds containing a diazirine group. This transforms the diazirine into a highly reactive carbene intermediate which will immediately react with any protein in close vicinity. The very short lifespan (≈5 ns) of the reactive free radicals (e.g., carbene) generated upon photoactivation allows for a tight (≈4 nm) labeling radius, ideal for labeling immediate interactors of membrane proteins of interest. By attaching a biotin to the diazirine containing molecule, researchers can tag and subsequently enrich physically adjacent proteins by streptavidin affinity capture as with other PL methods.

Originally used to study protein assemblies formed at immune-synapses,\cite{40} PPL is well suited for studying PPI occurring in plasma membrane of neurons, such as pre and postsynaptic interfaces, as well as astrocytes and glial cells. The technology is also potentially applicable to study interactions occurring in intact tissue or organoids or within intra-cellular locations. Whether used alone, or in combination with pre-existing techniques, PPL provides a promising complementary platform to study the microenvironments of neuronal proteins of specific interest with high spatiotemporal resolution.

**Data Analysis and Complementarity of Existing Methods**

The analysis of interactome data often requires specialized computational tools. Software packages like SAINT\cite{41} and ComPASS\cite{42} are publicly available tools to score and rank protein complex pulldown experiments based on data reproducibility, variance and enrichment relative to controls, to define more selective or probable associations. Ratiometric criteria,\cite{43} wherein the signal of each putative interactor is quantified relative to control samples, can establish specificity and eliminate non-specific associations. Receiver operating characteristic-based analyses, which benchmark results against a curated list of known interactions, are useful for establishing optimal cut-off scores that minimize the retention of false positives.

The lack of ideal controls for many situations, however, prompts the need for improved methods. For example, the results obtained for a PL study involving a tagged bait are usually compared against background labeling by free PL enzyme, which may show diffuse localization. Isotype controls can mitigate the loss of spatiotemporal resolution, while multivariate analyses using linear models can accommodate more complex study designs. Yet even with stringent analysis, the interactomes recovered by different experimental strategies do not always overlap for any given bait, reflecting intrinsic biases of each approach.

The relative pros and cons of different approaches are critical factors to consider when addressing a specific scientific question (Table 1). While IP-MS or AP-MS experiments are easier to set up, they can miss crucial transient interactions. Fusion protein in functional, peroxidase-based PL methods like APEX2 is perhaps best suited for PPI studies in cell culture (due to fast kinetics and labeling), while biotin ligase-based methods are preferred for in vivo experiments, as they do not require cytotoxic reagents.
Table 2. List of commonly used databases related to neuroscience and with relevance to neuroproteomics.

| Database         | Database URL                          | Features                                                                 |
|------------------|---------------------------------------|--------------------------------------------------------------------------|
| Allen Brain Map  | portal.brain-map.org/                 | Comprehensive gene expression datasets of various cells and tissue types from mouse and human brain. |
| BioGRID          | thebiogrid.org/                        | Archives and disseminates genetic, PPI, and PTM data from model organisms and humans. |
| BIOPLEX          | wren.hms.harvard.edu/bioplex/          | High-throughput AP-MS based human PPI data from 293T, HCT116 cell lines. |
| BrainMap         | www.bu.edu/dbin/cnsb/mousebrain        | CF-MS based data on protein complexes in adult mouse brain.              |
| CORUM            | mips.helmholtz-muenchen.de/corum/      | Resource of manually annotated protein complexes from mammals; includes complex function, localization, subunit composition, etc. |
| GeneCards        | www.genecards.org                      | Integrative database that provides comprehensive, user-friendly information on all annotated and predicted human genes. |
| HIPPE            | cbdm-01.zdv.uni-mainz.de/~mschaefer/hippie/ | Web tool to generate human PPIN with an integrated probability-based scoring system. |
| human base       | hb.flatironinstitute.org/              | Tissue-specific interaction, data-driven predictions of gene expression, function, regulation, and interactions in human. |
| Human Protein Atlas | www.proteinatlas.org                  | Useful information about protein localization and comparison about RNA expressions in cells and tissues. |
| Integrated Interaction Database | iid.ophid.utoronto.ca/ | Provide networks that are specific to tissues, sub-cellular localizations, diseases, and druggable proteins across 18 model species. |
| IntAct           | www.ebi.ac.uk/intact                   | Open source database and analysis tools for molecular interaction data, derived from literature curation or direct user submissions. |
| iRefIndex        | irefindex.vib.be/wiki/index.php/iRefIndex | Provides an index of protein interactions available in a number of primary interaction databases. |
| STRING           | string-db.org/                         | Database of known and predicted PPIs, include direct (physical) and indirect (functional) associations from computational prediction, and from interactions from primary databases. |
| SynaptosomeAtlas | synaptome.genes2cognition.org          | Database and visualization tool for different synapse types and subtypes of whole mouse brain. |
| SYNGO            | syngoportal.org                        | Knowledgebase for synapse related proteins, their function, and interactions |
| TissueNET        | netbio.bgu.ac.il/tissuenet/            | Provides quantitative tissue associations for human PPIs.                 |
| UniProt          | www.uniprot.org/                       | Comprehensive resource for protein sequence and annotation data; hub for the collection of functional information on proteins. |

