Molecular and Cellular Neuroscience xxx (2013) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Molecular and Cellular Neuroscience

journal homepage: www.elsevier.com/locate/ymcne

Predicting protein–protein interactions in the post synaptic density

Ossnat Bar-shira *, Gal Chechik

The Gonda Brain Research Center, Bar-Ilan University, Ramat Gan 52900, Israel

ARTICLE INFO

Article history:
Received 25 January 2012
Revised 9 April 2013
Accepted 19 April 2013
Available online xxxx

Keywords:
Classification
Computational biology
Network reconstruction
Postsynaptic density
Protein–protein interactions

ABSTRACT

The post synaptic density (PSD) is a specialization of the cytoskeleton at the synaptic junction, composed of hundreds of different proteins. Characterizing the protein components of the PSD and their interactions can help elucidate the mechanism of long-term changes in synaptic plasticity, which underlie learning and memory. Unfortunately, our knowledge of the proteome and interactome of the PSD is still partial and noisy. In this study we describe a computational framework to improve the reconstruction of the PSD network. The approach is based on learning the characteristics of PSD protein interactions from a set of trusted interactions, expanding this set with data collected from large scale repositories, and then predicting novel interaction with proteins that are suspected to reside in the PSD. Using this method we obtained thirty predicted interactions, with more than half of which having supporting evidence in the literature. We discuss in details two of these new interactions, *Lrrtm1* with *PSD-95* and *Src* with *Capg*. The first may take part in a mechanism underlying glutamatergic dysfunction in schizophrenia. The second suggests an alternative mechanism to regulate dendritic spines maturation.

Introduction

Understanding the structure and function of mammalian glutamatergic synapses has been a major focus of molecular neuroscience. Particular attention has been given to the postsynaptic density (PSD), a dense complex of proteins whose function is to detect and respond to neurotransmitter that is released from pre-synaptic terminals.

Broadly speaking, one set of these proteins is docked to the cell membrane, forming the "front end" of the post synaptic glutamate signaling cascades. These include AMPA (α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) and NMDA (N-methyl-D-aspartic acid) receptors, which convert the chemical signals from presynaptic terminal to electrical signal by allowing an influx of positive ions into the cell (Traynelis et al., 2010). Other proteins serve to constantly modulate these signals through several complex mechanisms. First, the flux of ions allowed into the cell is regulated by tuning the distribution and density of receptors or their subunit composition (*synaptic plasticity*) (Malenka and Bear, 2004; Matta et al., 2011; Sheng and Jong Kim, 2002). Second, the impact of the ion influx is regulated by tuning the morphology of dendritic spines (*structural plasticity*) (Bourne and Harris, 2008). Proteins within the PSD play a crucial ole in this regulation. For instance, the kinetics of NMDA receptors can be altered by *Src* tyrosine kinase (Ali et al., 2001), AMPAR trafficking is regulated in part by *Pick1* and *Grip1* (Kulangara et al., 2007; Volk et al., 2010), and changes in spine shape and size are mediated by proteins like *Cortactin*, *Actin*, *Src*, *Capg*, *Shank* and *Homer* (Fan et al., 2011; Huang et al., 1997; Sala et al., 2001).

These plasticity mechanisms tune the transmission of signals through the synapse using a carefully-orchestrated web of protein interactions. Mapping protein–protein interactions is necessary to gain insight into protein function (Legrain, Wojcik, and Gauthier, 2001), detect molecular pathways (Segal, Wang, and Koller, 2003), or identify potential drug targets (Archakov et al., 2003; Hormozdiari et al., 2010). Charting the protein–protein interactions is particularly important in context of the PSD, where modifications in protein conformation and interactions have been linked to neuropsychiatric and neurodegenerative disorders. Known examples include the association between autism and mutations in PSD proteins such as CNTNs, NRXNs or *Shank3* (Bourgeron 2009), and between hypofunction of NMDAR receptors and schizophrenia (Kristiansen et al., 2007; Stephan et al., 2006).

