UNIVERSITA' DEGLI STUDI DEL PIEMONTE ORIENTALE

"Amedeo Avogadro"

PhD Program

Biotechnology for Human Health

Cycle XXVI

HIGH-THROUGHPUT ANTIBODY VALIDATION PLATFORMS

SUPERVISOR:

Prof. Daniele Sblattero

PhD STUDENT:

Silvia Saragozza

TABLE OF CONTENTS

1. **ABSTRACT**

Antibodies are molecules able to specifically bind a particular antigen. Thanks to this capability they are extremely useful in many application. Indeed, antibodies are utilized in different routinary experiments, such as: Western blot, Enzyme-linked immunosorbent assay (ELISA), Immunohystochemistry, Immonofluorescence, Flow-cytometry, Immunoprecipitation, etc. Obviously, to carry out a good experiment a good antibody is needed. Indeed, many commercial available antibody aren't efficient and leads to wrong results. This problem derives from a wrong antibody selection, production and validation. This project is aimed at setting up a selection, production and validation platform for polyclonal antibodies. In particular, the polyclonal antibodies were selected, in form of scFv, with a phage/yeast display combination, produced in yeast cells as minibody with a C_{H2}-C_{H3} of rabbit IgG, validated with high-throughput technologies. A list of 78 proteins based on proteins structure and solubilization was drafted. Of these proteins only the first seven were used for scFv selection with two rounds of phage display and two additional rounds of yeast display. The selected scFvs were cloned in a yeast expression vector that allows the production of scFv fused to a rabbit IgG constant part. The collected yeast supernatants, containing the polyclonal, were sent us for the validation step. The polyclonal antibodies validation consists of three different analysis: Western blot, ELISA and protein microarray. From this validation arose that five polyclonal antibodies were highly specific and sensitive to the target protein, revealing that the entire platform is efficient. Furthermore, one of them was compared to the commercial ones using ELISA and protein microarray technologies. Only with protein microarray it could be possible observed that the commercial polyclonal presented cross-reactivity to other proteins. Instead, the yeast polyclonal resulted specific only to the target protein. This result confirmed that the selection with phage/yeast display combination is extremely specific avoiding the selection of sticky or unspecific antibodies. Moreover, protein microarray resulted a very useful and sensitive antibody validation technology. The second aim of this project concerned the use of a new technology to perform high-throughput antibody validation. Indeed, protein microarray is a successful tool, but present some drawbacks, such as production and purification of all the tested proteins and possible degradation of them once immobilized. These problems are overcome by a new technology called in situ protein array. This

technology consists of: cDNA immobilization on the slide; expression of the proteins of interest by *in vitro* transcription and translation system; capture the nascent proteins; reveal the proteins. Different types of in situ protein array have been set up, but the one used in this project is called protein array on demand. The peculiarity of this array is the presence of the capture agent directly on the template DNA. In this way only the DNA has to be printed on the slide. First, a DNA plasmid suitable for this system was constructed and subsequently tested with four different proteins available in our laboratory. Subsequently, a protein array on demand was performed, following the published method. The analysis showed that protein array on demand was less efficient and reproducible. For this reason a new protein array on demand was set up. The array slide, template and IVTT system were analyzed and tested to perform the best protein array on demand conditions. After different experiments the array on demand was performed using: CodeLink slide; biotinylated PCR DNA mixed with streptavidin; IVTT based on rabbit reticulocytes lysate. Since that the array was performed with an anti-tag antibody, an additional array was performed using an anti-GFP antibody direct against the eGFP. Even if the signal wasn't high it could be affirmed that this protein array on demand was functional and better than the previous ones. In conclusion, a new protein array on demand was set up and can be potentially used for antibody validation.

2. INTRODUCTION

2.1 THE ANTIBODY MOLECULE

Antibodies, also called immunoglobulins, are molecules produced by B cells of the immune system. The main activity of an antibody is to disrupt the pathogen presents in the organism through two different mechanisms: opsonization and neutralization. These mechanisms involves the antibody binding to a particular antigen in a strong manner. The main type of immunoglobulin used in research is the G (IgG) and it is a protein composed of two polypeptide chains link together by covalent binding [1] (Fig. 1). In every polypeptide an heavy and a light chain is present. The heavy chain is composed by three constant domains called C_{H1}, C_{H2} and C_{H3} and a variable domain called V_H. The C_{H1} and C_{H2} domains are linked together by covalent binding, forming a flexible hinge region. On the other hand, the light chain is composed only by two domains, one variable (V_L) and one constant (C_L). The antibody's part that confers the antigen specificity is the variable region (V_H and V_L), in particular three loops, called complementarity determining region (CDRs) 1,2 and 3, are responsible of the antibody diversity [2]. Finally, the constant region is responsible for glycosylation which is important for the recognition of effector molecules, such as immunoglobulin receptors or C1 complement complex.

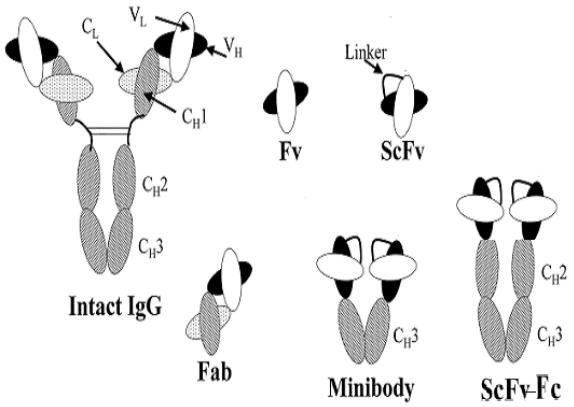


Figure 1. IgG structure and other antibody format. Structure of a full-length IgG and the possible smaller formats used in research: Fragment variable (Fv); single-chain fragment variable (scFv); Fragment-antigen-binding (Fab); minibody.

The scientific interest in antibodies arising from their ability to specifically recognize a determined target. The production of these molecules is not simple due to the higher molecular weight and post-translational modifications in the constant part. For this reason other antibody formats have been constructed maintaining the antibody specificity region (Fig. 1) [3]. The most used in research are:

1. Fragment-antigen-binding (Fab), composed by C_{H1} and V_H link with random and flexible peptide linker to V_L and C_L (50kDa);
2. single-chain fragment variable (scFv) molecules, formed by V_H and V_L link with a linker peptide of 15aa (25-30kDa);
3. fragment variable (Fv), only V_H and V_L are linked without a linker between them (25kDa);
4. minibody, formed by scFv-C_{H3} protein that self-assembled into a bivalent dimer (80kDa).;
5. ScFv-Fc, bivalent molecule composed by two scFvs fused to C_{H2}-C_{H3} IgG region (120kDa).

2.2 THE IMPORTANCE OF ANTIBODIES SPECIFICITY IN RESEARCH

The Human Genome Project, in 2000, has revealed a great number of genes that were previously unknown and subsequently shifted the attention from the genome to the proteome. The major approach to explore the whole human proteome is the use of affinity reagents, among which antibodies are the most widely used. Antibodies are able to recognize specifically a selected antigen. It is possible to produce antibody directed to a particular epitope (monoclonal) or different epitopes of the same antigen (polyclonal). Furthermore, antibodies can be easily marked with many kind of molecules such as: enzymes, fluorophores, biomolecules and drugs. All these features have lead to the use of antibodies in various research fields. Indeed, they are necessary for many standard analysis such as:

- Western blot (WB), where antibodies allow the recognition of denaturated proteins present in a lysate or protein mix;
- Enzyme-linked immunosorbent assay (ELISA), in which antibodies reveal the presence of full-lengh proteins on plastic plates;

- Immunoprecipitation (IP), where antibodies are not only used to detect the protein presence, but also to precipitated protein complexes;
- Flow-cytometry (FC), in which marked antibodies allow proteins recognition on cells or different marked antibodies can allow the separation of different cells;
- Immunohistochemistry (IHC), where tissues are stained with antibodies allowing to study tissue structure and cells distribution;
- Immunofluorescence (IF), in which tissues are depicted with antibodies conjugated with a fluorophore allowing us to compare not only protein distribution in different cells types but also cancerous cells with normal ones.

Antibodies are also employed in many new techniques. For example, they are very important in protein microarray technology, where they are used to detect the expression of proteins in different samples or cell lysates [4; 5; 6]. Recently, they have been used to address in specific tissue nanoparticles carrying drugs or other molecules [7; 8; 9]. Furthermore, antibodies are also utilized for the isolation of specific cells, such as stem cells [10; 11]. Obviously, all these analysis are based on the antibody specificity. Indeed, antibodies that recognize the wrong molecules or cross-react against other antigens will cause an incorrect result. These problems occur in the case of non-specific validation. Moreover, some analysis can't be carried out because an antibody, against the specific antigen, is not available. For this reason, another technology has been introduced for proteome analysis: mass spectrometry (MS)[12]. The main ability of MS is the analysis of proteins without affinity reagents use. On the other hand, MS is not high-throughput and can't be used in different applications, such as immunofluorescence and immunohystochemistry, for this reason technologies using antibodies are preferred. In the last decade, many articles have reported the inefficiency of different commercial antibodies [13; 14; 15]. This is the result of a validation tested only with few experiments. Indeed, many antibodies can have diverse sensibility in different analysis. To overcome this problem the Human Protein Organization (HUPO) have started a Human Antibody Initiative (HAI) that have the aim to [16] :

1. create a catalogue of validated antibodies produced by companies or academic research laboratories;
2. create a protein atlas that collect the differences of protein expression in normal and disease tissue.

For this purpose two public database have been accomplished. The first is called antibodypedia (www.antibodypedia.org) in which antibodies directed to 20767 gene

products are collected. The website provides data about all antibodies recognizing a certain protein and also the different methods used for each analysis (WB, ELISA, IP,IHC, FC, IF). The second portal is the protein atlas (www.proteinatlas.org) that collects millions of high-resolution images showing the spatial distribution of proteins in 44 different normal human tissues and 20 different cancer types, as well as 46 different human cell lines. The data is combined with validation performed for each antibody, including WB and IM analysis and, for a large fraction, a protein array assay and immunofluorescent based confocal microscopy (Fig. 2). For every protein searched information about expression profiles, protein classes and chromosome location are supplied. At the moment, 20329 gene products are collected in that portal.

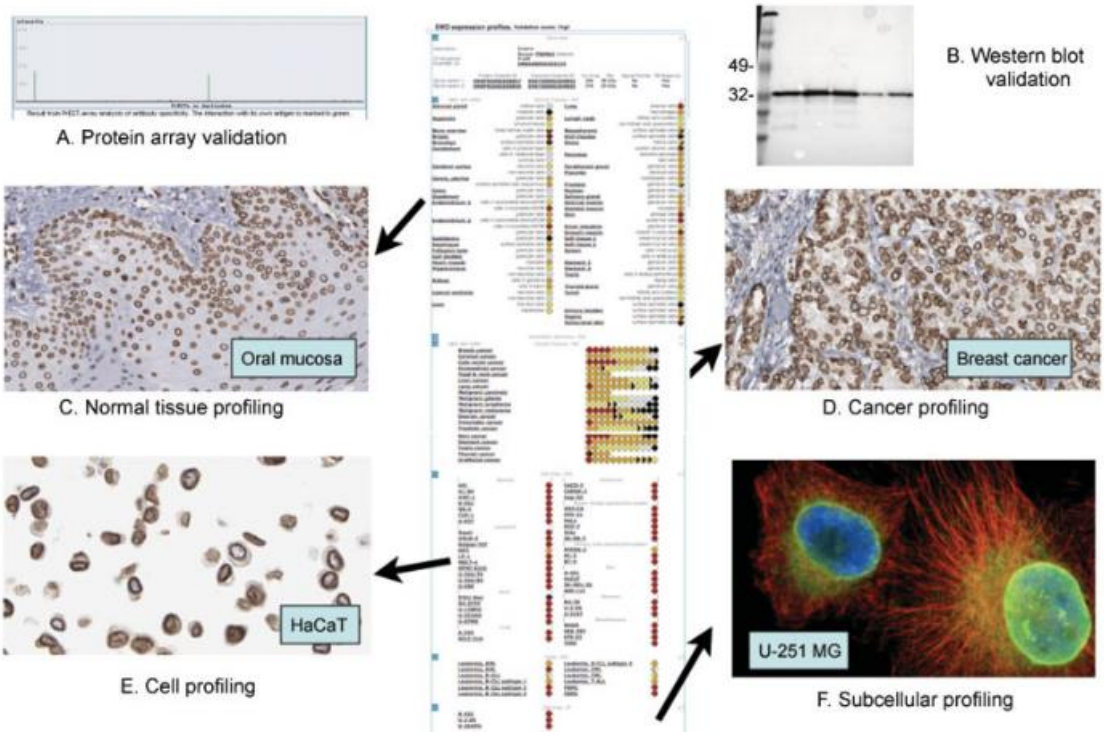


Figure 2. The human protein atlas [17]. Example of the data gives by the protein atlas database.

These two database are very useful for antibody research and utilization, even if many target remain without a specific antibody. For this purpose high-throughput antibody production, selection and validation is needed.

2.3 IN VIVO AND IN VITRO ANTIBODIES PRODUCTION TECHNOLOGIES

Antibodies are very powerful molecules in research, therapy and diagnosis. For this reason, it has been necessary to develop efficient methodologies to produce them. In 1975, Kohelr and Milstein have sewn up the "hybridoma method" (Fig. 3), based on B lymphocytes and tumoral cells fusion [18]. In particular, this method consists of the injection of a putative antigen and adjuvants in animals (mouse, rat or rabbit) to induce an immune rensponse that lead to antibodies production by B cells. When the animal sera is positive for antibodies against the antigen, the lymphocytes are collected and fused to tumoral cells to allow the formation of hybridomas. These cells are able to produce antibodies, such as B lymphocytes, and can reproduce itself *in vitro* for a long period, as a tumor cell line.

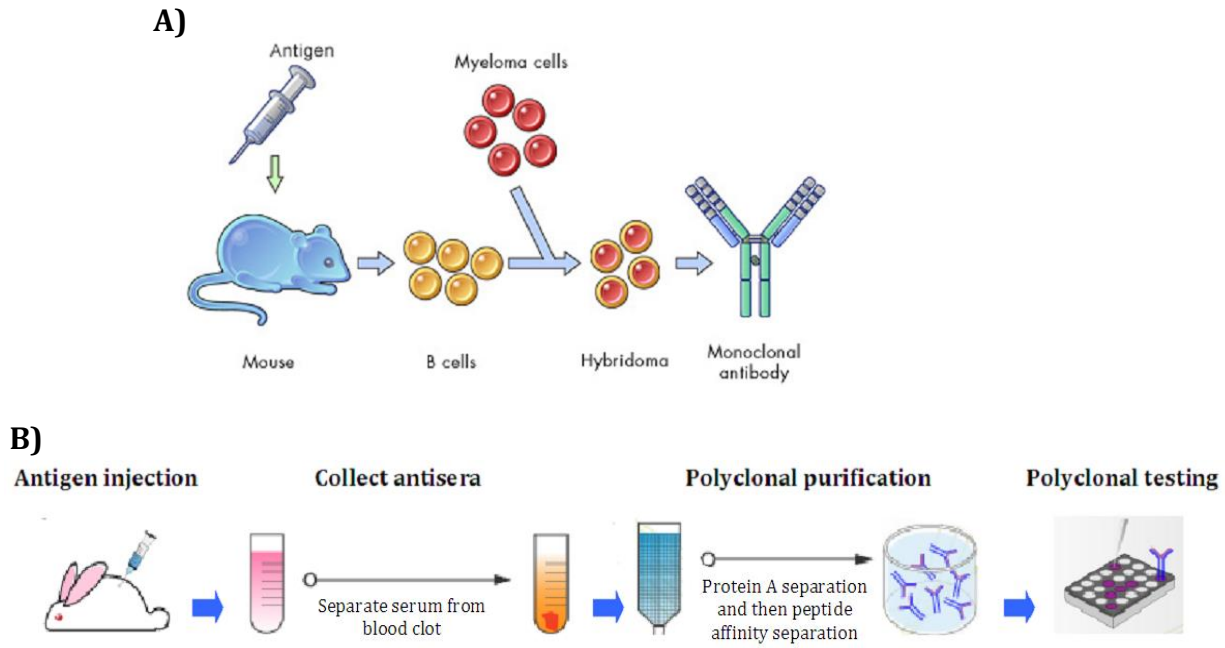


Figure 3. *In vivo* antibody production technology. A) The "hybridoma method" involves the use of animals (mouse, rat or rabbit). The hybridoma is a cell derived from the fusion of animal B cells, stimulated for antibody production through antigen injection, and tumor cells (myeloma). This cell type is able to produce antibody for long time *in vitro* allowing the antibody selection. B) Polyclonal antibodies production: animals (normally rabbit) are injected with the antigen supplemented with adjuvant and their blood is collected and the sera separated. Subsequently, the antisera can be purified with protein A, for IgGs separation, and peptide affinity column, for antigen specific IgGs separation. The polyclonals set can be subsequently tested.

This method is normally used for monoclonal antibodies production. In case of polyclonal antibodies, the same animals are injected with the antigen, supplemented with specific

adjuvants, and after 7-14 days the animal sera is collected and used directly as polyclonals antibodies or successively purified [19]. These methods are the well known methodologies for antibodies production, but they still to be very laborious, expensive and time-consuming. For these reasons *in vitro* methods have been developed. These technologies don't involve the antibodies production by animals, but utilize antibody libraries. The library consists in a collection of antibodies cDNA that can be screen to select the one of interest. There are different types of antibody libraries differing each other in how they are made [20]:

1. naïve;
2. immune;
3. semi-synthetic;
4. synthetic.

The naïve library derives from Ig of human donors or animals that are not immunized to a particular antigen. The light and heavy chains are amplified by polymerase chain reaction (PCR) and collected [21]. For this type of library the diversity is directly dependent to the number of the donors. On the opposite side the immune antibody libraries consist in the RNA amplification of immunoglobulins derived from immunized individuals or animals [20]. The semi-synthetic and synthetic libraries are based on random oligonucleotides synthesis to construct CDRs regions. In the semi-synthetic only a part of the antibody, normally the CDR3, is synthetic [22]. These last types of library are used to increase the library diversity. All these libraries contain at least from 10^7 to 10^{11} different antibody clones that can be used in all the *in vitro* antibody production technologies. In general, these methods are called "display technologies" and allow the linkage between genotype and phenotype. In this way the cDNA of the selected antibody is immediately avaliable for further applications. *In vitro* technologies, differently to the hybridoma method, allow a faster and easier selection and a major specificity of the antibodies selected has been shown [23]. Furthermore, *in vitro* antibody production allows to select antibodies against toxins and compounds that can't be injected in the animals. In the last decade, different display technologies have been developed, most used are: phage display and yeast display [24; 25; 26]. These two techniques are based on the display of the antibody on the cell surface, fusing it with a cell surface protein. In case of phage display the glycoproteins III or VIII are used; for yeast display mannoproteins are the used ones. The corresponding antibody cDNA is present inside the yeast cell or phage particle.

2.3.1 PHAGE DISPLAY TECHNOLOGY

Phage display is a technology that allows the selection of molecules through the exposition of peptides, proteins or antibodies on phage particles. The idea to use filamentous phages as an expression vector was introduced by Smith in 1985 [27]. Filamentous phages considered to be a good expression vector because [28]:

1. their genome can be modified, with insertion, without causing packaging problems;
2. their genome can be isolated;
3. they can tolerate stringent conditions necessary for selection steps;
4. coat proteins can be fused to other molecules, without loosing their infective role;
5. they can be accumulated in large amount in bacterial cells thanks to their non-lytic propagation.

The main phage utilized for phage display is the filamentous phage M13 [29](Fig. 4).

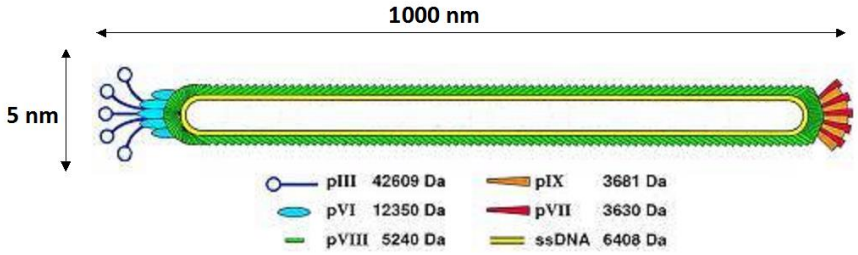


Figure 4. Filamentous phage M13 structure. The M13 genome is a single stranded DNA of 6407 bases. The DNA molecule is covered by the major coat protein gpVIII of which 2700 copies are present. On the opposite sides of the phage particle there are minor coat proteins: gpIII, gpVI, gpVII and gpIX. The proteins III and VI are located at one side and at the other there are gpVII and gpIX. The minor coat proteins are present in 5 copies each. A single phage particle is 1000nm in length and 5nm in diameter.

It is composed by a capsid that holds a single stranded DNA (ssDNA) encoding for capsid, infection and cell cycle proteins. The capsid is formed by five proteins: gpIII, gpVI, gpVII, gpVIII and gpIX. The gpVIII, called major coat protein, is present in 2700 copies for each phage particle and covers the entire ssDNA molecule. The other proteins are called minor coat proteins because only five copies each are present in a single phage. They are situated at the double-ended of a phage particle: gpIII and gpVI at one side; gpVII and gpIX at the opposite one. In phage display technology the gpIII and gpVIII proteins are generally used

for proteins exposition. A cDNA library can be directly inserted in the phage genome fused to gpIII or gpVIII gene [28] (Fig. 5). This allows the formation of phages containing each: 1. a DNA molecule containing the cDNA of a protein fused to phage protein; 2. the corresponding protein fused to all the copies of the phage protein. This method is potential only with peptides or small proteins because when big proteins, such as antibodies, are fused to gpIII or gpVIII the phage packaging or infection is compromised. To overcome this problem a new method was introduced. This new method uses a different vector called phagemid for protein display [30](Fig. 5). The phagemid is a plasmid containing:

1. phage replication origin;
2. packaging signal;
3. gpIII or gpVIII gene fused to the cDNA library.

The phagemid is transformed in the bacteria cell that will be infected by a helper phage containing wild-type phage genome. The resulting phages will have internalized the phagemid and will present only some copies of gpIII or gpVIII fused to the displayed protein [24].

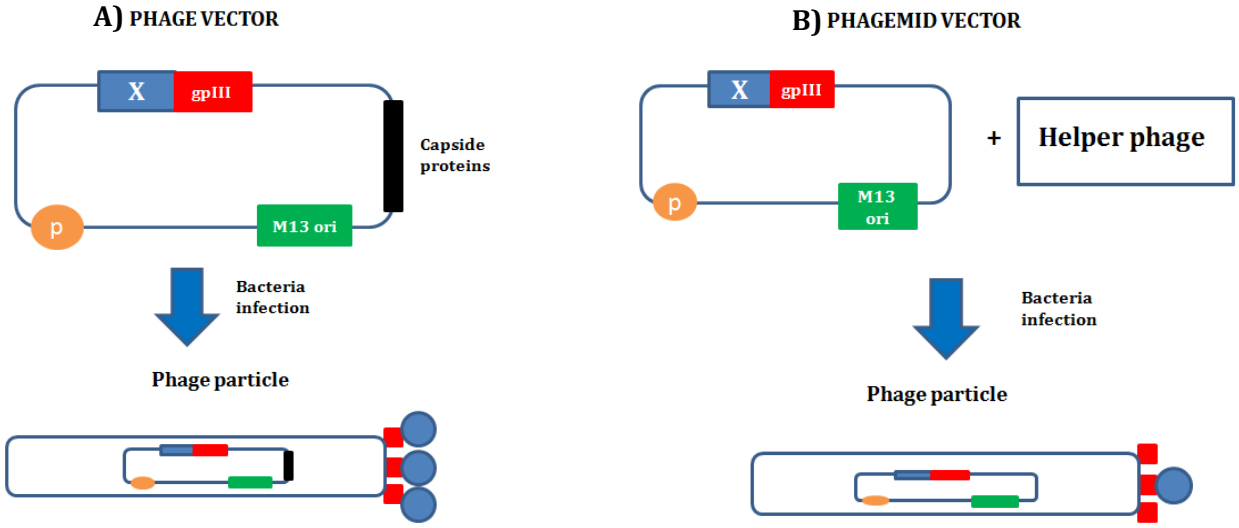


Figure 5. Phage display vectors. Phage display technology can be carry out with two different strategies. A)Phage vector: the cDNA library (blue) is directly inserted in the phage genome, fused to the gpIII gene (red). The resulting phages will present as genome the phage vector containing the cDNA of a protein X fused to the gpIII gene; as phenotype, the same protein X (blue) fused to all the gpIII copies (red). B) The phagemid vector is a plasmid containing the gpIII gene (red), the packaging signal (yellow) and the M13 ori (green). The cDNA library is inserted in this plasmid fused to the gpIII gene. To infect the bacteria a wild-type phage, called helper phage, is needed. Since that also a wild-type gpIII gene is present during the infection, only some gpIII copies (red) of the nascent phages, containing the phagemid vector, will be fused to the X protein (blue).

When a library is expressed with phage display it has to be selected to find the cDNA fragments of interest. The selection stage is called biopanning and consists of three steps (Fig. 6)[28]:

1. target immobilization: the antigen is bound on a solid support;
2. phage binding: exposure of the phage presenting antibody to the antigen;
3. removing unbound phages: washes to eliminate aspecific phages;
4. phage elution: the phage positive for antigen binding are eluted using free antigen or a competitor.

Normally, these steps are repeated three times with more stringent washing conditions to eliminate the major amount of unspecific antibodies.

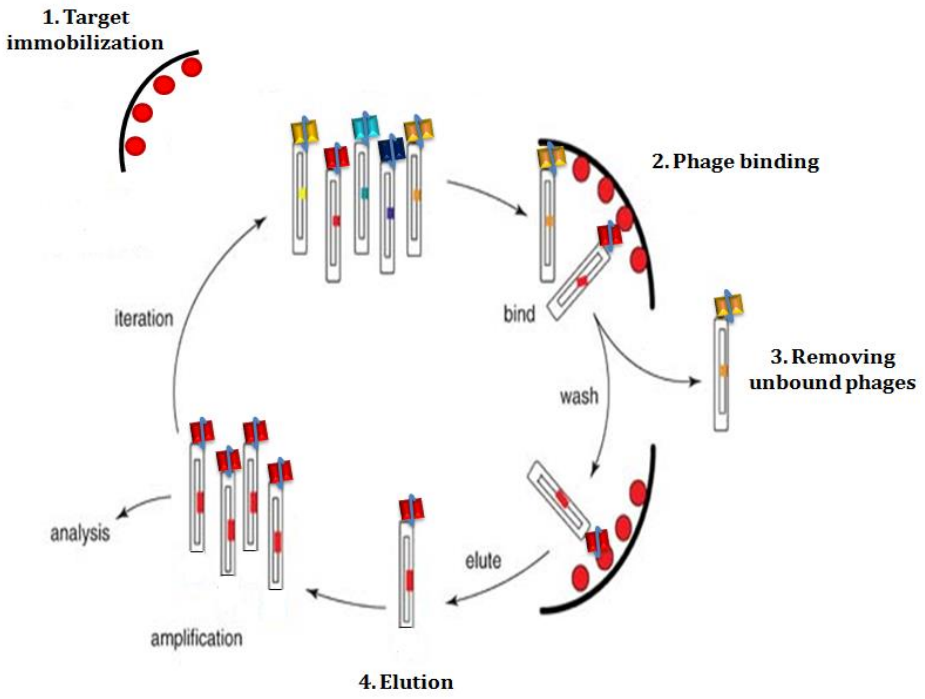


Figure 6. Biopanning of a scFv phage library. To select the phage presenting the antibody (scFv) of interest is necessary to perform rounds of selection. The selection consist in: 1. target immobilization; 2. phage binding; 3. removing of the unbound phages; 4. elution. This allow the enrichment of the phages carrying the scFv recognizing the target.

Phage display technology is a potential tool that can be used for different applications, but the main remains protein-protein interactions, in particular antibody-antigen. In antibody phage display a Fab or scFv can be expressed. Antibody libraries, containing 10^{11} clones, can be immediately use for display selection to find antibody with high specificity [30]. After the bacteria infection the phages displayed antibodies are collected and subsequently

selected against the specific antigen. It has been also demonstrated that with phage display specific antibodies against protein, haptens and complex antigens, as carcinoma cells, can be isolated [31]. Furthermore, monoclonal antibodies optimized with phage display technology are now used for clinical purpose, such as Motavizumab and Palivizumab, humanized antibody direct against respiratory sinchythial virus (RSV) [32]. Its has been reported that approximately 30% of all human antibodies now in clinical trial are derived from phage display [33]. This confirms the great specificity of phage display derived antibodies.

2.3.2 YEAST DISPLAY TECHNOLOGY

Another display technology that utilized organism as display system is the yeast display. Yeasts are unicellular eukaryotic organisms that have a generally-regarded-as-safe status (GRAS). The yeast cell is characterized by all the typical eukaryotic compartments such as: mitochondria, nucleus, endoplasmic reticulum, lisosomes, vesicles and Golgi apparatus (Fig. 7).