Like H$_2$O$_2$, to function. TurboID is advantageous over BioID by having much faster labeling with higher yields but sometimes exhibits background activity even without exogenous biotin which can complicate data analysis. Due to its lower background labeling rate, miniTurbo is better suited for capturing transient interactions in time-course experiments. Assuming access to suitable antibodies, PPI is a method of choice for researchers interested in characterizing membrane protein assemblies. CF-MS and CX-MS provide topological information but require more technical expertise and resources. In summary, researchers should choose the most relevant model and approach suited to the question(s) at hand, and preferably cross-validate findings using complementary methods in follow-up experiments.

**PPIN as a Framework to Navigate the Human Brain Proteome**

The human brain is an exceptionally complex structure, reflecting specialized spatial, structural, and cell-type organization across different cells and anatomical regions. Knowledge of the molecular properties of these diverse components can provide mechanistic insights into their functional connectivity in normal and pathogenic conditions. Documenting changing interactions during embryonic, juvenile, and adult stages can reveal systems supporting normal brain development or dysfunction.

Public curation databases such as BioGRID,[45] IntAct,[46] STRING,[47] and BioPlex[13] act as repositories for PPI information. Tissue- and brain-specific PPI have been reported in specialized databases, such as TissueNet,[48] IID,[49] and Brain Map.[17] Commonly used PPIN databases along with important specialized resources are listed in Table 2. Analysis of data in these sources reveals that, while far from complete, existing brain PPIN display characteristic properties. Network topology and modularity, that reflect the centrality of certain neuronal proteins as highly connected nodes, form interconnected clusters that underlie emergent systems level behaviors such as neurotransmission.[42] PPIN maps therefore inform on sets of proteins potentially associated with a particular process or clinical disorder.

In the systems biology domain, PPIN are studied by means of graphical representations, with proteins referred to as nodes, and their physical (functional) associations as edges (Figure 4A–D). Unlike signal transduction pathways, such as protein kinase phosphorylation cascades, PPIN are undirected and often follow a scale-free (i.e., power-law) distribution, in that, only some...
Figure 4. Protein interaction networks as a framework to study the molecular basis of neurological processes and diseases. PPINs are often visualized as 2D graphs, wherein each protein is shown as a node and the association between proteins is represented with lines or edges; highly interconnected sets of proteins, or clusters, often represent functionally as coherent modules such as multi-protein complexes or biochemical pathways. As illustrated, this connectivity can become perturbed in NDs, NDDs, and during viral infection (A–D), reflecting alterations in the physical organization and functional properties of modules normally found in healthy brain (E–H), leading to phenotypic changes and impaired brain functions that become more severe during disease progression (I–L). Protein complexes enriched for various synaptic functions are shown (adapted from[17]) and their potential links to NDs, NDDs, and ZIKV infection are highlighted (M).
nodes are highly connected. One outcome of analyzing PPIN as graphs is identifying key nodes that have high connectivity often referred to as hubs,[30] which serve as major conduits whose functional integrity is more central to cellular function. Computational analysis of primary (nearest) neighbors and secondary (indirect) interactions can therefore reveal unexpected relationships, suggest the roles of unannotated components, and reveal subsystems modulated by pathobiological stimuli.[17]

