Published in final edited form as:

Curr Protoc Cell Biol. ; 74: 15.21.1–15.21.14. doi:10.1002/cpcb.14.

Discovering Protein-Protein Interactions Using Nucleic Acid Programmable Protein Arrays

Yanyang Tang¹, Ji Qiu¹, Matthias Machner², and Joshua LaBaer^{1,*}

¹Virginia G. Piper Center for Personalized Diagnostics, Biodesign Institute, Arizona State University, Tempe, Arizona 85287, United States

²Division of Molecular and Cellular Biology, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, United States

Abstract

We have developed a protocol enabling the study of PPIs at the proteome level using *in vitro* synthesized proteins. Assay preparation requires molecular cloning of the query gene into a vector that supports *in vitro* transcription/translation (IVTT) and appends a HaloTag to the query protein of interest. In parallel, protein microarrays are prepared by printing plasmids encoding GST-tagged target proteins onto a carrier matrix/glass slide coated with antibody directed against GST. At the time of experiment, the query protein and the target proteins are produced separately using IVTT. The query protein is then applied to NAPPA arrays that display thousands of freshly produced target proteins captured by the antibody. Interactions between the query and immobilized target proteins are detected through addition of a fluorophore-labeled HaloTag ligand. Our protocol allows the elucidation of PPIs in a high-throughput fashion using proteins produced *in vitro*, obviating the scientific challenges, high cost, laborious work, as well as concerns about protein stability present in protocols using conventional protein arrays. Chaperone proteins present in the human cell lysate further assists in protein folding during IVTT. Furthermore, the specific covalent bond between the HaloTag and its added ligand allows the detection of interactions without the need for anti-tag antibody or chemical labeling of the query protein. The “programmability” of our NAPPA protocol supports the study of query proteins with different tag configurations and/or refined studies of protein domains or regions important for the observed interactions with target proteins.

Keywords

protein-protein interactions; protein arrays; NAPPA; cell-free gene expression; Halo Tag

*Correspondence: Joshua.LaBaer@asu.edu.

Conflicts of interest

The authors indicate no potential conflicts of interest.

Unit title and unit introduction

Discovering Protein-Protein Interactions Using Protein Arrays

Protein-protein interactions (PPIs) are fundamental to understanding biological systems. Proteomic techniques such as protein arrays allow profiling of PPIs at the proteome level [1–5]. Most conventional methods for studying PPIs on protein arrays require purified proteins, both query and target [3], and the detection of interactions on arrays is usually achieved by either chemically labeling the query protein (with a fluorophore or with biotin), or using anti-tag or query protein-specific antibodies [4, 5]. Despite demonstrated feasibility of these methods, protein arrays have not gained ground mainly due to methodical challenges [1, 2]. For instance, the purification of full-length human or eukaryotic proteins is technically challenging and requires extensive optimization. Some proteins are extremely difficult to produce in surrogate hosts such as bacteria, yeast or insect cells. Also, the storage and/or chemical labeling can denature or mask epitopes or domains that are important for protein interactions. In addition, protein-specific antibodies may not be available or lack specificity, leading to false positives or low signal-to-noise ratios [6].

To address some of these deficiencies, we have developed a protocol to study PPIs using *in vitro* synthesized query and target proteins. Our protocol begins with cloning of the query-encoding gene into the pJFT7_nHALO_DC expression vector which upon incubation with a HeLa cell lysate-based protein production system, enables the transcription and subsequent translation of a HaloTagged-query protein [2].

Our protocol was developed to work with NAPPA, which is an innovative protein platform that displays thousands of human target proteins without the need for protein purification [1, 2, 7, 8]. The human target genes are cloned into the pANT7_cGST expression vector which encodes for a carboxy-terminal GST-tag. Once the plasmids have been spotted on the array, the GST-tagged target proteins are freshly produced through IVTT using cell-free lysate, captured on the array by co-spotted anti-GST antibody, and displayed for the query to bind. Detailed protocols for the manufacturing of NAPPA arrays have been published before [7–9].

The detection of interactions between HaloTagged query proteins and putative target proteins on the NAPPA arrays is achieved by addition of a fluorophore-conjugated HaloTag ligand. The HaloTag, used as a solubility enhancing fusion tag, is a modified haloalkane dehalogenase engineered to form a covalent bond with haloalkanes, including the fluorescent ligand [10]. This ligand-based approach overcomes many of the detection challenges discussed above. Furthermore, the flexibility of using plasmid DNA and the on-demand protein production capability of NAPPA enables easy manipulation of the query proteins, including alterations in the tag configuration and fine mapping of interaction sequence or domains by analyzing protein variants. Although our protocol uses NAPPA as an example, the concept is generally applicable to other protein platforms.

