A comprehensive strategy to identify stoichiometric membrane protein interactomes

Avanti Gokhale,1 Patricia Perez-Cornejo,2 Charity Duran,1 H. Criss Hartzell1,* and Victor Faundez1,*
1Department of Cell Biology; Emory University School of Medicine; Atlanta, GA USA; 2School of Medicine; Universidad Autónoma de San Luis Potosí; San Luis Potosí, SLP Mexico

Keywords: membrane protein, Ano1, interactome, SILAC, epithelia, salivary gland

Submitted: 07/26/12
Revised: 09/03/12
Accepted: 10/30/12

http://dx.doi.org/10.4161/cl.22717
*Correspondence to: H. Criss Hartzell and Victor Faundez; Email: criss.hartzell@emory.edu and vfaunde@emory.edu

There are numerous experimental approaches to identify the interaction networks of soluble proteins, but strategies for the identification of membrane protein interactomes remain limited. We discuss in detail the logic of an experimental design that led us to identify the interactome of a membrane protein of complex membrane topology, the calcium activated chloride channel Anoctamin 1/Tmem16a (Ano1). We used covalent chemical stabilizers of protein-protein interactions combined with magnetic bead immuno-affinity chromatography, quantitative SILAC mass-spectrometry and in silico network construction. This strategy led us to define a putative Ano1 interactome from which we selected key components for functional testing. We propose a combination of procedures to narrow down candidate proteins interacting with a membrane protein of interest for further functional studies.

Biological membranes are interfaces between environments of differing composition. This membrane function is dynamic as membranes are conduits that interpret and modify their surroundings. Nowhere is this clearer than in the plasma membrane where fundamental cell processes critical for preserving cellular homeostasis take place. In both unicellular and multicellular organisms, plasma membrane proteins play integral roles in cell signaling, cell-cell and cell-matrix adhesion, transport of molecules and ions, cellular metabolism and maintenance of cellular architecture. All of these cellular functions critically depend on the unique lipid and protein composition of membranes. Here we focus on the problem of how to determine protein-protein associations that membrane proteins establish in their transmembrane, exofacial and endofacial domains in order to perform their functions. We use as a case study the calcium activated chloride channel Anoctamin 1 (Ano1).1

Genomic studies suggest that membrane proteins make up one quarter or more of all the predicted proteins in an organism. This fraction seems conserved from prokaryotes to Homo sapiens.2-4 Apart from their abundance, the extensive and diverse roles that membrane proteins play in cell-cell adhesion, intracellular and intercellular transport, cell-cell recognition and anchoring to cytoskeleton showcase their significance. Mutations in membrane proteins are implicated in a wide number of pathologies including heart disease, kidney dysfunction, myotonia, tumorigenesis, cystic fibrosis and an array of neuropathologies.5

Nearly 60% of all FDA-approved drugs target membrane proteins. In contrast, therapeutic drugs that target nuclear or cytoplasmic proteins account each for a mere 6%.6,7 Prominently represented in “druggable” membrane proteins are G-protein-coupled receptors and ion channels.6,7 G-protein-coupled receptors are the largest family of drug-susceptible targets.7 The numbers of G-protein-coupled receptors are estimated to be ~700–800 out of ~2,600 genes in the human genome [Homo sapiens build 37.3 consensus coding sequence (CCDS)
Most of these “druggable” membrane proteins adopt complex membrane topologies making their purification and the identification of their interactors challenging. Ano1, the focus of this addendum, belongs to the channel family and its activity is susceptible to chemical inhibitors. Moreover, Ano1 possesses a complex membrane insertion topology making it a good candidate to test new approaches for membrane protein biochemistry scalable to other membrane proteins (Fig. 1).