**PPIN Alterations Underlie Disease, including Neuropathological Disorders**

Knowledge of PPI can inform on the molecular basis of diseases of the CNS (Figure 4E–M). Dysregulation of PPI in particular biochemical pathways can disrupt signaling cascades, rewire protein homeostasis pathways, and cause accumulation of misfolded protein aggregates that ultimately leads to pathological states. Thus, identification of disease linked PPIN can provide mechanistic insights plus potential therapeutic targets. For example, NDs such as AD, Huntington’s disease (HD), and Parkinson’s disease (PD) are characterized by aggregate formation in different parts of brain[51] (Figure 4G.K). By changing patterns of translation and post-translational modification (PTM),[52] these aggregates not only sequester diverse types of proteins, including RNA binding proteins (RBP), they perturb cell function via alterations in the global protein networks of afflicted cells and tissue. Interestingly, many of these aggregation-prone proteins, such as α-synuclein, a hallmark protein aggregate for PD, also accumulate as aggregates in other NDs such as HD and Diffuse Lewy Body disease.[31]

The genetic and clinical heterogeneity of human specimens is compounded by technical difficulties in isolating intact endogenous protein assemblies from post-mortem samples prior to analysis. One solution has been to perform systematic PPIN analysis using a combination of cell- and animal-based models where access to material is less constrained. These include induced pluripotent stem cell-derived neurons and 3D organoid cultures. These models are increasingly being adopted to study the pathology of NDs/NDDs in specific cell-type contexts, as exemplified for developmental stage-specific protein dynamics in cortical neurons and neural progenitors.[53]

**Case study - AD**

Identifying PPIN impacted by AD can reveal mechanistic basis of this most common form of dementia, which affects ≥1 in 10 people over 65 and predicted to have nearly 50 million cases worldwide.[61] While AD is ultimately fatal, manifestation of clinical symptoms occurs years, even decades, after initial development of brain pathology,[62] motivating the search for early molecular markers, diagnostics enabling precision medicine, and targets for novel therapeutics. The accumulation of extracellular neuritic plaques of insoluble amyloid-β (Aβ), as well as vascular dysfunction, frequently precede symptom onset by 20 years or more; this pathology is followed by activation and hypertrophy of microglia, neuro-inflammation, and the accumulation of tau pathology, all of which occur 5–10 years ahead of the clinical phase of AD.[63] These resulting changes correlate with decreases in synaptic fields and loss of network integrity.[63]

Pathways regulating Aβ (formation, aggregation, clearance), tau (PTM, aggregation, and catabolism), and inflammation impact AD progression.[64] The pathological diagnosis of AD requires the presence of neuritic plaques and neurofibrillary tangles (NFT).[63] Pathological, genetic, imaging, and biochemical studies all provide strong support for an essential role of Aβ accumulation in pathogenesis of AD.[63–65] However, the accumulation of microtubule-associated protein tau into NFT and neuprol
threads within neocortical neurons (≥5 years before symptoms) correlates best with cognitive loss, reflecting the inherent correlation between neuronal function and cognitive abilities.\(^{[63,64]}\)

To better understand the complexity of AD, large-scale initiatives such as Accelerating Medicines Partnership-AD and synapse.org have compiled high dimensional multi-omics datasets across >2,000 individuals, providing a landscape of proteomic, transcriptomic, epigenomic, and metabolomic profiles of afflicted cell-types and tissues. Functional networks consistently altered during AD pathogenesis in prodromal versus late-stage AD have been described.\(^{[6,66,67]}\) For example, using qualitative LC-MS to compare dorsolateral prefrontal cortex and precuneus samples from AD and control individuals, Seyfried’s group identified AD-affected PPINs,\(^{[67]}\) including several linked to loss of synaptic plasticity and neurons before onset of disease symptoms. Similarly, glial cell protein networks were perturbed in symptomatic cases, highlighting that synaptic dysfunction may result from dysfunctional glia-mediated homeostatic control. Other studies provide complementary sets of functional modules (correlational networks) that also track with pathology.\(^{[9,10]}\)