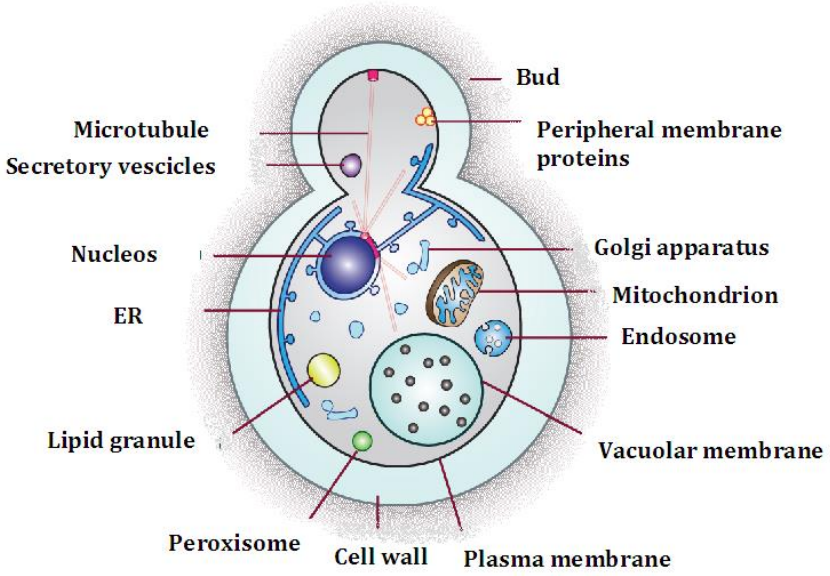


Figure 7. Yeast cell. Structure of a yeats cell. In the citosol are present: nucleos, endoplasmic reticulum, lipid granules, microtubules, Golgi apparatus, vacuolar membrane, vesicles and mitochondria. Al this compartments are covered by two layer: plasmamembrane and the cell wall.

All these elements form the cytoplasm of a yeast cell that is cover by a plasma membrane. The plasma membrane is in turn covered by a thick layer called cell wall and the space between them is called periplasmic space. The cell wall is 200nm thick and is principally composed by mannoproteins and β-linked glucans [34]. The proteins normally choose for yeast display are mannoproteins. Two types of them are present in yeast:

1. Extractable with SDS (associated with glucans through non-covalent binding);
2. Extractable only with glucans digestion (associated with glucans through covalent binding).

Only the second type of mannoproteins is used for yeast display, in particular the most used are: α-agglutinin, a-agglutinin and flocculins [35]. All these proteins have a glycosylphosphatidylinositol (GPI) anchor that is important for cell-surface protein expression. In yeast display the cDNA library cloned in yeast plasmid contains [36]: an inducible promoter, a secretion signal and the cDNA library fused to the mannoprotein. If the display is carried out with α-agglutinin or flocculin 1p only the C-term half part, correspondent to the 3' half of the gene, is present in the plasmid, because in this part is present the GPI anchor signal that is essential for the cell wall attachment (Fig. 8 A,B). When the display is carried out with the a-agglutinin the Aga2 gene is fused the library instead Aga1 gene is present in the yeast genome (Fig. 8 C). The Aga1 and Aga2 domains are linked together through covalent binding.

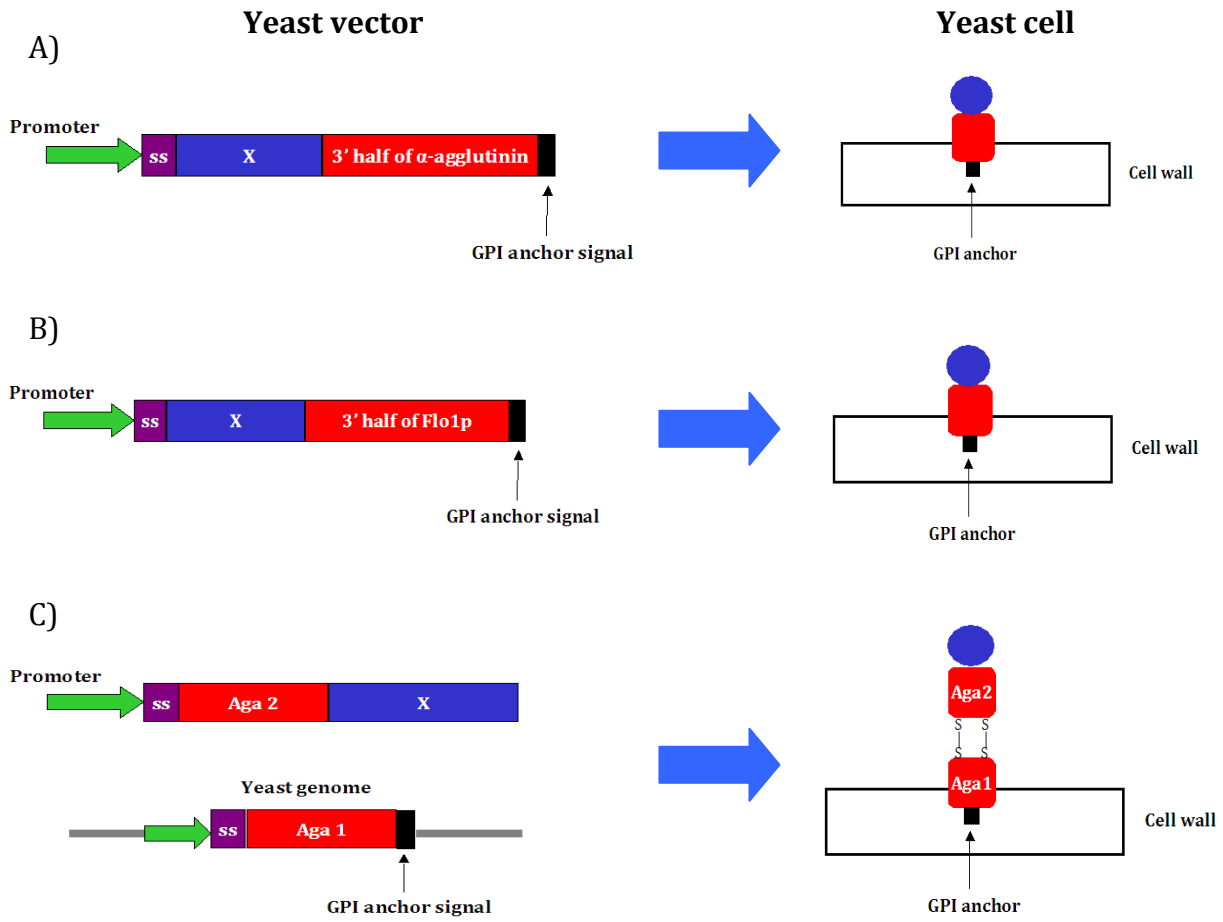


Figure 8. Yeast display vectors. Different proteins are used for yeast display selection. The most utilized are: a-agglutinin, α-agglutinin and flocculin. In case of α-agglutinin (A) and flocculin (B) the library is fused to the 3' half part of the corresponded gene, allowing the expression of the GPI anchor the C-terminal part of the mannoproteins and the displayed protein. The a-agglutinin protein is composed by two domains translate by two different gene Aga1 and Aga2. in this case the library is fused to the Aga2 gene and the Aga1 is present in the yeast genome (C).

The plasmid is present in multiple copies in the yeast cell and it has been demonstrated that almost 10^5-10^6 copies of mannoprotein are presenting the target gene [35]. After the library expression a selection step is necessary. In yeast display the selection is achieved with magnetic bead and flow-cytometric sorting. The magnetic beads are used to decrease the size of the library, that is too much for flow-cytometric sorting. The employment of high-speed flow-cytometry allows the detection of the amount of protein expressed by yeast and the affinity to the target using a two-color labeling scheme. The resulting graph shows (Fig. 9) [37]:

1. yeast displaying protein with less affinity to the target (black line);
2. yeast displaying protein with high affinity to the target (green line);

3. yeast displaying no express of proteins (blue line);
4. yeast displaying proteins with no binding activity to the target (yellow line).

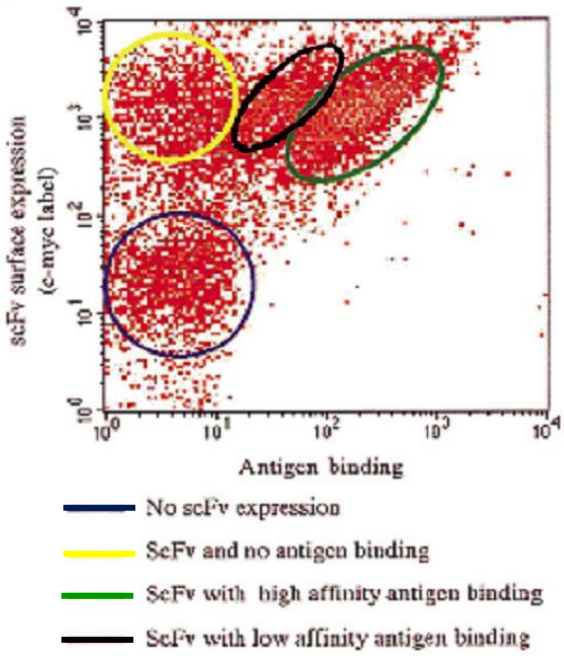


Figure 9. Yeast displaying scFv selection with flow-cytometry cell sorting [37]. Flow-cytometric analysis of yeast displaying scFv. The double-color detection allows to distinguish yeast for antigen binding and scFv expression on the yeast surface through anti-tag antibody (c-myc).

Since its introduction 10 years ago [38], yeast display has been used to express a variety of proteins for improved affinity, specificity, expression, stability, and catalytic activity. More recently, it has been widely used for antibody selection and engineering [37; 39]. It has been demonstrated that the distribution of affinity of antibodies selected by yeast display is similar to the antibodies isolated with phage display. Furthermore, yeast display is suitable for antibody affinity maturation [40].

2.3.3 DIFFERENCES BETWEEN PHAGE AND YEATS ANTIBODY DISPLAY

Either phage and yeast display are powerful tools for high specific antibody selection and maturation. Phage display was introduced a decade before yeast display and many paper have been published using these technologies [41] (Fig. 11).

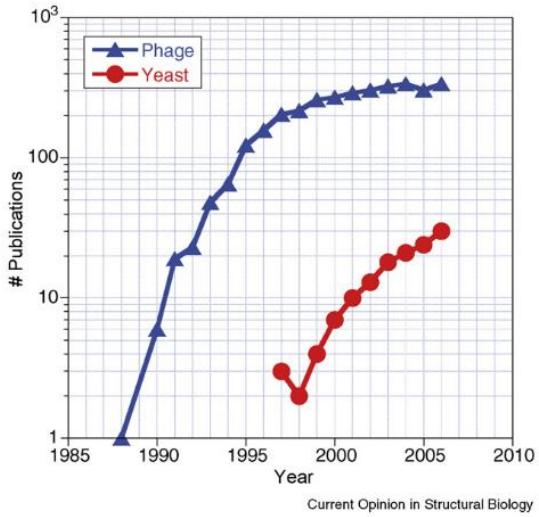


Figure 11. Phage and yeast display publications [41]. Number of publications regarding phage and yeast display by their discovery to 2008.

The main differences between phage and yeast display are:

1. type of organism: phage is a prokaryotic and doesn't allow the expression of complex proteins or post-translational modifications. Yeast is an eukaryotic organism.
2. safety: phage is a virus, and because of this not a safe organism; yeast has a GRAS.
3. library dimension: phage allows the use of largest libraries till 10^{11} clones; yeast uses smaller library. In yeast display additional rounds of magnetic-cell sorting are needed for the selection of largest library.
4. number of protein copies express by a single cell: using a phage vector is possible to have a maximum of 2700 copies displayed on a single phage, but it is true only for small proteins. With yeast display till 10^5-10^6 protein copies can be present on a single cell.

Recently, to exploit the advantage of the two technologies, researchers try to couple them. In 2011, Patel et al. have introduced a cross-species display system in which phage and yeast display are used in parallel to antibody library selection [42]. In this paper an adapter-direct display platform have been used. In this platform two different vectors are used: display vector and helper vector. The display vector avoids the anchoring sequence and can be used to display proteins on different species in combination with the helper vector that is species-specific. Using this technology Patel and coworkers have selected antibodies with a Kd in the nanomolar order. Furthermore they observed that the first round of selection with phage panning allows the enrichment of the library to 10^6; this permits the directly use of FACS for yeast selection. Another methodology, that combines

phage and yeast display, was published in 2012 by Ferrara et al. [43]. In this paper the researchers have selected antibodies against the antigen 85 of tuberculosis bacteria through two round of phage selection followed by one or two round of yeast display selection using FACS analysis. The passage between phage and yeast vector has been achieved by homologous recombination. In conclusion the combination of the two display method seems to improve the potential of both technologies.

2.4 ANTIBODY VALIDATION

The great use of antibodies in research has brought to the development of high-throughput production, selection and validation methods. If the antibody production is well characterized, the validation has been underrated. Many articles take a lot of attention in the antibodies production and selection and don't perform a real antibody validation. Indeed, as often as not the validation is carried out with Western blot and ELISA test only. For this reason many antibodies can be not functional in immunohistochemistry or show cross-reactivity if used in more sensible technique such as protein microarray. To validate antibodies there isn't a standard protocol because they have to be choosen depending on the experiment that has to be performed. From this point of view, Rimm' s laboratory has developed an algorithm for antibody validation for immunohistochemistry/quantitative immunofluorescence that can be used also for other application [15](Fig. 12).

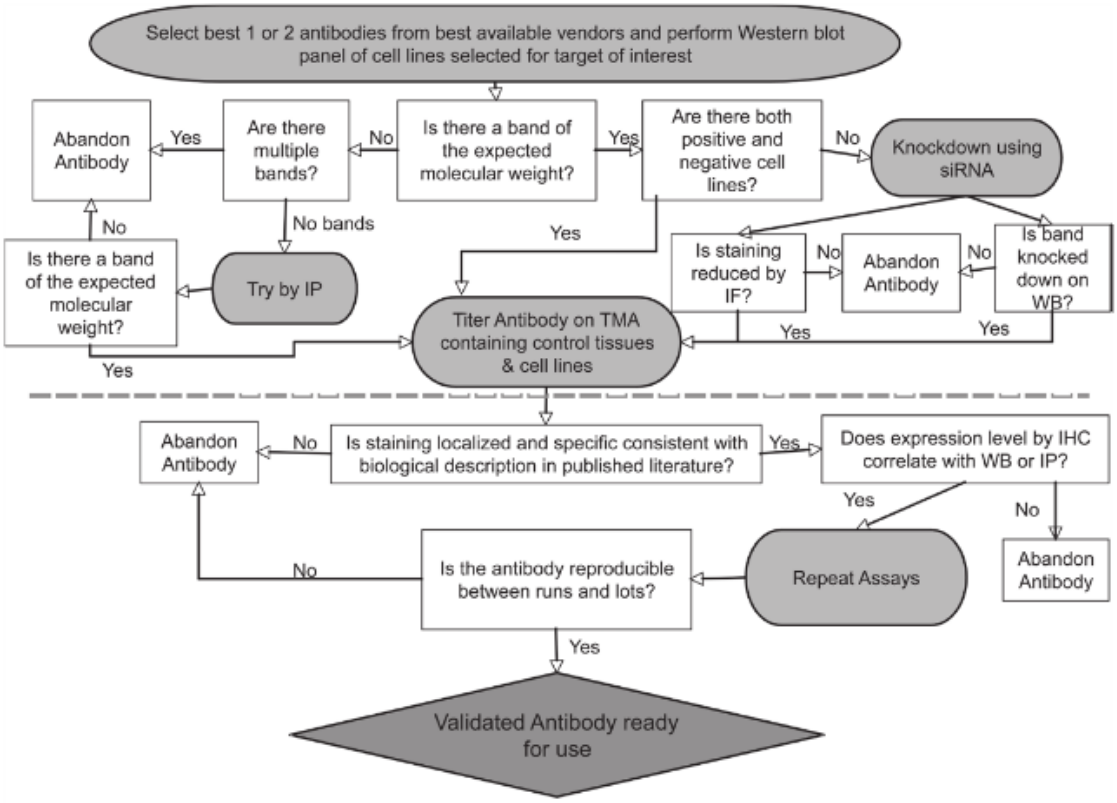


Figure 12. The Rimm's algorithm. The part upper the dot line concern the antibody validation cell lines for antibody specificity testing. The lower part consists in further validation with tissue microarray for localization of the antibody target and to test different antibody lots.

This algorithm starts from the validation in Western blot of different cell lysates that express the target protein. This part allows to test the specificity of the antibody using techniques such as IP or short interfering RNA (siRNA). Subsequently the antibody will be used for in immunohistochemistry test to verify their reproducibility and the capability to localize the target in every tissue expressing it. Tissue microarray and protein microarray are emerging as important tool for antibody validation [44; 45]. Indeed, in 2012 Sjöberg at al. have shown the validation of monoclonal, polyclonal and scFv antibodies against protein containing SH2 domain using high-throughput protein microarray [46]. This technology is very useful because is sensitive and allows to test an antibody on thousand of characters in a single experiment.

2.5 PROTEIN MICROARRAY TECHNOLOGY

The great interest arisen by proteome studies have led to development and improvements of different highly innovative technologies. In particular, there was a need to developed a method that increase the number of characters analyzed in one experiment, with less consumption of materials and solutions. One of the major tool that has been set up for proteomic studies is protein microarray. This technology allows the immobilization of thousands of proteins on a solid surface, in a miniaturized format, and their analysis, in parallel, in a single experiment. Every kind of proteins can be immobilized on a microarray slide: peptides, full-length proteins, antibodies, receptors, enzymes or aptamers. As solid support, slides made of glass, polymer or plastic are used. This support is subsequently treated to allow the protein immobilization through different types of binding (Fig. 13):

- Non-covalent (hydrophobic or entrapment: nitrocellulose, polystyrene; positive charges: poly-lysine);
- Covalent (active groups such as: aldehyde, epoxy, esters, etc.);
- Biomolecular interactions (streptavidin-biotin, His tag-nichel ions).

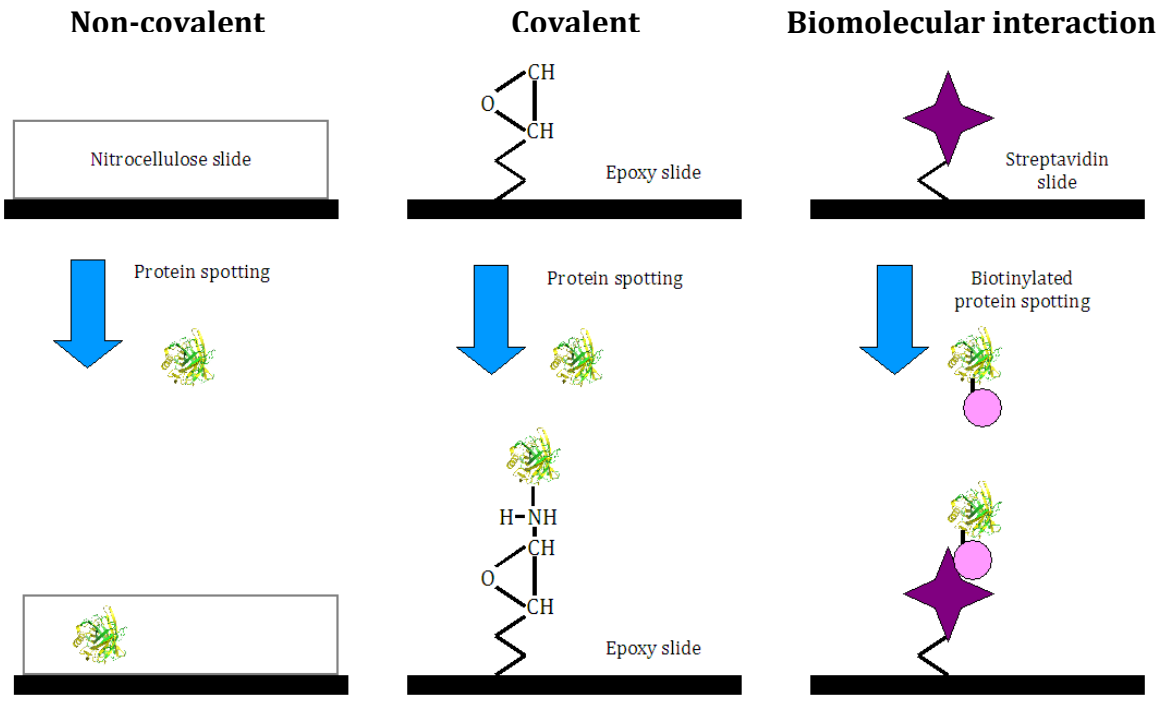


Figure 13. Proteins immobilization on different microarray slides. The proteins can be immobilized in different way on a solid surface. Non-covalent binding: proteins are capture through hydrophobic groups or entrap in a matrix (e.g. nitrocellulose). Covalent binding: proteins are immobilized through active groups such as epoxy, ester or aldhehyde that are able to form covalent bounds with all kind of proteins. Biomolecular interactions: the proteins are fused to tags or biomolecules such as biotin that permit the proteins capture by known groups or other molecule such as streptavidin.

The deposition of proteins on the slide is called “printing” or “spotting” and can be achieve through particular instruments called array printer. These instruments utilized pins to permit the deposition of nanolitre of protein. The pins are divided in two type:

1. contact pins: place nanolitre of sample directly on the slide surface;

2. non-contact pins: deposit drops of sample using capillary or ink jet technology.

Once the proteins are spotted they can be directly analyzed treating the slide with the desired solutions. The mainly detection method use with protein microarray is fluorescence, that is sensitive, safe and can be read with charge coupled device (CCD) camera or laser scanners. Although, other label methods are used, such as chemiluminescence and radioactivity. In recent years, also label-free methods have been used such as: mass spectrometry, surface plasmon resonance imaging (SPRI) and surface-enhanced laser desorption (SELDI)[47]. Protein microarray technology allows also the quantification of the signals using positive and negative controls and a calibration scale.

2.5.1 PROTEIN MICROARRAY APPLICATIONS

Protein microarrays are subdivided in three main groups :

1. analytical;
2. functional;
3. reverse-phase.

The analytical arrays are used to profile protein affinity and expression in different cells. The most common analytical array is antibody array, where this type of protein is immobilized on the slide and a proteins mixture is the probe (Fig. 14). Functional arrays allow the detection of proteins function and interactions. Normally full-length or domains are spotted and analyzed with proteins or other molecules such as: DNA, RNA, lipids, carbohydrates and drugs (Fig. 14). Finally, reverse-phase array are used to detect altered proteins or to highlight proteome differences between two cell types or healthy and diseased cells. Indeed, this kind of array consists in the printing of a cell lysate that can be characterized using a specific antibody or proteins (Fig. 14).

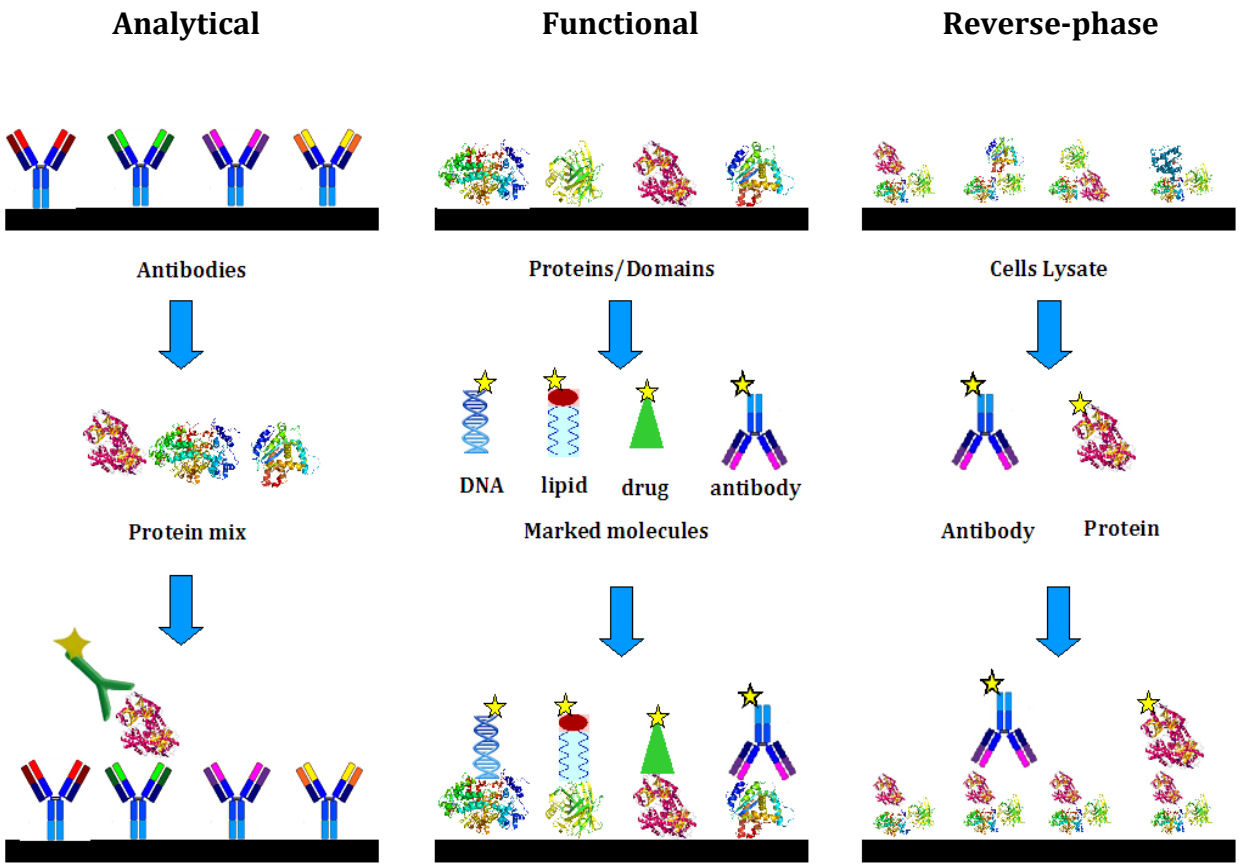


Figure 14. Protein arrays subtypes. Protein microarrays are divided in three subgroups: analitical, functional and reverse-phase. Analytical arrays are normally used in protein expression analysis using antibodies as capture molecules. Functional arrays allow the analysis of protein function, interaction (DNA, lipid, drug, protein, RNA) or antibody validation and consist in the immobilization of full-length protein or domain. Reverse-phase array are characterized by immobilization of cells lysate instead of single molecules and are used in proteome analysis.

It is immediately clear that protein microarray can be a powerful tool for many different applications. First it has been widely used for protein interaction studies that allow to understand signaling pathways [48] and to discover binding proteins [49]. Protein microarray are also important for clinical purpose because allows biomarkers and autoantigens discovery [50; 51], where healthy and diseased sera are used to screen proteins/peptides library. Furthermore, host-pathogens interface studies that lead to vaccine development can be carry out [52]. Another important fields in which protein microarray are widely used is antibody validation. Many times commercial antibodies seems to be less specific for their target and cross-reactivity can also be detect. These problems lead to a wrong or less accurate validation. Protein microarrays are potential tool for this purpose because provide multiplex analysis of different antigens revealing specificity and selectivity of the antibody. In 2008, A. Lueking and coworkers have shown

the characterization of a monoclonal antibody against the variant 6 of CD44 molecule [53]. Sjöberg R. et al. have used protein microarray to validate selectivity and specificity of affinity reagents direct against SH_2-domain[46]. In this study monoclonal antibodies, single chain fragment variable and polyclonal antibodies have been utilized as affinity reagents demonstrating that all these type of antibody format are feasible for protein microarray validation. Furthermore, they have shown that with protein microarray is possible to validate antibodies against proteins with 89% of similarity determine the exclusion of some aspecific monoclonals. Finally, it is possible to assure that protein microarray can be one of the major tool for antibody validation.

2.5.2 PROTEIN MICROARRAY IMPROVEMENTS

As described before, to perform a protein microarray is necessary to produce and purify all the proteins that have to be printed on the slide. This become an issue when you think to produce thousands of proteins in the same time. The development of protein libraries has allowed a more easily search of proteins cDNA but their production and purification remain a time-consuming issue. Furthermore, protein microarrays are not stable for long period so it is necessary to proceed with the experiment immediately after the spotting. Many groups of researchers have try to overcome these two problems link to protein microarray technology. The first innovation for protein microarray improvement was the use of cell-free protein expression systems, that are able to transcribed a DNA template and translate the correspondent mRNA in a protein [54]. The combination of this two tools has brought to a new technology called "*in situ* protein array" [55]. This new type of protein microarray consists in the immobilization of DNA instead of proteins eliminating the stability issue. Once the DNA is spotted it can be translate in proteins directly just when is needed, erasing the protein production and purification processes. Even if standard protein array continue to be widely used, many groups took interest in this "*in situ* protein array" developing different strategy to use it.

2.6 PROTEIN PRODUCTION WITHOUT CELL BOUNDERIES

Protein production is one of the key steps in biotechnology and functional proteomics. Conventional protein expression systems rely on prokariotic cells as bacteria, where the most used is *Escherichia coli* (*E. coli*) or eukaryotic cells as yeast, insect and mammalian cell systems [56; 57; 58]. The prokaryotic expression systems are easily adaptable to high-throughput expression methodologies, but are less efficient in the expression of complex mammalian proteins, in particular those which require post-translational modification. At the opposite site, systems based on eukaryotic organisms are capable of expressing post-translationally modified proteins but are difficult to be integrated into high-throughput methodologies. In order to find a new method that allows an easier, cost-effective and high-throughput application, cell-free based systems were developed [59]. In 1961 Niremberg and Matthaei have demonstrated the possibility to translate *in vitro* proteins with an *E. coli* extract using RNA or synthetic nucleotide as a template [59]. Subsequently, Zubay et al. has shown the coupled transcription and translation of a β-galactosidase chain using an *E. coli* extract and a virus DNA [60]. Compared to cell-based methods, cell-free expression systems are considerably faster since they don't require cell transfection, cell culture and extensive protein purification procedures (Fig. 16A). Moreover, they don't involve cell lysis steps that could denaturate proteins. In addition, cell-free systems allow the production of toxic, insoluble or membrane proteins, that are difficult to produce using cell-based methods. Finally, they can be used in a microlitre scale, making them ideal for high-throughput applications [61; 62].