**The Paucity of Membrane Protein Interaction Networks**

Although membrane proteins such as G-protein-coupled receptors and ion channels are abundant and critical, our knowledge of how these proteins are organized in the membrane with other proteins and how they couple to extracellular and intracellular proteins is far from comprehensive. Although great progress has been made in recent years in purification of membrane proteins for crystallography, these approaches involve disruption of the native membrane environment and replacement with artificial lipids. To understand how a membrane protein co-exists with other proteins, one of the main hurdles is to reproduce a milieu that preserves the protein’s inherent function, structure and protein-protein interactions. These difficulties have led to a major underrepresentation of membrane proteins in protein-interaction networks or interactomes. Here we use a wide definition of interactome as the network of molecular, genetic and/or metabolic associations aiming at describing contents, structure, function, behavior, or combinations thereof either of a protein, pathway, organelle or cell.

Most protein-protein interactions that have been discovered for G-protein-coupled receptors and channels have been defined painstakingly one-by-one or few at a time. Comprehensive genome- or proteome-wide strategies have been used relatively little for membrane proteins. Of nearly 700 human G-protein-coupled receptors that have been curated, fewer than 100 interacting proteins have been identified for less than 10% of them. The picture is even more fragmented for vertebrate ion channels except for few exceptions. In contrast, large-scale screening for pairwise interactions among 705 *Saccharomyces cerevisiae* proteins annotated as integral membrane proteins has provided nearly 500 protein interactors that participate in ~2,000 putative protein-protein interactions. This simple comparison tells us that protein interactions networks of vertebrate membrane proteins are under-explored. Except for *Saccharomyces cerevisiae*, this is a global problem that spans phyla. Analysis of curated protein-protein interactions gathered in the BIOGRID database of protein-protein interactions indicates that less than one-third of all reported interactions include membrane proteins. Moreover there is a significant under-representation of GO categories linked to the membrane: “intrinsic to membrane,” “integral to membrane,” “Golgi apparatus” and “membrane part.” Under-representation is irrespective of the method used to identify protein-protein interactions such as affinity capture followed by mass-spectrometry or immunoblot or yeast-two hybrid system. These systemic problems with membrane proteins have been approached successfully in several ways. For example, yeast-two hybrid strategies have been used where the bait and prey association are either excluded from the nucleus (protein fragment complementation) or constrained to a membrane such as in the split-Ubiquitin membrane yeast-two hybrid or more recently by tandem affinity purification of tagged membrane proteins from yeast detergent extracts. Alternatively, identifying interactions at the gene level provide a proxy for protein-protein associations since one-fifth to one tenth of protein-protein interaction pairs have a correlate in a genetic interaction pair.

**Philosophy of the Approach Presented Here**

There are diverse challenges facing the investigator in pursuit of a membrane protein interactome. Among them are: (1) the low relative abundance of membrane proteins, (2) the need for membrane solubilization with detergents, which may only partially solubilize membrane proteins or alter membrane protein function and their interactions and (3) the challenge of selecting candidates for study after protein identification by mass spectrometry, which routinely provides hundreds of proteins for one experiment. The approach presented here intends to minimize some of these problems and provide a rational approach, based on quantitative proteomics and in silico tools, to select candidates for study among the hundreds proteins detected by sensitive mass spectrometry methods. Even with these goals in mind, researchers should be aware that solving one problem might create others. For example, to overcome the low abundance of the membrane protein that we focused on, Ano1, we resorted to recombinant expression. This opens the door for new problems such as potential mislocalization due to over-expression and biases in defining protein stoichiometries.

As with the interpretation of any large data set, the approach presented here provides a prioritized list of potential interactors. Claims of “physiological relevance” ought to be founded in multipronged confirmation of each individual interaction. Moreover, it is central to emphasize that there is no substitution for confirmatory functional studies of a putative protein-protein interaction pair using endogenously expressed proteins in a tissue of interest.

**Identification of a Membrane Protein Interactome: the Ano1 Case Study with its Advantages and Caveats**

We chose to focus on the calcium activated chloride channel Ano1 (Fig. 1). Ano1 is expressed in the apical domain of exocrine cells like salivary glands and is essential for salivary exocrine secretion (Fig. 2) and a growing list of other functions. Members of the anoctamin family play central physiological roles in a multitude of cells including but not limited to epithelial secretion, signal transduction in sensory systems, modulation of smooth muscle contraction and control...
proteins and interact within the plane of the lipid bilayer. We assembled a sequence of well-established approaches to identify the putative interactome of Ano1. This approach could potentially be used to identify putative interactome of other membrane proteins.

