

Characterization of membrane protein interactomes by Co-interacting Protein Identification Technology (CoPIT)

Sandra Pankow (✉ pankows@scripps.edu)

Yates lab, The Scripps Research Institute, 10550 N Torrey Pines Road, La Jolla, CA 92037

John R Yates (✉ jyates@scripps.edu)

Yates lab, The Scripps Research Institute, 10550 N Torrey Pines Road, La Jolla, CA 92037

Casimir Bamberger

Yates lab, The Scripps Research Institute, 10550 N Torrey Pines Road, La Jolla, CA 92037

Diego Calzolari

Yates lab, The Scripps Research Institute, 10550 N Torrey Pines Road, La Jolla, CA 92037

Andreas Bamberger

Albert-Ludwigs-University Freiburg, Hermann-Herder-Strasse 3a, 79104 Freiburg i.Br., Germany

Method Article

Keywords: protein complex, protein interactions, network dynamics, ion channel, cystic fibrosis, immunoprecipitation, deep interactome

Posted Date: November 30th, 2015

DOI: <https://doi.org/10.1038/protex.2015.101>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Finding interaction partners of proteins is an important step in understanding their function. While co-immunoprecipitation of proteins followed by mass spectrometric identification (AP-MS) has become a standard approach to identify interaction partners, characterization of membrane protein interactomes is still challenging because of the typically low abundance and high hydrophobicity of membrane proteins. Here, we describe CoPIT (for Co-interacting Protein Identification Technology), an approach to comprehensively identify and characterize the interactome of membrane proteins. The CoPIT protocol includes experimental and computational methods to establish membrane protein interactome networks and to monitor interactome changes upon perturbation. While the approach in particular improves membrane protein interactome analysis as demonstrated by using a low abundant membrane protein (Cystic Fibrosis Transmembrane Conductance Regulator (CFTR)) as a model, it is applicable to all types of proteins. The time required to complete the protocol varies from 3 days to several weeks depending on whether all or only part of it is carried out.

Introduction

CoPIT (for Co-interacting Protein Identification Technology) is an experimental and computational framework that allows the comprehensive characterization of endogenous membrane as well as non-membrane protein interactomes using co-immunoprecipitations and mass spectrometry. The protocol contains detailed experimental methods for successful membrane protein Co-IP, preparation of the sample for mass spectrometry as well as a description of the statistical methods used for identification of specific interactions and identification of differentially regulated interactions between samples. CoPIT consists of three individual workflows that are executed sequentially as shown in Figure 1. First, an optimized experimental protocol provides enhanced sensitivity and efficiency for membrane protein interactome determination. To obtain meaningful results, it is absolutely critical to include proper controls in the experimental design (Figure 2). Depending on the experimental setting, appropriate controls are (ideally isogenic) cell lines that do not express the bait protein to identify the background caused by the antibody and “mock-IPs”, which are performed in the absence of specific antibodies against the bait, but in the presence of the beads that are used to couple the antibodies and which identify background caused by non-specific binding of proteins to beads. These controls must not be skipped as subsequent identification of the interactome relies on the results obtained in the control experiments. In the second part (step 5), CoPIT then determines the specificity of protein interactions in comparison to control experiments for unbiased discrimination of highly confident from less confident interactors. Additionally, interactome changes between different experimental settings can be determined. In the third part, the resulting interactomes are visualized with the Radial Topology Viewer (step 6), whereby the arithmetic distance of interactors to the bait protein in the center reflects a dynamic variable of the user's choice, and can be used to assess for example confidence of interaction or change in strength of interaction.