As is true for most protein array technologies, the NAPPA protocol presented here does not eliminate all possible drawbacks. For example, the display of properly folded membrane proteins in the absence of a lipid bilayer is challenging [11]. Also, cDNA for all query and

target proteins is needed in order to generate IVTT-compatible vectors. Moreover, cellular proteins within the cell-free expression system may result in false negatives by interfering with protein interactions. And lastly, proteins produced *in vitro* using the cell-free expression system might either lack critical post-translational modifications (PTMs) or display atypical ones [2]. Despite these potential downsides, the fact that NAPPA is a highly flexible platform makes it possible to counteract some of these issues, for example by adding missing enzymes and substrates required for PTMs or by including detergent and other critical reagents. In the past, we have been actively improving our protocol to specifically address these challenges [12, 13].

Introduction

This protocol describes the detailed procedures for proteome-level PPI profiling using freshly produced proteins without the need for protein purification. Basically, NAPPA arrays displaying thousands of human candidate target proteins are probed with a query protein fused to a HaloTag that allows the query to be visually detected by binding to fluorescently labeled ligand probes.

This protocol consists of 5 components: (1) Query gene cloning: insertion of the gene encoding the query protein into the pJFT7_nHALO_DC expression vector (Fig. 1A); (2) Query protein synthesis: IVTT of the expression vector to produce HaloTagged query protein (Fig. 1B); (3) Target protein synthesis: expression of target gene DNA spotted on NAPPA using IVTT (Fig. 1B); (4) PPI profiling: incubation of HaloTagged query protein with immobilized target proteins on the array and detection of interactions using an Alexa Fluor 660-conjugated HaloTag ligand (Fig. 1B); (5) Imaging and data analysis: Extraction of fluorescence intensity information from each spot on the array and identification of candidate target proteins using appropriate quality criteria.

Materials List

Reagents

- SuperBlock blocking buffer (Pierce, cat. no. 37535)
- Sodium chloride (NaCl, Sigma-Aldrich, cat. no. S3014)
- Sodium phosphate dibasic (Na₂HPO₄, Sigma-Aldrich, cat. no. S3264)
- Potassium chloride (KCl, Sigma-Aldrich, cat. no. P9541)
- Potassium phosphate monobasic (KH₂PO₄, Sigma-Aldrich, cat. no. P5655)
- Magnesium chloride (MgCl₂, EMD, cat. no. MX0045)
- Hydrogen chloride (HCl, Sigma-Aldrich, cat. no. 295426–227G)
 - NOTE: HCl is a strong corrosive. Wear a safety face shield, gloves and a lab coat. Prepare the solution in a chemical fume hood.
- Bovine Serum Albumin (Sigma-Aldrich, cat. no. A7030–100G)
- Tween 20 (Sigma-Aldrich, cat. no. P1379)

- DTT (Sigma-Aldrich, cat. no. 43819–5G)
- DEPC water (Ambion, cat. no. 9906)
- S.O.C medium (Thermo Fisher Scientific, cat. no. 15544034)
- Ampicillin sodium salt (Sigma-Aldrich, cat. no. A9518–25G)
- Gateway LR Clonase Enzyme mix (Thermo Fisher Scientific, cat. no. 11791019)
 - *NOTE:* The Gateway LR Clonase should be stored at –20 °C. Always keep on ice when in use.
- • MAX Efficiency *E. coli* DH5-alpha competent cells (Thermo Fisher Scientific, cat. no. 18258012)
 - *NOTE:* The MAX Efficiency *E. coli* DH5-alpha competent cells should be stored at –80 °C.
- QIAprep Spin Miniprep kit (Qiagen, cat. no. 27106)
- Nucleobond xtra Maxi kit (Macherey-Nagel, cat. no. 740 414.50)
- Human HeLa lysate-based cell-free expression system (1-step Human Coupled IVT Kit, Thermo Fisher Scientific, cat. no. 88881)
 - *NOTE:* The human cell-free expression system should be stored at –80 °C and limited to three freeze/thaw cycles. Aliquot if needed. Always keep on ice when use.
- 4X Laemmli sample buffer (Bio-Rad, cat. no. 1610747)
- Alex Fluor 660-HaloTag ligand (Promega, cat. no. G8471)
- Mouse anti-GST antibody (26H1) (Cell Signaling Technology, cat. no. 2624)
- Alexa Fluor 647-goat anti-mouse IgG (Thermo Fisher Scientific, cat. no. A-21235)
- NAPPA arrays printed with target ORF sequences in pANT7_cGST expression vector are available from NAPPA Protein Array Core (non-profit) (<http://nappaproteinarray.org/>). Customized NAPPA arrays can also be fabricated by the users according to the published protocol [2, 7–9] and tutorial video (<https://www.youtube.com/watch?v=Ur1fg9jQv04>) [8].
 - *NOTE:* The NAPPA arrays can be stored under anhydrous conditions for up to 6 months at room temperature (23 °C).
- Individual human ORF clones in different vectors including pDONR221, pJFT7_nHALO_DC, pANT7_cGST can be obtained from DNASU (<https://dnasu.org/DNASU/>) [14].
- Entry/expression vectors including pDONR221, pJFT7_nHALO_DC, pJFT7_nHALO_empty, pANT7_cGST can be obtained from DNASU.
- Sequencing primers for pJFT7_nHALO_query_gene are shown as follows:

- Forward: AAGCCTGCCTAACTGCAA
- Reverse: TTTTGTTTAACTACCACTTT

Solutions

Use Milli-Q purified water or equivalent in all recipes and protocol steps.