Figure 2 briefly presents this.

Figure 1. Membrane topology of Ano1. Model of mouse Ano1 topology. This model is based on data from Yu et al. (2012). In this model the sequence 628–638 forms the outer vestibule of the channel and amino acids in the first and third intracellular loops are involved in Ca-dependent gating. See Yu et al. for further details.

Ano1 is an eight transmembrane domain channel (Fig. 1). Ano1 molecular interactions through its intramembrane, exofacial and endofacial domains are mostly unknown. The complex membrane topology of Ano1, with multiple loops connecting Ano1’s trans-membrane domains, makes the identification of its interactome challenging as each hydrophilic segment of the protein would need to be tested individually in solution to identify its associated proteins (Fig. 1). Furthermore, Ano1, like other membrane proteins, may associate with accessory subunits that are themselves membrane proteins and interact within the plane of the lipid bilayer.

We assembled a sequence of well-established approaches to identify the putative interactome of Ano1. This approach could potentially be used to identify putative interactome of other membrane protein. Figure 2 briefly presents this.
experimental sequence. Two cell lines are used. One expresses a functional tagged membrane protein while the control is identical except that it lacks the recombinant membrane protein. One of these cell lines is labeled with either regular medium (Fig. 3, gray plate) while the second is labeled with non-radioactive isotope-labeled lysine and arginine-containing medium (Fig. 3, red plate) to equilibrium (this typically requires growth in labeling medium for six generations or more). Intact cells from each condition are then exposed to a membrane-permeant cross-linker, such as [dithiobis(succinimidyl propionate)] DSP, whose concentration is controlled to stabilize only close-proximity interactions (Fig. 3, step 1). The utility of using DSP is illustrated by the observation that the proteins co-purifying with Ano1 in the presence of DSP are reduced 15-fold when DSP is omitted. However, the use of DSP or any other cross-linker has to be carefully weighted since their use may introduce new variables that could confound data. For example, cross-linkers used at high concentrations may lead to inefficient leading to low percentage of stabilized protein complexes and an apparently low stoichiometry of interaction. To overcome these limitations, we defined concentrations, time and temperature of cross-linker incubation in vivo so that cross-linking increases by a 2- to 3-fold the bait’s molecular size, as determined by sucrose sedimentation. This provides a biochemical indication that the cross-linker has not created extended networks of protein interactions. For example, Ano1 sediments at ~6S without cross-linking yet after DSP treatment the shift in the sedimentation represents an increase of at least ~3S. This increase in the sedimentation coefficient would roughly correspond to a doubling of the molecular size of a ~150 kDa protein such as Ano1. This has been a universal criterion used in all our work with cross-linkers to minimize the possibility of false-positives due to extended networks of cross-linked proteins.1,40

Detergent-soluble cell extracts are applied to immunoaffinity magnetic bead matrices, where they are retained due to the inclusion of antibodies directed against the tag on the recombinant protein. The bound complexes are eluted with a molar excess of a peptide corresponding to the tag (Fig. 3, step 1). Eluates from matrices incubated with transfected cell (“heavy” isotope labeled) and non-transfected cell (unlabeled “light”) extracts are mixed 1:1 and analyzed by MS/MS (Fig. 3, step 2). Mass spectrometry determines peptide identity, distinguishes peptides based on their isotope label and provides an enrichment of peptides between the two experimental conditions, which can be differentiated by the isotope tags in them. Putative interacting proteins are prioritized based on their fold enrichment as a proxy for stoichiometry of association (Fig. 3, steps 2–3). Prioritized proteins are organized into putative networks (Fig. 3, step 4), and the network nodes with the highest number of connections are selected for functional testing (Fig. 3, step 5). We asked that this experimental strategy satisfy the traditional biochemical criteria of minimizing false positive protein-protein interactions common to mass-spectrometry approaches.45,46 For example, we identified the proteome that binds to antibody-bead matrices and we subtracted it from the Ano1 proteome. The outcome of this subtraction is quite illustrative since the Ano1 proteome led to the identification of 509 proteins, which are reduced to 209 after subtracting the non-specific bead proteome. Of these 209 proteins only 93 are enriched more than 2-fold by SILAC enrichment. In addition, our approach also considers the following:

(1) Protein isolation and identification should be quantitative.

