LECTURE-25

Nucleic Acid Programmable Protein Arrays

Handout

PREAMBLE

Various important cell processes such as ligand-receptor reaction, hormonal activity, enzymatic catalysis, DNA replication, respiration, growth etc. occur because of the indispensible protein-protein interactions. To understand cellular processes, studying the protein-protein interactions is most important. This involves production of desired proteins and studying them *in vitro*, which is highly time-consuming and demanding process. Since protein production takes lot of time, very few proteins can be studied at a time. It is also difficult to maintain the activity of the protein produced. Moreover, the shelf life of proteins is very low. Hence, there is need to develop new methods for production of better quality proteins in high throughput manner.

One of the most revolutionizing technologies developed in the recent times is the protein microarrays. A protein microarray has a solid support, on which thousands of proteins can be spotted and studied at once. This has become possible due to the advancement of cell-free expression systems. The construction of protein arrays does not take a lot of reagents and therefore is cost-effective and large number of proteins can be screened at once. One of the cell-free expression based techniques developed for construction of protein microarrays is Nucleic Acid Programmable Protein Arrays (NAPPA), which is discussed in this lecture.

OUTLINE OF LECTURE

I. NAPPA arrays

II. Workflow of NAPPA array

III. Applications

IV. Advantages and challenges

V. Conclusions

BOX-1: Terminology

APTES (aminopropyltriethoxysilane) coating: APTES is used for coating of the glass slide so that it becomes amenable to attach DNA and proteins.

GST tag: Tag fused at C-terminal end of protein, for capture on anti-GST capture antibodies.

BS^3 Linker: A water-soluble, non-cleavable, linker with a spacer of 8 carbon atoms between two hydroxysuccinimde groups, which links the primary amine groups to help tether the antibody to the glass slide.

BSA: The bovine serum albumin enhances DNA binding in NAPPA chemistry.

BOX-2: Recombinational cloning

Recombinational cloning: A novel site-specific recombination technique for transferring DNA sequences, which allows one universal strategy to move DNA sequence to any vector. The recombination cloning is a promising approach for high throughput genomics and proteomics applications.

Site-specific recombination: A genetic recombination technique where DNA strand exchange takes place between regions possessing reasonable degree of sequence homology. Specific recombinase enzymes cleave the DNA backbone and carry out interchange of DNA helices between specific sites on two different molecules. The common site-specific recombination technologies currently in use are the GATEWAY Technology (Invitrogen) and the Creator Technology (BD Clontech).

BP reaction: A site-specific recombination reaction between the attB and attP sites leading to generation of the master or entry clones, which can further be used at any time for specific purposes (Fig 1).

LR reaction: The main reaction pathway of the GATEWAY system consisting of a recombination reaction between a master clone and a destination vector used for generation of the expression clones (Fig 1). These expression clones can be used for a variety of applications.