RBPs that function in RNA metabolism, the translational stress response, and spliceosome machineries co-aggregate with insoluble tau in AD-affected brains.\(^{[68,69]}\) As tau undergoes pathological changes, its PPIN dramatically shifts; dissociating from microtubules and recruiting ribosomes, RBPs, and heat shock chaperones,\(^{[69,70]}\) with critical consequences. Tau association with the ribosome impairs its function leading to repression of translation, including of key synaptic proteins such as PSD-95.\(^{[70]}\) The interaction with TIA1 (an RBP that nucleates stress granules in neurons) facilitates conversion and propagation of tau into toxic oligomers, which correlate with neuropathology and cognitive performance in transgenic mice.\(^{[52,71]}\)

To evaluate proteome modules modulated by AD, Xu et al., generated PPIN for six distinct regions of AD versus control regions of post-mortem brain (e.g., hippocampus),\(^{[72]}\) and identified changes in components linked to activation of immune signaling, metabolism, and cell cycle control leading to neuroinflammation and apoptosis in AD. Intriguingly, certain PPIs were altered in the prodromal phase, but the modest sample size (n = 9) limited study power. This constraint is common to proteomic studies, due to throughput limitations, and is relevant to other studies examining PPIN in other neurodegenerative conditions.\(^{[70]}\)

Innovative strategies for extrapolating experimental data with clinical metrics, in combination with CF-MS, PL, Micromapping, and powerful genetic tools, such as CRISPR-Cas9 and Optogenetics, to study synaptic dysfunction at molecular level are needed to better understand the complex etiology of AD.

**Integration of Genomic, Transcriptomic, and Proteomic Data using PPIN to Understand Multifactorial Neuronal Diseases**

Researchers have traditionally tried to solve the causal basis of complex neurological disorders using genomics approaches, like GWAS. Genomic resources related to adult brain are available,\(^{[73]}\) but knowledge about genetic variants associated with specific disorders is not sufficient to elucidate the underlying mechanisms. In most instances, the function of genetic polymorphisms or mutated gene products and their causal association with disease is poorly understood. This caveat is particularly relevant to late onset disease, where the effects of particularly polymorphisms are pleiotropic (such as ApoE) or weak (such as BIN1 or PICALM).\(^{[74]}\)

Transcriptomics (e.g., bulk and single-cell RNA sequencing) can provide insights, but its extrapolation to make protein level inferences is problematic due to low overall correlations between mRNA and protein levels,\(^{[75]}\) reflecting differences in relative synthesis dynamics, PTM, aggregation, and turnover.\(^{[76]}\) This is particularly salient in studies of the human brain because the exquisite spatiotemporal regulation of mRNA processing, transport, and localized translation in neurons is essential to synaptic plasticity, cognitive learning, memory, and behavior.\(^{[77]}\) Measuring transcripts is not sufficient to identify the functional state of multiprotein assemblies that support core neuronal functions, particularly those that depend on proper localization at the axon growth cone, dendritic spines, and pre-/post-synaptic compartments.\(^{[78]}\) Comparative analyses of affected hippocampus of AD patients showed that biochemical modules linked to impaired microtubule dysfunction, inflammation, and impaired nucleic-acid binding are preferentially detected by proteomics rather than transcriptomic profiling.\(^{[67]}\) Thus, a multi-pronged integrative strategy is essential to reveal the basis of multifactorial disorders. For instance, integration of PPIN with transcriptomic datasets can provide a more holistic predictive model of phenotypic responses in NDs.\(^{[79]}\)

In contrast to genomics and transcriptomics, PPI studies have not been commoditized, and single-cell analyses have not been reported on a large-scale. In their tool, SCINET, Mohammadi et al propose a statistical method to infer the presence of global protein interactions within sequencing-derived cell populations based on the transcriptional activation of the genes participating in the protein interactions.\(^{[80]}\) Their method is reliant on the presence of a reference global protein interaction network along with relevant single-cell RNA-Seq data. Meanwhile, Klimm et al proposed an approach to find functionally-active modules in sequencing-derived cell populations by using a reference protein interactome.\(^{[81]}\) They use the differential transcriptional gene profiles to infer the most likely protein modules present in cell types that are inferred from single-cell RNA-Seq data. These approaches show the promise of integrating single-cell sequencing data with protein networks to gain insight into single-cell biology.