- PBS, 10×
 - 1.37 M NaCl
 - 100 mM Na₂HPO₄
 - 27 mM KCl
 - 18 mM KH₂PO₄
 - Adjust pH to 7.4 with HCl
- PBST, 1×
 - 1× PBS
 - 0.2% (vol/vol) Tween 20
 - pH 7.4
- PPI blocking buffer
 - 1× PBS
 - 1% (vol/vol) Tween 20
 - 1% (wt/vol) BSA
 - pH 7.4
 - Prepare fresh before use
- PPI washing buffer
 - 1× PBS
 - 5 mM MgCl₂
 - 0.5% (vol/vol) Tween 20
 - 1% (wt/vol) BSA
 - 0.5% DTT
 - pH 7.4
 - Prepare fresh before use

Equipment

- HBW HybriWell sealing system (Grace Bio-Labs, cat. no. 440904)
- Corning 245 mm Square BioAssay Dish (Fisher Scientific, cat. no. 06-443-22)

- CELLSTAR FourWell plate (VWR, cat. no. 30617-596)
- Slotted rack for microscopic glass sides (Sigma-Aldrich, cat. no. Z710989)
- StainTray slide staining system (Sigma-Aldrich, cat. no. Z670146)
- EchoTherm programmable chilling/heating incubator (Torrey Pines Scientific, cat. no. 11-680-35)
- VWR Rocking Platform Shaker (VWR, cat. no. 4000-304)
- ELGA LabWater purification system (ELGA, cat. no. PURELAB Option S-R 7-15)
- Allegra X-12 Benchtop Centrifuge (Beckman Coulter, cat. no. Allegr X-12R)
- New Brunswick Innova 44 Incubator Shaker (Eppendorf, cat. no. M1282-0000)
- Thermal Shaker for Microtubes (BIOTANG, cat. no. MS-100)
- Tecan PowerScanner (Tecan)
- Typhoon FLA 9500 Scanner (GE Healthcare, cat. no. 28-9969-43)
- Array-Pro analyzer, version 6.3 (Media Cybernetics)
- Excel 2013 software (Microsoft)

Steps and Annotations

Construct the query gene expression vector

1. Obtain pDONR221_query_gene expression vector or construct by cloning the gene encoding the query protein into the Gateway entry vector pDONR221 [Curr. Protoc. Mol. Biol. UNIT 3.20 (Park, Throop and LaBaer, 2015)].

Besides pDONR221, other common entry vectors supporting recombinational cloning, such as pDONR201 and pDONR223, can also be used.

More than ~32,600 genes in pDONR221 vector are available from DNASU. Check availability before cloning.
2. Obtain the destination expression vector pJFT7_nHALO_DC from DNASU [10].

pJFT7_nHALO_DC adds an N-terminal HaloTag to recombinant proteins. It has an in-frame stop codon and can work with a query gene with or without a stop codon in pDONR221 entry vector. If using a different expression vector without an in-frame stop codon, the gene cloned into the pDONR221 entry vector must have its own stop codon. If using an expression vector with a C-terminal tag, the gene in the pDONR221 entry vector must not have a stop codon.

There are ~6,000 human genes in pJFT7_nHALO_DC vector in DNASU. Check availability before cloning.

3. Prepare Gateway LR reaction mix as follows in a chilled tube, and incubate for 1 hr at 25 °C.

Name	Volume (μL)
Nuclease-free water (cold)	6
pDONR221_query_gene (150 ng/μL)	1
pJFT7_nHALO_DC (150 ng/μL)	1
Gateway LR Clonase (2 μg/μL)	2
Total per query gene	10

Mix fresh just before use by pipetting up and down several times. Do not vortex the reaction mix, and be gentle when operating. Keep all reagents on ice.

4. Put the tube containing Gateway LR reaction mix on ice for at least 5 min.
5. Transfer 5 μL of Gateway LR reaction mix to a chilled tube, add 40 μL of *E. coli* DH5- α competent cells and keep on ice for 30 min.
6. Heat shock the competent cells in a 42 °C water bath for 45 s, and then keep on ice for at least 2 min.
7. Add 200 μL of S.O.C medium and incubate with agitation at ~700 rpm for 30 min at 37 °C.
8. Pre-warm the LB agar plate supplemented with 100 μg/mL ampicillin at 37 °C.
9. Spread the mixture on the LB agar plate using a sterile applicator and grow overnight at 37 °C.
10. Pick single colonies with pipette tips, inoculate in 5 mL of LB medium supplemented with 5 μL of 100 mg/mL ampicillin, and culture overnight at 250 rpm at 37 °C in a ground shaker.
11. Dispense 300 μL of culture into 700 μL of 80% glycerol and store at –80 °C as a glycerol stock for future use.
12. Miniprep the query gene using the QIAprep Spin Miniprep kit and quantify DNA concentration with the NanoDrop Spectrophotometer.