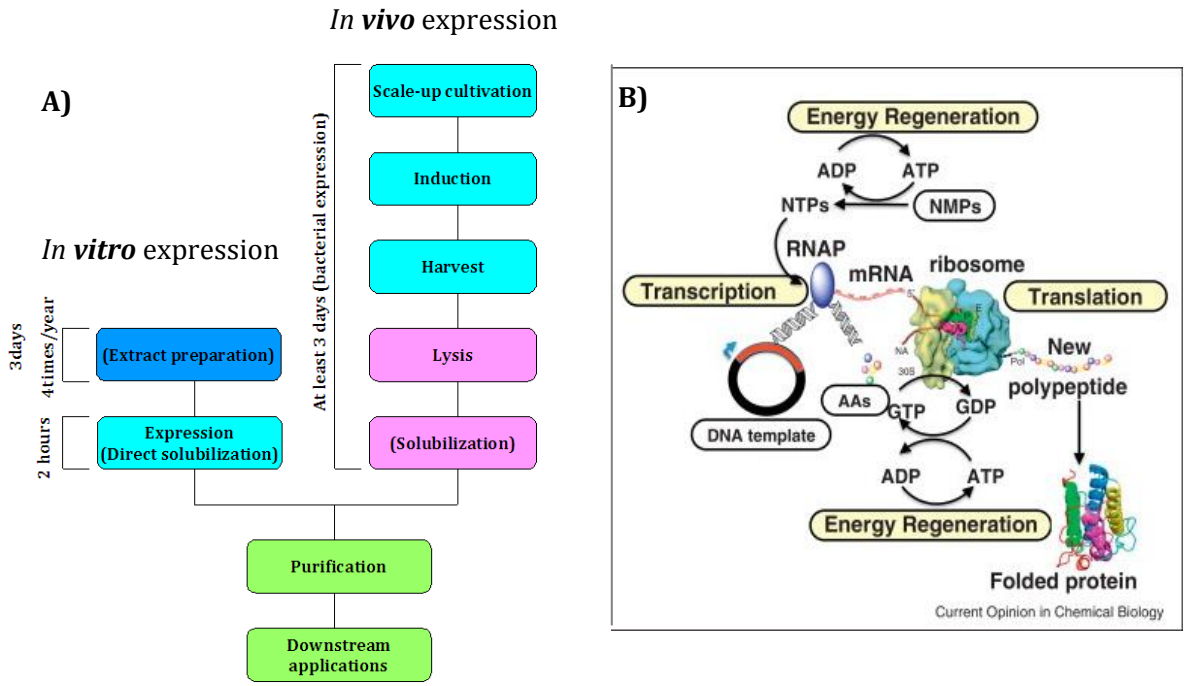


Figure 16. Protein production systems. A) Protein expression with *in vivo* or *in vitro* methods. In case of *in vivo* (bacteria) cells have to be cultivated, induced and collected; subsequently a cell lysis has to be carried out and in some cases also protein solubilization is needed. All these passages takes 3 days. When an in vitro expression system is used the protein production takes only 2 hours and the extract necessary for the expression has to be prepared only 4 times /year. B) Schematic representation of the passages occurring during an IVTT reaction [63]: DNA template has to be transcribed in mRNA with all the transcription complex, RNA polymerase and energy sources; the mRNA is translate in protein using amminoacids, tRNAs, ribosome and energy sources.

The most recent *in vitro* transcription and translation (IVTT) systems are composed of crude cell lysate supplemented of [64] (Fig. 16B):

1. energy sources (ATP, ADP);
2. RNA polymerase (T7, SP6 or T3) ;
3. amino acids;
4. tRNA;
5. enzymes;
6. NTPs ;
7. salts and ions.

Initially, only *E. coli* lysate was used, because it is easier to grow in a large scale and it is also less expensive. Successively, IVTT systems based on eukaryotic extracts, as rabbit reticulocytes and wheat germ, were generated [65]. These systems seem to be more stable than the *E. coli* one and more suitable for eukaryotic protein production. Recently, also insect and human cells lysates have been developed, showing an improvement in terms of protein quantity [66; 67]. The major advantage of cell-free systems is their flexibility that has allowed significant improvements in protein yields, folding and post-translational modifications through the addiction of chaperons or specific enzymes (glycosylation, phosphorilation, etc.) to the cells extract [68]. Thanks to all these improvements, IVTT systems are increasingly used in protein array technology, protein structural studies, large scale analysis of proteins, screening of antibody mutants, IVEC technology and display technologies such as ribosome display, mRNA display and *in vitro* compartimentalization [55; 62; 69; 70; 71].

2.7 IN SITU PROTEIN ARRAY

The great innovation of protein production in a cell-free mode and the possibility to transcribe and translate a DNA template in a single step has allowed the generation of high-throughput technologies, such as *in situ* protein arrays. These particular microarrays permit to build the protein array whenever is needed, starting from a DNA template. An *in situ* protein array consists of four different steps (Fig. 17): 1) immobilizing a DNA template on a solid support (Fig. 17A); 2) treating the support with an IVTT mixture (Fig. 17B); 3) capturing the nascent protein on the solid support (Fig. 17C); 4) revealing the protein of interest (Fig. 17D). In this way the standard protein array drawbacks of protein stability and long period of storage are overcome. The first *in situ* protein array which has been developed is the protein *in situ* array (PISA), also known as DiscernArray, by He and Taussing in 2001[55]. Afterwards, many other *in situ* protein arrays have been developed: Nucleic acid programmable protein array (NAPPA) [72], multiple spotting technique (MIST) [73], TUS-TER array [74], DNA array to protein array (DAPA)[75], *in situ* puromycin-capture from mRNA array [76].

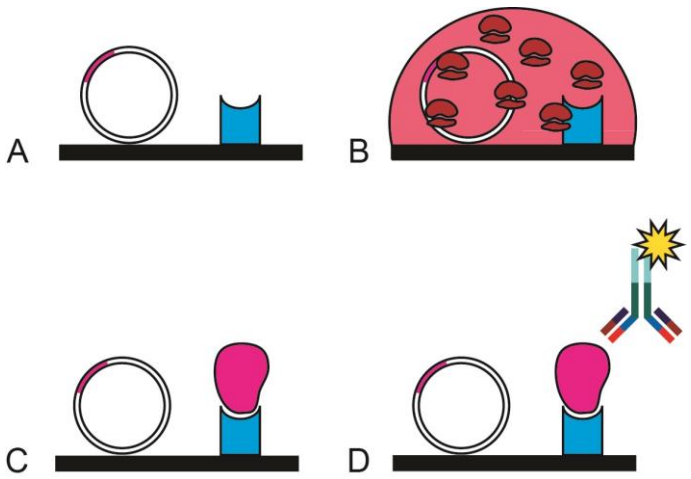


Figure 17. *In situ* protein microarray technology. Schematic resume of the main steps to perform an *in situ* protein microarray. A) printing of a DNA template and a capture agent; B) production of the proteins treating the slide with an IVTT mixture; C) capture of proteins from the capture agent; D) revealing of the protein of interest.

2.7.1 PROTEIN IN SITU ARRAY (PISA)

The PISA method (Fig. 18A) permits the generation of protein arrays in one step starting from PCR DNA fragments. In particular, these PCR fragments contain all the DNA sequence necessary for IVTT (T7 promoter, translation initiation site, poly(A)$_{28}$ tail, transcription and translation terminators) and the protein cDNA fused with a double (His)$_6$-tag sequence, which allows a stronger affinity binding to nickel ions (Ni^{2+}) than standard (His)$_6$-tag. In this way, the generation of the DNA template doesn't involve cloning passages, but different PCR cycles. As cell-free system a commercial kit based on rabbit reticulocytes lysate is currently used (T7 TnT Quick for PCR fragments).The template is added to the IVTT mixture and deposited on a Ni-NTA-coated microtiter plate or Ni-NTA magnetic agarose beads. After the reaction, the nascent protein is bound to the solid support thanks to the double (His)6-tag and can be subsequently analyzed. The authors tested the system using single-chain fragment variable (scFv) and experienced that in 25µl of reaction 120ng of protein was produced, of which 50% was bound on the microtiter [77].

2.7.2 NUCLEIC ACID PROGRAMMABLE PROTEIN ARRAY (NAPPA)

The Nucleic acid programmable protein array was designed by La Bear and colleagues in 2004 [72] (Fig. 18B). In this method a plasmid DNA is cross-linked to psoralen-biotin conjugate through UV light and subsequently printed with avidin on a glass slide covered with aminopropyltriethoxysilane (APTES). The DNA template encodes the protein/peptide of interest, fused at the C-terminal to glutathione S-trasferase (GST)-tag. The presence of the GST-tag allows the capture of the nascent protein by a monoclonal anti-GST antibody spotted on the slide concurrently with the plasmid DNA. After the deposition of DNA and antibody, the slide is treated with IVTT system based on rabbit reticulocytes lysate and the protein, produced and captured, can be immediately analyzed. Instead of PISA, this method permits a high-density format thanks to the type of support used. Indeed it has been shown that 512 spots, spaced 900μm from each other, can be printed in a single slide and all the proteins are produced using only 100μl of IVTT reaction. Furthermore, they have demonstrated that 675pg of protein can be produced and capture in each spot. In 2008, the same group of researchers have improved the NAPPA technology by the use of an high quality super-coiled DNA and adding bovin serum albumin (BSA) to the printing solution [78]. Finally, in recent years, they have shown that an IVTT system based on HeLa cells led up to a major quantity of protein produced by printed DNA [79].

2.7.3 MULTIPLE SPOTTING TECHNIQUE (MIST)

The multiple spotting technique is an high-throughput approach where a microarray slide is spotted a first time to allow the immobilization of proteins and a second time to permit the deposition of another compound, exactly in correspondence of the previous one [80]. Agenendt et al. have applied this method to build a new type of *in situ* protein array in which the first printing allows the spotting of DNA templates, plasmid or PCR fragments, and the second involves the deposition of the IVTT mixture [81] (Fig. 18C). After the expression, the proteins are immobilized on Nickel Chelate-coated or APTES slide and subsequently analyzed. The importance of this new approach is the limitation of protein diffusion and cross-talking, due to the deposition of the IVTT for every single spot. Furthermore, Angenendt et al. have demonstrated that only 35fg of unpurified PCR DNA are sufficient for the detection of full-length green fluorescent protein (GFP) in subnanolitre volume [81]. Moreover, the MIST allows to perform high-density protein

microarray with up to 13,000 spots in one slide. By now the MIST technology has not been used for particular studies, but its functionality is well characterized.

2.7.4 TUS-TER ARRAY

The most recent *in situ* protein array developed is the one called TUS-TER or Protein array on demand [74] (Fig. 18D). This method is based on the interaction between the *E. coli* protein TUS and a double stranded DNA sequence of 23bp called *Ter*. *In vivo* the binding of this protein to *Ter* allows the termination of the DNA replication. The peculiarities of this binding is that even if it is not covalent it is very strong (K_d= 3.4 x10^{-13}) and it occurs only when the protein encounter this particular DNA sequence [82]. Chatterjee et al. have exploited the capability of this protein fusing it to cDNA of target proteins that can be immobilized from the *Ter* sequence [74]. In particular a plasmid DNA is used as template, because contains transcription and translation sequence and target protein-TUS cDNA, but also as capture agent because it contains the *Ter* sequence. After the DNA printing on a nitrocellulose-coated slide, is only necessary to treat it with IVTT mixture to obtain the protein immobilization on the slide. The great advantage of this method is that there is no need for an additional molecule to capture the nascent protein, but only the presence of a plasmid DNA onto the microarray surface. To validate the system the Authors have also shown that the nascent proteins captured directly from the *Ter* site are present in the correspondent spot, therefore protein diffusion is avoided.

2.7.5 DNA ARRAY TO PROTEIN ARRAY (DAPA)

The DAPA methodology consists of the printing of a protein microarray using a DNA array as a mould [83] (Fig. 18E). In particular PCR fragments encoding for different proteins, fused to a tag, are spotted onto epoxysilane slide and subsequently the array is placed face to face with a second slide able to capture tagged proteins. Between the two slides a permeable membrane soaked with the IVTT mixture is present in order to allow the protein production. Afterwards, the proteins are produced starting from the DNA array, diffused through the membrane and are captured by the second slide. Using GFP as a control He and colleagues have evaluated that from 0,1ng of DNA template up to 30fmol of protein/spot can be captured. The great innovation of this technology is the possibility to

reuse the same DNA array to produce many protein arrays. It has been demonstrated that a single DNA array can be reused almost 20 times without significant variation [84]. Finally, it has been demonstrated that scFvs can be spotted on a slide using DAPA and they are also functional. These results are promising for the utilization of DAPA technology in proteome studies.

2.7.6 *IN SITU PUROMICYN-CAPTURE FROM mRNA ARRAY*

This type of microarray is the only one using RNA as a template. It is based on the assumption that when a ribosome encounters a region of double-stranded RNA or DNA-RNA hybrid it is not immediately released but it stalls on the mRNA. Tao and Zhu have hypothesized that this ribosome stalling can be long enough to provide the incorporation of a puromycin-grafted oligo immobilized on a solid phase, therefore the nascent polypeptide is directly captured by the puromycin [76] (Fig. 18F). This strategy consists of the immobilization, on a streptavidin-coated slide, of biotinylated puromycin-oligo and mRNA, which anneals at a 3' end to an RNA bumper oligo, so as to form a double-stranded RNA. After printing, the slide is treated with an *in vitro* translation mixture based on rabbit reticulocytes lysate and after the synthesis, the nascent peptide is captured by the puromycin-oligo. Finally, the slide is treated with a solution containing RNAse that degrades the mRNA, in order to have on the slide only the newly synthesized protein. After the assessment of the technology, the Authors have evaluated that, using this approach, 0,8 fmol/spot of protein are captured on the slide surface and this results are comparable with those of traditional protein microarray.

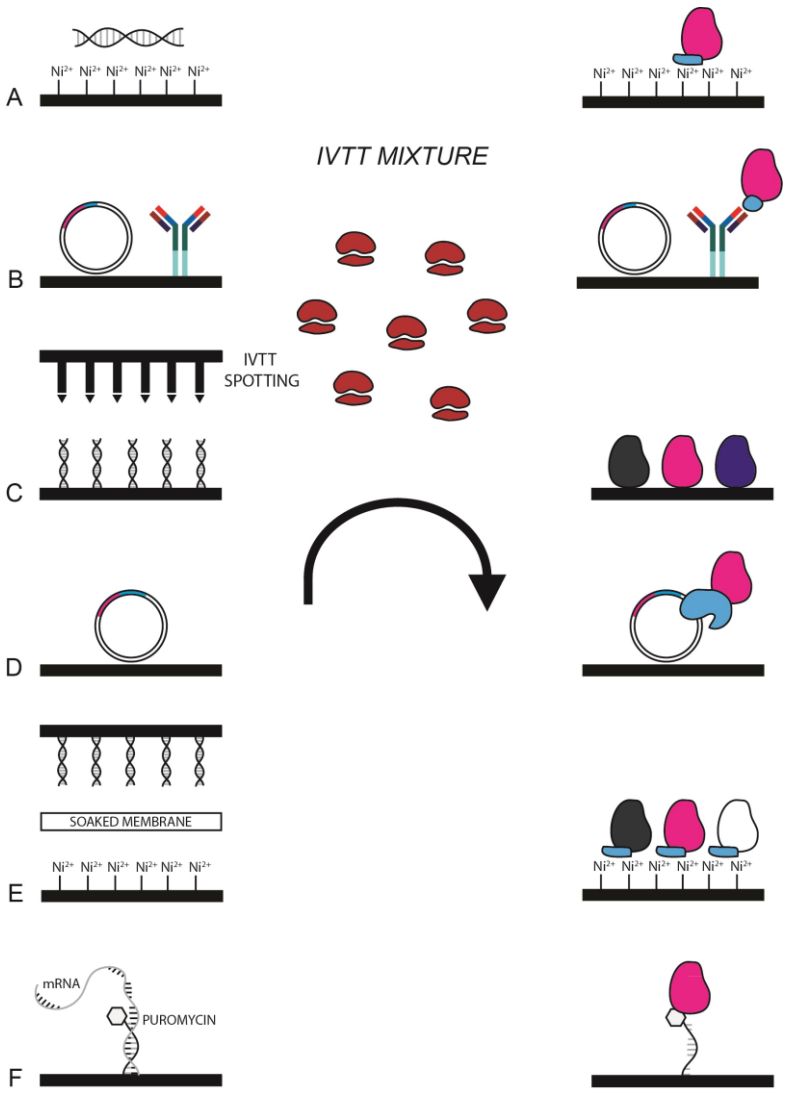


Figure 18. Schematic representation of all *in situ* protein microarrays. A) PISA: protein *in situ* array; B) NAPPA: nucleic acid programmable protein array; C) MIST: multiple spotting technique; D) TUS-TER array; E) DAPA: DNA array to protein array; F) *in situ* puromycin-capture from mRNA array.

2.7.7 IN SITU PROTEIN ARRAY APPLICATIONS

The main application of *in situ* protein microarrays is the individuation of protein-protein interactions. All the types of *in situ* protein array, previously described, can be used for interaction studies [55; 72; 74; 75]. In particular the NAPPA methodology is a great tool for this type of application due to the possibility of producing at same time query and target protein. La Bear et al. have demonstrated this capability using the Cdk inbhitor p16 as a query protein and Cdk4 and Cdk6 as positive targets, whereas Cdk2 has been used as negative control [72]. Furthermore, new protein-protein interactions were discovered thanks to *in situ* protein arrays [72; 85]. Besides interaction studies, the in situ protein

arrays which have been used for other kind of applications are NAPPA and PISA methods. The NAPPA was used in immunogenicity and immunoprofiling studies. In 2009, Montor et al., have directed a genome-wide study on *Pseudomonas aeruginosa* outer membrane proteins immunogenicity. According to this method, the reactivity of patients sera, positive or negative to cystic fibrosis, infected with *P. aeruginosa*, to bacteria outer membrane proteins was analyzed [86]. In this study the Authors identified a total of 48 antigens, 12 of which were detected in 10 patients and were considered as promising candidate for in-depth studies. Other similar studies were performed on different organisms, like Ornithodoros moubata [87], Plasmodium falciparum [88]and Coxiella burnetii [89], confirming that NAPPA is an important tool for immunogenicity studies and vaccine development. Moreover, NAPPA can also be considered suitable for immunoprofiling, since two different studies, in which this technology was used to analyze possible autoantibody biomarkers for early breast cancer and juvenile arthritis, were published [90; 91]. The usage of this novel methodology demonstrated that is possible to identify autoantibody biomarkers–that is–important not only for the diagnosis and the follow-up of the disease, but also to determine disease stages. In the matter of PISA method has been published the possibility to perform screening studies, functional assay and identification of mutants [92].

2.7.8 IMPROVEMENTS AND FUTURE PROSPECTIVES

In situ protein arrays have demonstrated to be a great tool for proteome studies and they have overcame the main drawbacks related to standard protein arrays. Under this aspect, many improvements regarding protein production, sensitivity and cross-talk were done in recent years. In 2010 a group of researchers have demonstrated that combining *E. coli* with wheat germ IVTT lysate, a major quantity of functional protein can be produced and captured on a glass slide, compared to rabbit reticulocytes lysate, that is the most IVTT used in *in situ* array technology [93]. Furthermore, based on NAPPA technology, these Authors have developed an autofluorescent microarray where the target protein is fused to a GFP, instead of GST-tag, and the nascent protein is captured by anti-GFP monoclonal antibody. With this approach, the protein is revealed immediately after the production, without the need of other treatments. Another important improvement of *in situ* protein arrays regards protein diffusion during the IVTT reaction that can cause cross-talking

between different spots. In a standard NAPPA slide the distance between two spot is 625μm that means approximately 2000 proteins analyzed on a single slide. This spot to spot space is the minimum distance necessary to avoid cross-talking events using NAPPA technology. To increase the density of NAPPA slide, without the presence of protein diffusion, a new type of slide was set up [94]. This slide is made of silicon and is composed of semispherical nanowells of 250μm in diameter, 75μm in depth and distanced each other 375μm. These nanowells are produced with photolithography and can be subsequently functionalized with different chemical groups. Moreover, every nanowell can contain 5nl of IVTT reaction mix indicating a minimal consume of cell-free IVTT systems that are expensive tools. The last challenge of *in situ* protein arrays is the use of label-based methods that are limiting and less sensitive, instead of label-free methods. For this reason, many groups have tried to combine this new type of arrays with label-free methods. Two different groups have demonstrated the possibility of combining NAPPA or MIST to matrix-assisted laser desorption/ionization (MALDI)[95; 96]. Furthermore, in 2012, an on-chip microfluidic protein microarray based on NAPPA technology was developed [97]. This method consists in a generator element where a dsDNA encoding for His6 tagged proteins and a detector element composed by a Cu(II)-NTA group are immobilized. When the microfluidic chip is treated with the IVTT mixture, the protein is produced from the DNA template and directly absorbed by the detector element and can be immediately used in surface plasmon resonance imaging (SPRI). Finally, the possibility to combine NAPPA methodology with other label-free methods like: nanogravimetry, atomic force microscope and anodic porous alumina, has been reported (APA)[98]. Thanks to all these improvements, *in situ* protein arrays are increasingly becoming a new technology able to analyze thousands of elements in almost one day and suitable for many other purposes.

3. AIM OF THE PROJECT

As described in the introduction, antibodies have become an important tool not only in scientific research, but also in proteome discovery, diagnosis and therapy. For this purpose, several improvements have been done in antibody production, selection and validation [23], but many proteins still remain unknown and without antibodies able to recognize and identify them. Moreover, many commercial antibodies present cross-reactivity or less specificity against their target [10; 11]. This leads to a not specific antibodies validation. It is became clear that an antibody has to be validated in relation with the experimental needs, but it can results laborious and cost-time expensive. For this reason the aim of this project is to develop new high-throughput technologies for antibodies validation. My laboratory has take part in a National Institute of Health (NIH) project, coordinated by Andrew Bradbury (National Laboratories of Los Alamos), that has the aim to produce, with high-throughput technologies, specific polyclonal antibodies against hundreds different proteins. In particular the project consists of (Fig. 19):

1. production of the target proteins;
2. selection of polyclonal antibodies through a phage/yeast display combination [43];
3. validation of the polyclonal antibodies with Western blot, ELISA and protein microarray;
4. set up of a new platform for high-throughput antibodies validation: protein array on demand.

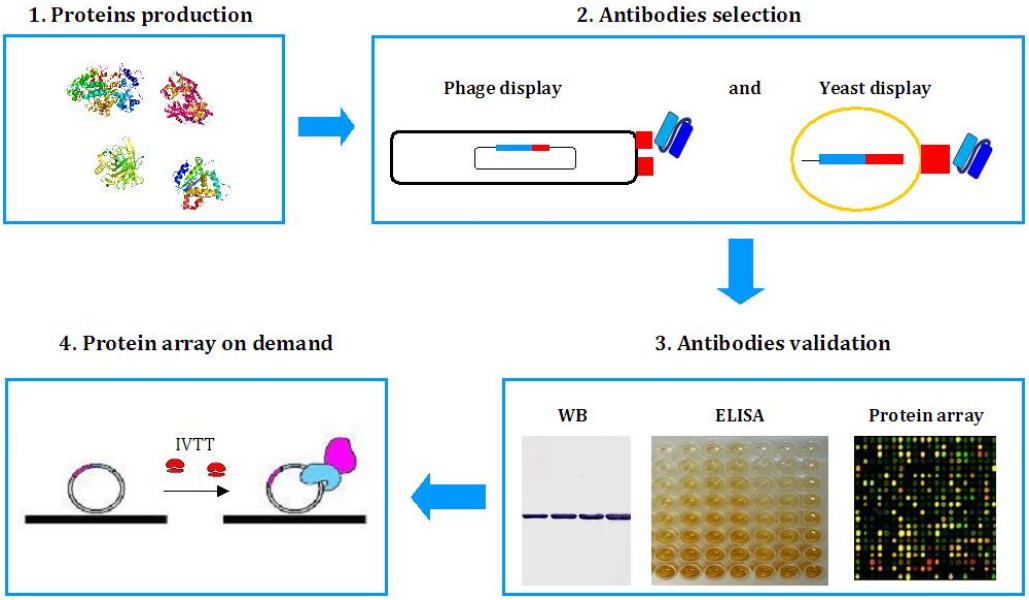


Figure 19. Experimental scheme of the project. The project consists of four different parts: 1. Protein production (antigens); 2. Selection of polyclonal antibodies against the putative antigens; 3. Validation of antibodies with WB, ELISA and protein microarray; 4. High-throughput antibodies validation using "protein array on demand".

The first part of the project (proteins production and antibodies selection) has been carried out in collaboration with Los Alamos Laboratories. After the scFvs selection, through phage/yeast display combination, and scFv-Fc rabbit polyclonals production in yeast cells, the polyclonals were sent us for validation with WB, ELISA and standard protein array. Subsequently, since protein microarray has some drawbacks, such as intensive protein production and immobilized proteins functionality, we want to set up an high-throughput protein array suitable for antibody validation. In particular, an *in situ* protein array called "protein array on demand" has been chosen for this purpose.

5. MATERIALS AND METHODS

Materials. As microarray slides FAST (Whatman) and Codelink® activated (Surmodics) were used. In ELISA, Western blot and protein microarray analysis the subsequent antibodies were employed: anti-SV5 (in house made), anti-mouse IgG horseradish peroxidase (HRP) conjugated (Dako), anti-mouse IgG alkalin phosphatase (AP) conjugated (Sigma), anti-mouse IgG Cyanine3 (Cy3) conjugated (Listarfish), anti-rabbit IgG HRP conjugated (Sigma), anti-rabbit IgG Cyanine5 (Cy5) conjugated (Listarfish) and polyclonal rabbit anti-CDK2A (full-length protein) (Santa Cruz). To reveal biotinylated proteins streptavidin HRP conjugated (BIOSPA), streptavidin AP conjugated (Sigma) and streptavidin phycoerythrin (PE) conjugated (Life Technologies) were used. As AP and HRP developer the nitro blue tetrazolium chloride/5-bromo-4-chloro-3-indolyl-phosphate (NBT/BCIP) (Roche) and 3,3',5,5'-Tetramethylbenzidine (TMB) (Sigma) were used. All PCR reactions were performed with GoTaq Flexi (Promega). In *vitro* transcription and translation kit used are: T7 TnT quick coupled transcription and translation system (Promega), T7 TnT quick for PCR DNA coupled transcription and translation system (Promega), TnT coupled Wheat Germ extract System (Promega) and 1-Step Human coupled in vitro expression kit (Pierce). Bacteria strains used: DH5α and NM522. Additional compounds: streptavidin (Biospa), bovine serum albumin (BSA) (Sigma), triton-X 100 (Sigma), tween-20 (Sigma), ampicillin (Sigma), 2xTY media (Sigma), select agar (Sigma), agarose (Sigma), TEMED and APS (Sigma), Page ruler Plus Protein marker (Fermentas), Quant-it™ PicoGreen® (Life Technologies).

Buffers and solutions. Codelink print buffer 6x (300mM sodium phosphate, pH 8.5), Codelink saturation buffer (10mM Tris-Hcl, 50mM ethanolamine, pH 9), phosphate buffered saline (PBS) (NaCl 0.137M, Na_2HPO_4 4,3mM, KCl 2,7mM, KH_2PO_4 1,5mM), TE buffer (Tris 10mM, EDTA 1mM pH 7.4), sample buffer 2x (Tris-HCl pH6,8 125mM, SDS 4%, glycerol 20%, β-mercaptoethanol 10%, blue bromophenol 0,004%), coomassie staining (acetic acid 10%, methanol 45%, coomassie blue R-250 0,1%), coomassie destaining (acetic acid 10%, methanol 10%).

Western blot antibodies validation. Biotinylated proteins were denaturated with SB 2x and loaded on a polyacrylamide gel for SDS-PAGE. After gel transfer, the nitrocellulose membrane was saturated in 4% milk in PBS plus tween-20 0,01% (PBST) for 1 hour at room temperature. The nitrocellulose was treated with polyclonal supernatants at 1:5 dilution in 2% milk in PBST and subsequently with anti-rabbit IgG-AP at 1:5000 in the same buffer. The membrane has been developed with NBT BCIP.

Sanger sequencing. PCR fragments were quantified by agarose gel and a sequencing reaction was carried out with: primer at 3,2µM concentration, 1µl of Big Dye Terminator (Applied Biosystems) mix and distillated water till 10µl final volume. The amount of DNA used depends on its length, according with the Big Dye datasheet. Susequently, the sequences reaction was purification on columns containing sephadex resin (Princeton Separation). After resin hydratation with 800µl of water, two centrifugation at 3600 rpm for 2' were carried out, to completely eliminate the water. The sequencing reaction was loaded on the resin and another centrifugation was performed, on order to eluate the purified reaction. This was mixed with formammide, denaturated for 2' at 96°C and analyzed with ABI 3130xl Genetic Analyzer.

***Antibodies validation with ELISA*.** Yeast supernatants, produced by National Laboratories of Los Alamos, containing rabbit polyclonal antibodies were validated using ELISA test. The corresponding proteins were coated on 96 well plate, at 10ng/µl in PBS and stored at 4°C overnight. The saturation was carried out with 2% milk in PBS and supernatants were used at 1:5 dilution in saturation buffer. The antibodies were revealed with anti-rabbit IgG-HRP at 1:5000. To verify the antigen coating a streptavidin-HRP at 1:200 were used.

Antibodies validation with protein microarray. All the biotinylated proteins, sent us from the Los Alamos National Laboratories, were brought to 50ng/µl concentration and put in a 384 well plate for microarray printing. The proteins were print on FAST slide, at 15°C and 50% humidity using the Biodeassay Calligrapher (BioRad). When printing finished, the slide was stored overnight at room temperature in a less humidity chamber. The day after, it was saturated with 4% milk in PBST at room temperature, with shacking, for 1 hour. Subsequently, the slide was treated with polyclonal supernatants at 1:5 dilution in 2% milk PBST for 1hour at room temperature. As secondary antibody an anti-rabbit IgG-Cy5, at 1:200 dilution, was utilized. To verify the protein spotting a streptavidin-PE, at 1:400

dilution, was used. To read the spots fluorescence and quantify them a ScanArray Gx (Perkin Elmer) was employed.