Knowledge of the landscape of brain-related proteome alterations, in particular PTMs such as phosphorylation, methylation, acetylation, glycosylation, and ubiquitination, that impinge on neuronal pathways can also highlight processes disrupted by pathology.\(^{[82]}\) Phosphorylation of synaptic proteins, such as vesicle trafficking factors, as well as epigenetic marks on chromatin, such as acetylation, methylation, and ubiquitination of histone proteins, have been shown to modulate core brain functions such as memory, learning, and other cognitive and behavioral processes.\(^{[83]}\) These marks are mediated by multi-protein complexes, such as writers and erasers of chromatin-associated PTMs, whose function is normally tightly regulated.

In summary, whereas mutation mapping and gene expression studies alone cannot ascertain the functional circuits altered during pathogenesis, PPIN information can provide a valuable vantage to interpret genomic analyses.
PPIN Alterations During Host-Pathogen Interactions

Proteins and peptides from pathogens are known to alter, disrupt, and hijack host cell PPIN across kingdoms. Here we present an illustrative case of viral-host protein dynamic interactions focusing on emerging Zika virus (ZIKV) infection that primarily affects the brain. ZIKV replicates in adult human brain tissue and impairs synapses and memory in mice. ZIKV seems to have neurotropism as it is vertically transmitted from pregnant women to the developing fetus causing microcephaly and possibly Guillain-Barré syndrome, a neuro-inflammatory disease of the peripheral nervous system that causes cholinergic dysfunction and muscular weakness.

Like other Flaviviridae, such as Dengue virus (DENV), ZIKV is mosquito-borne. ZIKV binds with host cells via receptors AXL and Tyro3 and inserts its single-stranded genomic RNA. This RNA encodes a single polyprotein that is processed into three structural proteins (envelope, membrane, capsid) and seven non-structural effector proteins, which perform key replicative functions. ZIKV infection causes neuronal cell death that ultimately leads to brain malfunction and spinal cord dysfunction. ZIKV hijacks the biochemical networks of the cellular secretory pathway and organelles for virion assembly, maturation, and release. Understanding how host machinery is commandeered by ZIKV requires a mechanistic re-wiring of global PPIN rewiring of host cells during infection.

Recently a detailed proteomics study revealed how ZIKV infection of human neural stem cells perturbs fetal brain development. Cultured neurons were converted into neurospheres prior to LC/MS analysis, revealing altered expression of ∼500 proteins during infection. Pathways affected by ZIKV centered on RBP networks related to RNA processing and splicing, microRNA biogenesis, and DNA damage, repair, and chromosomal instability.

An AP-MS study identified 386 putative ZIKV-protein interacting human host factors, several linked to neuronal development, retinal defects, and infertility. Subsequent phosphoproteomic analysis of ZIKV-infected human neuroblastoma cells showed dysregulation of 1216 sites on proteins associated with brain development, cell cycle, and cell organization. These include p38 MAPK, MARCKS, and DPYSL2, which regulate neurite outgrowth and brain development, as well as the DNA damage checkpoint regulators and their substrates.

Using BioID, Coyaud et al. reported 1224 cellular neighbors of 10 ZIKV polypeptides that localize to diverse host compartments, including the centrosome, lysosome, and peroxisome, consistent with viral-mediated disruption of the centriolar satellites, lysosome-endosome fusion, and reduction of lipid homeostasis, respectively. A recent study from the Krogan group used AP-MS to demonstrate ZIKV and DENV exploit human and mosquito cell machineries to complete replication. Using AP-MS, comparative analysis identified 28 proteins significantly overlapping between the PPIN of DENV and ZIKV. These include a conserved interaction of viral NS5 with the PAF1C transcriptional elongation complex, inhibiting expression of interferon-stimulated genes. The study reported ZIKV NS4A promotes microcephaly by engaging human ANKLE2, which is linked to autosomal recessive microcephaly.

