Mammalian display for high throughput antibody screening

A Thesis

submitted to

Indian Institute of Science Education and Research Pune in partial fulfilment of the requirements for the BS-MS Dual Degree Programme

by

Jatin Bedi

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

Pashan, Pune 411008, INDIA.

April, 2023

Supervisor: Prof. Christoph Merten

Department of Bioengineering, EPFL

Certificate

This is to certify that this dissertation entitled ***"Mammalian display for high throughput antibody screening"*** towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Jatin Bedi at École polytechnique fédérale de Lausanne, under the supervision of Prof. Christoph Merten, Department of Bioengineering, during the academic year 2022-2023.

Prof. Christoph Merten

Committee:

Prof. Christoph Merten

Prof. Kundan Sengupta

This thesis is dedicated to my family at home and IISER both.

DECLARATION

I hereby declare that the matter embodied in the report entitled Mammalian display for high-throughput antibody screening are the results of the work carried out by me at the Department of Bioengineering, École polytechnique fédérale de Lausanne, under the supervision of Dr. Christoph Merten and the same has not been submitted elsewhere for any other degree.

Jatin Bedi

10th April, 2023

ABSTRACT

Antibodies have great potential for their use in immune therapies. Developing potent therapeutic monoclonal antibodies for the treatment of cancer and infectious diseases is a critical task. Although there have been advancements in display technologies to simplify the process, the major bottleneck in most existing methods is that they depend on examining one antigen against a pool of antibodies, significantly decreasing the overall throughput. In this study, we report a high-throughput assay for screening of antibody library against antigen library. It leverages a mammalian display system to express proteins of interest on the surface of cells to enrich specific cell doublets, while eliminating the formation of unspecific binding events. The pairing information is retrieved through a fusion PCR post droplet encapsulation of the cell-doublets. As a proof of concept, cells expressing a nanobody against GFP on their membrane were successfully enriched in the form of cell doublets along with GFP expressing cells utilizing this binding assay. This was demonstrated even at lower ratios of protein-specific cells in the mixture. These results suggest that different protein display libraries can be deployed to discover new protein interactors using this approach. Using this technology, we offer a strategy that could allow for high throughput screening of antibodies which can be beneficial for immunological research.

ACKNOWLEDGEMENTS

I am grateful to a number of people without whom this project would not have been possible. I would first like to thank Danny Labes from UNIL- Flow Cytometry Facility AGORA, whose training in flow cytometry was invaluable to the project. I would also like to thank Luigi Bozzo from the Cellular Imaging Facility AGORA, who helped me every-time to troubleshoot the microscope.

I am grateful to members of the Laboratory of Biomedical Microfluidics, for without their support along with regular comments and feedbacks in group meetings I would be lost! Special thanks to Blandine Vergier, for her immense support and dedication to mentor me on this project and keeping me motivated even at my lowest. I am very thankful to my advisor, Dr. Christoph Merten, who kept me inspired throughout the course of this project and provided essential feedback at every step. He helped me see the bigger picture when I was astray among smaller narratives. I am also thankful to my external advisor Dr. Kundan Sengupta for his valuable guidance and support regarding my project.

My family and friends have been the biggest pillar of support throughout this journey of 5 years. Without their support I would not have reached this part of my life. I thank each and every one of them who always kept me motivated through thick and thin.

Finally, I would like to thank every individual person who has helped even with the smallest of things through the entire duration of this project. Whether it be borrowing some reagent or with some intellectual advice. Thank you all for the love and support!

CONTRIBUTIONS

Contributor name	Contributor role
Christoph Merten, Blandine Vergier	Conceptualization Ideas
Christoph Merten, Blandine Vergier, Jatin Bedi	Methodology
-	Software
Jatin Bedi	Validation
Jatin Bedi	Formal analysis
Jatin Bedi	Investigation
Christoph Merten	Resources
Jatin Bedi	Data Curation
Jatin Bedi	Writing - original draft preparation
Christoph Merten, Blandine Vergier	Writing - review and editing
Jatin Bedi	Visualization
Blandine Vergier	Supervision
Christoph Merten	Project administration
Christoph Merten	Funding acquisition