Protein array on demand vector design. A commercial plasmid optimized for cell-free expression in eukaryotic systems, called pTNT™ (Promega)(Fig. 20A), was used as template for rabbit reticulocytes lysate and wheat germ IVTT kits. The original Multiple Cloning Site (MCS) of pTNT™ was substituted with a new ones containing (Fig. 20B):

1. BssHII and NheI as restriction site for proteins of interest (POIs) insertion;
2. SV5 tag;
3. TUS protein cDNA with the E47Q mutation;
4. His_6 tag.

This new MCS was created through three different PCR cycles to obtain a DNA sequence that cuold be subsequently cloned in pTNT™, using XhoI and XbaI as cloning sites (Fig. 20B). The three PCR cycles were carried out with the following primers:

- TUS sense-1 (5'-TGCCGCTCATCTTGAGCAACACAAGCTATTGGTTGCCCGCGTGTTCTCTTTGCCGGAGTAAAAAACAGGATGAGCATAATCCGC-3'),
- TUS sense-2 (5'-GATGCGCGTTACGATCTCGTAGACCGACTCAACACTACCTTTCGCCAGATGGAACAAGAGCTGGCTATATTTGCCGCTCATCTTGAGCA-3');
- TUS sense-3 (5'-ACGTCTCGAGCCACCATGGGCGCGCATGCCAGGCTAGCGGCAAACCAATCCCAAACCCACTGCTGGGCCTGGATGCGCGTTACGATCTC-3')
- TUS anti (5'-ACGTTCTAGAACTTAATGATGGTGATGATGGTGATCTGCAACATACAGGTGCAGCCGTGG-3').

The new vector obtained will be called pTNT-TUS. All the proteins of interest were cloned using BssHII and NheI restriction sites.

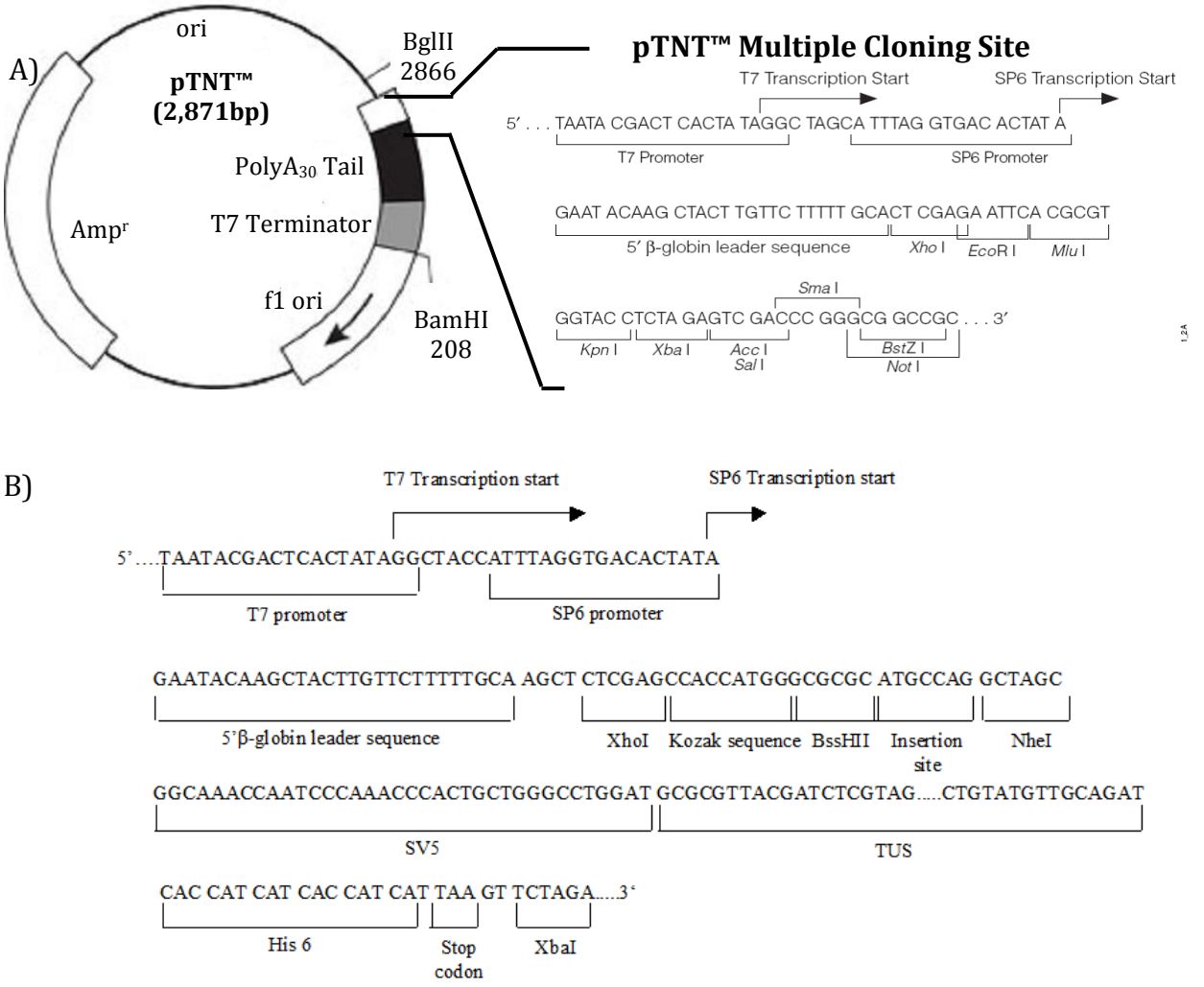


Figure 20. pTNT™ vector maps and Multiple Cloning Site. pTNT™ vector optimized for cell-free expression in eukaryotic lysate. A) pTNT™ vector map and MCS containing: ampicillin resistance, T7 and SP6 promoter, poly(A)$_{30}$ tail and T7 terminator. B) pTNT-TUS MCS.

Cell-free protein production. To allow the expression with HeLa cells IVTT system an encephalomyocarditis virus (ECMV) internal ribosome entry site (IRES) sequence was inserted in the pTNT-TUS vector. This sequence is essential for an efficient cell-free protein expression with lysate based on HeLa cells. The IRES was introduced in the pTNT-TUS vector through subcloning from pT7CFE1-His plasmid (Thermo Scientific), supplied with HeLa IVTT kit. The IRES sequence was amplified with a PCR reaction using the following primers:

- IRES XhoI sense: (AGCTCTCGAGGGTTATTTTCCACCATATTGCCG);
- IRES BSSHII anti: (AGCTGCGCGCCCATATTATCATCGTGTTTTTCAAAGG).

The amplified DNA was cut with XhoI and BssHII restriction enzymes and cloned in pTNT-TUS vector. This new plasmid will be called pIRES-TUS. All these plasmids were used with the corresponding IVTT lysate and the reaction was carried out following the kits data sheet (Tab. 1).

IVVT Lysate	**Rabbit reticulocytes (RRL)**	**Wheat germ (WG)**	**HeLa**
Compounds (50µl reaction)	• Master Mix 40µl • Methionine 1µl • Water till 50µl	• Extract 25 µl • Reaction buffer 2 µl • RNA Polymerase 1 µl • AA minus Met 1µl • AA minus Leu 1µl • Water till 50µl	• Lysate 15µl • Accessory protein 5µl • Reaction Mix 10µl • Water till 50µl
DNA (ng)	1000	1000	1000
Time and Temperature	1h30' at 30°C	1h30' at 30°C	From 90' to 6h at 30°C

Table 1. IVTT reaction kit data sheets. Description of the IVTT compounds, time and temperature needed for all the kits used.

All the kits could be used for protein expression starting from PCR fragments, but in case of IVTT based on rabbit reticulocytes lysate it was employed a kit optimized for PCR templates. The reaction compounds, time and temperature were the same of IVTT based on RRL plasmid kit. After the expression the proteins were loaded on polyacrylamide gel for Western blot analysis using the anti-SV5 at 1:5000.

TUS fusion protein binding to Ter sequence. The streptavidin at 10ng/µl was coated on 384 wells low volume plate (CORNING) at 4°C overnight. The day after the wells were washed briefly with PBS and subsequently incubated with 1000ng of 5'-biotinylated PCR obtained with primers pTNT seq bio (5'-bio- ACGGTTCCTGGCCTTTTGC-3') and pTNT anti or pTNT anti Ter (5'- CACTTTAGTTACAACAT ACTTATTAGGGAAGGGCGATCGGTG-3'), in which pTNT-TUS plasmid was the template. After 1h incubation, the DNA was discarded and the saturation with 2%milk in PBS was carried out. Subsequently, 10µl of the IVTT mix, already performed in vials, was added to the wells and incubated for 1h. After three 10' PBST washes, the anti-SV5 1:5000 was utilized to protein binding revealing. As secondary

antibody an anti-mouse IgG-HRP was used. All the incubation were performed at 30°C. The ELISA was read at 450nm with a Plate Reader.

Protein Array on demand. Biotinylated PCRs DNA at 400ng/µl were spotted on CodeLink activated slides. After the spotting, slides were incubated in a humidity chamber salt saturated overnight for covalent binding achievement. The day after, slides were saturated with CodeLink saturation buffer for 45' at room temperature and washed quickly with distillated water. Subsequently, for every 16 well PAD, 60µl of IVTT mix were deposited on the slides and incubated at temperature and time according to the IVTT kit used. After protein expression the slides were incubated at 15°C for 30' for TUS-*Ter* binding stabilization. After 15' PBST wash, the anti-SV5 at 1:5000 in PBS was used. Three 10' PBST washes were carried out before anti-mouse IgG-Cy3 at 1:200 in PBS addition. The previously washes were repeated and additional quickly washes with PBS and water were performed before slide drying. To control dsDNA printing the PicoGreen at 1:200 in TE buffer was used. Spot signals were read with the ScanArray Gx.

5. RESULTS

5.1 CHARACTERIZATION AND VALIDATION OF RABBIT POLYCLONAL ANTIBODIES

The first aim of the project was to select, with a phage/yeast display combination, scFvs produce them as scFv-Fc rabbit polyclonals, in yeast cells, and validate their specificity using Western blot, ELISA and protein microarray. In collaboration with the Los Alamos National Laboratories 78 different proteins were selected, for solubility and molecular structure, as target antigens of the NIH project (data not shown). All these proteins were produced as biotinylated proteins in *E. coli* and tested with Coomassie staining and Western blot analysis, carried out with streptavidin-AP staining (Fig. 21).

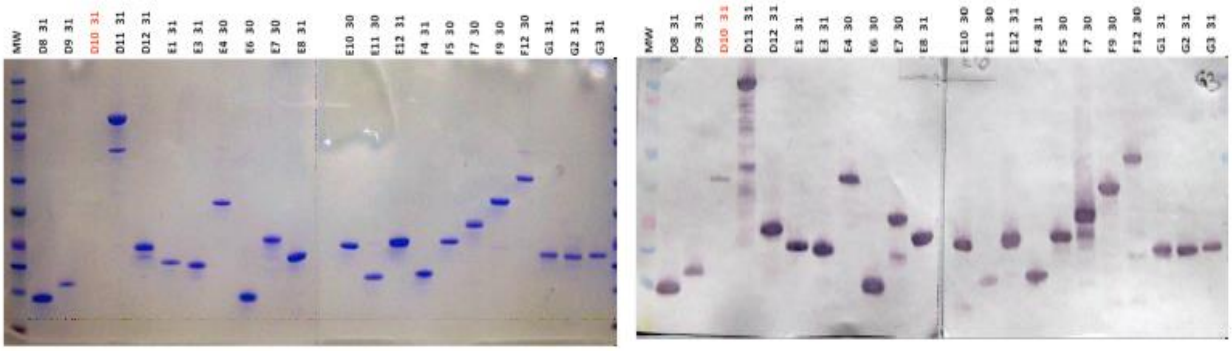


Figure 21. Putative antigens production. 78 proteins were produced as biotinylated proteins. One microlitre of each protein was loaded on a polyacrylamide gel and a Coomassie staining (left panel) and Western blot analysis(right panel) were carried out. The Western blot were performed with a streptavidin-AP 1:3000. Only 22 proteins are reported in this figure.

All proteins tested showed a good production yield using *E. coli* cells and were all soluble. Furthermore, all proteins were full-length and linked to the biotin. To evaluate the polyclonal antibodies selection protocol, only the first seven antigens of the list were selected to be used for the selection and production of polyclonal antibodies (Tab. 2).

PROTEINS (TOP 7)	ABBREVIATION	MOLECULAR WEIGHT (kDa)	LENGTH (bp)
Cyclin-dependent kinase 2 A	CDK2A	38,5	894
C-terminal binding protein 2a	CTBP2A	41,7	1005
MAP kinase 2A5	MAPK2A5	16,4	315
MAP kinase 8B	MAPK8B	46,4	1089
Phospholipase A_2 activating protein	PLAA	54,6	1374
Splicing factor 3A subunit 1	SF3A1	40,9	954
Ubiquitin specific peptidase 11	USP11	30,4	675

Table 2. TOP 7 antigens. List of the first seven proteins chose for the first round of polyclonal antibodies selection.

The strategy used was the following (Fig. 22):

1. a scFvs library [99] was cloned in a phagemid vector (pDAN) and two rounds of phage display selection were performed;
2. the enriched scFvs library was subsequently subcloned, with homologous recombination, in a yeast display vector, that used the a-agglutinin system for scFvs display. Two rounds of sorting with FACS were performed;
3. the selected scFvs were cloned, with homologous recombination, in a yeast expression vector to allow the production of scFv-Fc with a rabbit Fc.

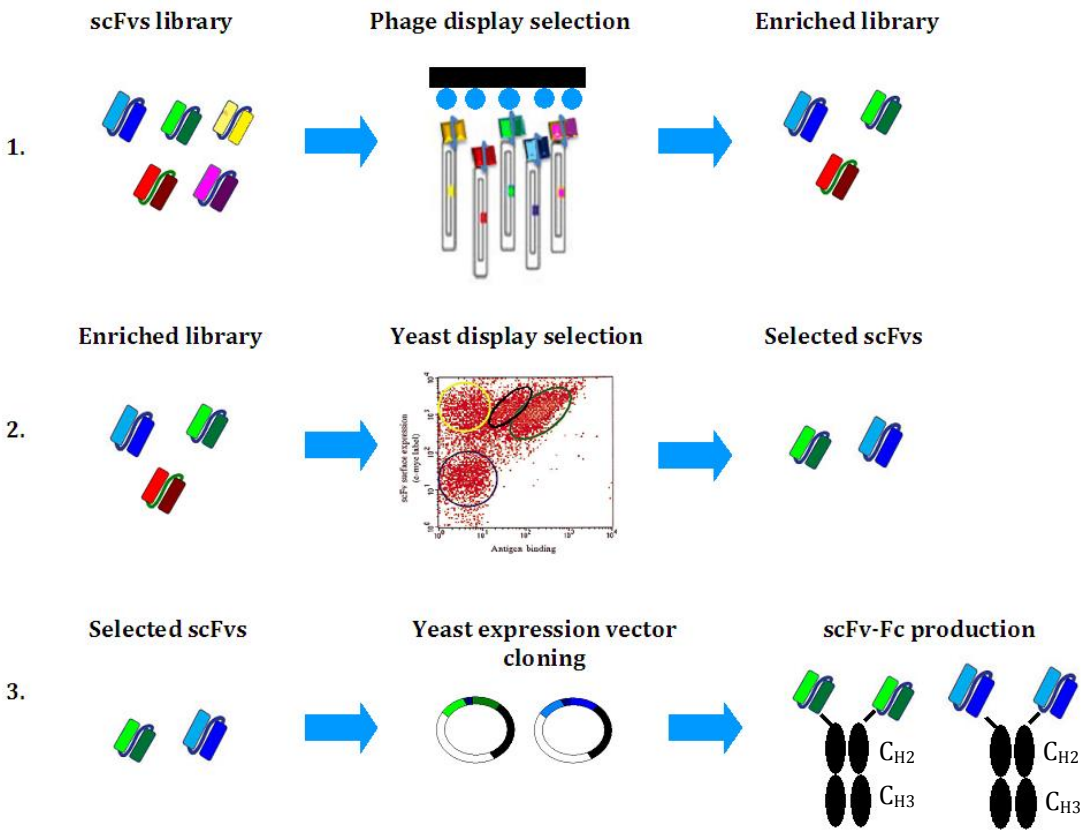


Figure 22. Polyclonal selection and production strategy. The strategy to select highly specific polyclonal antibodies consists of: 1. scFvs library cloning in a phage display vector and scFvs selection with two rounds of selection; 2. enriched scFvs library cloning in a yeast display vector and two rounds of selection with FACS analysis; 3. expression of the selected scFvs using yeast cells to obtain scFv-Fc polyclonal antibodies where the Fc is composed by rabbit IgG C_{H2}-C_{H3}.

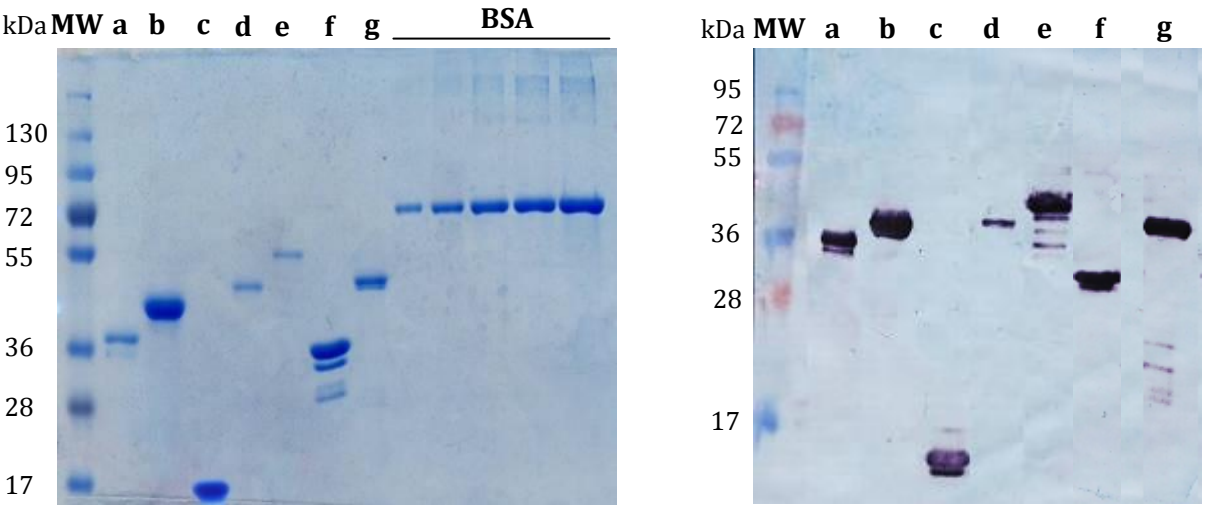
The scFv-Fc polyclonal antibodies were sent us as yeast cells supernatants for the validation analysis with: Western blot, ELISA, protein microarray. From the Los Alamos National Laboratory we received:

1. 78 antigens DNA in form of plasmid vector;
2. 78 biotinylated antigens.
3. polyclonal antibodies DNA in form of plasmid vector;
4. yeast supernatants containing the polyclonal antibodies;

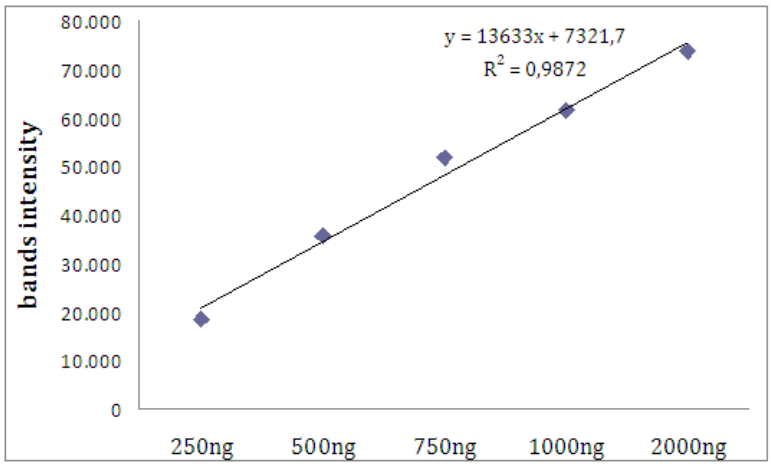
5.1.1 Western blot and ELISA polyclonal antibodies validation

After receiving all the samples from Los Alamos, the TOP 7 antigens were tested. To verify the grade of degradation and the protein concentration a Coomassie staining was performed (Fig. 23A). On the other hand, the proteins biotinylation were controlled with Western blot analysis using streptavidin-AP (Fig. 23A).

A)



B)



PROTEINS	CONCENTRATION (ng/µl)
CDK2A (A3)	375,4
CTBP2A (A7)	1632
MAPK2A5 (A10)	1426,5
MAPK8B (B5)	298,3
PLAA (C11)	265,9
SF3A1 (D2)	680,1
USP11 (D4)	1308,9

Figure 23 . TOP 7 antigens concentration and biotinylation test. The seven antigens sent us from Los Alamos National Institute were loaded on polyacrylamide gel to perform a Coomassie staining and a Western blot analysis. A) One microlitre of proteins was loaded in the following order: a) CDK2A; b) CTBP2A; c) MAPK2A5; d) MAPK8B; e) PLAA; f) USP11; g) SF3A1. To allow the protein quantification on the Coomassie gel a BSA calibration curve (250ng, 500ng, 750ng, 1000ng, 2000ng) was used. The Western blot was developed using a streptavidin-AP diluted 1:3000. B) Proteins were quantified by Imagej software; the BSA calibration curve was used as reference to construct a tendency line. The correspondent equation were used for proteins concentration calculation.

The Western blot analysis showed that all the proteins were full-length and biotinylated. Instead, in the Coomassie staining could be observed that the degradation was limited for all proteins tested. Furthermore, the Coomassie gel was used to estimate the protein concentration using the BSA calibration curve as reference. The signal intensity of bands was quantified by the ImageJ software and used to calculate the concentration of each protein (Fig. 23B). After antigen integrity and biotinylation evaluation, the yeast supernatants validation were performed initially with Western blot and ELISA analysis. In the first experiment, the polyclonals ability to recognize their specific antigens, in a denaturated form, compared to a control protein, were verified. Biotinylated antigens and the control protein (BSA) were loaded on a polyacrylamide gel and transferred on nitrocellulose membrane. Every polyclonal was tested on the correspondent antigen and the control protein (Fig. 24). All the yeast supernatants were used at 1:5 dilution and an anti-rabbit IgG-AP were used as secondary antibody, to allow the Western blot development.

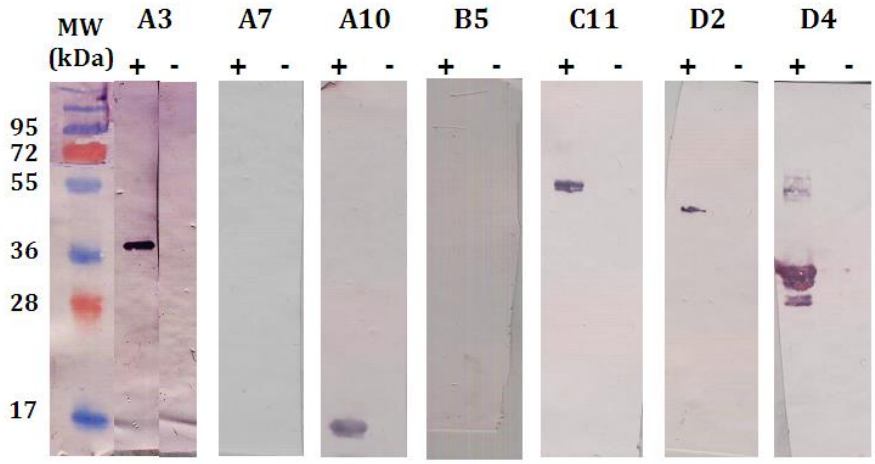


Figure 24. Western blot validation. The TOP7 antigens (400ng) were loaded on a polyacrylamide gel and transferred on a nitrocellulose membrane. Every protein (A3: CDK2A; A7: CTBP2A; A10: MAPK2A5; B5: MAPK8B; C11: PLAA; D2: SF3A1; D4: USP11) and a supplemented control protein (BSA) was treated with the correspondent yeast supernatant at 1:5 dilution. The anti-rabbit-AP 1:5000 was used for antibodies staining revealing.

From these Western blots it could be concluded that five polyclonal antibodies (anti-CDK2A, anti-MAPK2A5, anti-PLAA, anti-USP11, anti-SF3A1) were able to recognize the target in a denaturated form. On the other hand, only two supernatants didn't work in Western blot. This suggested that these two selected polyclonals could identify only conformational epitopes. Another hypothesis was that the yeast supernatant components

could interfered with the antigen recognition. An interesting aspect was that all the antibodies didn't show cross-reaction to the control protein, confirming high specificity of the selection protocol. After this preliminary results, a second validation test were performed using ELISA. In this analysis the yeast polyclonal supernatants were tested with all the seven biotinylated proteins, simultaneously, to control their specificity. The TOP7 antigens were coated on a 96 wells plate at 10ng/µl in PBS, at 4°C overnight. All the seven polyclonals were used at 1:5 dilution for ELISA development (Fig. 25A).

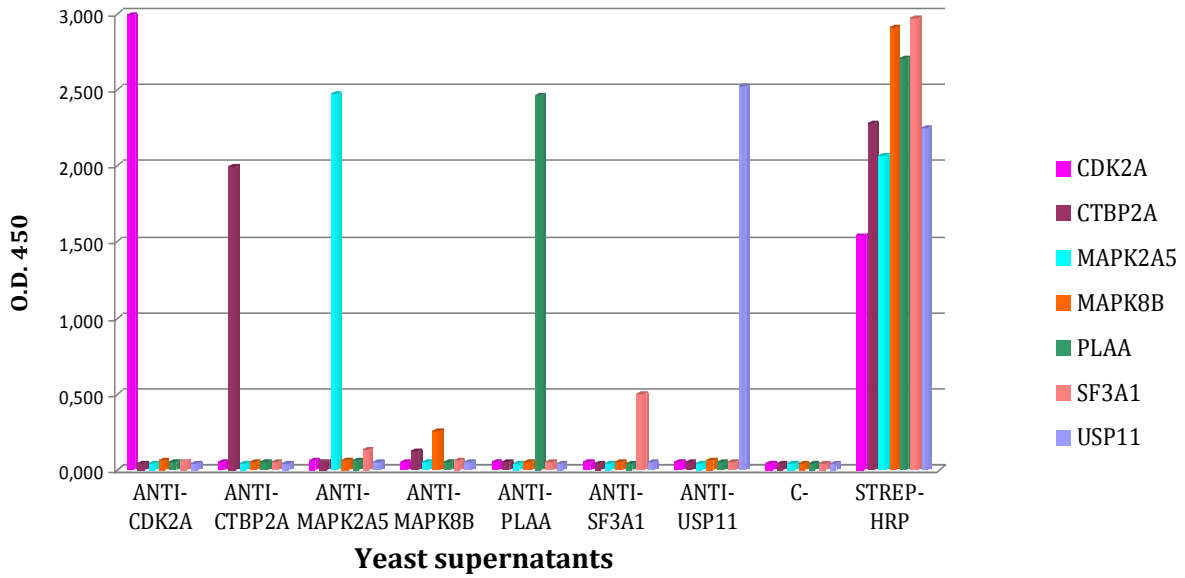


Figure 25. ELISA polyclonal antibodies validation. The seven antigens were coated at 10ng/µl concentration on ELISA plate. As primary antibody the yeast supernatants, at 1:5 dilution, were used. The proteins revealing was carried out with an anti-rabbit IgG-HRP. The values were read at an optical density of 450nm with a plate reader.

From these experiments could be observed that all the polyclonal supernatants were able to recognize specifically their own antigen. Furthermore, none of the antibodies showed cross-reactivity to the other proteins, confirming their high specificity. Despite that, the anti-SF3A1 and anti-MAPK8B polyclonals gave lower signals than the others. This problem could be related to a lower antibodies concentration in the yeast supernatant. In order to confirm this hypothesis antibodies concentration was measured by a second ELISA, with yeast supernatants coated directly on plate and developed with an anti-rabbit IgG-HRP (Fig. 26). In this way, it was possible to compare the amount of antibodies in the different supernatants.

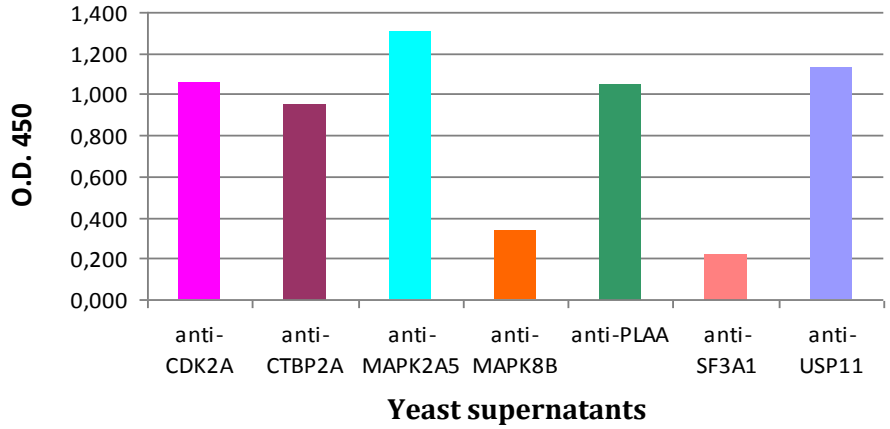


Figure 26. Antibodies concentration in yeast supernatants. Yeast supernatants, containing the polyclonal antibodies, were coated on ELISA plate at 1:10 dilution and developed with anti-rabbit IgG-HRP at 1:5000 dilution. The values were read at 450nm with a plate reader.