While both DNA miniprep and maxiprep are able to produce DNA of good quality as well as adequate concentration for sequencing or cloning purpose, the former is more time-efficient and can be done for a relatively large amount of DNA samples each time. Therefore, DNA miniprep is a better option for sequencing confirmation of query gene.

13. Confirm the sequence of the miniprep query gene using forward and reverse primers (see Reagents) for identity confirmation.
14. Maxiprep the query gene using the Nucleobond xtra Maxi kit, and quantify DNA concentration with the NanoDrop Spectrophotometer and store at –20 °C.

Compared to DNA miniprep, maxiprep produces DNA of higher concentration, which can be used for PPI assay on NAPPA and other downstream validation studies to ensure consistency in terms of DNA quality and concentration.

- 15.** Express the gene and check protein level using Alexa Fluor 660-HaloTag ligand [10].

Besides confirmation by sequencing, it is recommended to confirm the correct size as well as to assess the production of the query protein from the expression vector before performing the PPI assay.

- a.** Prepare 4 μL of IVTT mix as follows, gently mix with 1 μL of 200 ng/ μL pJFT7_nHALO_query_gene and incubate at 30 °C for 2 hr in the EchoTherm programmable incubator.

Name	Volume (μL)
HeLa lysate	2.5
Accessory proteins	0.5
Reaction mix	1
Total per query gene	4

The volume showed here is just for one query gene sample and can be scaled up accordingly. The reaction can be prepared in PCR tube for a very small volume.

Mix fresh just before use by pipetting up and down several times. Incubation at 30 °C is key to the efficient protein expression.

- b.** Add 2.5 μL of 4 μM Alexa Fluor 660-HaloTag ligand to the expression mixture and incubate for 20 min at room temperature.

Cover the tube with black shields to keep the light out. Alexa Fluor 660 is light sensitive.

- c.** Add 2.5 μL of 4X Laemmli sample buffer and boil the mixture for 5 min at 95 °C.

- d.** Run 10 μL of the resulting sample on SDS-PAGE until satisfactory separation. Include appropriate molecular weight markers.

Cover the gel apparatus with black shields to keep the light out.

- e.** Rinse the gel with Milli-Q water 5 times.

Minimize light exposure during washes.

- f.** Image the gel with the Typhoon FLA 9500 Scanner using 630/670 nm ex/em and 450 V PMT voltage.

- g. Compare the band mobility against the expected molecular weight of query protein plus HaloTag.

Ideally, a single band of the correct molecular weight is observed on the gel image. The molecular weight of HaloTag is ~34 kDa.

Express target genes on NAPPA

1. Obtain or prepare NAPPA arrays.

NAPPA with more than 10,000 sequence-verified, full-length human open reading frames in pANT7_cGST vector can be obtained from NAPPA Protein Array Core in standard or customized format. If homemade arrays are desired, refer to [2, 7–9] for NAPPA manufacture procedures. Follow Steps 1–15 to clone target genes of interest into the expression vector pANT7_cGST if not available in DNASU.

When working with a new set of NAPPA arrays, it is recommended to assess the target protein display with anti-GST antibody (CST, 26H1) [1, 2].

In addition to using individual arrays for each HaloTagged query protein (pJFT7_nHALO_query_gene), an additional array should also be included for the HaloTag only (pJFT7_nHALO_empty) as a negative control.

2. Transfer arrays to a CELLSTAR FourWell plate, add ~3 mL of SuperBlock and incubate for 1 hr at room temperature on a rocker for 60 tilts per min.
3. Place arrays in a slotted slide rack, briefly rinse with Milli-Q water, and remove excess liquid with centrifugation at 1,000*g* for 2 min at 4 °C.

Continue to the next step immediately after centrifugation.

4. Carefully align the HybriWell gasket to the slide surface at the top, slowly apply to the slide and seal by rubbing the adhesive areas with a wooden stick.

Avoid touching the HybriWell gasket to the printed spots on arrays.

Make sure the printing area is in the center of the chamber. Do not press down too hard, otherwise it will be difficult for the IVTT mix to flow through the HybriWell chamber.

5. Prepare 160 µL of IVTT mix per slide as follows, and slowly inject from the entry port on the HybriWell gasket into the chamber.