List of Figures

Table of Contents

CHAPTER 1 INTRODUCTION

Proteins, the fundamental building blocks of life, are complex macromolecules composed of a series of amino acids, the sequence of which is dictated by the corresponding genetic code. Playing a central role in many critical biological processes including molecular transportation, intra and intercellular signal transduction, catalysis, and immune responses, proteins are essential for the proper functioning of living systems. Significantly, a majority of the curiosities pursued in the fields of biology and biochemistry revolve around protein-protein interactions (PPIs), which encompass a wide range of cellular and metabolic phenomena, including cell-cell interactions and the regulation of developmental processes. PPIs were defined as “physical contacts with molecular docking between the proteins that occur in a cell or in a living organism *in vivo*” (De Las Rivas & Fontanillo, 2010). Many significant properties of PPIs have been laid out by Phizicky and Fields (Phizicky & Fields, 1995)

- Modification of kinetic properties of enzymes
- Allow substrate channelling
- Inactivate a protein
- Serve a regulatory role upstream or downstream
- create a new binding site for other effector molecules
- change the specificity of a protein for its substrate

The outcome of a majority of cellular processes can be studied through protein interactions (Farooq et al., 2021). PPIs can be identified via library screening-based approaches, where a collection of proteins are interrogated for their ability to interact with a target of interest. A genetic library is a collection of genes or DNA sequences cloned into a suitable vector. The screening process involves the identification of a specific clone that contains the target gene within the library. Typically, the products of a clone are examined, which is applicable only for expression libraries, where the DNA sequence directs the production of a protein recognized by an antibody or ligand. This method of library screening enables the identification of unique and valuable clones containing the desired gene, thereby enabling the isolation and characterization of specific proteins or peptides with desired properties. Many techniques have been developed to screen for these PPIs, including but not limited to Tandem Affinity Purification, Affinity Chromatography, Yeast two-hybrid, Yeast Display, and Phage display. There are various applications for each of these screens. A direct and relevant application of library screening is in the field of

immunology. Some of the most common surface-binding interactions in immunology include antibody-antigen binding, and T cell Receptor (TCR) binding to peptides presented on the surface of major histocompatibility complex (pMHC). (Sergeeva et al., 2006) High variability in structure among antibodies and TCRs allows for the generation of a library of diverse antibodies/receptors.

1.1 Antibody discovery

The discovery of high-affinity antibodies against important antigens with regards to their use in medicine is a crucial process. Despite being so relevant, there exist many challenges to isolate antibodies that target cell-surface proteins. A primary requirement is for the antigen to be conformationally stable and native. The soluble forms of surface-proteins are not always stable, and thus pose challenges to generate antibodies against them (Hutchings et al., 2010). To combat this issue, various different display technologies have been developed to identify suitable candidate proteins from a large library. Cell surface display has been used to express a protein of interest on the surface of prokaryotes (bacteriophage) or eukaryotic cells (yeast, mammalian cells) (Valldorf et al., 2022). This principle of fusing the protein of interest with a membrane anchoring motif in order to generate cell-surface display systems is a method of great appeal in the domain of protein bioengineering. By utilizing this approach, the displayed protein can be accessed readily by other targets, thereby providing numerous benefits. A noteworthy observation from multiple studies is that peptides which bind to cell-surface molecules selected through display technologies are frequently found to share significant domain similarity with their endogenous counterparts(Hartley, 2002; Smothers et al., 2002). Thus, these are powerful tools in a growing market. The field of monoclonal antibody generation is expected to be roughly 300 billion US dollars by the year 2025. (Lu et al., 2020)