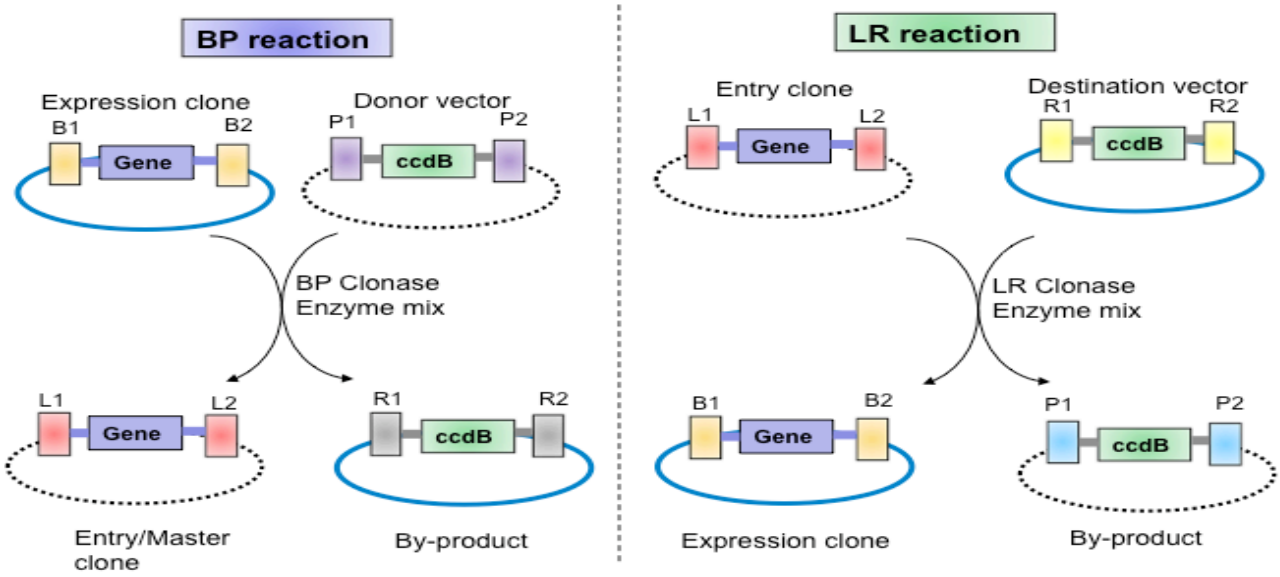


Fig 1. Recombinational cloning steps; BP and LR reactions

Gateway cloning: A powerful new recombinational cloning technology that facilitates protein expression and cloning of PCR products by using site-specific recombination enzymes rather than restriction endonucleases and ligases. This technique makes use of a master clone having a particular gene that can be rapidly transferred to desired destination vectors and thereby provides significant benefit over conventional cloning (Fig 2).

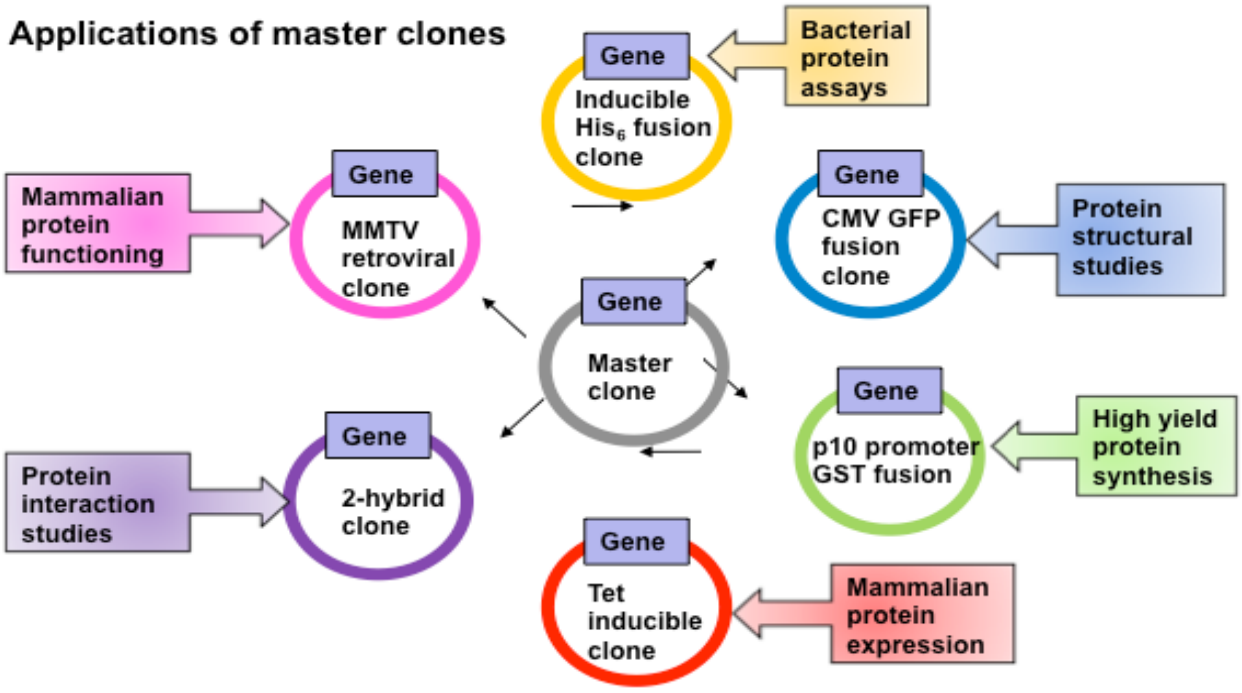


Fig 2. Master clones generated by Recombinational cloning can be used for several applications.