Name	Volume (µL)
HeLa lysate	80
Accessory proteins	16

Name	Volume (μL)
Reaction mix	32
Nuclease-free water	32
Total per slide	160

Both ports of the HybriWell gasket must be open for the IVTT mix to flow properly. Make sure the IVTT mix spreads over the entire array to get similar yields of all proteins. Remove air bubbles generated during injection out of the ports by gently tapping or massaging the HybriWell gasket. The presence of air bubbles may affect the production efficiency and display of proteins on NAPPA arrays.

Mix fresh just before use by pipetting up and down several times. Any remaining IVTT mix can be stored at -80°C for <1 month. However, the solution should be used within three freeze/thaw cycles to avoid a substantial loss in activity.

6. Seal both ports with the round stickers using tweezers.
Inadequate sealing will lead to significant artifacts on the arrays.
7. Place arrays on a Corning Square BioAssay Dish and incubate for 1.5 hr at 30°C , followed by 30 min at 15°C in the EchoTherm programmable incubator.
Incubation at 30°C is key for efficient protein production.
8. Remove the HybriWell gasket and wash arrays for 5 min with ~5 mL of PBST 3 times in a CELLSTAR FourWell plate on a rocker.
9. Place arrays in a slotted slide rack, thoroughly rinse with Milli-Q water 15 times.

Produce query protein in vitro

1. Mix 20 μL of 500 ng/μL pJFT7_nHALO_query_gene or pJFT7_nHALO_empty with 180 μL of IVTT mix (see Step 15 (a)) per slide and incubate for 2 hr at 30°C in the EchoTherm programmable incubator.
The control vector pJFT7_nHALO_empty supports in vitro production of HaloTag without the query protein. It will be used as a negative control to detect protein interactions mediated by the tag (see Step 29). Store the freshly synthesized query proteins on ice before use (at Step 29).

Profile protein-protein interactions on NAPPA

1. Transfer arrays (from Step 24) into a CELLSTAR FourWell plate, add ~3 mL of PPI blocking buffer and incubate for 2 hr at 4°C on a rocker for 60 tilts per min.
2. Place arrays in a slotted slide rack, briefly rinse with Milli-Q water, and remove excess liquid with centrifugation at $1,000g$ for 2 min at 4°C .
Continue to the next step immediately after centrifugation.

3. Carefully align the HybriWell gasket to the slide surface at the top, slowly apply to the slide and seal by rubbing the adhesive areas with a wooden stick.
4. Slowly inject 180 μ L of query protein or HaloTag IVTT mixture (from Step 25) from the entry port on the HybriWell gasket into the chamber.

Incubate one array with HaloTagged query protein and one with HaloTag only as a control. In our experience, the HaloTag interacts only weakly with human proteins. However, it is highly recommended to probe one array with HaloTag as a negative control when working with a new set of target-protein arrays in order to assess background signals for data analysis purpose (see Step 50 (b)).

Make sure the query fluid spreads over the entire array. Remove air bubbles generated during injection out of the ports by gently tapping or massaging the HybriWell gasket. The presence of air bubbles may affect the protein-protein interactions.

5. Seal both ports with the round stickers using tweezers.

Inadequate sealing will lead to significant artifacts on the arrays.

6. Place arrays on a Corning Square BioAssay Dish and incubate for 16 hr at 4 $^{\circ}$ C.
7. Remove the HybriWell gasket and gently wash arrays for 5 min with $\sim$ 1 mL of PPI washing buffer 3 times in a CELLSTAR FourWell plate on a rocker to remove unbound molecules.
8. Place arrays in a slotted slide rack, briefly rinse with Milli-Q water, and remove excess liquid with centrifugation at 1,000g for 2 min at 4 $^{\circ}$ C.

Continue to the next step immediately after centrifugation.

9. Carefully align the HybriWell gasket to the slide surface at the top, slowly apply to the slide and seal by rubbing the adhesive areas with a wooden stick.
10. Prepare 12.5 μ M Alex Fluor 660-HaloTag ligand in PPI washing buffer, and slowly inject 180 μ L from the entry port on the HybriWell gasket into the chamber.

Make sure the HaloTag ligand solution spreads over the entire array. Remove air bubbles generated during injection out of the ports by gently tapping or massaging the HybriWell gasket. The presence of air bubbles may affect the binding of HaloTag ligand to HaloTag and therefore protein interactions.

11. Seal both ports with the round stickers using tweezers.

Inadequate sealing will lead to significant artifacts on the arrays.

12. Place arrays on the StainTray slide staining system and incubate in the dark for 2 hr at 4 $^{\circ}$ C.

Cover the StainTray slide staining system with black shields to keep light out.

13. Remove the HybriWell gasket and gently wash arrays in the dark for 5 min with ~1 mL of PPI washing buffer 3 times in a CELLSTAR FourWell plate on a rocker.

Cover the CELLSTAR FourWell plate with black shields to keep light out. Minimize light exposure during washes.