1.2 Display Technologies

Display libraries are an ideal candidate for the selection of proteins to develop targeted therapeutics. One of the most common technologies in use today for the Ab discovery is Phage Display (PD). It has resulted in the production and sale of 14

therapeutic Abs that are in use today (Sellmann et al., 2020; Smith, 2019). The biggest advantage underlying this technology is in its construction of a large diversity of libraries (10^{11} clones) (Schwimmer et al., 2013). The antibody protein of interest is fused to the pIII minor coat protein of filamentous phage resulting in surface display of the protein on the phage, which can then be screened for specificity, and affinity (Fig. 1C). This is achieved by incubation with immobilized antigens on solid surfaces, or using soluble biotinylated antigens (Smith, 1985). The biggest challenge faced by PD lies in the expression of the protein itself. Prokaryotic translation machinery is not capable of performing post-translational modifications, such as formation of di-sulphide bonds. With this drawback, full length antibodies, as they would be expressed in-vivo, cannot be displayed on the phage surface (Jaroszewicz et al., 2022). Single-chain variable fragments (scFv), the region of the antibody encoding for the variable heavy and light chains, are used instead.

Yeast surface display (YSD) in *Saccharomyces cerevisiae* has been made possible through cell wall glycoproteins, namely, aga1p, aga2p, alpha agglutinin, and others (Cherf & Cochran, 2015). The method is very similar to the one used in mammalian cell display. The protein is fused at the C-terminus of a surface protein (for eg. aga2p) and lies between two epitope tags used for detection of the protein (Fig.1A). Once YSD was established for an scFv antibody against fluorescein (Boder & Wittrup, 1997), its usage in developing other scFv antibodies, Fab fragments and even whole IgG antibodies grew considerably. (Boder et al., 2012) This new eukaryotic model system quickly became a promising technology as compared to its phage and bacterial counterparts since it allowed selection of complex proteins that require post-translational modifications for the first time (Boder & Wittrup, 1997).

Despite the advances made in phage and ribosome display, its ability to detect binders relies on enzyme-linked immunosorbent assay (ELISA) screening, lowering the overall final throughput. YSD offered a novel eukaryotic screening platform. Its larger size also allowed for millions of clones to be analyzed via flow sorting. However, its major limitation remains in a difference in glycosylation patterns, expression and secretion machinery, and the codon usage from that found in mammalian systems (Parthiban et al., 2019).

Mammalian cells offer an ideal platform for the identification of proteins targeting cell-membrane protein. Compared to other techniques discussed above, it provides certain advantages as it employs cells that are suitable not only for selection but also

for production of the desired proteins. Mammalian expression and secretory apparatus are particularly favourable for producing properly folded full-length Ig molecules with glycosylation patterns and post-translational modifications similar to those found in humans (Valldorf et al., 2022). Another interesting benefit of utilizing a mammalian system is in taking advantage of somatic hypermutation through activation-induced cytidine deaminase in-vitro (Bowers et al., 2014; Di Noia & Neuberger, 2007). Surface expression of proteins can be achieved through transient transfection or stable cell-line generation through various transfection methods, or viral transduction technologies to generate a library (Beerli et al., 2008). The major limiting factor of mammalian display is smaller library size when compared to techniques like phage or bacteria display (Valldorf et al., 2022). Larger libraries can be constructed using random gene insertion techniques listed above, however this causes variability in gene expression for every gene insertion since it is a site-dependent phenomenon. Recently, with advances in genetic engineering tools like CRISPR, this issue can be circumvented to an extent. It allows for monoclonal display of proteins in a cell with library sizes that can reach up to 10^7 (Parthiban et al., 2019). To achieve protein display in a mammalian host, the protein to be targeted is inserted in-frame to two important sequences- N terminus IgK-leader sequence, a signal sequence targeting the final protein to the surface, and the C-terminus transmembrane domain of Platelet-derived growth factor receptor (PDGFR). This transmembrane domain allows for surface anchoring of the protein of interest in the mammalian cell (Fig.1B). The protein is also tagged with two different epitope tags for detection- Myc tag (EQKLISEEDL) and the HA tag (YPYDVPDYA).