While these studies demonstrate viral perturbations of host PPIN, important details are lacking. Research conducted in immortalized cell lines cannot recapitulate viral invasion of brain. Viral impact during the prodromal phase is unclear, as is the mechanism of host immunity evasion during vertical transmission. Comparative time-course PPIN studies of suitable animal models could address these questions.

Current Challenges in Network Neuroscience and Future Directions

Despite remarkable progress in technical and analytical tools for systematic PPIN mapping, considerable challenges need to be addressed to complete interactome mapping of the CNS especially with respect to neurodevelopment and neurodegeneration. Some obstacles are inherent to any PPIN study, while some are unique to neuroscience.

Experimental approaches identify a wealth of information but produce significant false positives and negatives results. CF-MS based approaches rely on stringent computational techniques such as machine learning to produce reliable interactions with sensitivity and precision. Alternate methods are needed to confirm putative interactions. Several criteria such as screen completeness, assay sensitivity, sample consistency, and confirmation with other methods can estimate and reduce error rates. Putative interactions that can be confirmed by multiple techniques or algorithms are more likely to be valid. Ideally, a combination of experimental and statistical criteria followed by benchmarking and ideally independent testing can be adopted and scaled up to verify novel associations.

Developing effective methods that minimize bias (false negatives) is another challenge. Although the “dark matter” missed by existing methods is unknown, integration of complementary techniques and optimizing experimental design can markedly improve coverage. We encourage the use of global (e.g., CF-MS) and targeted (e.g., PI) to address particular neurobiology research questions.

Most assays provide qualitative, rather than quantitative, descriptions of PPIs. Determining the extent or affinity of interactions is a slow, tedious process. The biophysics of PPI are largely studied one partner at a time. For example, surface plasmon resonance and isothermal calorimetry can only be performed using small numbers of purified proteins. Hence, there is a burgeoning need for HTP techniques to study native PPI in a quantitative manner. Ideally, quantitative interactome methods could be consistently applied to diverse cell-types, tissues, and models, while preserving native subcellular contexts.

Integration of genomic and transcriptomic data with PPIN represents an early opportunity, but the heterogeneous nature of biological systems confounds amalgamation. Many studies assume co-expressed genes predict protein interactions, but the poor correlation between transcripts and proteins levels mitigates such predictions. The brain PPI network is likely far more complex than simple in vitro model systems due to cellular heterogeneity and the complexity of neural networks. The CNS is formed from different types of astrocytes, glial cells, and neurons linked by specialized synaptic interfaces that are impacted differentially in NDD and ND in different brain regions. Current studies provide a static snapshot of this ensemble, yet even within a single neuron, PPIN can vary significantly between
compartment due to protein trafficking and turnover.\[94\] Another drawback is insufficient coverage of key transitions during disease progression, such as PTM. Functionally significant PPIN may exist only transiently in specific pathological stages and so missed in post-mortem samples. Collecting longitudinal PPI causally related to clinical outcomes is of the utmost value.

Single-cell sequencing technologies have enabled researchers to study the transcriptional state of biological systems at the single-cell level, leading to applications across a number of diverse applications such as studying development and tissue heterogeneity. While single-cell transcriptomics is gaining tremendous popularity now, single-cell proteomics is still in its nascent stage.\[95\] Systematic study of proteome and PPIN in a cell type specific/selective manner would certainly help us understand basic brain function. Integrating single-cell sequencing data with protein interaction networks is an active research challenge that presents an opportunity to gain insight into the hierarchical workings of the protein interactome.

Similarly, spatial proteomics should be integrated with PPIN at increasing resolution. Spatial information on PPIN can be gained by reducing the complexity of the sample using innovative subcellular omics as described.\[96\] Alternatively, organspecific PL coupled with imaging can provide valuable clues in this regard.\[29,30\] Local application of diverse PP detection methods can provide details about PPIN in various brain regions.