14. Place arrays in a slotted slide rack, briefly rinse with Milli-Q water, and remove excess liquid with centrifugation at 1,000g for 2 min at 4 °C.
15. Scan the slides using the Tecan PowerScanner with appropriate settings (Resolution 10 µm; Channel 2: 676/37; without autogain).

With 'autogain', an automatic gain will be assigned for each individual slide, and the scanning parameters may change from slide to slide.

Herein the 'autogain' is not recommended.

16. Save the array images in 16-bit TIFF format.

Select target candidates

1. Open array images in the Array-Pro Analyzer microarray software and examine the spot shape, dust and nonspecific binding to remove any possible false positive signals.

Most commercial microarray analyzing software can be used for this work, such as ScanArray Express (PerkinElmer) and GenePix (Molecular Devices).

2. Quantify the average fluorescence signal intensity from microarray spots in an array image and produce a data file (.csv) using the Array-Pro Analyzer.
3. Open the data file (.csv) using the program Microsoft Excel 2013.
4. Estimate the background from nonspecific binding of probes by the first quartile of the printing-buffer-only spots as a negative control.
5. Adjust the raw signal intensity of each spot by subtracting the nonspecific binding background.
6. Normalize signal by dividing the background-adjusted signal intensity of each spot by the median background-adjusted value of all proteins on array.

The aim of data normalization is to decrease the background variation between individual arrays.

7. Calculate Z-score using the normalized value for each spot on all arrays.
8. Select target candidates based on the following criteria:
 - a. Z-score greater than or equal to 3;
 - b. Z-score ratio of query to the negative control (HaloTag) higher than 1.5;

- c. The targets have to meet the previous criteria in two independent experiments;
- d. In addition, we also selected some potential candidates based on visual inspection of the luminous radiation (“ring/halo”) around the spots.

Previous work showed that these rings are best identified using the Array-Pro Analyzer by adjusting the contrast of microarray image, especially when the Z-score is low [6, 15].

Commentary

Background information

Proteins are the workhorses that undertake a variety of essential biological processes, including gene replication and expression, cell growth and proliferation, intercellular communications, and so on. While a single protein may function properly, the majority of these activities are executed and coordinated through a network of interactions, or *interactome*. Therefore, studying proteins in the context of their interacting partners is of great importance to fully understand a protein’s function and its role within the context of living systems.

Early attempts at experimental proteome-scale interactome network mapping started in the mid-1990s. In the past decade, significant steps have been taken towards the generation of comprehensive protein-protein interaction (PPI) network maps, including the first human interactome [16–22]. However, high quality protein-protein interactome datasets that emerged from these approaches were often low in coverage, and more than 80–90% of the interactions within the human proteome remain to be mapped [23, 24]. Yeast two-hybrid (Y2H) and immunoprecipitation/mass spectrometry (IP/MS) are two platforms that have been routinely applied for the high-throughput study of protein interaction networks. Despite their demonstrated feasibility, both Y2H and IP/MS technologies also have a number of disadvantages. Y2H is often troubled with false negatives due to the limitation of the repertoire of proteins that can be produced, properly folded, and transported into the nucleus, and the limited coverage of the search space of all possible pairwise interactions [16–20]. While IP/MS is invaluable to determine the constituents of protein complexes, it is not always the best method to detect direct interactions. IP/MS also suffers from the interference by abundant proteins and the failure to detect low abundance interactors or those that are not produced in a particular cell type [21, 22].

Protein arrays offer a third high-throughput platform to interrogate the interactions between the query protein and thousands of target proteins on a microscopic glass slide [1–5]. This technology has several advantages over Y2H and IP/MS. First, it will only detect binary interactions between a query protein and the displayed target proteins, whereas the detection of misleading bridging interactions is unlikely to happen. Second, because proteins are individually organized on the array, interaction results are interpreted immediately. Third, the experimental set-up to study query proteins of interest against thousands of target proteins is extremely fast and easy. Finally, it is worth noting that the same plasmids used to

produce both query proteins and target proteins on NAPPA can be immediately deployed for more detailed interaction studies once a new interaction has been detected on the array.

Critical parameters

Our generic protocol works well for many query proteins we have tested and should serve as the starting point for assay optimization. Critical parameters affecting PPIs include the buffer components, query protein concentrations, and incubation time and temperature. Optimization of these parameters requires a balanced approach and depends on specific experimental goals. High salt/detergent concentrations and excessive washing will reduce background, but also diminish weak signals. High incubation temperature and long incubation time may result in high background, but it also gives high signals that may be needed to detect weak interactions. Short incubation time at a low temperature may result in high signal to noise (S/N) ratio if only strong interactions are of interest. Similarly, high query protein concentration will lead to strong signals and, probably, high background. On the other hand, low concentration may only reveal strong interactions. To obtain a strong S/N ratio, a balance needs to be reached by optimizing the parameters discussed above. Apart from these general considerations of PPIs, individual query proteins may have unique requirements for buffer, temperature and tag configuration to preserve protein conformation for native interactions.