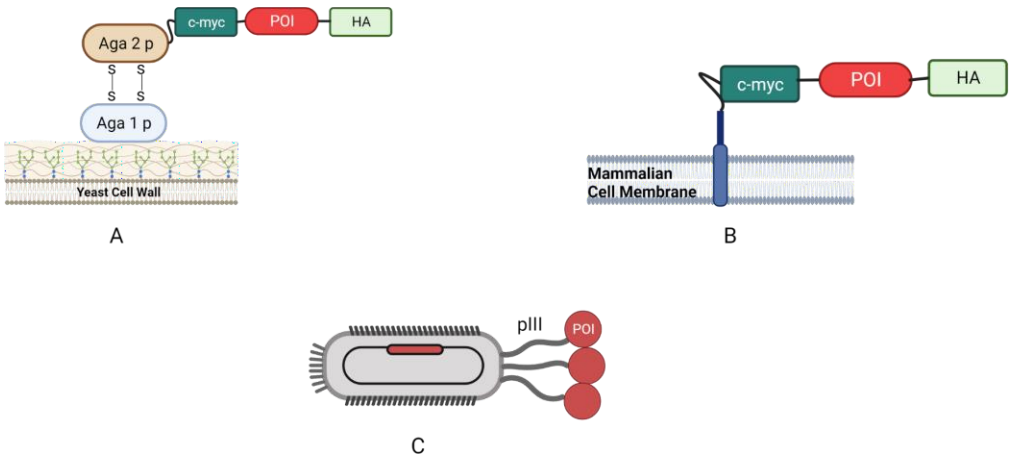


Figure 1:- Illustrations of various display methods that are in use today, namely (A) Yeast Cell Display (B) Mammalian Cell Display (C) Phage Display

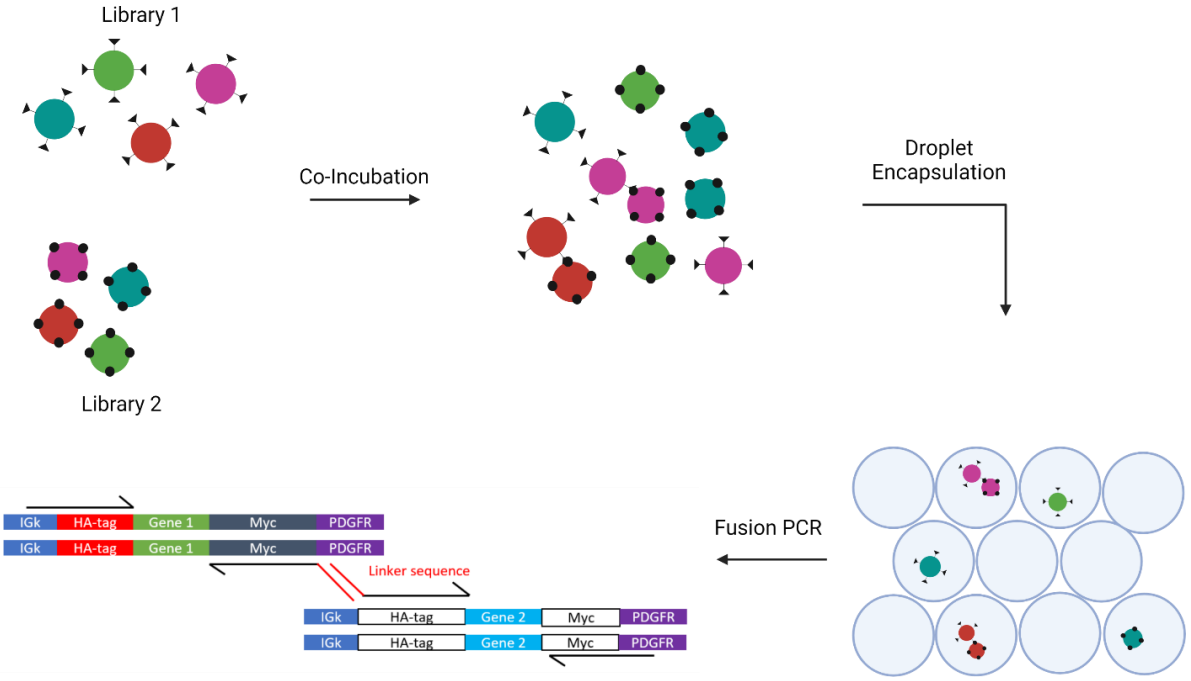


Figure 2: Schematic representing the proposed pipeline for analyzing cell-doublet interactions and obtaining a final readout. Two different display libraries of proteins are generated. The different types of monoclonal display cells are then incubated, followed by droplet encapsulation to isolate and analyse cell-doublet events. Finally, the cells are lysed in droplet and undergo an overlap extension PCR to retrieve the DNA sequence of the interacting proteins