Illustration: Recombinational cloning

The BP reaction of GATEWAY cloning is a site-specific recombination reaction between the attB site of an expression clone or a PCR product and attP site of a donor vector in the presence of BP Clonase enzyme master mix. The reaction is incubated for just an hour at 25°C to obtain the entry or master clones containing the gene of interest. Once this master clone, flanked by attL sites is produced, it can then be transferred into any destination vector to produce expression clones for a specific desired application. The reaction yields more than 90% correct clones.

The LR reaction is essentially the reverse of the BP reactions where the master clone, flanked by attL sites, recombines with a destination vector with attR sites. This reaction takes place in the presence of the LR Clonase enzyme mix and results into transfer of the gene from the master clone to the destination vector to produce an expression clone for a specific purpose. This reaction enables generation of several expression clones for various applications in very short time, thereby providing significant advantage over conventional cloning techniques.

The gene in the master clone can be transferred to various destination vectors by means of the LR reaction to produce expression clones for several applications. Proteins can be efficiently expressed in bacterial, yeast and mammalian systems and used for a variety of applications such as structural & functional studies, protein interaction studies, protein assays, producing high yields of proteins for experimentation etc. The rapid recombination between clones that is possible with the GATEWAY system cannot be done with conventional cloning techniques due to which the GATEWAY protocol is now being extensively adopted.

I. NAPPA – Nucleic Acid Programmable Protein Microarrays

Protein microarrays consist of solid surfaces on which the proteins of interest are spotted. Earlier, each protein used for spotting on the slide surface using chemical techniques. This required manual production of each protein, which was very tedious. The self assembling protein microarray or NAPPA approach introduced by Ramachandran et al. (2004) used spotting of expression plasmids containing cDNAs of interest on the array surface and expression of *proteins in situ* by a mammalian cell free expression system at the time of assay. All proteins were expressed with fusion tags, which correspond to capture agents printed along with the plasmid DNA and used to capture the protein immediately after translation. By producing the proteins just-in-time for assay the opportunity for its denaturation is significantly reduced, and the use of mammalian transcription/translation system encouraged natural protein folding for mammalian protein. The NAPPA microarray is a highly innovative cell-free expression based technology, which helps in production of thousands of proteins simultaneously, and also capturing them on the slide to form the array as and when they are formed (Ramachandran *et al.* 2004).

In NAPPA, cDNA (containing a tag, usually GST) of the desired proteins are spotted along-with BSA, BS^3 and capture antibody (anti-GST antibody) on functionalized slide (Fig 3). For activation of the array, a cell-free expression mixture containing *in vitro* transcription and translation mix (rabbit reticulocyte lysate, T7 polymerase, amino acid mixtures, RNase inhibitor etc.) is added onto the slide. This leads to the production of desired proteins from the cDNA. The protein produced

contains a tag, which binds to the capture agent coated on the slide. Thus, a protein replica is formed in place of the DNA array (Fig 4).

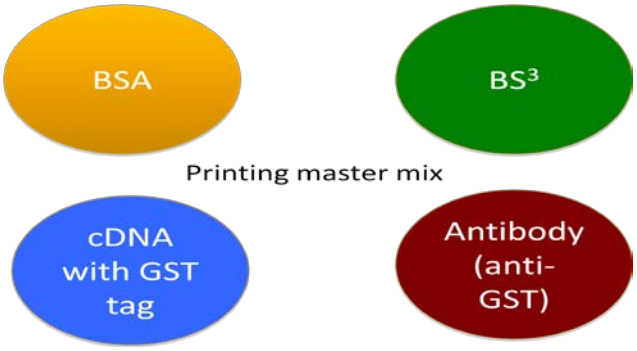


Fig 3. Master mix for NAPPA printing consists of cDNA with GST tag, BS3 cross linker, BSA and capture antibody (anti-GST).