As previously supposed, the amount of antibodies in the anti-SF3A and anti-MAPK8B was lower than the other polyclonal supernatants. In conclusion, the anti-SF3A and anti-MAPK8B supernatants contained a low amount of antibodies that caused low signals in the ELISA validation.

5.1.2 Protein microarray optimization

The Western blot and the ELISA validations have shown that all the polyclonals could identify their target protein in the conformational state and only two of them couldn't recognize the antigen in a denaturated form. Furthermore, the seven antibodies didn't reveal cross-reactivity to the other six targets, confirming their high specificity. As showed in the introduction these two validation are not sufficient for a good antibody validation test. For this reason, another technology, able to validate the antibodies on a largest scale, were used. This technology is the protein microarray. The polyclonal supernatants were tested on the 78 antigens listed at the beginning of the project, already produced as biotinylated proteins. First, a way to standardize the amount of protein spotted was needed. All the biotinylated antigens were spotted at 50ng/µl and, after saturation, treated with streptavidin-PE (Fig. 27A). As negative controls six non biotinylated proteins, available in our laboratory, were used. In Figure 27 it is shown the optimized protein

spotting that will be used in all subsequent protein arrays. The scan images were also quantified and the values are reported in a graph (Fig. 27B).

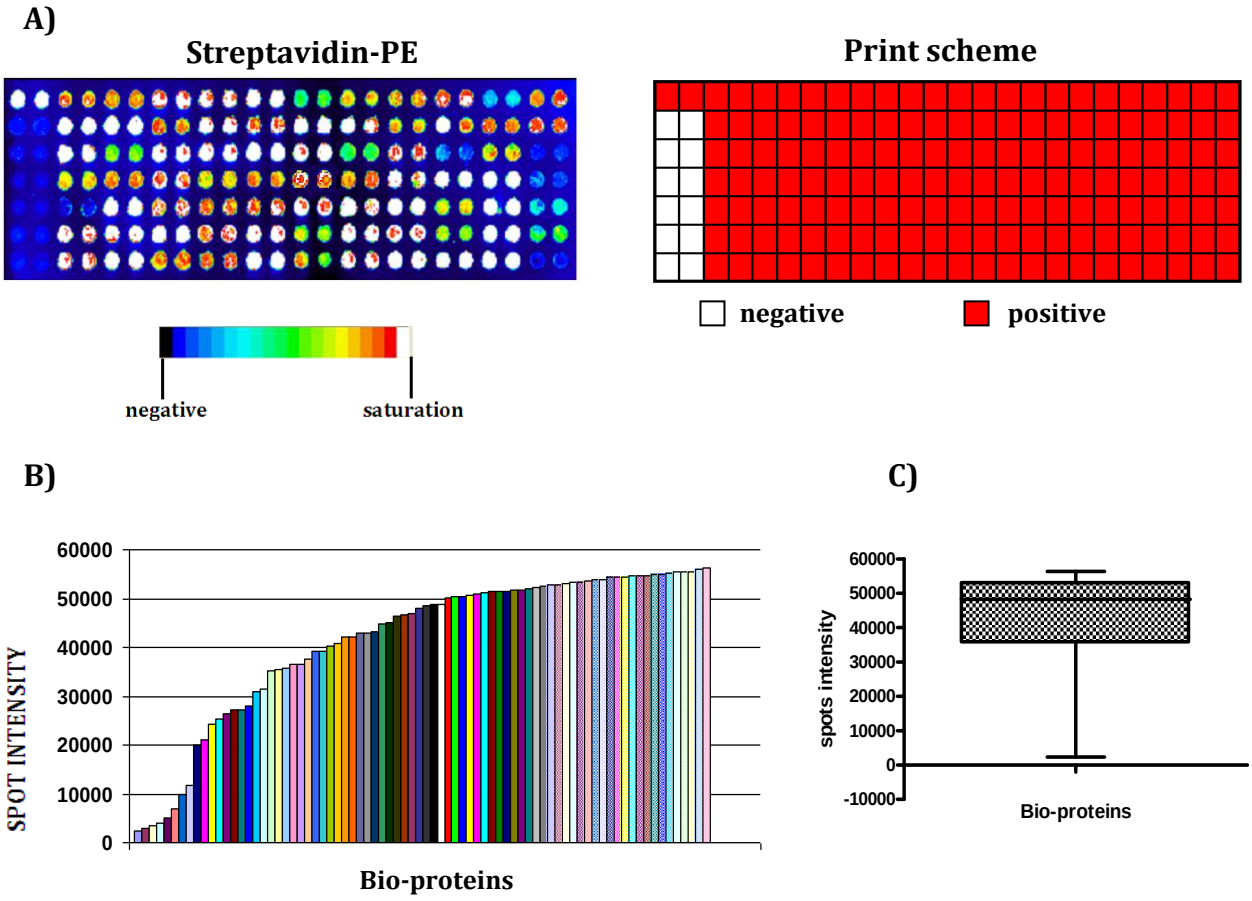


Figure 27. Protein microarray standardization. The 78 biotinylated antigens were spotted in duplicate, at 50ng/µl in PBS, on FAST slide. Other six random no biotinylated proteins were spotted as negative controls. A) Print scheme: biotinylated (red)and not biotinylated (white) proteins. After incubation with streptavidin-PE, the array was scanned with the ScanArrayGx. B) The spots intensity were calculated as media of the numbers of pixel minus background pixels by ScanArrayGx. For every proteins a media of the duplicates were calculated and reported in the graph. C) Box-plot of the quantified bio-proteins spots intensity.

As it could be observed in the array scan as well as in the histogram, although the level should be the same for all the proteins some were presented at a much lower concentration than expected. This could be due either to protein degradation, during storage, or to lower protein stability, when immobilized on the array slide. The distribution analysis, performed with a box-plot (Fig. 27C), revealed that the major part of the proteins was collected in the same intensity range, 35000 to 50000. Nevertheless, excluding the lower concentrated proteins, the majority of the spots intensities were comparable between each other and high enough for being detected with a specific antibody. To

perform a good protein array is also necessary to normalize the signals. Since that the protein array will be used to validate rabbit polyclonal antibodies, a rabbit IgGs calibration curve was used in the normalization. Rabbit IgGs were spotted at different concentrations on FAST slide and treated with the anti-rabbit IgG-Cy5 (Fig. 28). The array was scanned and quantified (Fig. 28).

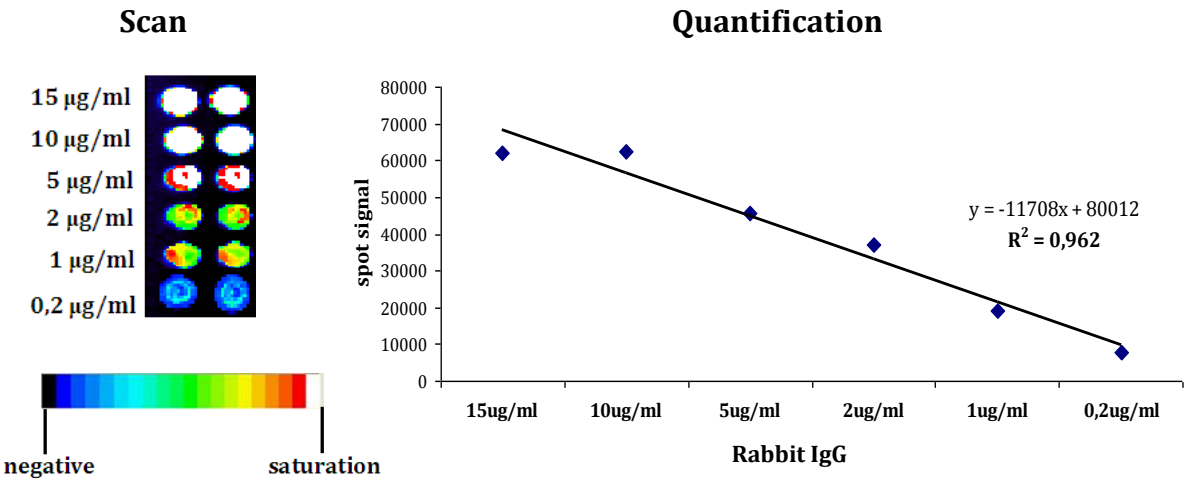


Figure 28. Rabbit IgGs calibration curve. To allow the polyclonal antibodies signals normalization a rabbit IgGs calibration curve was set up. Different concentrations of rabbit IgGs were spotted in duplicates and treated with anti-rabbit IgG-Cy5. In the left panel was shown the array scan with the correspondent IgGs concentrations. The quantification of these spots was reported in the graph (right panel) and a tendency line was designed. The values read by ScanArray Gx represent the media-background spots pixels. The media of the duplicates was calculated and put in the graph.

The IgGs calibration showed that the first two points of the curve (15µg/ml and 10µg/ml) corresponded to the saturation signal, demonstrating that 10µg/ml could be considered as the saturation point. Furthermore, from the graph could be observed that the IgGs tendency line was linear and with the R^2 value = 0,962. This analysis confirmed that this rabbit IgGs calibration curve could be used for all the further antibody validation by microarray.

5.1.3 Protein array validation of yeast polyclonal supernatants

Protein microarrays were set up and optimized for the polyclonal antibodies supernatants validation. The 78 biotinylated antigens and the rabbit IgGs calibration curve were spotted

as previously described. After overnight dehumidification, the slide was saturated and subsequently treated with yeast supernatants at 1:5 dilution (Fig. 29). As positive control a streptavidin-PE was utilized.

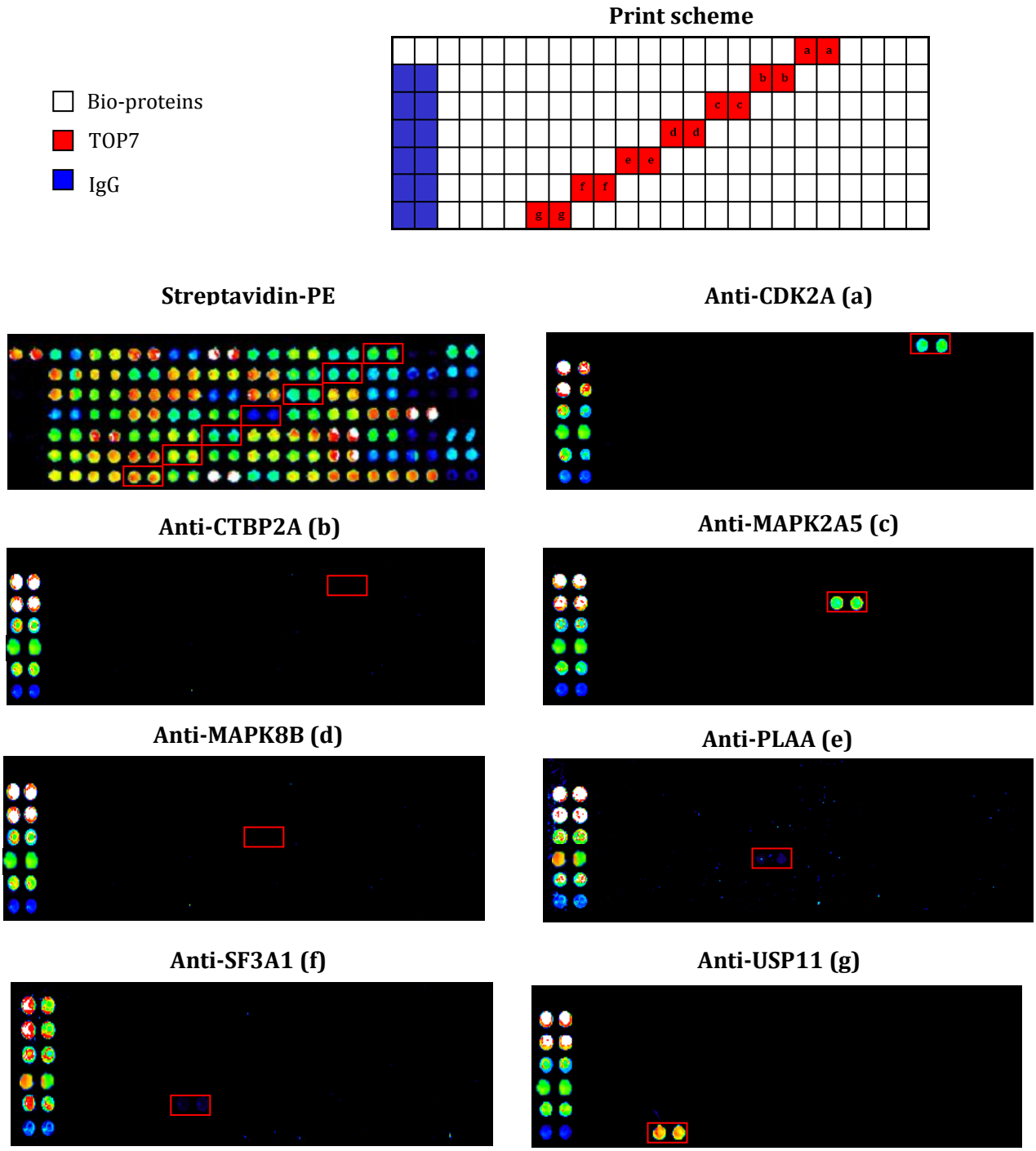


Figure 29. Yeast supernatants validation with protein array. The 78 antigens were spotted in duplicates on FAST slide at 50ng/µl. A rabbit IgGs calibrator (15µg/ml, 10µg/ml, 5µg/ml, 2µg/ml, 1µg/ml, 0,2µg/ml) was also printed. In the print scheme the TOP7 antigens are indicated as: a (CDK2A), b (CTBP2A), c (MAPK2A5), d (MAPK8B), e (PLAA), f (SF3A1) and g (USP11); whereas the rabbit IgGs are depicted in blue. A positive control with streptavidin-PE is shown in the top right panel. The other PADs were treated with all the different yeast supernatants at 1:5 dilution. The slide was scanned with the ScanArray Gx. For every scan the correspondent antigen is highlighted with a red rectangle.

The protein array validation (Fig. 29) showed that three polyclonal supernatants (anti-CDK2A, anti-MAPK2A5, anti-USP11) worked very well. They could recognized specifically their own target without showing cross-reactivity to the other proteins. The anti-PLAA and anti-SF3A supernatants were also able to recognize the specific target, but were less sensitive than the previous three. In case of anti-SF3A polyclonal, the problem could be related to the low antibody concentration in the supernatants shown in ELISA (Fig. 26). In the other case, the lower sensitivity of the anti-PLAA could be related to the presence of the yeast supernatant components, that could interfered with the antibody-antigen binding. Finally, the anti-MAPK8B and anti-CTBP2A seem to be unable to identify the target proteins. In case of anti-MAPK8B, analyzing the array with an higher laser power (excluding the rabbit IgGs calibration curve to avoid saturation) it was possible to get positive spots correspondent to the target protein (Fig. 30). This confirmed that anti-MAPK8B supernatant was functional and specific but gave a low signal out of the IgGs calibration curve. The array treated with the anti-CTBP2A resulted negative even if the IgGs curve was excluded (Fig. 30). Since that this supernatant worked in ELISA but not in Western blot, the problem in microarray analysis could be related to the degradation of the CTBP2A once spotted on the slide.

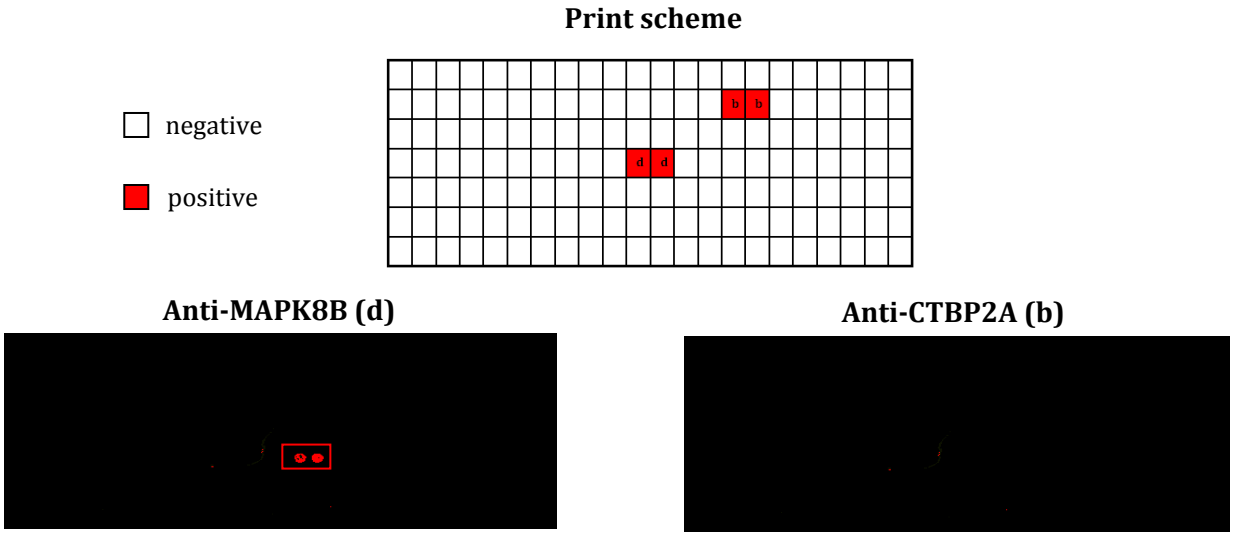
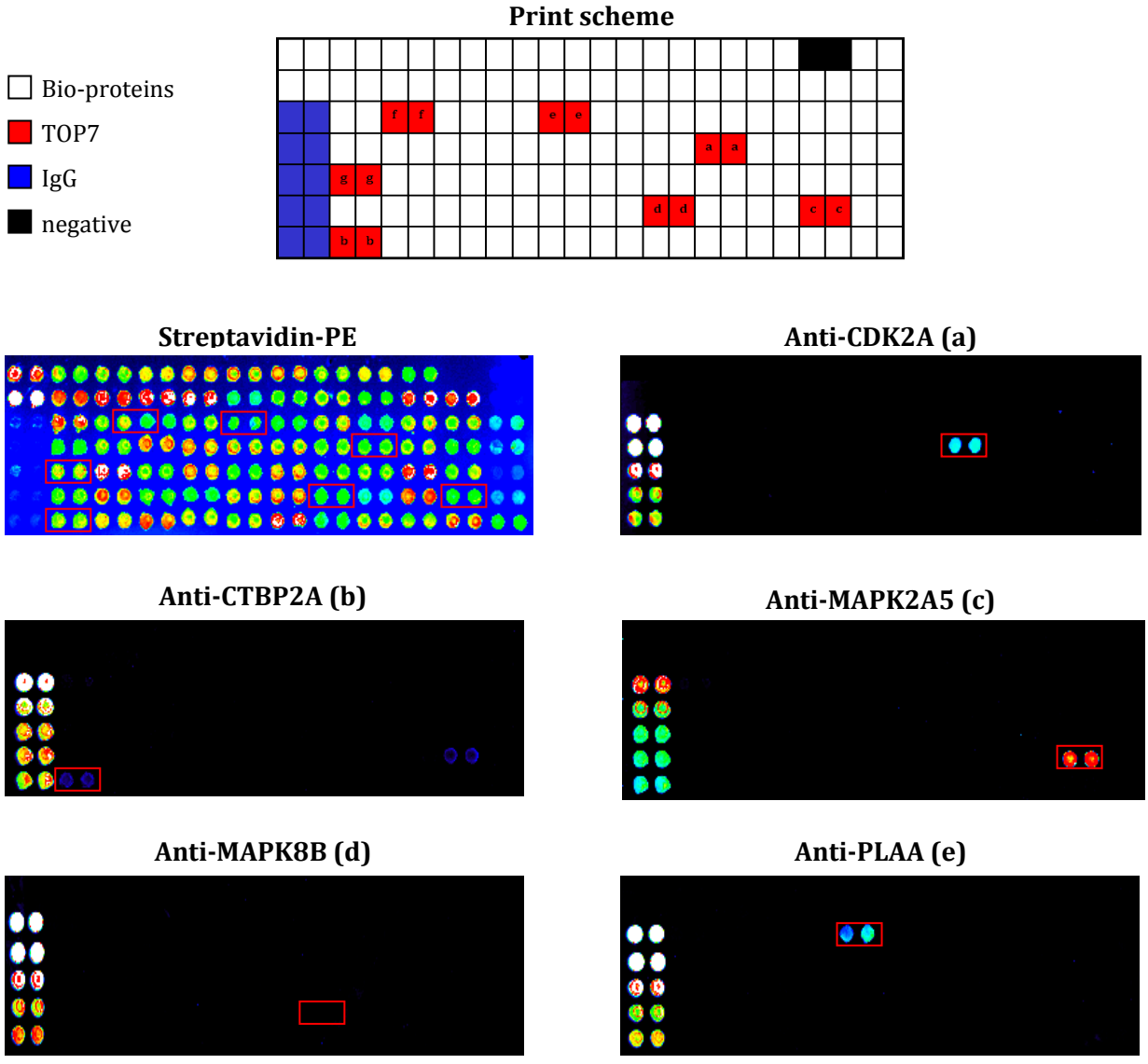


Figure 30. Array Scan without rabbit IgGs calibration curve. The PADs correspondent to anti-CTBP2A (b) and anti-MAPK8B (d) were re-scanned excluding the IgGs calibration curve.

5.1.4 Protein microarray validation of purified polyclonal scFv-Fc

From the previous protein array analysis arose that only three polyclonal supernatants were specific and sensitive to their target. The other polyclonals were resulted specific but with a lower sensitivity. One of the hypothesis, linked to this low sensitivity, was related to the antibodies concentration in the yeast supernatants. Indeed, the seven polyclonals presented different antibodies concentration that can leaded to different signals, independent to the antibody avidity. The other idea was related to the presence of the yeast supernatants components that could, in some cases, masked the antibody-antigen binding. For these reason the polyclonal supernatants were purified in order to obtain scFv-Fc polyclonals. These purified antibodies were sent us and tested with a new protein array (Fig. 31). All the purified polyclonal antibodies were used at 1ng/µl.



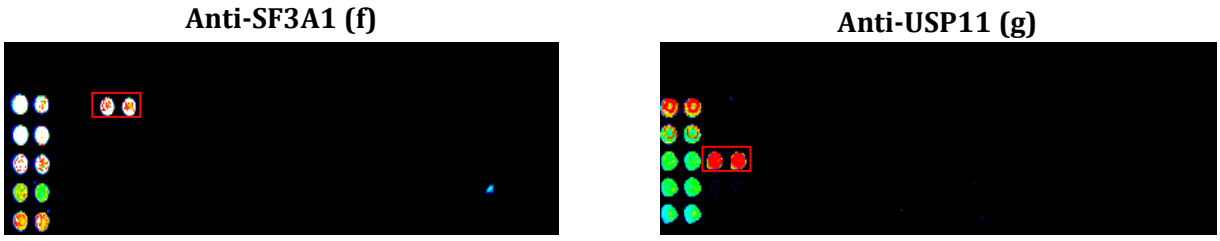


Figure 31. Purified polyclonal antibodies validation with protein array. The 78 antigens were spotted in duplicates at 50ng/µl. A rabbit IgGs calibration curve was printed for signal normalization (15µg/ml, 10µg/ml, 5µg/ml, 2µg/ml, 1µg/ml). In the print scheme are shown the TOP7 antigens position (a (CDK2A), b (CTBP2A), c (MAPK2A5), d (MAPK8B), e (PLAA), f (SF3A1), g (USP11)), the rabbit IgGs (bleu) and an empty spot (black). As positive control a streptavidin-PE was used. The other PADs of the same array were treated with 1ng/µl of purified polyclonal antibodies. The array scans were performed with ScanArray Gx. For every PAD the correspondent antigen is highlighted with a red rectangle.

In this new validation experiment it could be immediately noted that purified antibodies worked better that the not purified yeast supernatants. The three antibodies already positive as supernatants (anti-CDK2A, anti-MAPK2A5, anti-USP11), showed the same signals in the purified form. The anti-SF3A1 and anti-PLAA polyclonal, that in the previous validation were able to recognize their antigens with low sensitivity, in this case showed an enhanced specificity and sensitivity. As expected from previous results, all these five polyclonals didn't show a cross-reactivity to all the other proteins spotted on the slide. These results demonstrated that using the same amount of antibodies an higher signal can be achieved. Furthermore, the absence of the supernatant components might increase the antibody-antigen binding. The last two polyclonal antibodies (anti-CTBP2A, anti-MAPK8B) confirmed the low performance when used as purified as in supernatants. The purified anti-CTBP2A polyclonal was able to recognize its specific target, but reacted also to another protein: MAPK2A5. The purified anti-MAPK8B polyclonal didn't show any reactivity to its target protein and to the others. Even if the IgGs calibration curve was excluded from the scan the anti-MAPK8B resulted negative (data not show). In conclusion, five polyclonal antibodies resulted specific and sensitive to their own target, and could be used for research experiments. Only two of the seven polyclonal antibodies failed the protein array validation. To improve the anti-MAPK8B and anti-CTBP2A specificity some improvements will be performed in the selection protocol. The new polyclonals will be tested again with Western blot, ELISA and protein microarray.

5.1.5 Comparison with commercial polyclonal antibody: anti-CDK2A

Since that the validated polyclonal antibodies have to be used for research experiments, one of the yeast-purified polyclonals was compared to the correspondent commercial antibody available. The selected polyclonal was the anti-CDK2A. The yeast-purified and a commercial polyclonal rabbit antibodies directed to the full-length CDK2A protein were compared in ELISA and protein microarray. The following condition were used: commercial polyclonal at 0,2ng/µl (1:1000 dilution), commercial polyclonal at 0,07ng/µl (1:3000 dilution, recommended dilution) and yeast-purified polyclonal at 1ng/µl. Initially, the ELISA test was performed (Fig. 32). The biotinylated CDK2A and BSA were coated on a 96 wells plate and the two anti-CDK2A polyclonals were tested on the two proteins.

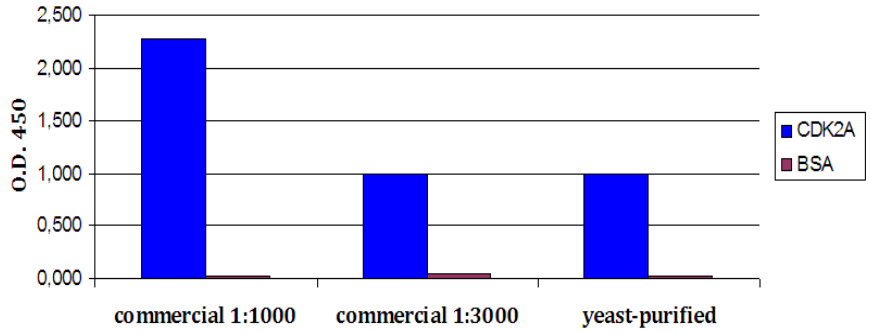
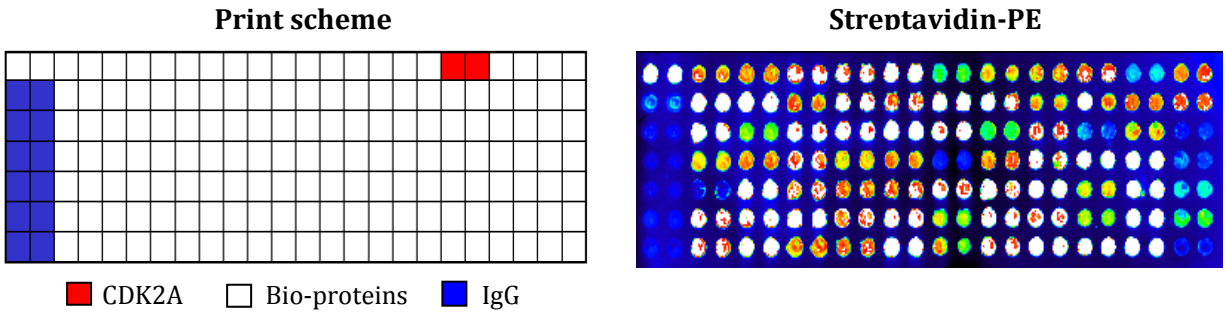


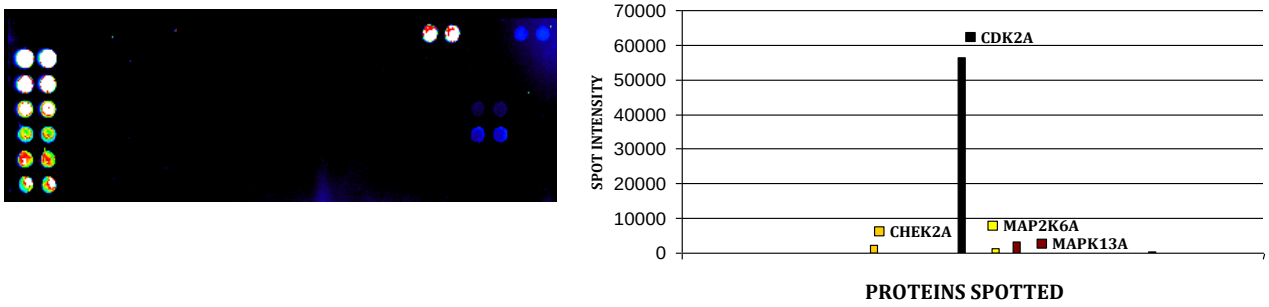
Figure 32. Anti-CDK2A polyclonals comparison with ELISA test. The biotinylated CDK2A was coated in a 96 wells plate and revealed with: commercial polyclonal antibody at 1:1000 dilution; commercial polyclonal antibody at 1:3000 dilution (recommended dilution); yeast-purified polyclonal antibody at 1ng/µl. As control protein the BSA was used. The values were read at 450nm, with a plate reader, and reported in the graph.