A significant challenge in the process of screening is the preparation of antigens, particularly when it comes to ion channels and G protein-coupled receptors (Hutchings et al., 2019). These targets can present multiple obstacles, such as poor expression and solubility requiring extensive optimization of purification conditions. Furthermore, preserving the native conformation of membrane proteins post-purification poses an additional hurdle. Conventional screens are all directed towards screening antibody libraries against one target antigen. To our knowledge, there are no methods in use today that enable multiplexed screening of two libraries of proteins against each other. These challenges underscore the need for further innovations in the development of techniques and technologies that enable more efficient preparation and characterization of membrane proteins. Researchers have tried to circumvent this issue by utilizing whole cell panning along with yeast display libraries (Hoogenboom et al., 1999). Yeast display library are co-incubated with adherent mammalian cells displaying a target protein. Yeast cells with no specific binding partners are washed off to enrich for specific binding antibodies (Tillotson et al., 2013). This approach is challenging, relying on yeast displayed libraries, targets the display library against only a single antigen, and may be hindered by unspecific binding among yeast cells itself. Thus, technologies need to be developed to allow

multiplexed screening for antigen antibody reactions in a mammalian background without the need of purification of antigens.

It is essential to retrieve back the pairing information of the interacting cell-complexes in the form of gene sequences. This can be achieved through fusion PCR. At the Laboratory of Biomedical Microfluidics, droplet microfluidic platforms have been established to screen antibodies for binding to proteins expressed on the membrane of target cells. In this system, small aqueous droplets are surrounded by oil and act as independent screening vessels to identify protein binding pairs. This technology enables highly multiplexed screening for the identification of antibodies or T-cell receptors binding to hundreds of different targets in a single screen (Fig. 2). (Shembekar et al., 2018) Droplet encapsulation allows for isolation of cell doublets, which can then be lysed to perform PCR in droplet for a final readout of sequences of interacting proteins.

1.3 Objectives

The goal of the study is to establish a method to screen for two protein libraries against each other in a mammalian system. As a proof of concept, we will establish the method with three candidate proteins (Fig.3). These consist of two antigens: GFP and mCherry, and one nanobody: GFP-nanobody (Pleiner et al., 2015). GFP-nanobody specifically interacts with GFP, while having no known affinity towards mCherry. These proteins will be fused along with the PDGFR-transmembrane domain and IgK leader sequence by cloning them in a suitable plasmid for their surface expression in HEK293T cells. Their surface expression will be characterized through fluorescence microscopy. We hypothesize that protein pairs with a known binding affinity will interact and form cell complexes at a greater ratio when compared to a protein pair without any binding affinity. After several rounds of optimization of conditions, we successfully showed that protein-specific cellular complexes could be enriched in our assay system.

Aim: -

1) To clone candidate recombinant proteins for their surface expression in mammalian cells.
2) To characterize surface expression of the mammalian display system in HEK293T cells for its subsequent use in screening libraries.
3) To optimize conditions, mainly the formation of specific cell doublets, while eliminating the formation of unspecific binding events.