Reagents

- Monoclonal Antibody or polyclonal antibody covalently coupled to Protein G or Protein A Sepharose 4 Fast Flow beads \ (Cat No. 17-0618-01, GE Healthcare)
- Dulbecco's Phosphate buffered Saline \ (DPBS) without CaCl_2 , and MgCl_2 \ (Cat No. 14190-144, Gibco, Invitrogen)
- Ultra-Pure Tris \ (Cat No. EC-406, National Diagnostics)
- Igepal CA-630 \ (Cat No. 18896-50ML, Sigma-Aldrich)
- Ultra-pure EDTA \ (Cat No. 3579, Gibco)
- Sodium Chloride \ (Cat No. S7653-5KG, Sigma-Aldrich)
- PhosStop \ (Cat No. 04906845001, Roche)
- Complete ULTRA, EDTA-free Protease Inhibitor cocktail \ (Cat No. 05892953001, Roche)
- Non-stick surface microcentrifuge tubes \ (e.g. Cat No. 20170-650, VWR)
- Glycine, electrophoretic grade \ (Cat. No. 100191, Roche)
- Methanol, HPLC grade \ (Cat. No. 9093-03, J.T.Baker)
- Chloroform, 99.8+%, ACS reagent, stabilized with ethanol \ (Cat. No. 423555000, Acros Organics)
- Rapigest SF Surfactant \ (Cat. No. 186001861, Waters)
- TCEP \ (Cat. No. 77720, Thermo)
- Iodoacetamide \ (Cat. No. 90034, Thermo)
- Sequencing grade recombinant trypsin \ (Cat. No. V5111, Promega)
- Formic acid \ (Cat. No. A117, Fisher Scientific)
- Acetonitrile, HPLC grade, $\geq 99.8\%$ purity \ (Cat. No. LC015-2.5, Honeywell)
- Ammonium acetate, HPLC grade \ (Cat. No. A2149, Spectrum)
- Kasil No. 1624 potassium silicate solution \ (PQ Corporation)
- Kasil No.1 potassium silicate solution \ (PQ Corporation)
- Fused Silica \ (Cat. No. 1068150047, i.d. 100 μm , Polymicro Technologies)
- Fused Silica \ (Cat. No. 160-2250-5, i.d. 250 μm , Agilent Technologies)
- Aqua C18 resin, 3 μm and 5 μm particle size \ (Cat. No 00B-4299-B0, Phenomenex)
- Strong cation exchange resin \ (e.g. Cat. No. 4621-0507, Hichrom Limited)

Equipment

- High-pressure column and sample loader \ (pressure bomb)
- HPLC \ (e.g. HP 1200; Agilent Technologies) with post-HPLC split for nanoflow
- P-2000 microcapillary laser puller \ (Sutter Instrument Co.)
- Mass spectrometer \ (LTQ Velos Pro Orbitrap or similar, Thermo Fisher)
- Water bath sonicator \ (55kHz, 80 kHz, Branson)
- Cooled Microcentrifuge
- Thermomixer \ (e.g. Eppendorf Thermomixer R)
- Insulin injection syringe \ (BD Bioscience)
- Rotator
- Cell scraper ****Software**** Xcalibur \ (Thermo Fisher)
- RawExtract or RawConverter SEQUEST or ProLuCID DTASelect2 Matlab incl. biostatistics package \ (www.mathworks.com)
- CoPITgenerator \ (www.proteomicswiki.com)
- Radial Topology Viewer \ (www.proteomicswiki.com)
- RawExtract, SEQUEST, ProLuCID, DTASelect2 and Census may be obtained from the Yates laboratory \ (www.fields.scripps.edu). Access to high performance parallel computing resources \ (cluster or cloud-based computing) is highly recommended. RawExtract, Sequest, ProLuCID, and DTASelect are also integrated into the Integrated Proteomics Pipeline, IP2 \ (www.integratedproteomics.com), which is a commercial platform for mass spectrometry data analysis.
- **Equipment**** ****setup**** Fritted micro-capillary column preparation and MudPIT setup has been described in detail elsewhere ¹⁻³. The CoPIT protocol has been successfully carried out in combination with MudPIT 3,4 using LTQ, LTQ Orbitrap, LTQ Velos Pro Orbitrap as well as Orbitrap Elite mass spectrometers \ (all Thermo Fisher). Briefly, purified peptides were re-suspended in 50 μl of buffer A \ (95% H_2O , 5.0% MeCN, 0.1% formic acid) and pressure loaded onto the back-end of a preparative MudPIT microcapillary column consisting of fused silica \ (i.d. 250 μm) packed in-house with 2.5 cm of 5 μm Aqua C18 resin and 2.5 cm of strong cation exchange resin. The preparative column was connected by a small union body to an analytical reversed-phase column \ (115 mm fused silica i.d. 100 μm) packed with