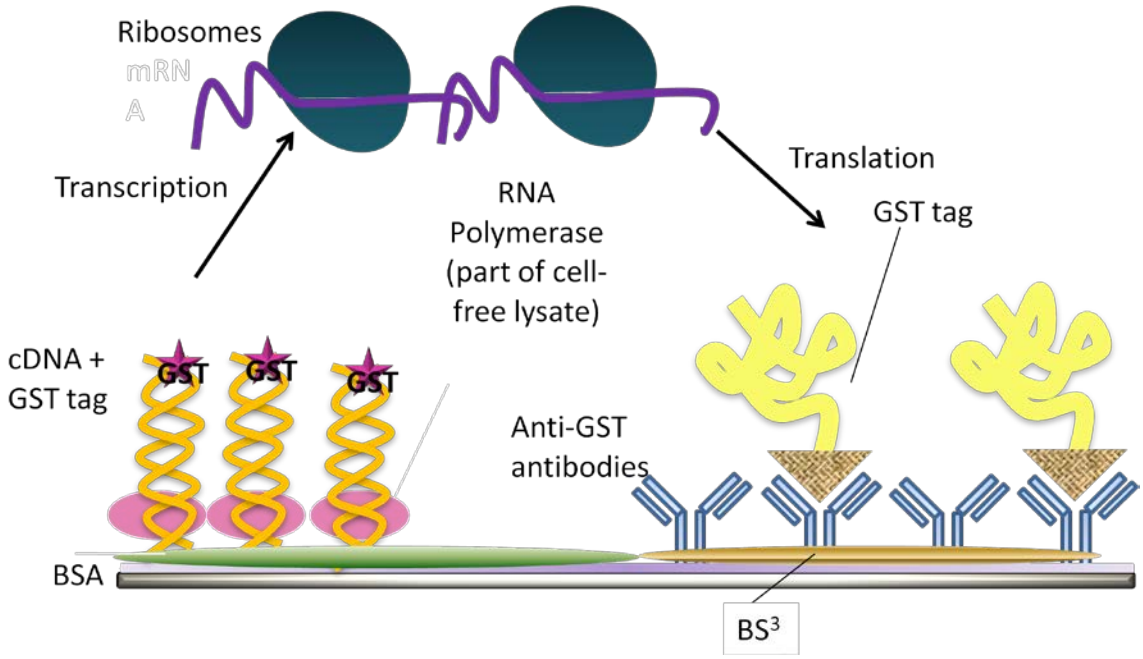


Fig 4. Working principle of Nucleic Acid Programmable Protein Arrays. The master-mix containing cDNA (with GST tag), BSA, BS^3 and anti-GST antibody are printed on array surface. Adding RRL to the arrays carries out protein expression. The newly synthesized proteins are captured by antibody through the GST tag and protein microarray is produced.

Illustration: Nucleic Acid Programmable Protein Array (NAPPA)

An aminosilane-coated glass slide forms the array surface for NAPPA. To this, the NAPPA master mix is added which consists of BSA, BS3, GST-tagged cDNA and the anti-GST capture antibodies. The BSA improves efficiency of immobilization of the cDNA onto the array surface while the BS3 cross-linker facilitates binding of the capture antibody. The cDNA is expressed using a cell-free extract to give the corresponding protein with its GST tag fused to it. This tag enables capture of the protein onto the slide by means of anti-GST antibodies. NAPPA technique can generate very high-density arrays but the protein remains co-localized with cDNA.

II. WORKFLOW OF NAPPA

To ensure construction of NAPPA protein microarray, one requires careful design of each of the components. The design and role of each of the components of NAPPA work-flow are as follows:

Preparation of master-mix for printing

The solid support for printing the array is generally a glass slide. But other supports such as gold, nitrocellulose, hydrogel also may be used. Depending upon the surface of the microarray, the surface chemistry for printing cDNA differs.

- The glass slide is coated with APTES (aminopropyltriethoxysilane or aminosilane) and it contains a large amount of positively charged amino groups, which bind to the negatively charged phosphate groups on DNA (Fig 5). The slides are exposed to UV or baked at 85°C enabling the strong covalent attachment of DNA on the silane coated slides.

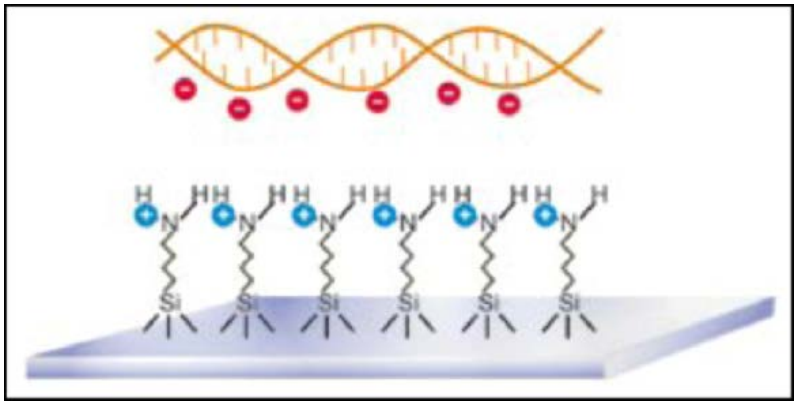


Fig. 5. Aminosilane coating of the glass slide. The positive groups of amine on the slide surface covalently bind to the negatively charged phosphate groups of DNA, which enables immobilization of DNA molecules onto the slide surface.