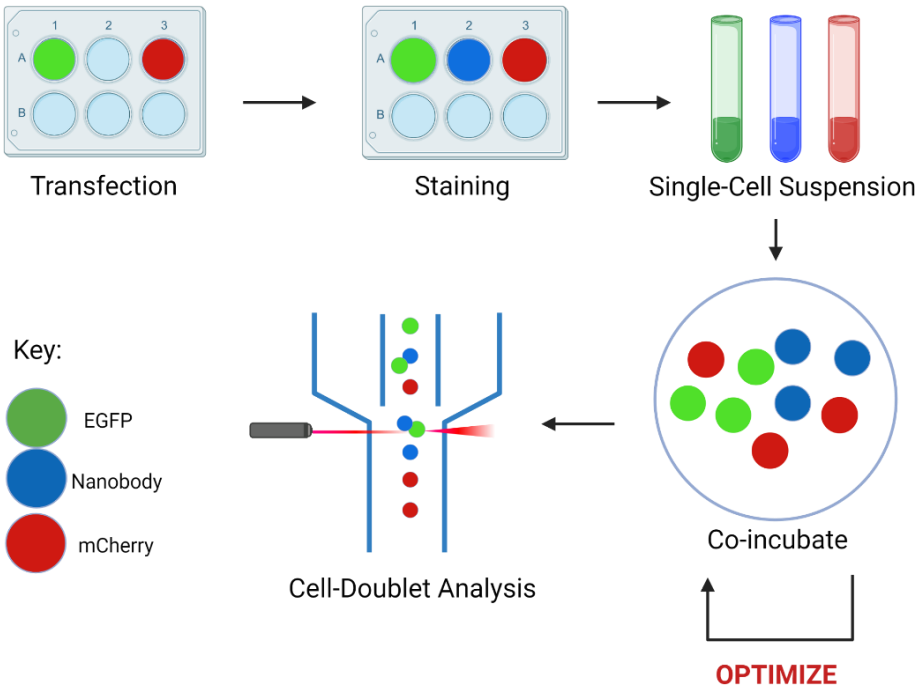


Figure 3: Schematic of the protocol. The cells are transfected with the constructs in a 6-well plate. GFP-nanobody cells are then stained for their detection. The cells are harvested as a single cell suspension and then co-incubated. This process of co-incubation undergoes various cycles of optimization to achieve an optimal readout in a flow-cytometer.

(The study below does not describe methods for the overlap extension PCR. The cells are analyzed through a traditional flow cytometer instead of droplet encapsulation to help in faster optimizations)

CHAPTER 2 MATERIALS AND METHODS

2.1 Vector Construct

pDisplay™ Mammalian Expression Vector (Invitrogen, V66020) was used as a backbone to clone in three genes- eGFP, GFP-nanobody, and mCherry. The genes were cloned using GeneArt™ Gibson Assembly HiFi Master Mix (Invitrogen, A46628). pDisplay was first digested with a single enzyme, SalI-HF (New England Biolab, R3138S) to linearize the backbone and allow for gene insertion (Fig. S3A). 1ug of the plasmid was incubated with 20 units of enzyme for 15 mins at 37C. The vector construct contains 3 important sequences of interest: - 1) The Leader sequence for correct signalling of the peptide on the cell membrane. 2) PDGFR transmembrane domain for anchoring the peptide on the cell membrane and displaying it extracellularly. 3) HA-tag for the detection of protein with antibody staining. The genes of interest were cloned in-frame using primers designed specifically for Gibson assembly (Table S1), downstream of the leader sequence and upstream of the PDGFR-TM domain. Insertion of the gene in the vector backbone was confirmed through PCR reactions of DNA extracted from transformed colonies (Fig. S3 B-D). Extracted DNA which contained the inserted gene showed a band at the size of the gene itself. All clones that showed a positive band were sent for sequencing to confirm correct and in-frame insertion in the vector(Fig. S4). In each case we obtained at least 1 clone that had the gene inserted in-frame. The colonies that contained the desired vector were then later expanded in larger culture volumes and a midi-prep was performed to extract larger quantities of DNA for future transfection experiments.

Figure 4: Vector map of pDisplay

PCR cycling conditions

Step	Temp (°C)	Time (s)
Initial Denaturation	98	30
Denaturation **Annealing** **Elongation**	98 60 72	7 20 25
Final extension	72	120
Hold	4	Infinite

2.2 Cell culture and transfection

HEK293T cells were provided by Alan Guichard from Michelle De-Palma group (EPFL) and cultured in RPMI (Gibco, 61870-010) media, supplemented along with 10% Fetal Bovine Serum (FBS), 5mM HEPES buffer, and Penicillin (100 units/ml)/Streptomycin (100 µg/ml). Cells were incubated in 5% CO2 at a constant temperature of 37°C and split every 3 days upon 80% confluency.