3 μm Aqua C18 resin. Peptides were analyzed by nano-ESI-LC/LC-MS/MS by placing the triphasic MudPIT column inline with a quaternary HPLC pump (Agilent) and separating the peptides in multiple dimensions with a modified 6-step gradient containing 0%, 20%, 40%, 60%, and 100% of Buffer C (500 mM ammonium acetate/5% MeCN/0.1% formic acid) over 12 h or a 10-step gradient (0% to 90% Buffer C) over 20 h as described previously⁴. Each full scan mass spectrum (400-2000 m/z) was followed by 6 (LTQ, LTQ-Orbitrap XL) or 20 (Orbitrap Elite, Velos or Velos Pro) data-dependent MS/MS scans at 35% normalized collisional energy and an ion count threshold of 500 counts (LTQ) or 1000 counts (LTQ-Orbitrap XL, Elite or Velos). Additionally a dynamic exclusion list with an asymmetric exclusion window of -0.51 and +1.5 Da, repeat count of 1 and list size of 500 was used.

Procedure

Step 1 Co-immunoprecipitation Timing: 1.5 days All steps should be performed as quickly as possible, on ice and with pre-chilled buffers. It is very important not to use previously frozen or otherwise manipulated samples, as this will result in loss of interactors. It is also strongly advised to test the specificity and suitability of antibodies for immunoprecipitation with Western blotting first, and verify the specificity of the antibody for the bait (for example using a gene knockout model). While mass spectrometry is exquisitely sensitive, it cannot correct for insufficient antibody specificity or affinity or for suboptimal performance of the immunoprecipitation. Thus the bait should be highly enriched as compared to input and the antibody should be highly specific for best results. In this regard, it is recommended to use only Sepharose beads that are sufficiently saturated with antibody and show better than 90% covalent coupling efficiency of antibodies to the beads. Saturation of ProteinA/G-Sepharose beads and covalent coupling efficiency can be easily checked by SDS-gel electrophoresis and Coomassie blue staining. Additionally, the optimal starting amount may need to be empirically determined for different bait proteins. A very general recommendation is 1×10^8 cells for bait proteins that are equal to or less than 100 molecules/cell.

- 1.1. Rinse cells with at least 10 volumes (v/v) of 1 x DPBS. Remove excess 1 x DPBS by decanting or by vacuum suction and tapping on a paper towel.
- 1.2. Immediately add 1.5 ml of ice-cold TNI lysis buffer per 15 cm cell culture dish (TNI: 0.5% Igepal CA-630, 50 mM Tris [pH 7.5], 250 mM NaCl, 1 mM EDTA and 1x Complete ULTRA EDTA-free Protease Inhibitor cocktail, 1x PhosStop).
- 1.3. Incubate for 20 min on ice on an orbital shaker.
- 1.4. Scrape off cell lysate with a large cell scraper and transfer to a 1.5 ml microcentrifuge tube.
- 1.5. Sonicate the cell lysate for 3 min in a water bath sonicator operating at 55 kHz. Critical: Do not use a probe sonicator, this will result in loss of interactions.
- 1.6. Centrifuge for 30 min. ($18,000 \times g$, 4°C).
- 1.7. Preclear the cell lysate with an appropriate amount of Protein G Sepharose 4 Fast Flow beads (100 μl of 50% slurry in lysis buffer) for 2 h at 4°C with head-over-head rotation. The volume of preclear beads is the same as the volume of antibody coupled beads. For volumes larger than 1.4 ml, pool the cell lysates and transfer into a 15 ml tube.
- 1.8. Centrifuge for 3 min, $500 \times g$, 4°C .
- 1.9. Transfer supernatant to a new microcentrifuge tube containing the appropriate amount of antibody-coupled sepharose slurry (e.g. 100 μl of 50% antibody slurry, equaling 200 μg of antibody for starting amount of ≈ 10 mg). Incubate over night with head-over-head rotation at 4°C .
- 1.10. Centrifuge the