A solid knowledgebase of the PSD proteome and interactome has a potential to lead to new targets for treating such disorders. In schizophrenia for example, a first approach to enhance NMDA receptor's activity is to target the NMDAR itself, and multiple compounds were proposed for modulating NMDA receptor activity. Unfortunately, despite continuous efforts, most NMDAR-targeting drugs were found ineffective or induce severe side effects (Kalia et al., 2008). Recovering the pathways that control NMDA receptors can provide a new perspective into this problem. NMDAR is regulated by a collection of PSD kinases, phosphatases, and other molecules through multiple regulatory pathways (Hunt and Castillo, 2012), In this paper, our computationally inferred interactions suggest a pathway by which *PSD-95*, *Lrrtm1* and *Neurexin* decrease *Src* induced NMDAR tyrosine phosphorylation. Such

* Corresponding author. Fax: +972 3 535 2185.
E-mail address: ossnat.barshira@gmail.com (O. Bar-shira).

1044-7431/$ – see front matter © 2013 Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.mcn.2013.04.004

pathways offer a potential explanation for NMDAR diminished activity, and suggest a set of proteins and their domains of interaction to be examined as possible drug targets.

Obtaining the full set of protein interactions in the PSD can also improve our understanding of the interplay between activity (synaptic efficacy) and structure (spine morphology). Dendritic spines change their size and shape in correlation with synaptic activity (Kasai et al., 2003), a transformation that involves changes in actin filaments length and organization (Hotulainen and Hoogenraad, 2010). Actin cytoskeleton regulation was already shown to involve multiple proteins that reside in the PSD, including NMDAR, CaMKII, and GTPases (Rho, Ras), and identifying the full list of protein interactions could reveal additional molecular mechanism connecting synaptic efficacy to structural plasticity.

Despite the significant effort to identify the full list of proteins of the PSD and their interactions (Bayés et al., 2010; Collins et al., 2006; Cheng et al., 2006; Fernández et al., 2009; Li et al., 2004; Peng et al., 2004; Yoshimura et al., 2003), the current reconstructions of the PSD networks are likely to be partial and noisy. For instance, a meta analysis of studies that detected PSD proteins shows that only 42% of the proteins were detected in more than one study (Collins et al., 2006). Mapping the protein interactions may be similarly noisy: A survey of small scale PPI studies of the NMDA receptor, found that 41% of the proposed PSD proteins (77 out of 186) had no known interactions with the rest of the network (Pocklington et al., 2006). High-throughput measures of protein–protein interactions (PPI) provide very valuable evidence on protein interactions, but they are also susceptible to under- and over-detection (Qi et al., 2006). This calls for developing methods that can combine evidence from multiple experiments and produce a high confidence reconstruction of the PSD network.

Here we describe a computational approach to reconstruct the PSD network based on learning the characteristics of PSD protein interaction, and predicting new interacting pairs. Similar approaches were successfully applied in other PPI networks (Skrabanek et al., 2008). We start with a "seed" network of PSD proteins built from high confidence interactions, and then expand that network repeatedly by adding edges from a list of suspected interactions. Finally, we further expand the network using proteins that are suspected to reside in the PSD, and predict how they interact with the network. The end result of this process is a PSD network with 25% more protein interactions than the initial network.

Results

Overview of the reconstruction approach

To predict novel interactions between PSD proteins, we follow a three-step procedure, illustrated in Fig. 1A. In the first phase, *seed network construction*, we construct a PPI network using evidence from high confidence interactions of PSD.

In the second phase, network *expansion with candidate interactions*, we expand the seed network using potential interactions proposed by high throughput experiments. To decide which interactions to add, we grow the network in an iterative process, layer by layer (Fig. 1B). At each iteration, we use the current network to train models of interactions (Fig. 1B(iii)), then rank the candidate interactions (Fig. 1B(iv)) and add the highest confidence interactions to the network (Fig. 1B(v)). These steps are repeated until no more interactions are found. The end result of this phase is an *expanded network*.

In the third phase, network *expansion with candidate proteins*, we predict *de novo* interactions between the expanded network and proteins that were experimentally pulled out of the PSD. We apply the same iterative expansion procedure as in the second phase, but here we consider all possible interactions between each candidate protein and the network reconstructed so far. The resulting *final network* is evaluated using the literature and using additional experimental assays that were not used during training.