- If the cDNA is biotinylated (using psoralen-biotin), the slide is coated with avidin along with aminosilane. This leads to a strong biotin-avidin binding and this interaction enables the DNA immobilization onto the slide surface.
- For protein capture, a BS^3 cross-linker and anti-tag antibody is also spotted on the silane coated glass slide. This linker helps in formation of amine-amine bonds, and thus is used for tethering the capture antibody on the glass slide.

The cDNA for the protein of interest is designed for two purposes; 1) the cDNA should be captured on the slide before protein production, and 2) it must have a tag, which helps in identification and capture of the expressed protein. The cDNA plasmids are designed to form a fusion tag at either the N or C terminal. Generally, plasmids are designed to add a C-terminal glutathione S-transferase (GST) tag.

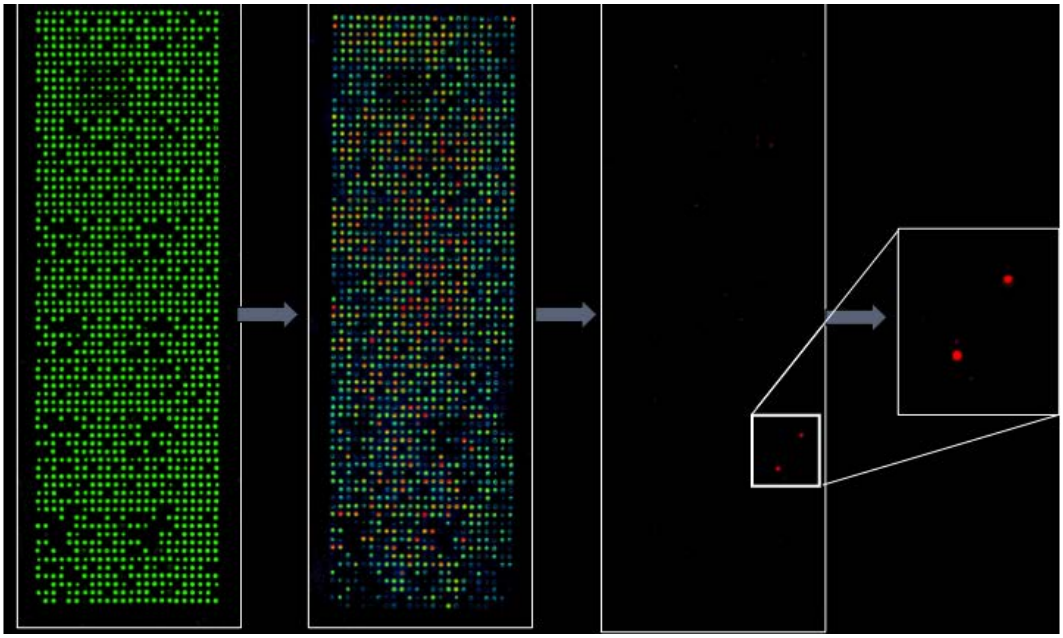
• Bovine Serum Albumin (BSA) is added to this master mix as it enhances the binding of the DNA to the slide.

• The master-mix is printed on the slide by manual spotting or using an automated microarrayer.

Protein production and capture

For protein production using *in vitro* transcription and translation system (IVTT), a cell-free expression mixture containing rabbit reticulocyte lysate, T7 polymerase, amino acid mixtures and RNase inhibitor is added onto the slide. The proteins expressed from the cDNA contain a tag (usually GST tag) at the C-terminal of the protein. The anti-GST antibody is coated on the aminosilane slide, with the BS3 cross-linker. The synthesized

protein gets immobilized to the slide by antigen-antibody reactions and protein array is formed (Fig 6).



DNA detection Protein expression Specific-protein expression Inset (expressed proteins)

Fig. 6 Quality control of NAPPA array: PicoGreen dye is used for testing DNA printing quality, Ani-GST antibody is used for testing protein expression. The protein specific antibody is used to test protein-specific expression.