binding reaction for 3 min at 500 x g, 4 °C. 1.11. Remove the supernatant, transfer the beads to a new tube and wash the beads three times with 20 – 100 bead volumes of TNI lysis buffer. 1.12. Centrifuge for 3 min at 500 x g, 4 °C, carefully remove supernatant, and wash beads two times with TN lysis buffer containing no Igepal CA-630. 1.13. Carefully remove all of the supernatant with an insulin syringe. Optional: Freeze beads for >1 h at –80 °C to increase yield. PAUSE POINT: beads can be stored at -80°C for up to two weeks. ****Step** **2** **Elution** **of** **protein** **complexes** **and** **sample** **clean** **up** **Timing**:** ****60** **min**** 2.1. Elute proteins twice with at least four to 10 bead volumes of elution buffer (0.2 M glycine, pH 2.3/ 0.5% Igepal CA-630) for 20 min, 37 °C, with shaking. 2.2. Combine the eluates and transfer to a new microcentrifuge tube. Critical: Make sure to get no beads into the eluate, they might clog microcapillary columns and produce background signal later in the mass spectrometer. 2.3. Neutralize eluate with 1/10 vol (v/v) freshly prepared 1 M NH₄CO₃. 2.4. Add 4 vol (v/v) Methanol to the eluate and vortex. 2.5. Add 1 vol (v/v) Chloroform and vortex well for 0.5 - 1 min. 2.6. Centrifuge for 10 min at 18,000 x g. 2.7. Remove the supernatant without disturbing the pellet. Note: The pellet may be very tiny or hardly visible at all. 2.8. Wash pellet with 3 vol (v/v) Methanol. 2.9. Centrifuge for 10 min at 18,000 x g. 2.10. Remove the supernatant without disturbing the pellet. PAUSE POINT: the pellet can be stored at -80°C for up to four weeks. ****Step** **3** **Digestion** **of** **eluted** **proteins** **Timing**:** ****15** **h**** 3.1. Re-solubilize the methanol / chloroform precipitated proteins (pellet) in 100 mM Tris, pH 8.5, 0.2% Rapigest and sonicate for 1 h in a water bath sonicator. 3.2. Reduce cysteine bonds with 5 mM TCEP for 20 min. 3.3. Alkylate with 10 mM Iodoacetamide or Chloroacetamide for 30 min shielded from light. 3.4. Digest proteins with recombinant trypsin (30:1 ratio protein/trypsin) over night at 37°C with shaking (e.g. in an Eppendorf Thermomixer). 3.5. Inactivate Rapigest by adding formic acid to 9% final concentration and incubate for at least 1 h, 37°C, with shaking. 3.6. Reduce samples to near dryness in vacuo (approximately 45 min). PAUSE POINT: Samples can be stored at -80°C for up to four weeks. ****Step** **4** **Mass** **spectrometric** **analysis** **of** **peptides** **Timing**:** ****1-2** **days**** This step may be carried out in a proteomics core facility or proteomics laboratory with MudPIT experience. Resolubilize sample in a small amount of buffer A, load onto a preparative MudPIT column and perform a MudPIT experiment as described in 2,3. ****Step** **5** **Data** **analysis** **Timing**:** ****2** **days** **to** **several** **weeks**** The task to confidently identify proteins from mass spectra is complex. Prior experience in mass spectrometry as well as in data analysis software is very strongly recommended. Convert raw data with an appropriate converter (e.g. RawExtract) into a file format that is compatible with the database search engine (e.g. Sequest or ProLuCID). Search MS/MS spectra against the appropriate protein database with a search engine such as ProLuCID to determine cross-correlation scores of acquired to in silico calculated spectra. Filter search engine results to a low false discovery rate (FDR), for example using DTASelect2. Search engine results from MS/MS spectra should be combined for all biological replicates before assigning peptides to proteins and filtering the results. We also strongly recommend to conservatively filter the search results and adjust the false positive rate to a peptide false positive rate of less than 0.5% and a protein false positive rate of less than 1.0%. Resolve protein identification ambiguities for example by using a gene centric approach. Redundant identifications can skew the statistical analysis towards genes identified with many protein isoforms. Critical: The bait must be among the top 10 proteins with the highest number