Some protocol steps where vulnerable reagents are involved should be performed with special care. For example, the HeLa cell lysate-based IVTT system will lose its activity due to repeated freeze-thaw cycles and needs to be aliquoted. Furthermore, cell-free protein production requires stringent control of temperature for maximal efficiency.

Troubleshooting

Troubleshooting advice can be found in Table 1.

Anticipated results

PPI profiling on NAPPA should ideally result in the detection of one or more interacting candidates. In other words, interactions between the query and select target proteins should lead to bright microarray spots on the array. The relative fluorescence intensities of spots correlates with the binding strength between interaction partners. A higher intensity, at times with the luminous radiation (“ring/halo”) around the spot, is observed primarily for a stronger interaction. A ring is observed when the query protein binds to target proteins that “bleed” into the neighboring areas around spots during protein production and immobilization. This signal information is subject to data quantification and analysis, eventually leading to identification of interacting partners for the query protein of interest.

Figure 2 shows the representative results of PPI profiling of a HaloTagged version of the *Legionella pneumonia* effector LidA as a query protein against ~2,000 human target proteins on NAPPA array. Interactions between Rab1B and Rab27B with LidA are visually distinctive and have the highest Z-scores among all targets on arrays (Fig. 2).

Time considerations:

The time required for the entire procedure depends on the experimental design, such as the number of query proteins to be studied and the number of NAPPAs to be probed. For construction of pJFT7_nHAO_query_gene expression vector, Gateway LR reaction and transformation require 2.5 hr, followed by plating on agar and overnight incubation. The *in vitro* generation of target proteins on NAPPAs, including blocking and IVTT mix incubation, takes ~4 hr. The PPI profiling on NAPPAs, including blocking, query protein binding, and Alexa Fluor 660-HaloTag ligand detection, takes ~21 hr, much of which is time for incubation. Scanning NAPPAs requires ~7 min per slide with one setting cycle. Data quantification and candidate selection requires 12 hr.

Literature cited

- [1]. Ramachandran N, Hainsworth E, Bhullar B, Eisenstein S, Rosen B, Lau AY, et al. Self-assembling protein microarrays. *Science* 2004;305:86–90. [PubMed: 15232106]
- [2]. Ramachandran N, Raphael JV, Hainsworth E, Demirkan G, Fuentes MG, Rolfs A, et al. Next-generation high-density self-assembling functional protein arrays. *Nat Methods* 2008;5:535–8. [PubMed: 18469824]
- [3]. Ramani SR, Tom I, Lewin-Koh N, Wranik B, Depalatis L, Zhang J, et al. A secreted protein microarray platform for extracellular protein interaction discovery. *Anal Biochem* 2012;420:127–38. [PubMed: 21982860]
- [4]. Jones RB, Gordus A, Krall JA, MacBeath G. A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature* 2006;439:168–74. [PubMed: 16273093]
- [5]. Kaushansky A, Gordus A, Budnik BA, Lane WS, Rush J, MacBeath G. System-wide investigation of ErbB4 reveals 19 sites of Tyr phosphorylation that are unusually selective in their recruitment properties. *Chem Biol* 2008;15:808–17. [PubMed: 18721752]
- [6]. Yu X, Decker KB, Barker K, Neunuebel MR, Saul J, Graves M, et al. Host-pathogen interaction profiling using self-assembling human protein arrays. *J Proteome Res* 2015;14:1920–36. [PubMed: 25739981]
- [7]. Qiu J, LaBaer J. Nucleic acid programmable protein array a just-in-time multiplexed protein expression and purification platform. *Methods Enzymol* 2011;500:151–63. [PubMed: 21943897]
- [8]. Miersch S, LaBaer J. Nucleic Acid programmable protein arrays: versatile tools for array-based functional protein studies. *Curr Protoc Protein Sci* 2011;
- [9]. Sibani S, LaBaer J. Immunoprofiling using NAPPAs protein microarrays. *Methods Mol Biol* 2011;723:149–61. [PubMed: 21370064]
- [10]. Saul J, Petritis B, Sau S, Rauf F, Gaskin M, Ober-Reynolds B, et al. Development of a full-length human protein production pipeline. *Protein Sci* 2014;23:1123–35. [PubMed: 24806540]
- [11]. Katzen F, Peterson TC, Kudlicki W. Membrane protein expression: no cells required. *Trends Biotechnol* 2009;27:455–60. [PubMed: 19616329]
- [12]. Yu X, LaBaer J. High-throughput identification of proteins with AMPylation using self-assembled human protein (NAPPAs) microarrays. *Nat Protoc* 2015;10:756–67. [PubMed: 25881200]
- [13]. Karthikeyan K, Barker K, Tang Y, Kahn P, Wiktor P, Brunner A, et al. A Contra Capture Protein Array Platform for Studying Post-translationally Modified (PTM) Auto-antigenomes. *Mol Cell Proteomics* 2016;15:2324–37. [PubMed: 27141097]
- [14]. Seiler CY, Park JG, Sharma A, Hunter P, Surapaneni P, Sedillo C, et al. DNASU plasmid and PSI: Biology-Materials repositories: resources to accelerate biological research. *Nucleic Acids Res* 2014;42:D1253–60. [PubMed: 24225319]
- [15]. Montor WR, Huang J, Hu Y, Hainsworth E, Lynch S, Kronish JW, et al. Genome-wide study of *Pseudomonas aeruginosa* outer membrane protein immunogenicity using self-assembling protein microarrays. *Infect Immun* 2009;77:4877–86. [PubMed: 19737893]