III. APPLICATIONS OF NAPPA

Autoantibody biomarker screening

Antibodies to tumor antigens have advantages over other serum proteins as potential cancer biomarkers as they are stable, highly specific, easily purified from serum, and are readily detected with well-validated secondary reagents. The antibodies directed at self-antigens are referred to as autoantibodies. NAPPA arrays have been used to identify autoantibody biomarkers in sera that can be readily used for the early detection of cancers. For autoantibody screening, samples such as serum or cell lysate is added on the chip. If target antigens are present in the sample, autoantibodies bind to their targets, and this can be detected using a labeled anti-IgG antibody (Fig 7).

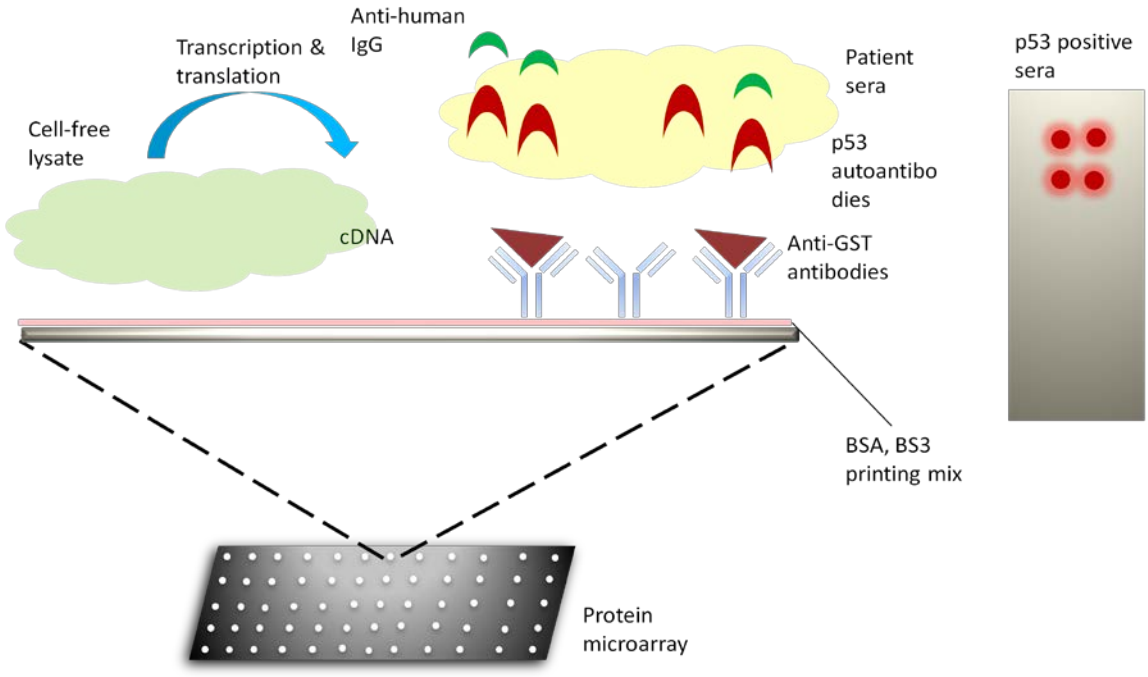


Fig 7. Detection of p53 autoantibodies using NAPPA microarrays

Illustration: NAPPA Application – biomarker discovery

The authors generated protein microarrays based on NAPPA expression, which they probed with diluted sera of breast cancer patients having p53 autoantibodies. Detection was carried out by means of HRP-linked anti-human IgG. This study detected p53 autoantibodies by means of NAPPA microarrays, which was confirmed by ELISA. The p53 levels were found to be directly related to tumour burden with serum antibody concentration decreasing after neoadjuvant chemotherapy.

Protein Interactions - One key application of NAPPA is to test protein-protein interactions (Fig 8). Typically, this is done by probing an array of proteins with a purified query protein. NAPPA is also able to employ co-expression of the target and the query protein by transcribing and translating them in the same extract. To do this, simply appropriate query DNA was added into the IVTT before applying it to the array. For fluorescent detection, the query proteins were expressed with an epitope tag that is different from the one used to capture the target proteins on the array. Following the co-expression, the binding of the query to a specific target protein was detected by using an antibody to the query epitope.