From the ELISA it could be seen that the commercial, at the recommended dilution (1:3000), and the yeast-purified polyclonals showed comparable specificity and sensitivity to the CDK2A protein, indicating that our polyclonal was comparable to the commercial ones. As expected, the commercial polyclonal, used at higher concentration (1:1000), showed an higher signal than the others. In this condition, with a simple control protein as BSA, both polyclonals didn't show cross-reactivity to the control protein. It has to be observed that the amount of the commercial antibody is lower than the yeast-purified. This suggested that the commercial seems more sensitive than the yeast-purified. To compare the antibodies on a larger number of proteins, a protein microarray was utilized. The 78 biotinylated proteins were spotted and treated with commercial, at different

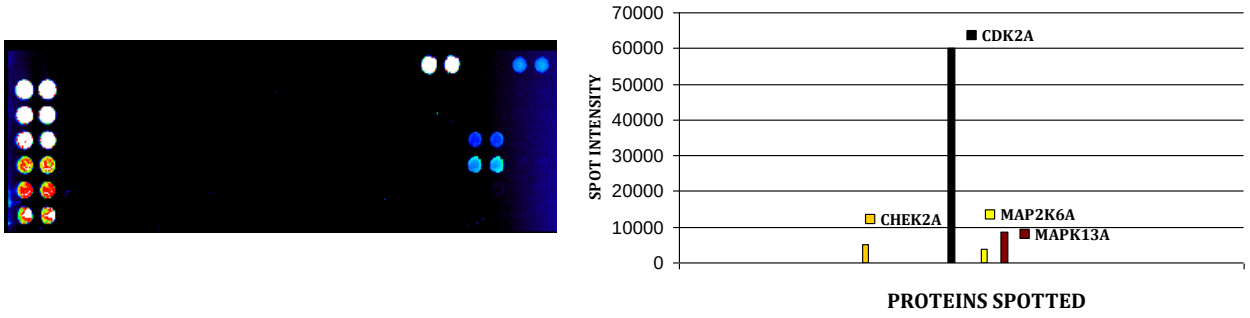
concentrations (0,07, 0,2, 1 ng/µl), and yeast-purified, at 1ng/µl, anti-CDK2A polyclonals. As positive control the streptavidin-PE was used (Fig. 33). An array scan and quantification were performed (Fig. 33).



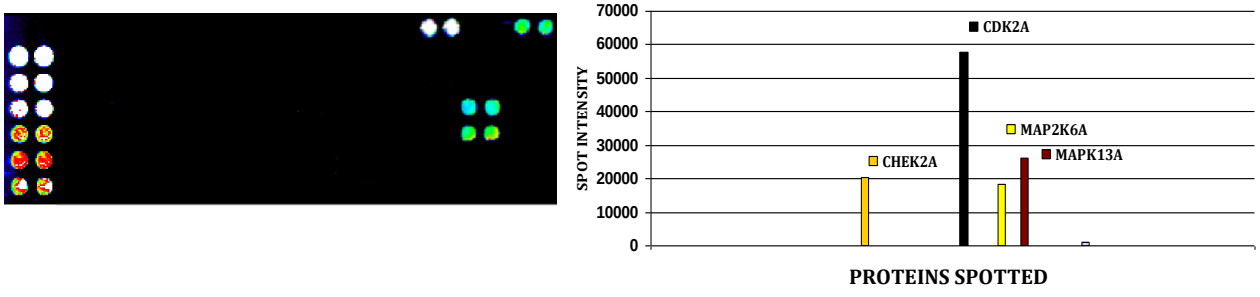
Commercial anti-CDK2A **0,07ng/µl**



Commercial anti-CDK2A **0,2ng/µl**



Commercial anti-CDK2A **1ng/µL**



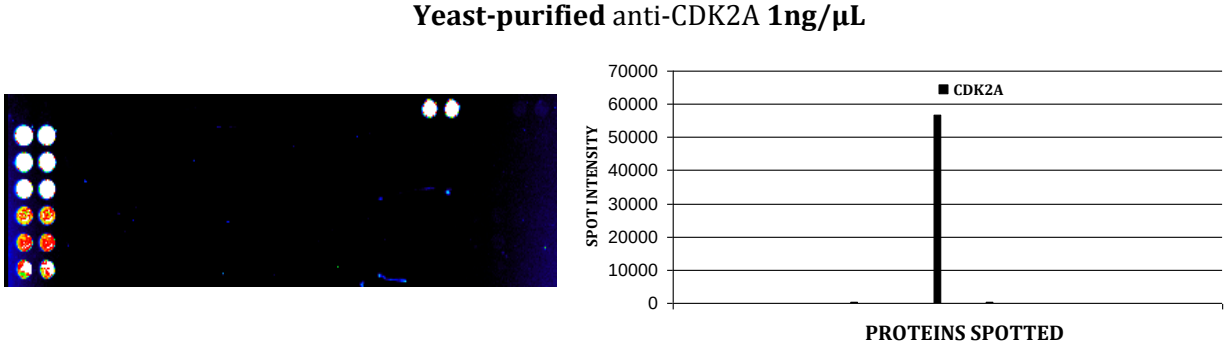


Figure 33. Commercial and yeast-purified anti-CDK2A polyclonal comparison with protein array. 78 biotinylated antigens were spotted in duplicates at 50ng/µl on FAST slide. In the print scheme the CDK2A is depicted in red and the rabbit IgGs calibration curve in blue (15µg/ml, 10µg/ml, 5µg/ml, 2µg/ml, 1µg/ml, 0,2µg/ml). After saturation, the PADs were treated with the commercial polyclonal anti-CDK2A, at 0,07, 0,2 and 1ng/µl, and the yeast-purified anti-CDK2A at 1ng/µl. Every PAD was developed and quantified with ScanArray Gx. The values correspondent to pixels media-pixels background were reported in a graph as media of the duplicates spots.

Figure 33 shows that the commercial polyclonal anti-CDK2A specifically recognizes its target protein, but also other three unspecific proteins: MAP Kinase 13A (MAPK13A), MAP 2 Kinase 6A (MAP2K6A) and Check point Kinase 2a (CHEK2A). The unspecificity of the commercial polyclonal was present in all the three dilutions, confirming that these three proteins could be potentially recognized in standard research experiments. On the other hand, the yeast-purified polyclonal didn't show any particular cross-reactivity to that three proteins or to the others. Furthermore, from the quantification analysis arose that the signal intensity of CDK2A spots in the yeast-purified PAD was comparable to the commercial ones. To investigate the reason why the commercial antibody recognized the three unspecific proteins, they were studied. First, the amminoacids sequences of CDK2A, MAPK13A, MAP2K6A and CHEK2A proteins were aligned, using ClastalW2 tool, to reveal possible common parts that could be recognized from the antibody. From this study arose that there weren't any similarity in the four different proteins sequences. Subsequently, it was performed a research to underline possible hits patterns that could be present in all the proteins. Using PRATT tool it was showed that all the four proteins contain this particular pattern: G-[ERS]-G-[AT]-x-G-x(0,1)-V-x(4)-[DENT]-[KR]-x-[STV]-[CGP]-x(3)-[AI]-[ILM]-[EK]-x-[AIL]. This pattern could possibly form in these proteins a conformational epitope that could be potentially recognized from the commercial polyclonal antibody. Since that the yeast-purified polyclonal anti-CDK2A didn't cross-react with the other three proteins, it could be supposed that the antibodies directed to this epitope are not present

or are present in a very low amount. At the end, from the protein microarray analysis it could be affirmed that the anti-CDK2A yeast-purified polyclonal antibody could specifically recognize the CDK2A protein without showing any cross-reactivity. As a conclusion, five polyclonal antibodies selected with phage/yeast display combination method were validated with Western blot, ELISA and protein microarray resulting with high specificity and sensitivity. When used in a direct comparison the anti-CDK2A resulted better than the correspondent commercial available polyclonal antibody. Taken together, this show that this technology had a potential for the production of specific antibodies that could be more efficient than the commercial ones available. Finally, protein microarray technology confirmed to be a very important tool in the antibody validation analysis.

5.2 PROTEIN ARRAY ON DEMAND ASSESSMENT

The second aim of this project was to set up an high-throughput technology that could be utilized for antibody validation. As showed in the introduction protein microarray technology presents some drawbacks such as: production and purification of all the proteins tested and possible degradation of the immobilized proteins. To overcome these problems a new type of protein microarray called *in situ* protein array has been set up (Fig. 17). In particular, this technology allows to spot cDNA that is directly transcribe and translate by IVTT system. To allow the capture of the newly synthetized protein a capture agent is also needed. With this new technology, the protein production passages are eliminated and the protein is tested immediately after its *in situ* production. The *in situ* protein array of our interest is called "protein array on demand" [74]. This technology is the only *in situ* protein array that used the DNA molecule as template and capture agent. In detail, "protein array on demand" is based on the capability of an *E. coli* protein, called TUS, to bind specifically a dsDNA sequence, called *Ter* (Fig. 34).

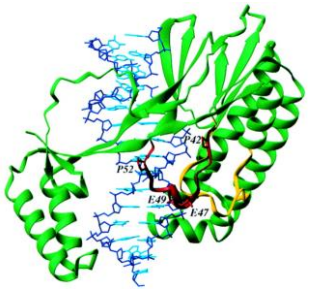


Figure 34. TUS-*Ter* binding. Interaction of the TUS protein (green) with the *Ter* sequence (blue).

This *in situ* protein array consists in the printing of a DNA encoding for proteins of interest (POIs) fused to the TUS protein. The immobilized DNA is transcribed and translated with an IVTT system and the nascent protein directly bind to the *Ter* sequence, located in the encoded DNA molecule (Fig. 35). So, only the DNA, in form of plasmid, has to be printed on the array slide to perform a protein array on demand. Therefore, for this type of array, the template DNA construction is essential.

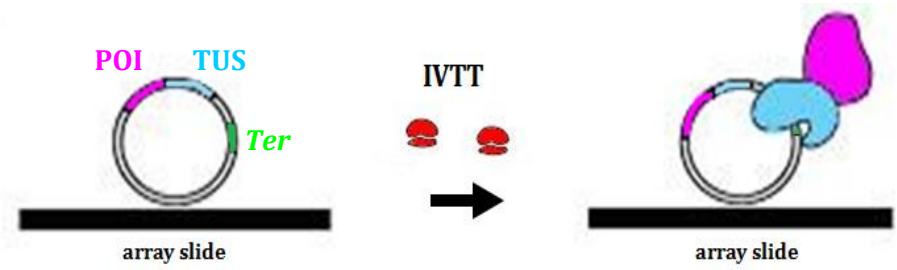


Figure 35. Protein array on demand. Schematic representation of the protein array on demand technology. A plasmid DNA containing protein of interest (pink) fused to TUS protein (light blue) and *Ter* sequence (green) is printed on an array slide. After transcription and translation with IVTT system of the immobilized DNA, the protein binds the DNA through TUS- *Ter* binding.

5.2.1 Protein array on demand template construction

As showed in figure 35, the protein array on demand is based on the DNA template immobilization on array slide, that encodes proteins of interest (POIs) fused to TUS and contains the *Ter* sequence. Furthermore, this DNA molecule has to be suitable for IVTT protein production systems. As the protein expression in the protein array on demand technology was carried out with an IVTT system based on rabbit reticulocyte lysate, a DNA vector optimized for IVTT protein production in this eukaryotic system was used. This particular vector is called pTNT™ (Fig. 7A) and contains:

- T7 promoter for transcription initiation;
- 5' β globin leader sequence and Kozak sequence as ribosome binding sites;
- multiple cloning site (MCS) for protein of interest insertion;
- stop codon;
- T7 terminator for transcription ending.

As a first step, it was necessary to modify the vector by inserting the TUS coding sequence. In particular, the TUS cDNA was inserted without the first methionin to allow its fusion with POIs. The strategy that was used for TUS cDNA insertion in pTNT vector was the following (Fig. 36): 1) extraction of *E. coli* genomic DNA containing the TUS cDNA; 2) amplification with PCR of the TUS CDS; 3) introduction with PCR of restriction sites (BssHII-NheI) for POIs insertion and tags for detection; 4) pTNT cloning. The resulting vector was called pTNT-TUS. This plasmid was transformed in DH5α cells and sequenced to confirm the correct DNA insertion.

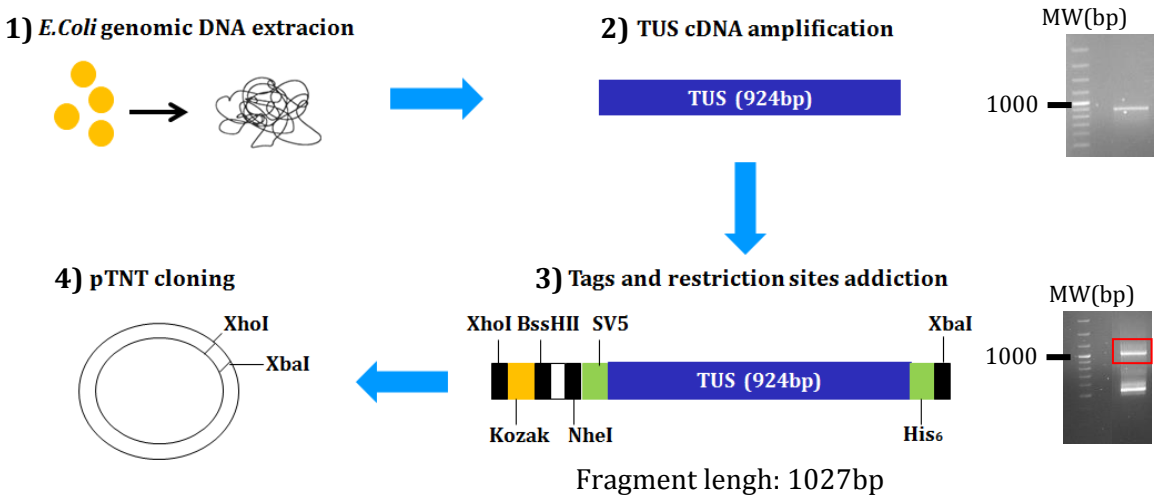


Figure 36. TUS cDNA insertion in the pTNT plasmid. The MCS of pTNT™ vector had to be modified to allow the insertion of TUS protein CDS, tags for protein detection and cloning sites for POIs insertion (BssHII-NheI). 1) *E. coli* NM522 genomic DNA extraction. 2) PCR amplification of TUS CDS (blue) without the first methionin codon. 3) PCR amplification of TUS CDS and addiction of Kozak sequence (yellow), restriction sites (black) and tags (green). 4) pTNT™ vector digestion with XhoI and XbaI enzymes of the PCR fragment obtained. All the PCR reactions were controlled on agarose gel.

To perform the protein array on demand a *Ter* sequence had to be added to the pTNT-TUS vector. To improve the avidity of TUS protein to the plasmid DNA a multiple *Ter* sequence was inserted. This multiple *Ter* contained three *Ter* sequences divided by random nucleotides (21-22bp) (Fig. 37). This DNA fragment of 230bp was synthesized by GeneArt® gene synthesis as plasmid vector. The sequence of interest was extracted through enzymatic cut with BamHI and inserted in the pTNT-TUS vector using the same restriction site. This new vector was called pTNT-TUS-*Ter*.

Figure 37. Multiple *Ter* sequence. A sequence containing three *Ter* sequences, divided by 21-22 random nucleotides and an AgeI restriction site, was constructed. At 5' and 3' end was put a BamHI site that allowed the sequence cloning in the pTNT-TUS vector.

Finally, the pTNT-TUS-*Ter* could be used for protein array on demand experiments.

5.2.2 Protein expression and Ter binding

Once obtained the suitable template for protein array on demand it was tested using four proteins available in our laboratory (Tab. 3). All these proteins were inserted in the pTNT-TUS vector using BssHII and NheI as restriction sites.

PROTEINS	kDa
DPP6 (Dipeptidyl-peptidase 6)	3,6
RON2 (Macrophage stimulating 1 receptor/2)	4,7
TG2 DM1 (Transglutaminase 2 domain 11)	47
eGFP (Enhancer green fluorescent protein)	23,6

Table 3. POIs cloned in pTNT-TUS. Proteins already available in our laboratory used to test the protein array on demand.

After POIs cDNA cloning, the plasmids were transformed in *E. coli* DH5α cells and sequenced. Before proceeding with the protein array on demand set up, the proteins expression and their *Ter* binding capability were tested. Proteins expression was analyzed with Western blot. The IVTT reaction was carried out using the T7 TnT Quick transcription and translation kit with 200ng of template in a final volume of 10µl. After 1h30' at 30°C the reactions were denaturated with sample buffer 2x and loaded on polyacrylamide gel. The

Western blot was carried out using the anti-SV5 antibody to reveal the TUS fusion proteins (Fig. 38).

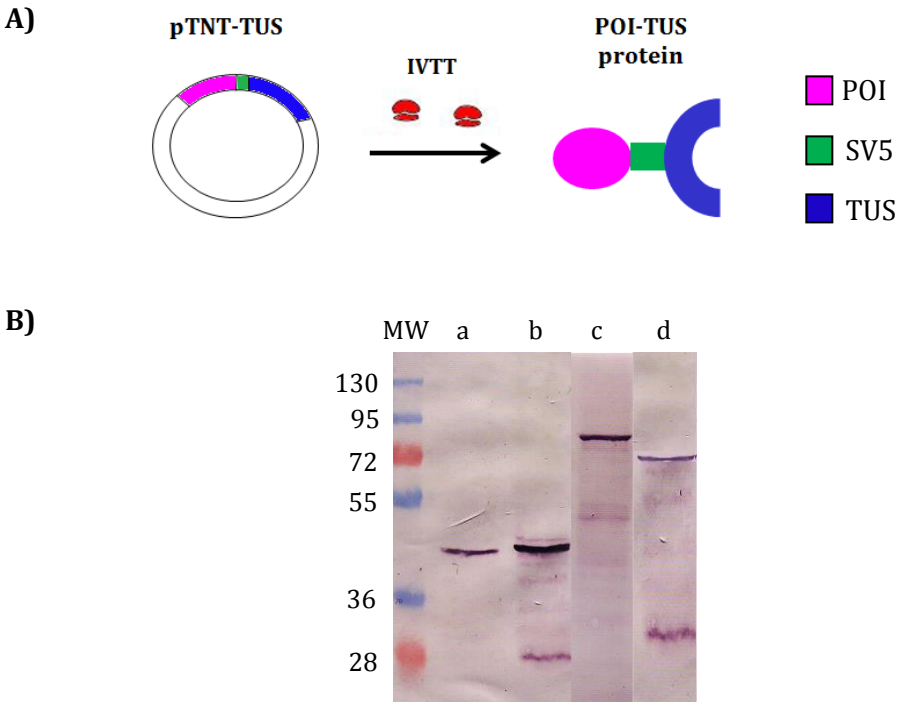


Figure 38. TUS fusion proteins expression with IVTT system. Schematic representation of the TUS fusion proteins expression by IVTT system (A). The TUS fusion proteins were produced in 10µl of IVTT reaction and loaded on a polyacrylamide gel at 10%: DPP6-TUS (a), RON2-TUS (b), TG2 DM11-TUS (c), eGFP-TUS (d). After gel transfer, the nitrocellulose was developed with anti-SV5 and anti-mouse IgG-AP (B).

From the Western blot analysis it could be observed that all the TUS fusion proteins were produced by IVTT system based on rabbit reticulocytes lysate. Furthermore, a very low protein degradation was visible, indicating that this type of protein expression system is efficient. Despite that, it was possible to observe that the amount of proteins were different even if the reaction volume and the template amount were the same. These differences could be related to the proteins sequence or length. After protein production assessment, it was necessary to confirm that this proteins were able to bind efficiently the *Ter* sequence. To test it, an ELISA, in a low volume 384 wells plate, was set up. The molecule coated in the ELISA well was a DNA fragment, able to express POIs fused to TUS as well as containing the *Ter* sequence for protein immobilization (Fig. 39A). This DNA fragment was obtained through PCR amplification using pTNT-TUS vector as template. This PCR DNA was immobilized wells through streptavidin-biotin binding on the ELISA. Since that the PCR DNA amount necessary for proteins expression and *Ter* binding was unknown, an ELISA

set up was performed. After streptavidin coating, different concentrations of PCR DNA, encoding for eGFP-TUS protein, (200, 400, 600, 800, 1000, 1500, 2000 ng) were used for DNA immobilization set up. After 1h binding, the DNA was recovered from the plate and loaded on agarose gel (Fig. 39B). To understand the amount of DNA bound on the plate a DNA calibrator was used. Subsequently, the eGFP-TUS protein, already produced with IVTT system, was added to the same ELISA wells and incubated for 1h. The anti-SV5 was used for protein detection (Fig. 39C). As negative control a PCR DNA lacking the *Ter* sequence was used.

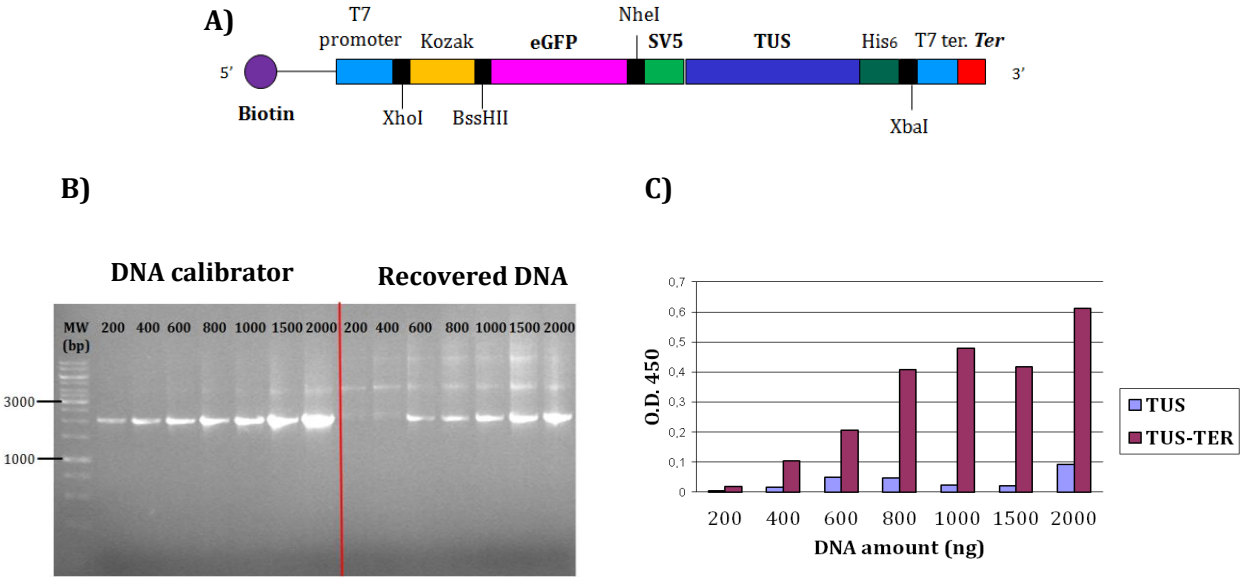


Figure 39. ELISA set up for TUS-*Ter* binding study. To verify the TUS-*Ter* binding an ELISA test was set up. A) PCR DNA fragment, immobilized on the ELISA wells, containing: all sequences necessary for IVTT (T7 promoter, Kozak, T7 terminator), eGFP-TUS coding sequence, *Ter* sequence at 3' end and biotin at 5' end. B) Different concentrations of PCR DNA (200, 400, 600, 800, 1000, 1500, 2000 ng) were incubated with streptavidin coated wells for 1h. After the incubation the DNA was recovered and compared on an agarose gel to a DNA calibrator. C) After PCR immobilization the eGFP-TUS protein, produced with IVTT system, was incubated in the same wells for 1h. The protein bound to DNA was detected with anti-SV5 at 1:5000 dilution. As negative control a PCR without *Ter* sequence was used.

From the agarose gel it could be observed that the higher amount of PCR DNA that can be captured from the streptavidin coated well was 400ng (Fig. 39B). Observing the protein detection results (Fig. 39C) it could be evident that the higher amount of protein was bound to the well correspondent to 2000ng PCR DNA. Despite that, the correspondent negative control was also high compared to the other tested conditions. For this reason, in the further experiments will be used 1000ng PCR amount. Indeed, this condition corresponded to the higher amount of protein detection and the lower background signal.

After calibration set up, the same ELISA was performed for all the other TUS fusion proteins (DPP6, RON2, TG2DM11) (Fig. 40). The experiment was repeated for three times in order to performed a statistical analysis.

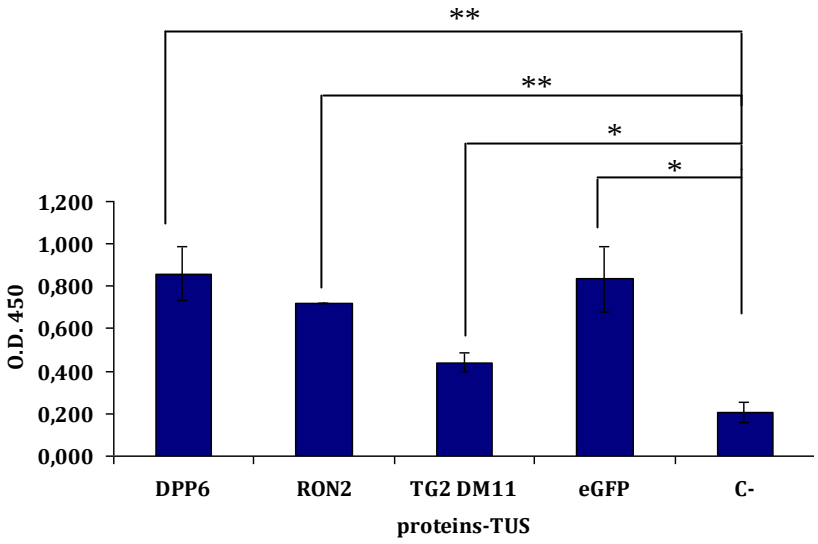


Figure 40. TUS-*Ter* binding test. For all the proteins cloned in pTNT-TUS was performed the ELISA test previously showed. As negative control an IVTT mix without template was used. The results were expressed as the media of three different experiments. A statistical analysis with *t-Student* test was performed (* = p<0,01, ** =p<0,001).

This analysis confirmed that different TUS fusion proteins could efficiently bind to *Ter* sequence. Furthermore, this binding was resulted statistically significant for all the proteins tested. Finally, could be confirmed that our vector allowed the expression of TUS fusion proteins, with IVTT system, that were able to bind specifically the *Ter* sequence.

5.1.3 Protein array on demand

After TUS fusion proteins production and *Ter* binding assessment, the protein array on demand was performed [74]. The plasmid DNA encoding for TUS-fusion proteins with and without multiple *Ter* sequence were printed on FAST slide, heated at 80°C, for DNA entrapment, and dried overnight at room temperature. Three different analysis were performed: I) PicoGreen staining, for immobilized dsDNA detection (Fig. 41A) II) POIs-TUS binding to immobilized pTNT-TUS-*Ter* detection(Fig. 41B); III) POIs-TUS expression from immobilized DNA and their subsequent binding to pTNT-TUS-*Ter* detection (protein array on demand) (Fig. 41C). The POIs-TUS binding to pTNT-TUS-*Ter* was assessed through

POIs-TUS expression from a not immobilized plasmid DNA. The proteins was detected with the anti-SV5 antibody.

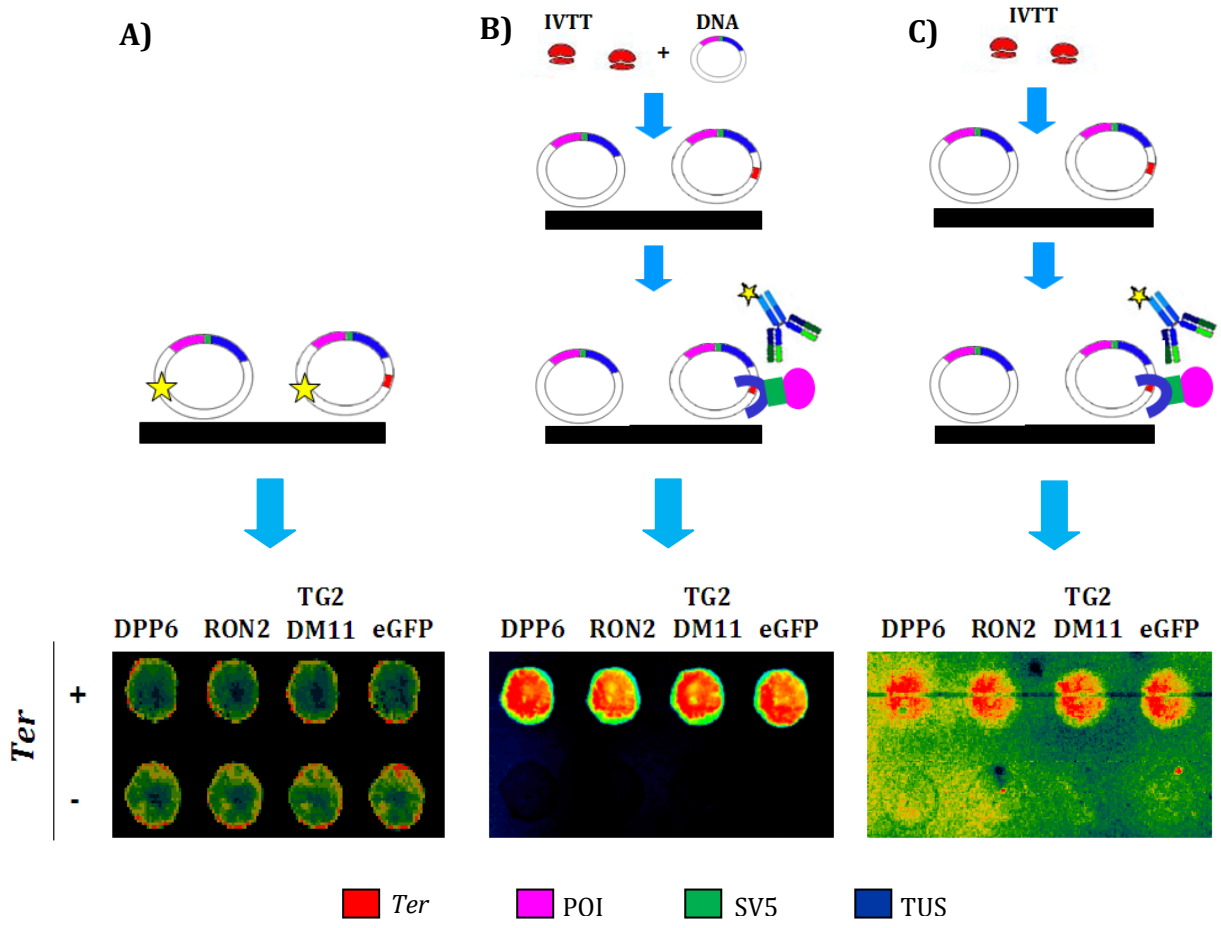


Figure 41. Protein array on demand. pTNT-TUS and pTNT-TUS-*Ter* plasmids, encoding for DPP6-TUS, RON2-TUS, TG2 DM11-TUS and eGFP-TUS, were printed on FAST slide. On the same slide, three experiments were performed: A) PicoGreen staining; B) expression of eGFP-TUS with TnT Quick IVTT from 200ng of pTNT eGFP-TUS in solution; C) expression of all the four tested proteins with TnT Quick IVTT from immobilized DNA (protein array on demand). In the upper part a schematic representation of the experiment is shown. Proteins were detected with the anti-SV5 antibody.