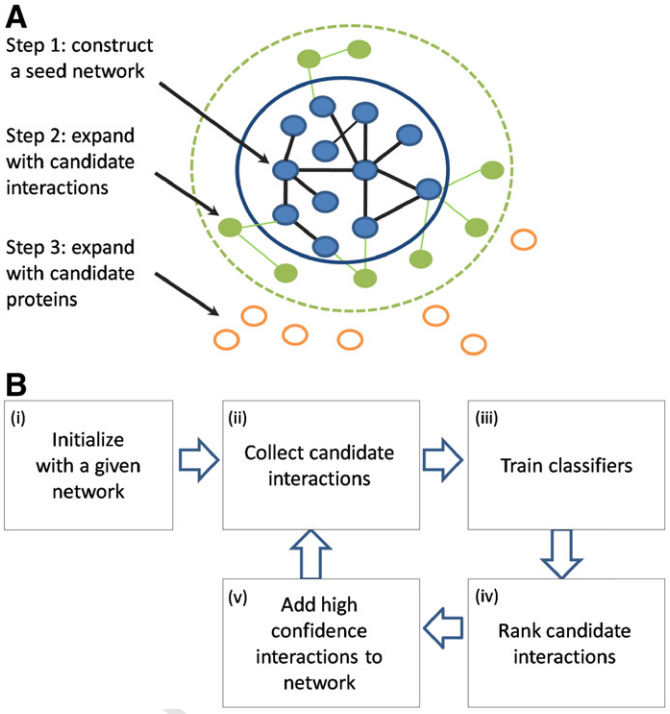

Fig. 1. Network reconstruction is performed in three iterative steps. (A) Starting with a seed network curated interactions, we consider a set of candidate interactions, which have evidence to connect the proteins in the seed network. We then further consider interactions from candidate proteins that are not known to connect to the seed or expanded network. (B) Expanding a network (by either candidate interactions or proteins) takes place in iterations. (i) The network is initialized to be the seed network. (ii) A set of candidate interactions is proposed. (iii) Classifiers are trained on interactions from the existing network. (iv) The candidate interactions are ranked by the trained classifiers. (v) The most likely interaction are validated and added to the current network.

Training models of interactions using the high confidence PSD network

We start with creating a trusted seed network by collecting PPIs from small scale experiments that report interactions within the mouse PSD (Cho et al., 2007; Dong et al., 1997; Jackson and Nicoll, 2009; Leonard et al., 1998; Leonoudakis et al., 2004; Nishimune et al., 1998; Saglietti et al., 2007; Sato et al., 2008; Schulz et al., 2004; Schwenk et al., 2009; Setou et al., 2002; Silverman et al., 2007; Song et al., 1998a,b; Stegmüller et al., 2003; Terashima et al., 2004; Torres et al., 1998; Torres et al., 2001; Uchino et al., 2006; Von Engelhardt et al., 2010; Wang et al., 2006; Xia et al., 1999), and from a comprehensive study that collected PPI from 190 studies (Pocklington et al., 2006). This set of interactions served as a training set for learning binary classifiers that detect repeating patterns that can be used to predict new interactions.

In our supervised learning framework, each protein is represented by a vector of measurements from various sources, which are called 'features'. We selected features that are expected to be highly correlated in pairs of proteins that interact, and less correlated in non-interacting proteins. Fig. 2 shows examples of three features that follow this pattern for one pair of proteins that are known to interact (Gria2, Gria3, Fig. 2A–C) and one pair of proteins that is believed not to interact (Grb2, Actg1, Fig. 2D–F). Fig. 2A depicts the expression profile of Gria2 and Gria3 across a large compendium of microarray experiments (see Experimental methods), showing that the profiles Q5 of the two proteins are highly correlated across the compendium. At the same time, the expression profiles of another pair of genes (Grb2,

Please cite this article as: Bar-shira, O., Chechik, G., Predicting protein–protein interactions in the post synaptic density, Mol. Cell. Neurosci. (2013), http://dx.doi.org/10.1016/j.mcn.2013.04.004