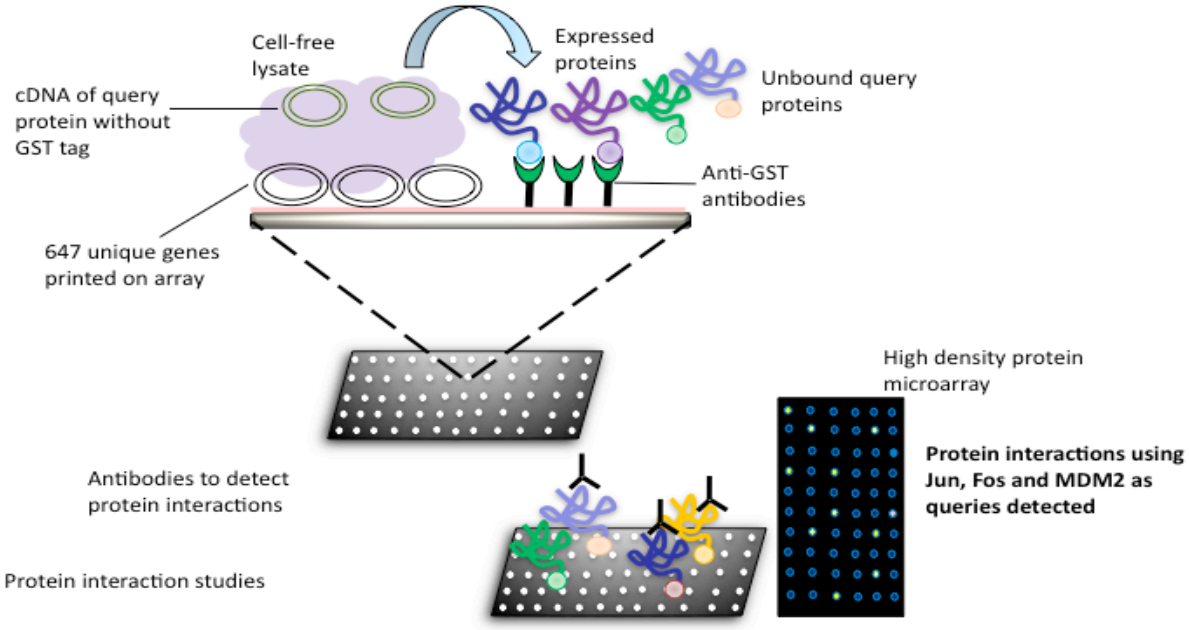


Fig 8. Protein-protein interaction study using NAPPA arrays.

Illustration: NAPPA Applications – Protein-protein interactions

Ramachandran et al., tested the use of NAPPA microarrays by immobilizing 29 sequence-verified human genes involved in replication initiation on the array surface and then expressing them in duplicate with RRL. The expressed proteins bound to the anti-GST antibodies present on the array surface. The authors made use of each of these expressed proteins to probe another duplicate array of the same 29 proteins thereby generating a 29 x 29 protein interaction matrix. 110 interactions were detected between proteins of the replication initiation complex, of which 63 were previously undetected ones.

IV. ADVANTAGES AND CHALLENGES

NAPPA is a promising cell-free expression based protein microarray technology. This technique uses mammalian expression systems and thus allows efficient protein folding. Access to a wide variety of cloned cDNAs, allows spotting of almost any protein on the array. One of the most prominent advantages of NAPPA is long shelf-life of the arrays. NAPPA arrays are very cost-effective as the volume of the cell-free lysate and the template is very low. NAPPA is very effective process as 95% of the proteins are usually expressed and captured on the slide (Fig 9).