The PicoGreen staining confirmed that dsDNA could be efficiently bound to nitrocellulose slides (Fig. 41A). When the TUS protein was expressed from not immobilized DNA, it could be efficiently captured from the *Ter* sequence (Fig. 41B). Furthermore, where the *Ter* sequence was absent the protein was completely undetected. From the protein array on demand could be observed that all the TUS fusion proteins were produced from immobilized DNA and subsequently bound the DNA carrying the *Ter* sequence (Fig. 41C). Finally, it could be affirmed that the Chatterjee's protein array on demand is functional. Despite that, the signals were very low and sometimes not visible. These results suggested

that the protein array on demand wasn't well established and enough sensitive to be used for our purpose: antibodies validation analysis.

5.3 SET UP OF A NEW PROTEIN ARRAY ON DEMAND PROTOCOL

Since that the published protein array on demand seems not reproducible and enough sensitive, a new protein array on demand protocol, suitable for high-throughput antibody validation, was set up. Our hypothesis was that the reduced reproducibility of protein array on demand was dependent on DNA accessibility for the IVTT system. It was possible that IVTT's enzymes and ribosomes couldn't reach the spotted DNA and perform its transcription and translation. For this reason, we looked for improved conditions for an efficient IVTT from immobilized DNA by evaluating different array slides and DNA templates.

5.3.1 Array slide and DNA template set up

Different slides able to bind covalently the DNA were evaluated. These slides are coated with particular reactive groups, such as: epoxysilane, aldehydesilane, aminosilane and N-hydroxysuccinimide (Fig. 42). All these groups are able to form covalent binding with other chemical groups:

- Epoxysilane binds: Amino, thiol and hydroxyl groups;
- Aldehydesilane binds: amino groups;
- Aminosilane binds: positively charged groups;
- N-hydroxysuccinimide binds: amino groups.

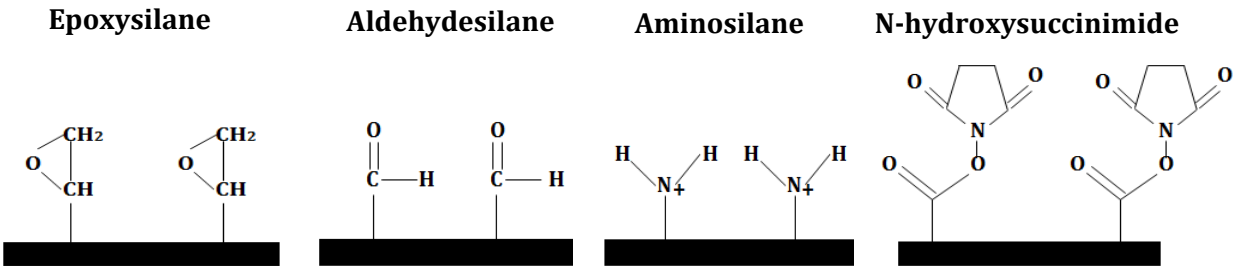


Figure 42. DNA binding slides. Different types of slides able to bind modified or not modified DNA molecules.

Since that epoxysilane and aminosilane slides bound randomly the DNA molecule itself, causing possible modifications, they were excluded. The remaining slides were able to bind the same chemical group (NH_2). As slides covered with N-hydroxysuccinimide were already used in another type of *in situ* protein array [93], it was decided to use these slides for our new protein array on demand. In particular, the Authors had used CodeLink® slides which present the reactive groups included in a 3D gel matrix. To immobilized the DNA on CodeLink® slide an NH_2 group had to be added. Three different DNA printing strategies were carried out to allow DNA immobilization (Fig. 43):

- Amino-modified PCR products;
- Biotinylated PCR products;
- Biotinylated PCR products mix with streptavidin.

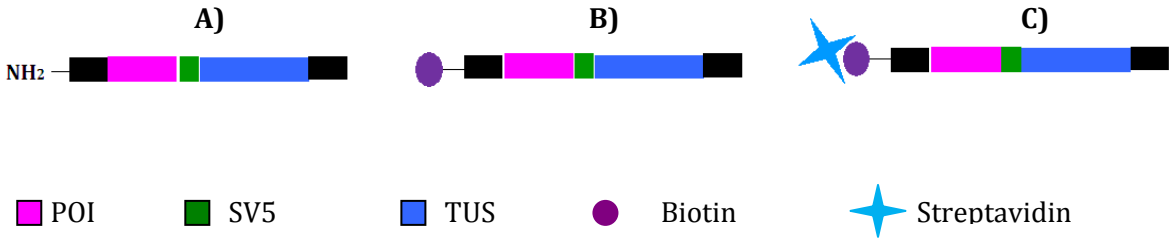


Figure 43. modified DNA molecules spotted on CodeLink® slides. Three different DNA molecules were printed on CodeLink slides: A) Amino-modified PCR DNA; B) Biotinylated PCR DNA; C)Biotinylated PCR DNA mixed with streptavidin. The PCR DNA was compatible for IVTT expression of TUS fusion proteins.

The same PCR product was produced adding at 5'end an NH_2 group or a biotin through amino-modified or biotinylated primers (pTNT seq). The amplified PCR DNA was suitable for IVTT expression of TUS fusion proteins. In case of biotinylated PCRs a preliminary incubation of 10' with 5µg of streptavidin was carried out. PCRs encoding for DPP6, RON2, TG2 DM11 and eGFP fused to TUS, containing or not the *Ter* sequence, were spotted on CodeLink® slide. For all the three printing conditions were tested (Fig. 44): I) dsDNA immobilization with PicoGreen staining; II) the TUS fusion proteins expression and subsequent *Ter* binding.

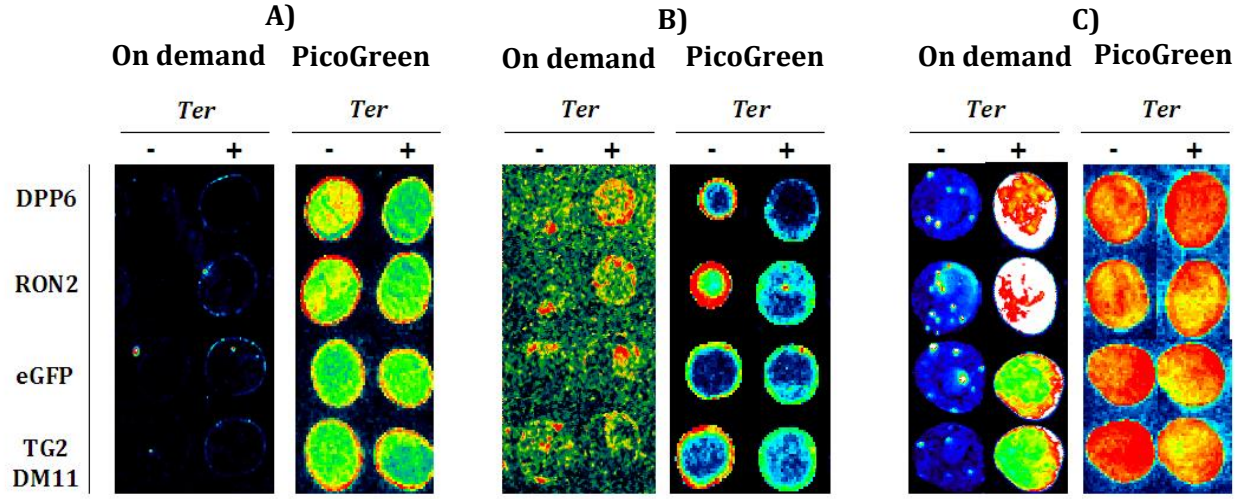


Figure 44. DNA template set up. Three different modified PCR DNA, encoding for DPP6, RON2, TG2 DM11 and eGFP fused to TUS, were printed on CodeLink® slide: A) Amino-modified PCR DNA; B) Biotinylated PCR DNA; C) Biotinylated PCR DNA preincubated with streptavidin. For every type of template a PicoGreen staining and a protein array on demand with TnT Quick for PCR DNA were performed. The proteins were detected with the anti-SV5 antibody.

From this experiment it could be observed that all the three types of PCRs DNA were efficiently immobilized on CodeLink® slides. In particular, the biotinylated PCR DNA previously incubated with streptavidin showed the better spots shape and DNA homogeneous distribution than the other conditions (Fig. 44C). Observing the three protein arrays on demand (Fig. 44), the biotinylated PCR DNA mixed with streptavidin was the only condition that allowed the transcription and translation of proteins from immobilized DNA. Furthermore, the expressed proteins were efficiently captured from the DNA carrying the *Ter* sequence. To confirm its reproducibility the experiment was repeated for three times. In conclusion, the new protein array on demand, in which

biotinylated PCRs were incubated with streptavidin and printed on CodeLink® slide, was set up.

5.3.2 Protein array on demand improvements

After protein microarray on demand set up some improvements were performed, in order to reach the best conditions. In particular, the printing conditions and the protein production with IVTT were analyzed. To increase the amount of immobilized DNA, different printing conditions were tested. The same 5'-biotinylated PCR DNA, encoded for eGFP-TUS, was mixed with:

1) 3µg streptavidin;
2) 3µg streptavidin, 0,005% triton-X;
3) 1µg streptavidin, 0,005% triton-X;
4) 1µg streptavidin, 10µg BSA;
5) 1µg streptavidin, 10µg BSA, 0,005% triton-X.

All these five printing conditions were spotted on CodeLink® slide together with the CodeLink printing buffer at 1x or 2x concentration. After DNA spotting the slide was stained with PicoGreen to assess spots shape and amount of printed DNA (Fig. 45). A protein array on demand was carried out to compare the amount of protein produced and bound for all the printing conditions (Fig. 45).

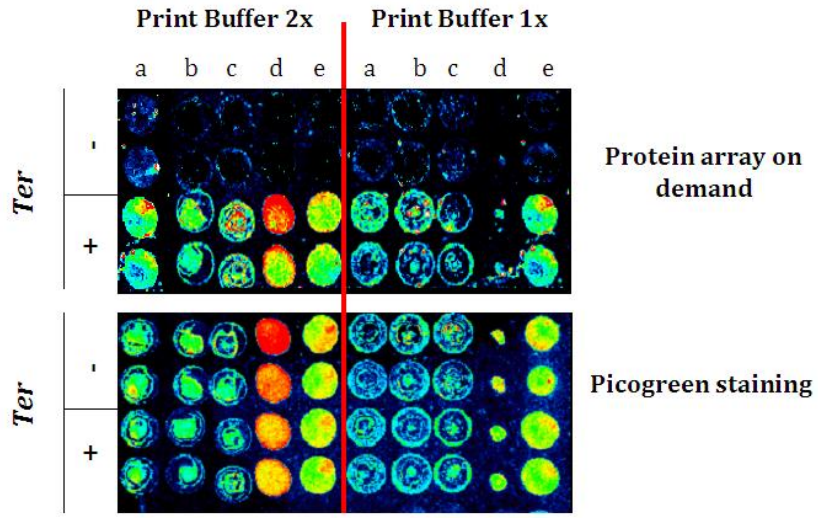


Figure 45. DNA printing improvement. The same biotinylated PCR DNA with or without *Ter* sequence, encoding for eGFP-TUS protein was printed on CodeLink® slide with different printing strategies: a) 1µg streptavidin, 10µg BSA, 0,005% triton-X; b) 3µg streptavidin, 0,005% triton-X; c) 1µg streptavidin, 0,005% triton-X; d) 3µg streptavidin; e) 1µg streptavidin, 10µg BSA. All these conditions were diluted in CodeLink printing buffer 1x or 2x. PicoGreen staining and protein array on demand, with anti-SV5 staining, were carried out.

The PicoGreen staining revealed that the use of triton-X didn't improve the DNA immobilization on this kind of slides. Indeed, the DNA resulted less concentrated and non-homogeneously distributed where 0,005% of triton-X was used. On the other hand, the mix of streptavidin and BSA showed a good DNA immobilization with an homogeneous distribution as in presence of 1x or 2x printing buffer. Finally, the streptavidin alone resulted less efficient in DNA immobilization when the printing buffer 1x was used, but was efficient when used with 2x buffer. Furthermore, observing the correspondent protein array on demand the condition that allowed the major amount of protein production and binding was the one with biotinylated PCRs mixed with streptavidin and printing buffer 2x. In conclusion, the 5'-biotinylated PCRs mixed with 3µg of streptavidin and diluted in CodeLink print buffer 2x was the best achieved condition for DNA printing and protein array on demand development. Another evaluated aspect was the IVTT protein expression. Till now all the protein arrays on demand were performed using an IVTT system based on rabbit reticulocytes lysate, but other eukaryotic IVTT systems, based on other cell type, are today available, such as wheat germ and HeLa cell lysate. To understand the capability of these IVTT systems an analysis were done. First, it was necessary to construct a DNA template suitable for HeLa cells IVTT system. For this type of IVTT the T7 prometer and the Kozak sequence are insufficient for protein expression. The insertion of an

encephalomyocarditis virus (ECMV) internal ribosome entry site (IRES), downstream of the T7 promoter, was necessary. This IRES sequence was cloned into the pTNT vector. The correspondent plasmid was called pIRES-TUS. To compare the efficiency of the different IVTT systems a Western blot analysis was done. The same protein, eGFP-TUS, was expressed with the three different IVTT systems, in 10µl of reaction, following the kit protocols. The reactions were denaturated, loaded on polyacrylamide gel and analyzed with Western blot (Fig. 46A). The eGFP-TUS protein was revealed with the anti-SV5 and the resulting bands were quantified with the Imagej software (Fig. 46B).

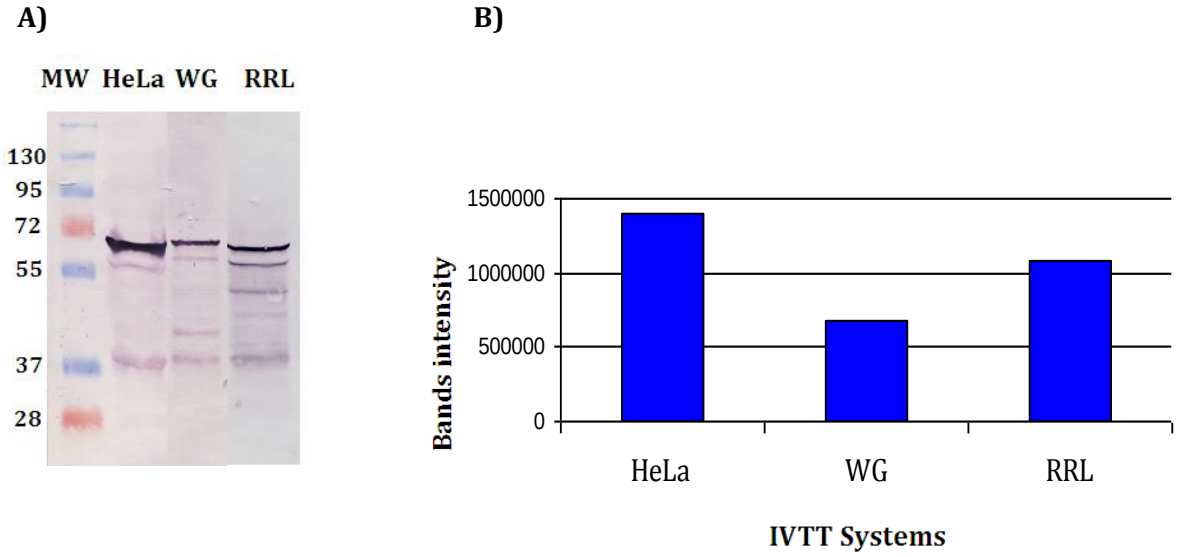


Figure 46. Protein production improvement. The eGFP-TUS protein was produced with IVTT base on: HeLa cell, wheat germ (WG) and rabbit reticulocytes (RRL) lysates. The protein produced in 10µl of reaction was denaturated and loaded on polyacrylamide gel. A) Western blot analysis performed with anti-SV5 staining and anti-mouse IgG-AP. B) Western blot bands quantification with Imagej software.

Looking at the Western blot analysis seems that the HeLa cell lysate gave the major amount of proteins and also the less amount of protein degradation, compared to the other two systems. To verify the capability of HeLa cell lysate in our protein array on demand PCRs, encoding for DPP6, RON2, TG2 DM11 and eGFP fused to TUS, containing or not the *Ter* sequence, were spotted at the previously described conditions. Two different protein arrays on demand were performed, one with T7 TnT Quick for PCR DNA and the other with 1-Step Human coupled *in vitro* expression kit (Fig. 47).

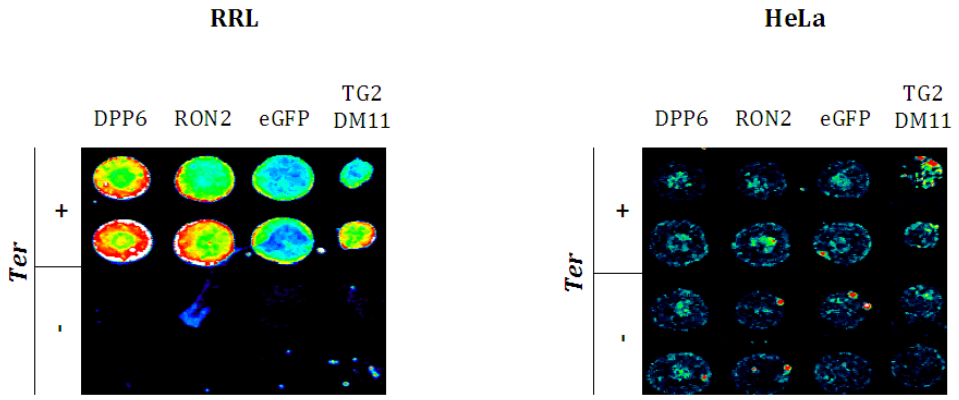


Figure 47. RRL/HeLa protein array on demand. Biotinylated PCRs with or without IRES sequence were spotted at 400ng/µl with streptavidin and printing buffer 2x. The PCRs encoded for DPP6, RON2, eGFP, TG2 DM11 fused to TUS and contained or not the *Ter* sequence. The DNA containing the IRES sequence was treated with HeLa cell lysate IVTT, whereas the ones containing only the Kozak sequence was treated with rabbit reticulocytes lysate IVTT. After *in situ* protein production an anti-SV5 staining was performed. The arrays were scanned with ScanArray Gx.

The protein arrays showed that the HeLa cell lysate didn't allow the expression of the eGFP-TUS protein from the immobilized DNA, impeding the protein array on demand production (Fig. 47). On the other hand, the rabbit reticulocytes lysate confirmed its capability to perform an *in situ* protein production from spotted DNA. In conclusion, the final condition to produce a protein array on demand are:

- Template: 400ng/µl of biotinylated PCRs DNA mixed with 3µg of streptavidin in CodeLink buffer 2x;
- Support: CodeLink® slide;
- IVTT system: rabbit reticulocytes lysate.

5.2.4 Protein array on demand

Since that the protein array on demand was set up an antibody validation test was done. The PCRs DNA encoding for DPP6, RON2, TG2 DM11 and eGFP fused to TUS, containing or not the *Ter* sequence, were spotted at the optimized conditions. Two protein arrays on demand were performed (Fig. 48):

1. with anti-SV5 staining, where all the proteins had to be recognized;

2. with anti-GFP staining, where only the eGFP-TUS had to be detected.

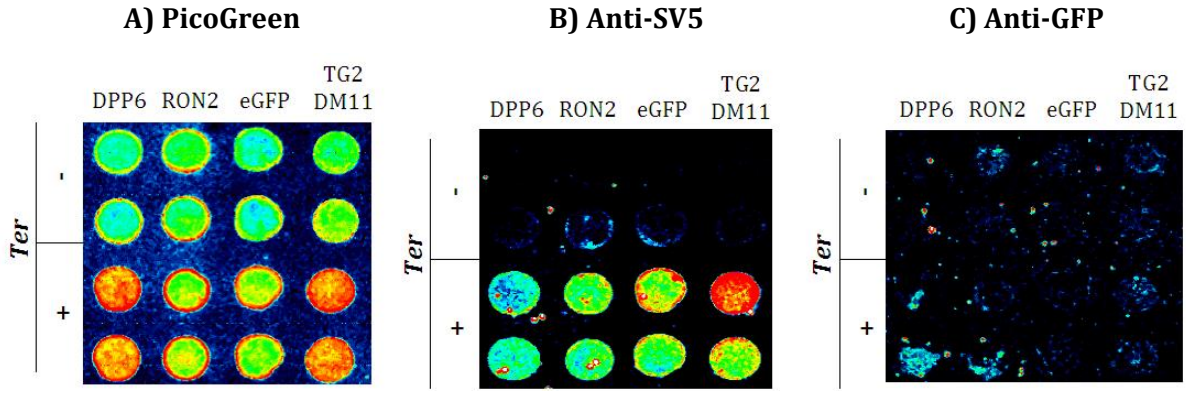


Figure 48. Protein array on demand with anti-GFP. PCRs DNA, with or without *Ter* sequence, encoding for DPP6, RON2, TG2 DM1 and eGFP fused to TUS were spotted on CodeLink slide. A) PicoGreen staining; B) Protein array on demand with anti-SV5 staining; C) Protein array on demand with anti-GFP staining.

From these arrays could be observed that all the four proteins were produced and bound specifically to the DNA containing the *Ter* sequence. On the other hand, when the anti-GFP was used there wasn't seen any signal in correspondence of the eGFP-TUS protein. To investigate if the problem was related to the anti-GFP binding to eGFP-TUS protein, an IVTT mix plus 200ng of eGFP-TUS PCR DNA in solution was used on the same array slide to allow a higher amount of eGFP-TUS protein production and binding to *Ter* (Fig. 49).

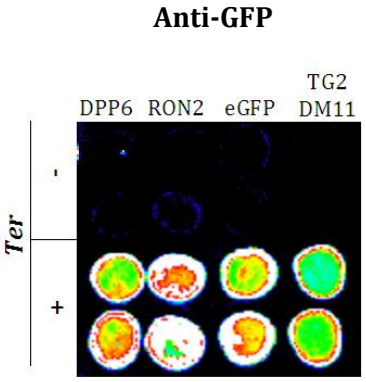
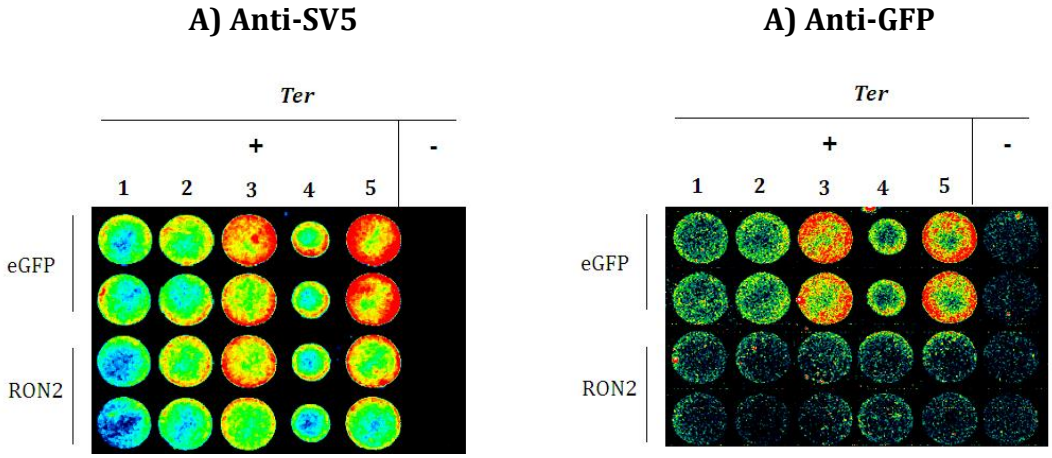


Figure 49. eGFP-TUS protein array on demand with DNA in solution. The previously tested slide was treated with IVTT mix containing 200ng of PCR DNA, encoding for eGFP-TUS protein, in solution. After protein expression and binding the anti-GFP antibody was used.

As expected, the eGFP-TUS could be specifically recognized from the correspondent antibody. The second aspect that was studied the steric hindrance. Indeed, it was possible that the protein, once produced and captured on the slide, remained enclosed through the PCR DNA molecules, leading to an unavailability of the protein to the correspondent antibody. To increase the space between the immobilized DNA molecules different concentrations of BSA were added to the PCRs. Indeed BSA could compete with streptavidin for the N-hydroxysuccinimide binding. This competition could lead to a DNA molecules spreading. The PCR DNA encoding for eGFP-TUS protein was spotted on CodeLink slide with different concentration of streptavidin and BSA (Fig. 47).



47. Protein array on demand with BSA addiction. The PCRs DNA encoding for eGFP-TUS or RON2-TUS proteins were printed on CodeLink slide in different printing conditions: 1. 3µg streptavidin, 25µg BSA; 2. 3µg streptavidin, 20µg BSA;3. 2µg streptavidin, 10µg BSA;4. 2µg streptavidin, 20µg BSA; 5. 1µg streptavidin, 15µg BSA. Two protein arrays on demand were performed and depicted with: A) anit-SV5 staining; B)anti-GFP staining. As negative control a PCR without *Ter* was used.

The anti-SV5 stained protein array on demand revealed that both eGFP-TUS and RON2-TUS were produced. Furthermore, using the anti-GFP staining, only the spots correspondent to eGFP-TUS protein were positives. In particular, two printing conditions (3 and 5) seem better that the others tested. Obviously, it could be also seen that the signal due to the anti-GFP antibody was lower than the anti-SV5 ones. This aspect suggested that this wasn't the best condition to perform a protein array on demand, but showed that the printing conditions were responsible for the achievement of a good protein array on demand.

6. DISCUSSION

Antibodies are crucial molecules in scientific research in particular for proteomic studies. To perform a good study a very specific and sensitive antibody is needed. These antibodies features depend on a good antibody validation. For this reason our purpose is to validate different polyclonal antibodies, in form of scFv-Fc, with high-throughput methods.

6.1 POLYCLONAL ANTIBODIES VALIDATION

In collaboration with Bradbury's group a list of 78 proteins was filled out, based on proteins solubility and structure. The first seven proteins of the list (Tab. 2) were used for polyclonals selection with two rounds of phage display and two rounds of yeast display. The selected polyclonal antibodies were cloned in a yeast vector and produced as scFv-Fc rabbit. Three different analysis were performed for antibodies validation: Western blot, ELISA and protein microarray. From the Western blot analysis five polyclonal supernatants resulted able to recognize specifically their target protein without reacting to the control ones (Fig. 24). Only two antibodies (anti-CTBP2A, anti-MAPK8B) have shown no binding activity to their antigens. This result was explained by an antibodies specificity directed only to conformational epitopes. To confirm this hypothesis an ELISA validation was performed (Fig. 25). In the ELISA, all antibodies were tested simultaneously with all the seven antigens to validate their specificity on an higher number of proteins. The values reported in the graph demonstrated that all the polyclonals were specific to their own antigen and didn't present cross-reactivity to the other proteins tested (Fig. 25). As expected, the two polyclonals not functional in Western blot, resulted positive in the ELISA, confirming their ability to recognize the target only in the conformational form. Furthermore, from this validation it could be observed that the polyclonal antibodies presented different sensitivity to the target. Evaluating the antibodies concentration in the supernatants it was observed that the antibodies sensitivity was proportional to the supernatants concentration (Fig. 26). Finally, the ELISA validation has revealed that all the tested polyclonals were positive even if presented different sensitivity. As previously described, the Western blot and the ELISA analysis are not sufficient for a good antibody validation. For this reason, the protein microarray technology was used as major sensitive and largest samples screened toll (Fig. 29-30). The protein array showed that six polyclonals supernatants were able to recognize specifically their own antigen, even if with

different sensitivity. Indeed, anti-CDK2A, anti-MAPK2A5 and anti-USP11 polyclonals gave high signals, instead anti-PLAA, anti-SF3A1 and anti-MAPK8B showed weak signals. This could be related to a lower antibodies concentration, indeed, two of them presented a lower amount of antibodies in ELISA compared to the others. To validate in a correct way the polyclonal antibodies they were purified and tested with a second protein array analysis (Fig. 31). This new protein array showed that the five polyclonals anti-CDK2A, anti-MAPK2A5, anti-USP11, anti-PLAA and anti-SF3A presented high specificity and sensitivity to the target proteins. Furthemore, none of them showed cross-reactivity to other proteins. Unfortunately, the anti-MAPK8B and anti-CTBP2A were not efficient even in the purified form. The anti-CTBP2A recognized its target but also the MAPK2A5; the anti-MAPK8B didn't react to its antigen. The problem of these polyclonal antibodies could be related to the selection method or to a not efficient antibody production. To study this aspect new select and production will be performed for this two antibodies. The new polycloanls will be tested on protein array. In conclusion, five polyclonal antibodies produced with the phage/yeast display combination method resulted specific and sensitive in Western blot, ELISA and protein microarray. Only two antibodies, anti-CTBP2A and anti-MAPK8B, seem unfunctional in Western blot and protein array, but recognized the target in the ELISA validation. To confirm the potential of this new antibody selection method, a yeast-purified polyclonal antibody was compared to the commercial ones. The yeast-purified anti-CDK2A was tested in ELISA and protein array, together with its correspondent commercial antibody. The ELISA showed that the yeast-purified and the commercial polyclonals present comparable specificity and sensitivity to CDK2A protein (Fig. 32). Looking at the protein array analysis it was observed that the yeast-purified antibody was better than the commercial ones (Fig. 33). Indeed, the commercial anti-CDK2A bound other three unspecific proteins: MAP Kinase 13A (MAPK13A), MAP 2 Kinase 6A (MAP2K6A) and Check point Kinase 2a (CHEK2A). Studying the sequences of these proteins arose that they don't show an alignment, but watching at the proteins patters with Expasy prosite tool they contain all the following pattern: G-[ERS]-G-[AT]-x-G-x(0,1)-V-x(4)-[DENT]-[KR]-x-[STV]-[CGP]-x(3)-[AI]-[ILM]-[EK]-x-[AIL]. This pattern can allow the formation of a conformational epitope that can be possibly recognized by the commercial anti-CDK2A antibody. On the other hand, the yeast-purified antibody didn't show cross-reaction to these three proteins or to the others, maintaining an high sensitivity to the target protein. This final analysis confirmed that the phage/yeast display combination is a very specific and successful method for high specific polyclonal antibodies selection.