of spectral counts. Recommended minimum spectral counts for the bait protein is between 50 and 100 spectral counts per experimental replicate (SpC). To distinguish specific protein interactors from non-specifically binding proteins, run CoPITgenerator according to the instructions (www.proteomicswiki.com). CoPIT generator reports a rank-ordered list of interactors with the bait protein that can be further analyzed. A confidence score (P-value) is reported for every potential interactor that reflects its likelihood of interaction. A user defined P-value cutoff can be set to differentiate weak from strong interactors. To compare interactomes, the P-value cutoff can be adjusted according to the ratio of bait in the different experimental conditions:  To identify differentially binding proteins in two different experimental conditions, calculate ratios r_p for individual proteins (p) from the sum of all intensities per protein I_p and experiment condition e_1 and e_2 and normalize to the sum of all bait intensities (SpC) according to:  Calculate errors according to random error of measurement:  in order to find significantly different proteins for example with a cutoff of at least 2-fold difference plus one or two standard deviations. **Step 6 Network representation and data presentation Timing: 2 days** The core interactome can be loaded into the Radial Topology Viewer (www.proteomicswiki.com) for visualization. The Radial Topology Viewer displays the bait at the center of the interactome network and arranges the interactors radially according to a user-defined distance. Interactors can be further grouped in multiple spokes based on a user-defined classification scheme, for example according to GO-terms. Interactions between the identified interactors can be obtained with the GeneMANIA 2.2 Plugin in Cytoscape 2.8.2 ^{5,6}. Export connectivity information to a .txt or .csv file to load into Radial Topology Viewer. A step-by-step instruction how to obtain networks from GeneMania and Cytoscape is available in Morris et al., 2014 ⁷.

Timing

5 days at minimum to several weeks for the whole procedure. Timing for the different parts of the protocol is given below each individual step.

Troubleshooting



Anticipated Results

CoPIT identifies and quantifies the interaction between a bait and interacting proteins using a label-free MudPIT approach. The protocol is performed in three separate steps comprising 1) performing the experimental Co-IP and mass spectrometry, 2) analyzing the mass spectrometric data and carrying out statistical analyses to identify high confidence interactors and differential interactions if more than one experimental condition was analyzed, and 3) constructing interactome networks. Typically, performing a CoPIT experiment results in an increase in sensitivity by 30-100 fold (see Extended Data Figure 2, Pankow, Bamberger et al., 2015 ⁸) and should result in several hundred to thousands of Spectral Counts for the bait per experiment. The actual result is depending on a number of factors and will vary with