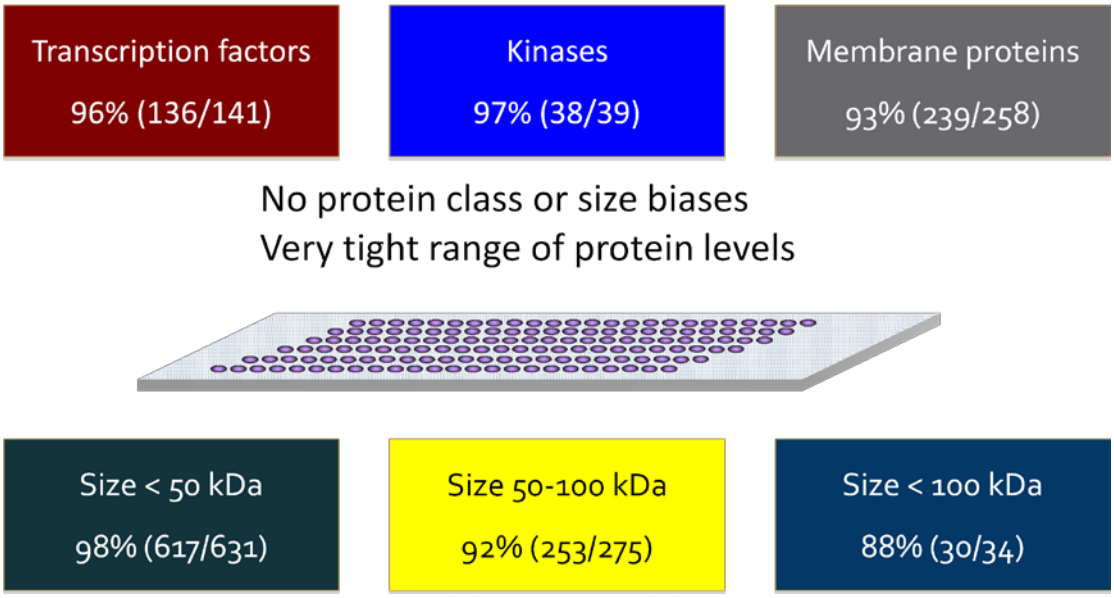


Fig 9. Success of protein expression using NAPPA protein microarrays. Protein expression in kinases, transcription factor and membrane proteins (>95%). There is no bias of low or high molecular weight protein (Ramachandran et al. 2008).

NAPPA approach offers following advantages over traditional methods:

- Replaces printing proteins with the more reliable process of printing DNA
- Avoids the need to express and purify proteins
- Avoids concerns about protein shelf life because the proteins are made fresh at the time of assay
- Protein integrity: uses mammalian machinery to synthesize proteins
- No tedious gene introduction into mammalian cells
- Interaction is not limited to context of nucleus
- Gene toxicity and auto activation of reporter genes are not an issue
- Expensive mass spectrometers are not required
- Analysis of multimeric complexes
- Post translational modifications
- Protein production and obtaining data in real time
- Sequence-verified plasmid templates provide good quality controls
- Attachment/detection scheme changeable

However, like any other technology, NAPPA has certain limitations. The process is time consuming, as cloning of each cDNA template needs to be done before generating array, or alternately it depends on availability of clones. Each of the cDNA needs to be cloned with fusion tags (e.g. GST). Another problem associated with NAPPA is that the production of pure protein arrays is not possible, as the expressed protein remains co-localized with cDNA. The peptide tags may produce steric hindrance while studying

protein interactions. Finally, It is difficult to assess the functionality of expressed proteins after *in vitro* transcription and translation steps.

V. CONCLUSIONS

NAPPA is very effective technique for production of protein microarrays. It has already been used in various studies for immunological screening, screening of biomarkers and to study protein-protein interaction. The arrays can be stored for long, as NAPPA uses DNA for printing, and proteins can be produced on demand. Integration of NAPPA microarrays with label-free detection techniques such as surface Plasmon resonance imaging would be interesting future direction for several high throughput proteomic applications.

References

- Park, J. & LaBaer, J. Recombinational Cloning. Current protocols in molecular biology, 2006, 3.20.1 – 3.20.22.
- Hartley, J.L., Temple, G.F. & Brasch, M.A. DNA cloning using in vitro site-specific recombination. Genome Res. 10, 1788–1795 (2000).
- Goshima, N. et al. Human protein factory for converting the transcriptome into an in vitro-expressed proteome. Nat methods 2008, 5(12), 1011-1017.
- Aguiar, J. C. et al. High-throughput generation of P. falciparum functional molecules by recombinational cloning. Genome Res. 2004, 14(10B), 2076-82.
- Chandra, H. & Srivastava, S. Cell-free synthesis-based protein microarrays and their applications. Proteomics 2010, 10, 1-14.
- Anderson, K. A., Ramachandran, N., Wong, J., Raphael,J. V. et al., Application of protein microarrays for multiplexed detection of antibodies to tumor antigens in breast cancer. J. Proteome Res. 2008, 7, 1490–1499.
- Ramachandran, N., Hainsworth, E., Bhullar, B., Eisenstein, S. et al., Self-assembling protein mircoarrays. Science 2004,305, 86–90.
- Ramachandran, N., Raphael, J. V., Hainsworth, E., Demirkan,G. et al., Next-generation high-density self-assembling functional protein arrays. Nat. Methods 2008, 5, 535–538.