- [16]. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, et al. A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 2000;403:623–7. [PubMed: 10688190]
- [17]. Calderwood MA, Venkatesan K, Xing L, Chase MR, Vazquez A, Holthaus AM, et al. Epstein-Barr virus and virus human protein interaction maps. *Proc Natl Acad Sci U S A* 2007;104:7606–11. [PubMed: 17446270]
- [18]. Shapira SD, Gat-Viks I, Shum BO, Dricot A, de Grace MM, Wu L, et al. A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell* 2009;139:1255–67. [PubMed: 20064372]
- [19]. Rozenblatt-Rosen O, Deo RC, Padi M, Adelmant G, Calderwood MA, Rolland T, et al. Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. *Nature* 2012;487:491–5. [PubMed: 22810586]
- [20]. Tafforeau L, Rabourdin-Combe C, Lotteau V. Virus-human cell interactomes. *Methods Mol Biol* 2012;812:103–20. [PubMed: 22218856]
- [21]. Krogan NJ, Cagney G, Yu H, Zhong G, Guo X, Ignatchenko A, et al. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* 2006;440:637–43. [PubMed: 16554755]
- [22]. Jager S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, Clarke SC, et al. Global landscape of HIV-human protein complexes. *Nature* 2012;481:365–70.
- [23]. Venkatesan K, Green J, Shapiro SR, Steinhardt GF. Correlation of hydronephrosis index to society of fetal urology hydronephrosis scale. *Adv Urol* 2009;960490.
- [24]. Tyagi M, Hashimoto K, Shoemaker BA, Wuchty S, Panchenko AR. Large-scale mapping of human protein interactome using structural complexes. *EMBO Rep* 2012;13:266–71. [PubMed: 22261719]

Significance Statement

Proteins are the biomolecular machines that drive all important life functions. Proteins orchestrate the cellular processes of life through their interaction with and/or manipulation of other biomolecules. Identifying the interactors of proteins is essential to deciphering their biological functions. This unit provides a protocol for elucidating protein-protein interactions (PPIs) at the proteome scale on Nucleic Acid Programmable Protein Array (NAPPA) in a high-throughput multiplexed fashion without the need of purifying target or query proteins.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Flow scheme of protein interaction assay using NAPPA. (A) Construction of query vector. Homologous recombination-based cloning of the query gene of interest into the expression vector. Size, amount, and stability of query proteins can be assessed by SDS-PAGE. (B) PPI profiling on NAPPA. DNA printing and immobilization is validated with Pico-Green staining (green); production and display of GST-tagged target proteins is examined with anti-GST antibody (red) after IVTT; binding of HaloTagged query protein to its interactors on NAPPA is detected using Alexa Fluor 660-conjugated HaloTag ligand.

Figure 2.

A representative image of a NAPP array probed with HaloTagged LidA for identification of host targets for *L. pneumophila*. Spots of interactors are zoomed in and marked with an arrow.

Table 1.

Troubleshooting table.

Step	Problem	Possible reason	Solution
43	No/low signals	Array slides were loaded upside down in the hybridization chamber, resulting in no target hybridization of probes (Steps 19, 28 and 34).	Repeat the experiment and ensure that the hybridization chamber is covered on the slide surface with microarray spots.
		Low levels of the query protein (Step 25).	Check the human cell-free expression system for protein production <i>in vitro</i> , as it may lose its activity during shipment and storage.
			Check protein amounts using fluorophore-conjugated HaloTag ligand.
		Low production level of target proteins on NAPPA (Step 22).	Check the human cell-free expression system for protein production <i>in vitro</i> , as it may lose its activity during shipment and storage.
			Check the protein amount on NAPPA using monoclonal mouse GST-specific antibody.
		Array slides are loaded upside down in the Tecan PowerScanner (Step 40).	Load the slide again in the Tecan PowerScanner with the microarray spots facing up.
		A wrong laser is chosen for array scanning (Step 40).	Choose the laser with appropriate excitation and emission wavelengths.
45	High background	The concentration of query protein may be too high (Steps 25 and 29)	Decrease the concentration of query proteins.
			Increase the washing cycles using PPI washing buffer.
		The query protein is sticky.	Optimizing assay parameters such as a lower incubation
			temperature and a more stringent washing buffer.