affinity, specificity and suitability of the antibody used for Co-IP, the abundance of the bait protein in the experimental system, the amount of starting material available as well as the type of mass spectrometer used. For example, a Co-IP performed for Δ F508 CFTR as bait, (abundance: 100-200 molecules per cell) with a starting amount of 5 mg of protein from a whole cell lysate and using an LTQ Orbitrap Elite or LTQ Orbitrap Velos Pro can result in over 3500-4500 spectral counts for Δ F508 CFTR after stringent filtering to 0.2 % peptide FDR. High sequence coverage of the bait (often greater than >60% per experiment) reflects the increased sensitivity of CoPIT. In a successful experiment the bait will also usually be included in the list of top ten proteins in terms of spectral counts and will be highly enriched over input in a western blot experiment (see Pankow, Bamberger et al., 2015⁸). High enrichment of the bait and broad coverage of the interactome in the Co-IPs is crucial for sufficient discriminative power of the subsequent data analysis. In a successful experiment, the bait and strong interactors will be in the outmost tail of the right descending flank of the normal distribution calculated for potential baits in CoPITgenerator. The maxima of the two normal distributions can be well discerned with one being close to zero. In addition to a good confidence score (P-value, typically $P \geq 0.9$), we normally require that proteins were detected in at least two independent biological replicates of the same experimental condition to be considered high-confidence interactors. However, proteins with lower P-values may represent weaker or transient interactors, and may still be true interactors. Which P-value threshold is set depends on the experiment and can be dynamically adjusted according to the ratio of bait protein between two experimental conditions. The approach can be tested by comparing scores for putative background proteins such as antibody sequences and human hair keratins, which should have P-values well below $P \leq 0.5$ (if the bait is not a keratin or other protein known to interact with it, of course). Finally, the complexity of a specific protein interactome will depend on the specific bait protein. As CoPIT not only identifies direct interactors, but also enables the detection of indirect second and third degree interactors, a large but specific interactome may be rationalized by an estimation of the degrees of interaction.

References

1. Abelin, J. G. et al. Complementary IMAC enrichment methods for HLA-associated phosphopeptide identification by mass spectrometry. *Nat Protoc* 10, 1308-1318, (2015).
2. Schieltz, D. M., Washburn, M. P. & Hays, L. G. Analysis of Complex Protein Mixtures Using Nano-LC Coupled to MS/MS. *CSH protocols* 2006, (2006).
3. Wolters, D. A., Washburn, M. P. & Yates, J. R., 3rd. An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* 73, 5683-5690, (2001).
4. Washburn, M. P., Wolters, D. & Yates, J. R., 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19, 242-247, (2001).
5. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27, 431-432, (2011).
6. Montojo, J. et al. GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. *Bioinformatics* 26, 2927-2928, (2010).
7. Morris, J. H. et al. Affinity purification-mass spectrometry and network analysis to understand protein-protein interactions. *Nat Protoc* 9, 2539-2554, (2014).
8. Pankow S., Bamberger C., Calzolari D., Martínez-

Acknowledgements

We would like to thank Dr. Daniel Cociorva and Dr. Tao Xu for suggestions and comments on data analysis strategies. This work is supported by NIH grants 5R01HL079442-08 (J.R.Y.), P01AG031097 (J.R.Y.), P41 RR011823 (J.R.Y.), HHSN268201000035C (J.R.Y.), and CFF mass spectrometry fellowship BALCH050X6 (S.P. and J.R.Y.).