(2) Protein isolation and identification should consider the contribution of lipid bilayers to the native state and interactions of a membrane protein.

(3) Experimental strategy should be designed with the realization that the bait membrane protein may reside in different membrane compartments either because of protein trafficking or overexpression artifacts.

(4) The strategy should be adaptable for the identification of sub-proteomes.

Protein isolation and identification should take advantage of quantitative approaches to binding partner identification and sorting. This criterion is a critical step in selecting and narrowing down candidates after a high-throughput screen has delivered a large number of “hits” (Fig. 3). This step is central for subsequently designing experiments to test the functionality of individual protein-protein interactions. We prefer SILAC (stable isotope labeling with amino acids in culture) to quantitatively label the whole proteome of a cell45-48 followed by immuno-affinity purification (using the membrane protein of interest as the bait protein) with magnetic beads and the use of antigenic...
peptides for elution (Fig. 3). This method offers three major advantages. First, isotope labeling occurs in vivo and the whole proteome is uniformly tagged while the native state of proteins and their interactions is maintained. Second, the proteomes can be distinguished by the isotopic tags in lysine and arginine allowing a contrast between two experimental conditions (Fig. 3). In our Ano1 experiments, we used this to define those proteins that spuriously bind to magnetic bead immuno-affinity matrices, but the same approach could be used to compare the interactomes of different splice variants, for example. Finally, the fold enrichment prioritizes components in the putative interactome (Fig. 3). This is of great advantage as we can estimate apparent stoichiometries among proteins in the interactome when compared with the bait. We qualify the stoichiometry as “apparent” because the bait, Ano1, is recombinantly expressed likely above the levels of the endogenous protein, a fact that necessarily biases stoichiometries. For example the recombinant Ano1 interactome identified 93 proteins enriched greater than 2-fold. Ano1 itself was enriched ~30-fold and 4 of the identified proteins were similarly enriched ~30-fold. These top proteins co-enriching with Ano1 are the actin-binding proteins Ezrin, Moesin, Radixin, which we functionally confirmed as modulators of Ano1 channel activity. In contrast, a protein in the interactome data set that is found enriched only 2-fold may be lower due to its expression in cells at levels below the recombinant target (Ano1-FLAG), localization to distinct cellular places at steady-state, lower affinity, lack of optimal cross-linking, or other reasons, but may still represent a potential biologically relevant interactor.

Stoichiometrically prioritized interactomes can still be composed of dozens or hundreds of potential interactors. This makes the selection of candidate proteins for experimentation problematic. Thus, we coupled the Ano1 protein interactome to an in silico data mining algorithm to generate a proposed protein interaction network (Fig. 3). The proposed interaction network built in silico is then used to select high connectivity proteins for biochemical and functional testing. It is important to highlight that this putative network does not represent the actual map of experimentally verified functional protein-protein interactions. Rather, it takes advantage of an independently generated data set that can be used to refine the actual map of experimentally verified functional protein-protein interactions. Rather, it takes advantage of an independently generated data set that can be used to refine the interactome data set to select those candidates most likely to represent physiologically relevant interactors. To construct a network, we preferred the Dapple algorithm due its predictive value, its statistical definition and ranking of network nodes and the stringency of the protein-protein interaction databases selected (www.broadinstitute.org/mpg/dapple/dapple.php). DAPPLE builds protein interactions networks gathering significant physical connectivity among proteins from protein-protein interactions databases (MINT, BIND, IntAct, PPrel, ECrel, Reactome and others). The data set contains nearly half a million interactions.