Improvements in metabolomics and lipidomics techniques have evolved rapidly to enable researchers to gain important biological insights on the disease relevance of these important, chemically diverse molecules in neurons.\[97\] Additionally, there is growing excitement around the feasibility of routine combined protein–metabolite profiling, in recognition of the orthogonal insights gained from profiling diverse biomolecular types.\[98\] However, while interactions between lipids and proteins have long been recognized as functionally important, they are difficult to map globally due to the chemical properties of lipids and existing labor-intensive techniques. The development of high-throughput lipid–protein interaction mapping techniques is beginning to enable researchers to interrogate these molecular interactions in an increasingly global capacity,\[99\] though routine profiling in parallel with other high-resolution techniques is not yet routine. In addition to the technical data acquisition side, advances are needed to better interpret lipidomic data in conjunction with other data types, where interactions can occur with varying degrees of specificity and multiple, related lipid species may have overlapping functional roles. Due to the critical, functional role lipids play in neurons, in particular, it is likely advancements in profiling and interpreting lipid–protein interactions may play an especially important role in understanding the molecular basis of neurobiological processes.

Conclusion

As understanding of the connectivity of healthy and diseased brain circuits evolves, it is apparent that aberrant accumulation, trafficking, misfolding, and aggregation impacts cellular PPI which ultimately impairs neuronal function. The usual trajectory for PPIN analysis is obtaining high-confidence data, analyzing the results using systems biology approaches, validating prioritized interaction via alternate approaches, and potentially discovering new therapeutics targets. For greater impact in neuroscience, it is crucial to identify key functional modules altered during development or disease progression, understand their underlying molecular relationships, and ascertain markers that inform on pathogenesis. As the field of neuroproteomics expands, it is important to remember that rigorous study design and statistical criteria are equally important as well-defined questions. With steady technical advancements, PPIN studies are poised to provide transformative insights into neuronal development and neuropathology, ushering in a golden era of neurobiology.

Acknowledgements

The authors sincerely thank Sadhna Phanse, Benjamin Blum, and Yusuf Ahmed for their valuable inputs, and members of the Emili and Wolozin groups for helpful discussion and feedback. This work was supported in part by funds from the U.S. NIH (RF1AG056318, the BrightFocus Foundation and the BU Kilachand Award to B.W. and 1R01AC061706, 1R01AG064932, RO1AG061705 to B.W. and A.E).

Conflict of Interest

B.W. is the chief scientific officer of Aquinnah Pharmaceuticals. The remaining authors declare no conflict of interest.

Keywords

Alzheimer’s disease, mass-spectrometry, neuroscience, protein–protein interaction network, systems biology

Received: August 7, 2020
Published online: December 29, 2020
[88] P. P. Garcez, J. M. Nascimento, J. M. de Vasconcelos, R. Madeiro da Costa, R. Delvecchio, P. Trindade, E. C. Loiola, L. M. Higa, J. S. Cassoli, G. Vitoria, P. C. Sequeira, J. Sochacki, R. S. Aguiar, H. T. Fuzii, A. M. de Filippis, J. L. da Silva Goncalves Vianez Junior, A. Tanuri, D. Martins-de-Souza, S. K. Rehen, Sci. Rep. 2017, 7, 40780.

[89] P. Scaturro, A. Stukalov, D. A. Haas, M. Cortese, K. Draganova, A. Pichlmair, Nature 2018, 561, 253.

[90] E. Coyaud, C. Ranadheera, D. Cheng, J. Goncalves, B. J. A. Dyakov, E. M. N. Laurent, L. Pelletier, A. C. Gingras, J. H. Brumell, P. K. Kim, D. Safronetz, B. Raught, Mol. Cell. Proteomics 2018, 17, 2242.

[91] P. S. Shah, N. Link, C. M. Jang, P. P. Sharp, T. Zhu, D. L. Swaney, J. R. Johnson, J. Von Dollen, H. R. Ramage, L. Satkamp, B. Newton, R. Huttenhain, M. J. Petit, T. Baum, A. Everitt, O. Laufman, M. Tassetto, M. Shales, E. Stevenson, G. N. Iglesias, L. Shokat, S. Tripathi, V. Balasubramaniam, L. G. Webb, S. Aguirre, A. J. Willsey, A. Garcia-Sastre, K. S. Pollard, S. Cherry, A. V. Gamarnik, I. Marazzi, J. Taunton, A. Fernandez-Sesma, H. J. Bellen, R. Andino, N. J. Krogan, Cell 2018, 175, 1931.