Figures

Figure 1

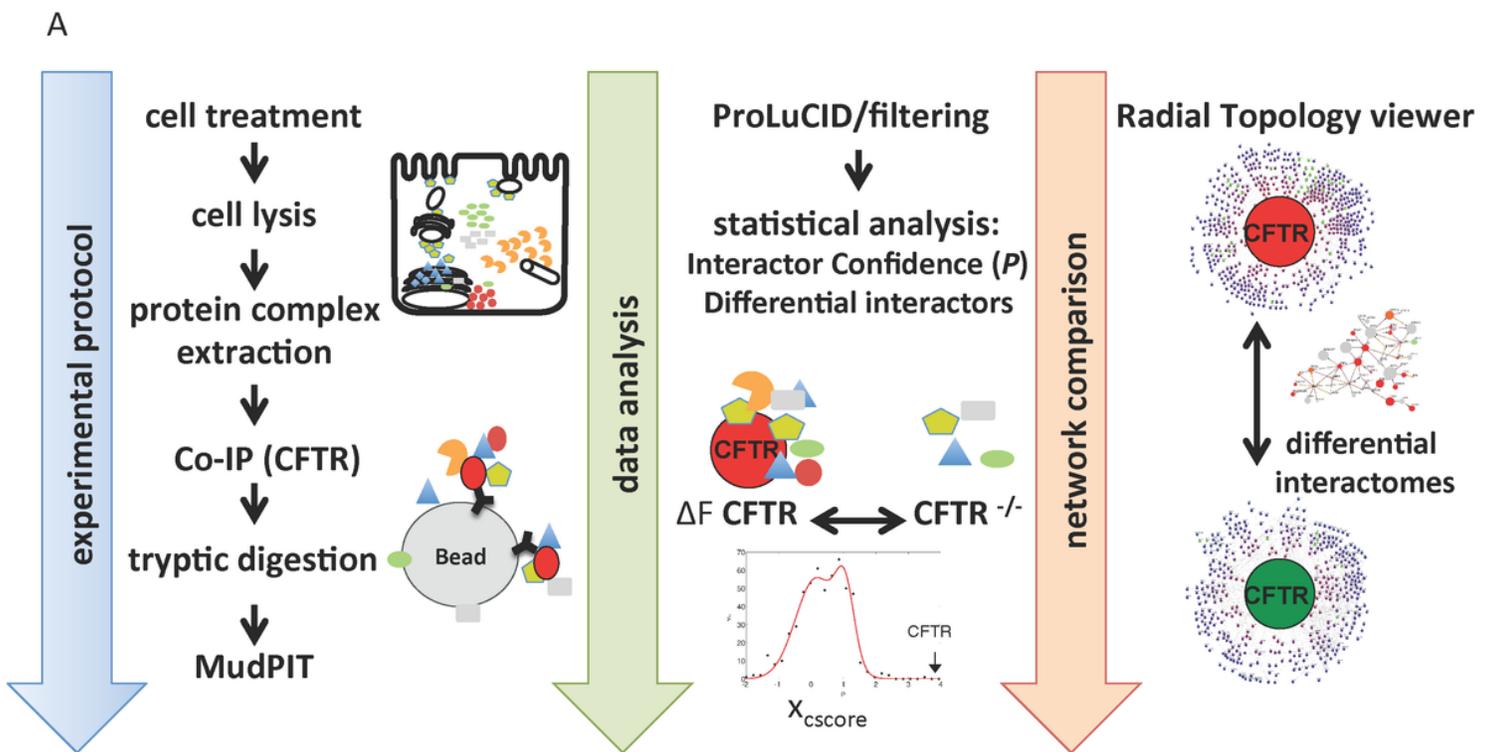


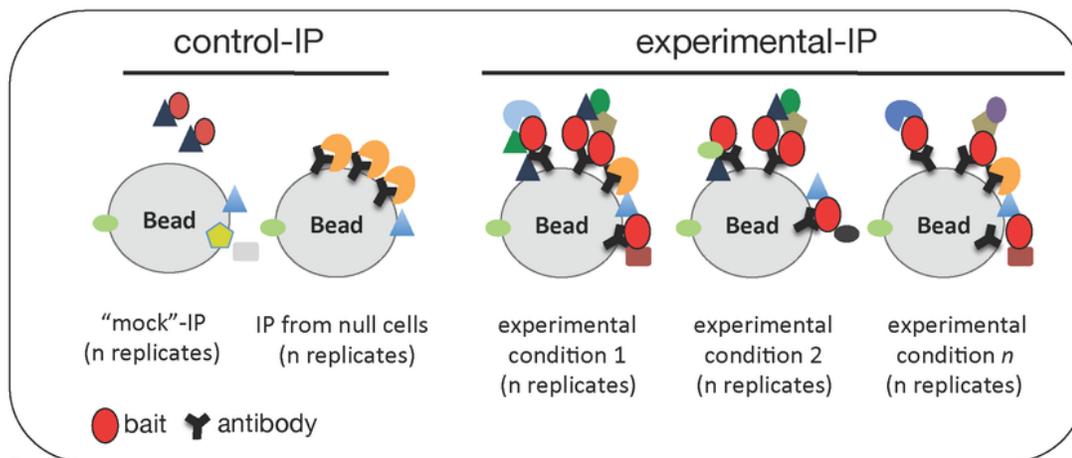
Figure 1

Schematic CoPIT workflow. The three individual workflows performed in CoPIT are shown. Experimental Protocol: Experimental protocols were optimized for high bait and interactor recovery by adjusting cell

lysis-, Co-IP and elution conditions. Data analysis: Resulting Raw-data are searched with ProLuCID and uniformly filtered to a peptide FDR of $\leq 0.5\%$. Signal-to-noise discrimination is carried out by including different negative controls, such as experiments from CFTR null cells, and significantly regulated interactions are identified using a counting statistics. Network comparison: Protein IDs and results are stored in a relational database, which contains also further ontological information. The Radial Topology Viewer visualizes ranking, functional annotation and connectivity of interactors in the resulting networks as well as differences between conditions.