Furthermore, protein array technology confirming its efficiency as antibodies validation tool. Obviously, other technologies as immunohystochemistry and immunofluorescence are needed to confirm the real specificity of these antibodies.

6.2 PROTEIN ARRAY ON DEMAND

Protein array technology is very important and useful tool for many different applications, such as antibodies validation. Despite that, protein array has also some drawbacks: 1. proteins production and purification processes have to be performed for all the printed proteins; 2. proteins can go on degradation after their immobilization on slides. To overcome these problems a new type of protein microarray has been set up, called *in situ* protein array [75]. This new technology consists of (Fig. 17): 1. printing of the proteins cDNA on the array slide; 2. transcription and translation of the printed DNA through IVTT systems; 3. capturing of the nascent proteins; 4. revealing of the proteins. In the last decade, many types of *in situ* protein arrays, differing each other for DNA template and capture agent used, have been developed [72; 73; 74; 75; 76; 77; 100]. The *in situ* protein array of our interest is called "protein array on demand" [74]. This array exploits the role of an *E. coli* protein called TUS to bind specifically a dsDNA sequence called *Ter* [82]. The peculiarity of this array is the presence of the template and the capture agent in the same DNA molecule (Fig. 35). Our aim was to use this protein array for high-throughput antibodies validation. The pTNT™ commercial vector, optimized for protein expression with rabbit reticulocytes lysate IVTT systems, was modified to insert (Fig. 36): I) TUS protein cDNA without methionin; II) two enzyme restriction sites (BssHII-NheI), for protein of interest insertion; III) SV5 and His_6 tags, for protein detection. After plasmid modification, a set of four proteins, available in our laboratory, were inserted in this new plasmid called pTNT-TUS (Tab. 3). The expression of these TUS fusion proteins, with rabbit reticulocytes lysate IVTT, were tested by Western blot (Fig. 38). This analysis confirmed that the pTNT-TUS plasmid was suitable for TUS fusion proteins expression through rabbit reticulocytes lysate IVTT system. Furthermore, a low proteins degradation were observed. After proteins expression assessment the ability of that proteins to bind the *Ter* sequence was investigated. An ELISA test for TUS-*Ter* binding evaluation was set up using eGFP-TUS as control protein (Fig. 39). Subsequently, the ELISA was performed for all the four TUS fusion proteins (Fig. 40). The results showed that all the TUS fusion proteins could bound specifically to the *Ter* sequence. After the protein production and binding evaluation a protein array on demand was performed. The pTNT-TUS and pTNT-TUS-*Ter* plasmid were

spotted on FAST slide. Three different analysis were performed (Fig. 41): 1. PicoGreen staining, for dsDNA printing evaluation (Fig. 41A); 2. protein expression with T7 TnT Quick plus eGFP-TUS plasmid in solution, for TUS-*Ter* binding evaluation on array slide (Fig. 41B); 3. protein production with T7 TnT Quick using the immobilized DNA as template, for protein array on demand assessment (Fig. 41C). All the four proteins were detected with anti-SV5 antibody. The PicoGreen staining showed that dsDNA could be efficiently immobilized on nitrocellulose. The figure 41B confirmed that TUS fusion proteins could be efficiently captured by DNA carrying *Ter* sequence. Finally, the protein array on demand showed the expression of TUS fusion proteins from immobilized DNA and their cupture by *Ter* sequence. Despite that, the proteins signals were weak and sometimes not detectable. These results showed that the protein array on demand wasn't enough sensitive for our purpose. For this reason a new protein array on demand protocol was set up.

6.3 NEW PROTEIN ARRAY ON DEMAND SET UP

To perform a new protein array on demand was necessary set up the array slide and the DNA template. Between the slides able to immobilized DNA the one cover with N-hydroxysuccinimide group were used. In particular, CodeLink® slides were used for the following experiments. Three different templates were tested on this slides (Fig. 43): 1. Amino-modified PCR DNA; 2. Biotinylated PCR DNA; 3. Biotinylated PCR DNA mixed with streptavidin. For these three conditions a PicoGreen staining and a protein array on demand were performed (Fig. 44). This experimet showed that the biotinylated PCR DNA preincubated with streptavidin was the only condition that allowed a protein array on demand production. Further improvements were performed: DNA printing and IVTT protein production. From the printing optimization arose that the best DNA immobilization and protein array on demand were reached in presence of biotinylated PCR DNA premixed with 3µg of streptavidin (Fig. 45). In the protein expression optimization the best IVTT system for our protein array on demand was the one based on rabbit reticulocytes lysate (Fig. 47). Once optimized the protein array on demand, its antibody validation ability was tested. An anti-GFP monoclonal antibody were used to reveal the presence of the eGFP-TUS protein on the array (Fig. 48). The protein array on demand was developed with anti-SV5 and anti-GFP. Surprisingly, only the anti-SV5 staining was positive. Our idea regarding this problem was related to steric hindrance. Indeed, it was possible that the protein was unavailable to its antibody because the PCR DNA molecules surround it. Based on this idea, the BSA was added to the printing solution to increase the space between the PCR DNA

molecules (Fig. 49). Different concentrations of streptavidin and BSA were tested with the same PCR DNA template. Protein arrays on demand with anti-SV5 and anti-GFP staining were performed. The arrays showed that the addition of BSA allowed a binding between eGFP-TUS and its antibody. Even if this signal was lower than the anti-SV5 one, it represented an improvement in the protein array on demand. In conclusion, it could be asserted that a new protein array on demand has been set up and it is functional and reproducible.

7. **REFERENCES**

[1] S. Kipriyanov, and F. Le Gall, Generation and production of engineered antibodies. Molecular biotechnology 26 (2004) 39-60.

[2] O. Brekke, and I. Sandlie, Therapeutic antibodies for human diseases at the dawn of the twenty-first century. Nature reviews. Drug discovery 2 (2003) 52-62.

[3] A. Frenzel, M. Hust, and T. Schirrmann, Expression of recombinant antibodies. Frontiers in immunology 4 217.

[4] W. Jiang, Y. Mao, R. Huang, C. Duan, Y. Xi, K. Yang, and R.-P. Huang, Protein expression profiling by antibody array analysis with use of dried blood spot samples on filter paper. Journal of immunological methods.

[5] A. Gehring, C. Barnett, T. Chu, C. DebRoy, D. D'Souza, S. Eaker, P. Fratamico, B. Gillespie, N. Hegde, K. Jones, J. Lin, S. Oliver, G. Paoli, A. Perera, and J. Uknalis, A high-throughput antibody-based microarray typing platform. Sensors (Basel, Switzerland) 13 5737-5748.

[6] A.-C. Andersson, S. StrÃ¶mberg, H. BÃ¤ckvall, C. Kampf, M. Uhlen, K. Wester, and F. PontÃ©n, Analysis of protein expression in cell microarrays: a tool for antibody-based proteomics. The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society 54 (2006) 1413-1423.

[7] H. Tekin, and M. Gijs, Ultrasensitive protein detection: a case for microfluidic magnetic bead-based assays. Lab on a chip 13 4711-4739.

[8] D. Shahbazi-Gahrouei, and M. Abdolahi, Detection of MUC1-expressing ovarian cancer by C595 monoclonal antibody-conjugated SPIONs using MR imaging. TheScientificWorldJournal 2013 609151.

[9] M. Firer, and G. Gellerman, Targeted drug delivery for cancer therapy: the other side of antibodies. Journal of hematology & oncology 5 70.

[10] D. Ag, R. Bongartz, L. Dogan, M. Seleci, J.-G. Walter, D. Demirkol, F. Stahl, S. Ozcelik, S. Timur, and T. Scheper, Biofunctional quantum dots as fluorescence probe for cell-specific targeting. Colloids and surfaces. B, Biointerfaces 114 96-103.

[11] S.-I. Fukada, Y. Ma, T. Ohtani, Y. Watanabe, S. Murakami, and M. Yamaguchi, Isolation, characterization, and molecular regulation of muscle stem cells. Frontiers in physiology 4 317.

[12] Y. Ahmad, and A.I. Lamond, A perspective on proteomics in cell biology. Trends Cell Biol 24 257-264.

[13] R. Hafko, S. Villapol, R. Nostramo, A. Symes, E. Sabban, T. Inagami, and J. Saavedra, Commercially available angiotensin II Atâ,, receptor antibodies are nonspecific. PLoS One 8.

[14] C. Saper, An open letter to our readers on the use of antibodies. The Journal of comparative neurology 493 (2005) 477-478.

[15] J. Bordeaux, A. Welsh, S. Agarwal, E. Killiam, M. Baquero, J. Hanna, V. Anagnostou, and D. Rimm, Antibody validation. BioTechniques 48 197-209.

[16] M. Uhlen, Mapping the human proteome using antibodies. Molecular & cellular proteomics : MCP 6 (2007) 1455-1456.

[17] M. UhlÃ©n, and S. Hober, Generation and validation of affinity reagents on a proteome-wide level. Journal of molecular recognition : JMR 22 (2009) 57-64.

[18] G. KÃ¶hler, and C. Milstein, Continuous cultures of fused cells secreting antibody of predefined specificity. 1975. Journal of immunology (Baltimore, Md. : 1950) 174 (2005) 2453-2455.
[19] H. Cooper, and Y. Paterson, Production of polyclonal antisera. Current protocols in neuroscience / editorial board, Jacqueline N. Crawley ... [et al.] Chapter 5 (2009).
[20] H. Zhou, Y.L. Zhang, G. Lu, H. Ji, and C.P. Rodi, Recombinant antibody libraries and selection technologies. N Biotechnol 28 448-52.
[21] A. Frenzel, J. Kugler, S. Wilke, T. Schirrmann, and M. Hust, Construction of human antibody gene libraries and selection of antibodies by phage display. Methods Mol Biol 1060 215-43.
[22] M. Silacci, S. Brack, G. Schirru, J. Marlind, A. Ettorre, A. Merlo, F. Viti, and D. Neri, Design, construction, and characterization of a large synthetic human antibody phage display library. Proteomics 5 (2005) 2340-50.
[23] A. Bradbury, S. Sidhu, S. DÃ¼bel, and J. McCafferty, Beyond natural antibodies: the power of in vitro display technologies. Nature biotechnology 29 245-254.
[24] G. Winter, A. Griffiths, R. Hawkins, and H. Hoogenboom, Making antibodies by phage display technology. Annual review of immunology 12 (1994) 433-455.
[25] X. Wang, and E. Shusta, The use of scFv-displaying yeast in mammalian cell surface selections. Journal of immunological methods 304 (2005) 30-42.
[26] M. He, and M. Taussig, Ribosome display of antibodies: expression, specificity and recovery in a eukaryotic system. Journal of immunological methods 297 (2005) 73-82.
[27] G.P. Smith, Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 228 (1985) 1315-7.
[28] J. Pande, M.M. Szewczyk, and A.K. Grover, Phage display: concept, innovations, applications and future. Biotechnol Adv 28 849-58.
[29] J. Kehoe, and B. Kay, Filamentous phage display in the new millennium. Chemical reviews 105 (2005) 4056-4072.
[30] H.M. Azzazy, and W.E. Highsmith, Jr., Phage display technology: clinical applications and recent innovations. Clin Biochem 35 (2002) 425-45.
[31] P. Pansri, N. Jaruseranee, K. Rangnoi, P. Kristensen, and M. Yamabhai, A compact phage display human scFv library for selection of antibodies to a wide variety of antigens. BMC Biotechnol 9 (2009) 6.
[32] N. Clementi, N. Mancini, L. Solforosi, M. Castelli, M. Clementi, and R. Burioni, Phage Display-based Strategies for Cloning and Optimization of Monoclonal Antibodies Directed against Human Pathogens. Int J Mol Sci 13 8273-92.
[33] T. Kretzschmar, and T. von Ruden, Antibody discovery: phage display. Curr Opin Biotechnol 13 (2002) 598-602.
[34] A. Kondo, and M. Ueda, Yeast cell-surface display--applications of molecular display. Applied microbiology and biotechnology 64 (2004) 28-40.
[35] K. Kuroda, and M. Ueda, Cell surface engineering of yeast for applications in white biotechnology. Biotechnol Lett 33 1-9.
[36] M. Ueda, and A. Tanaka, Genetic immobilization of proteins on the yeast cell surface. Biotechnol Adv 18 (2000) 121-40.
[37] M.J. Feldhaus, R.W. Siegel, L.K. Opresko, J.R. Coleman, J.M. Feldhaus, Y.A. Yeung, J.R. Cochran, P. Heinzelman, D. Colby, J. Swers, C. Graff, H.S. Wiley, and K.D. Wittrup, Flow-cytometric isolation of human antibodies from a nonimmune Saccharomyces cerevisiae surface display library. Nat Biotechnol 21 (2003) 163-70.
[38] E.T. Boder, and K.D. Wittrup, Yeast surface display for screening combinatorial polypeptide libraries. Nat Biotechnol 15 (1997) 553-7.

[39] G. Chao, W. Lau, B. Hackel, S. Sazinsky, S. Lippow, and K. Wittrup, Isolating and engineering human antibodies using yeast surface display. Nature protocols 1 (2006) 755-768.
[40] A. Koide, and S. Koide, Affinity maturation of single-domain antibodies by yeast surface display. Methods Mol Biol 911 431-43.
[41] S.A. Gai, and K.D. Wittrup, Yeast surface display for protein engineering and characterization. Curr Opin Struct Biol 17 (2007) 467-73.
[42] C.A. Patel, J. Wang, X. Wang, F. Dong, P. Zhong, P.P. Luo, and K.C. Wang, Parallel selection of antibody libraries on phage and yeast surfaces via a cross-species display. Protein Eng Des Sel 24 711-9.
[43] F. Ferrara, L. Naranjo, S. Kumar, T. Gaiotto, H. Mukundan, B. Swanson, and A. Bradbury, Using phage and yeast display to select hundreds of monoclonal antibodies: application to antigen 85, a tuberculosis biomarker. PLoS One 7.
[44] O. Stoevesandt, and M. Taussig, Affinity reagent resources for human proteome detection: initiatives and perspectives. Proteomics 7 (2007) 2738-2750.
[45] F. PontÃ©n, K. JirstrÃ¶m, and M. Uhlen, The Human Protein Atlas--a tool for pathology. The Journal of pathology 216 (2008) 387-393.
[46] R. SjÃ¶berg, M.r. Sundberg, A. Gundberg, A. Sivertsson, J. Schwenk, M. UhlÃ©n, and P. Nilsson, Validation of affinity reagents using antigen microarrays. New biotechnology 29 555-563.
[47] M.F. Templin, D. Stoll, M. Schrenk, P.C. Traub, C.F. Vohringer, and T.O. Joos, Protein microarray technology. Trends Biotechnol 20 (2002) 160-6.
[48] C. Woodard, C. Goodwin, J. Wan, S. Xia, R. Newman, J. Hu, J. Zhang, S. Hayward, J. Qian, J. Laterra, and H. Zhu, Profiling the dynamics of a human phosphorylome reveals new components in HGF/c-Met signaling. PLoS One 8.
[49] H. Zhu, M. Bilgin, R. Bangham, D. Hall, A. Casamayor, P. Bertone, N. Lan, R. Jansen, S. Bidlingmaier, T. Houfek, T. Mitchell, P. Miller, R.A. Dean, M. Gerstein, and M. Snyder, Global analysis of protein activities using proteome chips. Science 293 (2001) 2101-5.
[50] C.-J. Hu, G. Song, W. Huang, G.-Z. Liu, C.-W. Deng, H.-P. Zeng, L. Wang, F.-C. Zhang, X. Zhang, J. Jeong, S. Blackshaw, L.-Z. Jiang, H. Zhu, L. Wu, and Y.-Z. Li, Identification of new autoantigens for primary biliary cirrhosis using human proteome microarrays. Molecular & cellular proteomics : MCP 11 669-680.
[51] R. Ummanni, H. Mannsperger, J. Sonntag, M. Oswald, A. Sharma, R. KÃ¶nig, and U. Korf, Evaluation of reverse phase protein array (RPPA)-based pathway-activation profiling in 84 non-small cell lung cancer (NSCLC) cell lines as platform for cancer proteomics and biomarker discovery. Biochimica et biophysica acta.
[52] I. Uzoma, and H. Zhu, Interactome mapping: using protein microarray technology to reconstruct diverse protein networks. Genomics, proteomics & bioinformatics 11 18-28.
[53] A. Lueking, J. Beator, E. Patz, S. MÃ¼llner, G. Mehes, and P. Amersdorfer, Determination and validation of off-target activities of anti-CD44 variant 6 antibodies using protein biochips and tissue microarrays. BioTechniques 45 (2008).
[54] J. Andreadis, and L. Chrisey, Use of immobilized PCR primers to generate covalently immobilized DNAs for in vitro transcription/translation reactions. Nucleic acids research 28 (2000).
[55] M. He, and M.J. Taussig, Single step generation of protein arrays from DNA by cell-free expression and in situ immobilisation (PISA method). Nucleic Acids Res 29 (2001) E73-3.
[56] C.P. Chou, Engineering cell physiology to enhance recombinant protein production in Escherichia coli. Appl Microbiol Biotechnol 76 (2007) 521-32.

[57] I. Hunt, From gene to protein: a review of new and enabling technologies for multi-parallel protein expression. Protein Expr Purif 40 (2005) 1-22.
[58] D.L. Hacker, M. De Jesus, and F.M. Wurm, 25 years of recombinant proteins from reactor-grown cells - where do we go from here? Biotechnol Adv 27 (2009) 1023-7.
[59] M.W. Nirenberg, and J.H. Matthaei, The dependence of cell-free protein synthesis in E. coli upon naturally occurring or synthetic polyribonucleotides. Proc Natl Acad Sci U S A 47 (1961) 1588-602.
[60] J.K. DeVries, and G. Zubay, DNA-directed peptide synthesis. II. The synthesis of the alpha-fragment of the enzyme beta-galactosidase. Proc Natl Acad Sci U S A 57 (1967) 1010-2.
[61] A. Chandrasekaran, R. Bharadwaj, J. Park, R. Sapra, P. Adams, and A. Singh, A microscale platform for integrated cell-free expression and activity screening of cellulases. Journal of proteome research 9 5677-5683.
[62] A. Jackson, J. Boutell, N. Cooley, and M. He, Cell-free protein synthesis for proteomics. Briefings in functional genomics & proteomics 2 (2004) 308-319.
[63] C.J. Murray, and R. Baliga, Cell-free translation of peptides and proteins: from high throughput screening to clinical production. Curr Opin Chem Biol 17 420-6.
[64] J. Swartz, Developing cell-free biology for industrial applications. Journal of industrial microbiology & biotechnology 33 (2006) 476-485.
[65] H. Mingyue, H. Yongzhi, L. Qin, and W. Mingrong, From DNA to protein: No living cells required. Process Biochemistry 46.
[66] O. Kovtun, S. Mureev, W. Johnston, and K. Alexandrov, Towards the construction of expressed proteomes using a Leishmania tarentolae based cell-free expression system. PLoS One 5.
[67] S. Mikami, T. Kobayashi, and H. Imataka, Cell-free protein synthesis systems with extracts from cultured human cells. Methods in molecular biology (Clifton, N.J.) 607 43-52.
[68] A. Spirin, High-throughput cell-free systems for synthesis of functionally active proteins. Trends in biotechnology 22 (2004) 538-545.
[69] M. He, and F. Khan, Ribosome display: next-generation display technologies for production of antibodies in vitro. Expert review of proteomics 2 (2005) 421-430.
[70] O. Miller, K. Bernath, J. Agresti, G. Amitai, B. Kelly, E. Mastrobattista, V.r. Taly, S. Magdassi, D. Tawfik, and A. Griffiths, Directed evolution by in vitro compartmentalization. Nature methods 3 (2006) 561-570.
[71] R. Roberts, and J. Szostak, RNA-peptide fusions for the in vitro selection of peptides and proteins. Proceedings of the National Academy of Sciences of the United States of America 94 (1997) 12297-12302.
[72] N. Ramachandran, E. Hainsworth, B. Bhullar, S. Eisenstein, B. Rosen, A.Y. Lau, J.C. Walter, and J. LaBaer, Self-assembling protein microarrays. Science 305 (2004) 86-90.
[73] P. Angenendt, J. Kreutzberger, J. Glokler, and J.D. Hoheisel, Generation of high density protein microarrays by cell-free in situ expression of unpurified PCR products. Mol Cell Proteomics 5 (2006) 1658-66.
[74] D. Chatterjee, K. Sitaraman, C. Baptista, J. Hartley, T. Hill, and D. Munroe, Protein microarray on-demand: a novel protein microarray system. PLoS One 3 (2008).
[75] M. He, O. Stoevesandt, E. Palmer, F. Khan, O. Ericsson, and M. Taussig, Printing protein arrays from DNA arrays. Nature methods 5 (2008) 175-177.
[76] S.C. Tao, and H. Zhu, Protein chip fabrication by capture of nascent polypeptides. Nat Biotechnol 24 (2006) 1253-4.
[77] M. He, Generation of protein in situ arrays by DiscernArray technology. Methods Mol Biol 264 (2004) 25-31.

[78] N. Ramachandran, J. Raphael, E. Hainsworth, G. Demirkan, M. Fuentes, A. Rolfs, Y. Hu, and J. LaBaer, Next-generation high-density self-assembling functional protein arrays. Nature methods 5 (2008) 535-538.
[79] F. Festa, S. Rollins, K. Vattem, M. Hathaway, P. Lorenz, E. Mendoza, X. Yu, J. Qiu, G. Kilmer, P. Jensen, B. Webb, E. Ryan, and J. Labaer, Robust microarray production of freshly expressed proteins in a human milieu. Proteomics. Clinical applications.
[80] P. Angenendt, J.r. GlÃ¶kler, Z.n. Konthur, H. Lehrach, and D. Cahill, 3D protein microarrays: performing multiplex immunoassays on a single chip. Analytical chemistry 75 (2003) 4368-4372.
[81] P. Angenendt, J. Wilde, G. Kijanka, S. Baars, D.J. Cahill, J. Kreutzberger, H. Lehrach, Z. Konthur, and J. Glokler, Seeing better through a MIST: evaluation of monoclonal recombinant antibody fragments on microarrays. Anal Chem 76 (2004) 2916-21.
[82] M. Valjavec-Gratian, T. Henderson, and T. Hill, Tus-mediated arrest of DNA replication in Escherichia coli is modulated by DNA supercoiling. Molecular microbiology 58 (2005) 758-773.
[83] O. Stoevesandt, M. He, and M. Taussig, Protein microarrays printed from DNA microarrays. Methods in molecular biology (Clifton, N.J.) 671 95-106.
[84] O. Stoevesandt, M. Vetter, D. Kastelic, E.A. Palmer, M. He, and M.J. Taussig, Cell free expression put on the spot: advances in repeatable protein arraying from DNA (DAPA). N Biotechnol 28 282-90.
[85] K. Sitaraman, and D.K. Chatterjee, Protein-protein interactions: an application of Tus-Ter mediated protein microarray system. Methods Mol Biol 723 185-200.
[86] W. Montor, J. Huang, Y. Hu, E. Hainsworth, S. Lynch, J.-W. Kronish, C. Ordonez, T. Logvinenko, S. Lory, and J. LaBaer, Genome-wide study of Pseudomonas aeruginosa outer membrane protein immunogenicity using self-assembling protein microarrays. Infection and immunity 77 (2009) 4877-4886.
[87] R. Manzano-RomÃ¡n, V. DÃ-az-MartÃ-n, M. GonzÃ¡lez-GonzÃ¡lez, S. Matarraz, A. Alvarez-Prado, J. Labaer, A. Orfao, R. PÃ©rez-SÃ¡nchez, and M. Fuentes, Self-assembled Protein Arrays from an Ornithodoros moubata Salivary Gland Expression Library. Journal of proteome research 11 5972-5982.
[88] D.L. Doolan, Y. Mu, B. Unal, S. Sundaresh, S. Hirst, C. Valdez, A. Randall, D. Molina, X. Liang, D.A. Freilich, J.A. Oloo, P.L. Blair, J.C. Aguiar, P. Baldi, D.H. Davies, and P.L. Felgner, Profiling humoral immune responses to P. falciparum infection with protein microarrays. Proteomics 8 (2008) 4680-94.
[89] P.A. Beare, C. Chen, T. Bouman, J. Pablo, B. Unal, D.C. Cockrell, W.C. Brown, K.D. Barbian, S.F. Porcella, J.E. Samuel, P.L. Felgner, and R.A. Heinzen, Candidate antigens for Q fever serodiagnosis revealed by immunoscreening of a Coxiella burnetii protein microarray. Clin Vaccine Immunol 15 (2008) 1771-9.
[90] K. Anderson, S. Sibani, G. Wallstrom, J. Qiu, E. Mendoza, J. Raphael, E. Hainsworth, W. Montor, J. Wong, J. Park, N. Lokko, T. Logvinenko, N. Ramachandran, A. Godwin, J. Marks, P. Engstrom, and J. Labaer, Protein microarray signature of autoantibody biomarkers for the early detection of breast cancer. Journal of proteome research 10 85-96.
[91] D. Gibson, J. Qiu, E. Mendoza, K. Barker, M. Rooney, and J. LaBaer, Circulating and synovial antibody profiling of juvenile arthritis patients by nucleic acid programmable protein arrays. Arthritis research & therapy 14.
[92] M. He, and M. Taussig, DiscernArray technology: a cell-free method for the generation of protein arrays from PCR DNA. Journal of immunological methods 274 (2003) 265-270.

[93] X. ZÃ¡rate, D. Henderson, K. Phillips, A. Lake, and D. Galbraith, Development of high-yield autofluorescent protein microarrays using hybrid cell-free expression with combined Escherichia coli S30 and wheat germ extracts. Proteome science 8 32.
[94] B. Takulapalli, J. Qiu, D. Magee, P. Kahn, A. Brunner, K. Barker, S. Means, S. Miersch, X. Bian, A. Mendoza, F. Festa, K. Syal, J. Park, J. LaBaer, and P. Wiktor, High density diffusion-free nanowell arrays. Journal of proteome research 11 4382-4391.
[95] M. Kimzey, X. Zarate, D. Galbraith, and S. Lau, Optimizing microarray-based in situ transcription-translation of proteins for matrix-assisted laser desorption ionization mass spectrometry. Analytical biochemistry 414 282-286.
[96] A. Oleinikov, M. Gray, and J. Zhaoâ€¦, Self-assembling protein arrays using electronic semiconductor microchips and in vitro translation. Journal of proteome â€¦ (2003).
[97] T. Seefeld, A. Halpern, and R. Corn, On-chip synthesis of protein microarrays from DNA microarrays via coupled in vitro transcription and translation for surface plasmon resonance imaging biosensor applications. Journal of the American Chemical Society 134 12358-12361.
[98] C. Nicolini, and E. Pechkova, An overview of nanotechnology-based functional proteomics for cancer and cell cycle progression. Anticancer research 30 2073-2080.
[99] D. Sblattero, and A. Bradbury, Exploiting recombination in single bacteria to make large phage antibody libraries. Nat Biotechnol 18 (2000) 75-80.
[100] M. He, O. Stoevesandt, E.A. Palmer, F. Khan, O. Ericsson, and M.J. Taussig, Printing protein arrays from DNA arrays. Nat Methods 5 (2008) 175-7.

ACKNOWLEDGEMENTS

I want to thank Prof. Claudio Santoro and Daniele Sblattero to give me the possibility to work and improve my knowledge in their laboratory. Furthermore, I thank Diego Cotella for his support and help in all the types of scientific doubts and problems. Finally, I thank all the people that have been part of the Applied Biology laboratory in these three years: Valentina Sedini, Marta Armilli, Eleonora Rizzato, Sara Zonca, Laura Patrucco, Frank Antony, Maria Felicia Soluri, Zsilvia Bakò, Omar Sabry and Andrea Chiesa.