[92] K. Karagoz, K. Y. Arga, Comput. Biol. Chem. 2013, 45, 1.

[93] H. N. Chua, L. Wong, Drug Discovery Today 2008, 13, 652.

[94] H. Suzuki, J. Physiol. 2006, 575, 373.

[95] S. Heo, G. H. Diering, C. H. Na, R. S. Nirujogi, J. L. Bachman, A. Pandey, R. L. Huganir, Proc. Natl. Acad. Sci. U. S. A. 2018, 115, E3827.

[96] N. Slavov, Curr. Opin. Chem. Biol. 2020, 60, 1.

[97] (a) A. K. Tharkeshwar, K. Gevaert, W. Annaert, Proteomics 2018, 18, 1700113; (b) A. K. Tharkeshwar, J. Trekker, W. Vermeire, J. Pauwels, R. Sannerud, D. A. Priestman, D. Te Vruchte, K. Vints, P. Baatsen, J. P. Decuyper, H. Lu, S. Martin, P. Vangheluwe, J. V. Swinnen, L. Lagae, F. Impens, F. M. Platt, K. Gevaert, W. Annaert, SciRep 2017, 7, 41408.

[98] P. L. Wood, Schizopr Res 2019, 212, 107.

[99] B. C. Blum, F. Mousavi, A. Emili, Mol. Omics 2018, 14, 307.

[100] A. E. Saliba, I. Vonkova, S. Deghou, S. Chesca, C. Tischer, K. G. Kugler, P. Bork, J. Ellenberg, A. C. Gavin, Nat. Protoc. 2016, 11, 1021.

Avik Basu is a research scientist at Center for Network Systems Biology, Boston University. His previous works include application of various proteomics and protein interaction techniques in erythrocyte biology. His current research focuses on application of various proximity labeling techniques in neuroscience especially for studying neurodevelopmental as well as neurodegenerative disorders. Avik’s future research interests lie on combining proteomics, posttranslational modifications information and protein interaction network data integrated from various technologies, to better understand the neuronal circuits and brain function in health and diseases.

Peter Ash is a research assistant professor at Boston University. His previous works include discovery of a novel neuropathology consisting of dipeptide repeats that develop in individuals with ALS and FTD with expansion mutations in C9ORF72, and characterization of macromolecular complexes in the adult murine brain including an ALS-relevant assembly of RNA binding proteins that have antagonistic roles in modulating RNA splicing. His future research interests lie in understanding the composition and dysfunction of RNA binding protein complexes and resulting RNA abnormalities in neurodegenerative disease.
Benjamin Wolozin, M.D., Ph.D., is a professor of pharmacology and neurology, as well as a member of the Center of Systems Neuroscience and the Center of Neurophotonics at Boston University. His research focuses on the pathophysiology of neurodegenerative disease. He has made numerous discoveries over his career, including generating the first conformationally sensitive antibody (Alz-50), identifying the potential benefit of statins for subjects at risk for Alzheimer’s disease, identifying the role for stress granules in the pathophysiology of amyotrophic lateral sclerosis and discovering the role of tau in regulating the translational stress response. He is a fellow of the American Association for the Advancement of Science.

Andrew Emili is a professor in biochemistry and biology at Boston University (BU) and Founding Director of the Center for Network Systems Biology (CNSB). Prior to 2017, he was a founding member and Principal Investigator for 18 years at the Donnelly Center for Cellular and Biomolecular Research at University of Toronto. He is an internationally recognized leader in functional proteomics, integrative systems biology, and precision mass spectrometry. His group develops and applies innovative technologies to map protein interaction networks and macromolecular complexes of cells and tissues on a global-scale, publishing “interactome” maps of unprecedented quality, scope, and resolution.