Figure 2

A



B

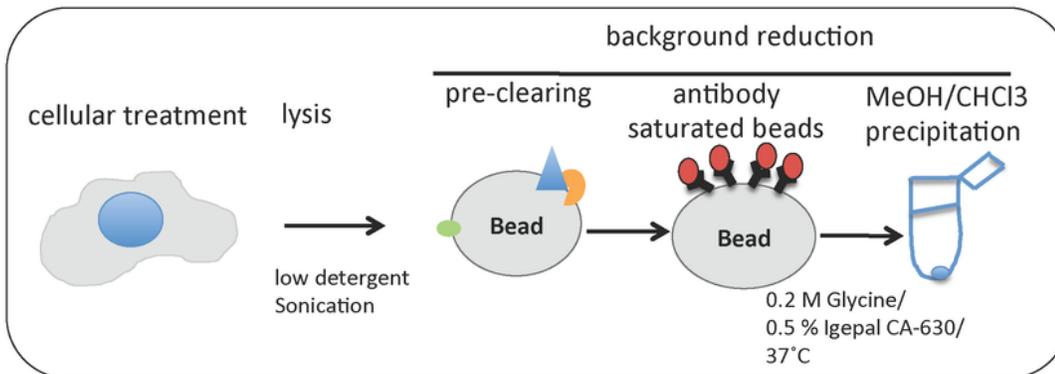


Figure 2

CoPIT experimental design *A*. Schematic of a typical CoPIT experiment with different experimental conditions and including control-IPs from cells not expressing the bait as well as bead controls (mock-IP).
B. Schematic of experimental steps for background reduction.

$$\frac{\sum_{i=1}^n I_{bait} e_i}{\sum_{i=1}^n I_{bait} e_{1i}}$$

Figure 3

Formula 1 Formula

$$r_p = \log_{10} \left(\frac{\frac{\sum_{i=1}^n I_p e_2 i}{\sum_{i=1}^n I_p e_1 i}}{\frac{\sum_{i=1}^n I_{bait} e_2 i}{\sum_{i=1}^n I_{bait} e_1 i}} \right)$$

Figure 4

Formula 2 Formula

$$\sigma_{r_p} = \frac{1}{2} \left(\log_{10} \left(1 + \sqrt{\frac{1}{\sum_{i=1}^n I_p e_1 i} + \frac{1}{\sum_{i=1}^n I_p e_2 i}} \right) + \left| \log_{10} \left(1 - \sqrt{\frac{1}{\sum_{i=1}^n I_p e_1 i} + \frac{1}{\sum_{i=1}^n I_p e_2 i}} \right) \right| \right)$$

Figure 5

Formula 3 Formula

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplement0.docx](#)