Figure 3. Flow diagram of the strategy to generate an in vivo membrane protein interactome using SILAC mass spectrometry. See text for details and Perez-Cornejo et al. (2012).
We have shown that DSP is suitable for a labile membrane protein-clathrin adaptor interaction. Moreover, we have validated the use of DSP by genetically testing putative interactors. Critical to the use of DSP is to control the extent of the chemical crosslinking reaction to ~10–20% of the total bait protein present in the sample. Alternatively, we limit the increase of the bait original molecular size to 2- to 3-fold after cross-linking. This criterion decreases, but does not eliminate, the chance of false-positive protein identification due to extensively crosslinked complexes expanding beyond proteins in the near vicinity of the bait. DSP is one of few cross-linker options whose chemistry is amenable for in vivo controlled protein crosslinking. Other cross-linkers can be tested for any membrane protein of interest yet their concentrations, temperature and time of cross-linking must be determined empirically and cannot be assumed similar between cross-linkers (www.piercenet.com/browse.cfm?fidID=0203). Finally, while reversible crosslinking of proteins in situ is a powerful method to stabilize transient, weak, or low abundance protein-protein interactions it is central to keep in mind that the final definition of a protein-protein interaction pair requires multipronged experimental scrutiny.

**Experimental strategy should be designed with the realization that the bait membrane protein may reside in different membrane compartments either because of protein traffic or overexpression artifacts.** At the steady-state, membrane proteins usually are concentrated in one subcellular compartment, but during their life cycle they usually traffic along the endomembrane system of eukaryotic cells. Proteins resident in the plasma membrane can be internalized and recycled back to the cell surface via carriers bound to the cell surface either from endosomes or the Golgi complex. Thus, protein interactions engaged by a membrane protein will likely be distinct in different compartments where they inherently reside or transit. Controlled in vivo cross-linking allows the identification of proteins irrespective of the compartment where they traffic or reside as suggested by the identification of VAMP-3, an endosomal SNARE that cycles between plasma membrane and endosomes, which is co-enriched nearly 20-fold with Ano1. However, if the membrane protein is recombinantly expressed subcellular localization should not be inferred from bait interactors as they may reflect artifacts of overexpression. Subcellular localization should be determined from experiments performed with endogenous proteins.

The strategy should be adaptable for the identification of sub-proteomes. Figure 2 depicts an approach designed to identify the widest possible interactome for Ano1. However, the approach can be adapted for the identification of compartment-specific interactomes. Two examples illustrate this point. First, a tag placed in the exofacial domain of a membrane protein would allow antibody application to intact cells prior to cross-linker treatment. Thus, it would be possible to purify just the surface pool of membrane proteins or those membrane proteins endocytosed after specific times. Second, the use of different cross-linker chemistries could be used to separate interactomes. For example, using membrane-impermeant cross-linkers could differentiate the exofacial interactome. In general, sulfo-derivates of cross-linkers such as DSP are excellent options.

**Conclusions**

The arsenal of tools to explore soluble protein interactions and their functionality is vast, yet options for membrane proteins are limited. Here we discussed the logic of an approach used by us to identify the interactome of a membrane channel, Ano1. We combined the best available tools of immuno-affinity chromatography, quantitative mass-spectrometry, in silico analysis and the use of reversible covalent chemical stabilizers of protein-protein interaction to define a high confidence list of membrane protein interactions to be experimentally validated one-by-one by alternate means (Fig. 3). Although this is one of several approaches to identify membrane protein interactomes, its reproducibility, comprehensive nature, multiple built-in controls and the rapid turn-around time from deciding to do the experiment to obtaining prioritized putative interactors compel us to recommend this experimental design.
Acknowledgments

Supported by grants from the NIH GM60448 (H.C.H.), EY014852 (H.C.H.), NS42599 (V.F.), GM077569 (V.F.), Emory University Research Committee (H.C.H.), NEI training grant 5T32EY00792-25 (C.D.) and NIH FIRST program fellowship K12 GM000608 (A.G.). Additional support was provided by the Microscopy Core of the Emory Neuroscience NINDS Core Facilities Grant P30NS055077 and NEI Core Grant P30EY006360.

References


47. Mann M. Functional and quantitative proteomics using SILAC. Nat Rev Mol Cell Biol 2006; 7:952-8; PMID:17139335; http://dx.doi.org/10.1038/nrm2067.