Chemical transient transfection was performed in 6-well plates. Multiple wells of a 6-well plate were seeded in 2 ml complete media with cells (~0.5 × 10^6 cells per well)

one day prior to transfection. The next day the media was replaced with fresh 1.5ml complete media. They were transfected with ~3 µg of plasmid vector using 9 µL of 1 µg/µL linear (PEI) polyethylenimine hydrochloride (PolySciences, 49553-93-7). Two tubes were prepared for the transfection of 1 well in the 6-wells plate. The first tube contained PEI diluted in 250uL OptiMEM (Gibco, A41248-01) while the other tube contained the DNA diluted in 250uL OptiMEM. The tubes are then mixed, vortexed well and incubated for 30 minutes at room temperature to allow for PEI-DNA complexes to form. This mixture is then added dropwise in the well. The ratio of PEI:DNA used was 3:1 (9 µg of PEI to 3µg of PEI). Cells are incubated at 37C, 5% CO2 for 12hours, then the transfection media is removed and replaced with 2ml complete media. Cells are incubated for an additional 36h before analysis.

To analyze transfection efficiency of mCherry and EGFP transfected cells, they were suspended in 1x FACS buffer (1x PBS, 1% FBS, 2mM EDTA) at a density of 106/ml. For the analysis of GFP-nanobody transfected cells, 10^6/ml cells were collected in 100uL medium and incubated with the primary antibody against the HA-tag for 30 minutes at 4C. They were then washed with 1xFACS buffer, followed by incubation with a conjugated secondary antibody for 30minutes at 4C. The cells were washed again with 1xFACS buffer before the final analysis. Primary antibodies used was against HA-tag (Invitrogen, 26183) at 1:1000 dilution in FACS buffer. Secondary antibody was diluted in FACS buffer as well: Goat anti Mouse-Alexa 568 (Invitrogen, A11031) at 1:2000 dilution.

2.3 Cell Binding Assay

48h after transfection, cells are prepared for the binding assay. pDisplay-GFP nanobody transfected cells and untransfected cells (for control) are stained by incubation with CellTrace Violet dye (Invitrogen, C34557). The cells are first washed with 1xPBS to remove any excess media in the sample. 5uM working stock of the dye is prepared by adding 1uL per 1ml of 1xPBS. This solution is added on top of the cell culture for 20 mins and the plate is incubated at 37C protected from light. The solution is then aspirated out and the cells are washed once with 1xPBS solution containing 1%FBS. The cells are then incubated in complete media for at least 10 minutes before proceeding with further analysis. Cells are detached using only 3mM Ethylenediaminetetraacetic acid (EDTA) dissolved in 1xPBS instead of trypsinizing

them so as to not cleave cell-membrane proteins. Cells are then pelleted and a single cell suspension is created by passing them through a 70um filter. They are counted using trypan blue to seed the right ratio of cells desired. For the assay, cells transfected with the three different plasmids were mixed and co-incubated in FreeStyle-293 media (Gibco, 12338-018) supplemented with 0.1% Plurionic F-68 (BioConcept, 5-75F02-H) and 10U/ml DNase I (Invitrogen, 18047019) and 2.5mM MgCl2 in 1.5ml MCTs (Eppendorf) at 37°C on a thermomixer for 1 hour at a constant shaking speed of 400 RPM. For the control, cells transfected with pDisplay-GFP nanobody were replaced with wild type HEK293T cells. Post incubation the cells were pipetted and analyzed directly by flow-cytometry.

2.4 Flow Cytometry Analysis

The experiments were performed on a 5 laser BSD LSR-II cytometer at the FCF-UNIL facility at AGORA. The lasers in use were of the wavelength 488nm, 405nm, and 561nm. Data analysis of the FCS files was performed using FlowJo software (BD Biosciences).

To analyze cell doublets, cells were directly transferred into FACS tubes and analyzed. The cells were gated based on scatter properties and doublets were analyzed based on a double positive fluorescence readout. First the live cells were gated based on Forward Scatter (FSC) vs Side Scatter (SSC) followed by selection of all cells positive for blue, and then finally were plotted on a mCherry vs GFP dot plot to assess percentage of blue-green doublets vs blue-red doublets (Fig. 5).

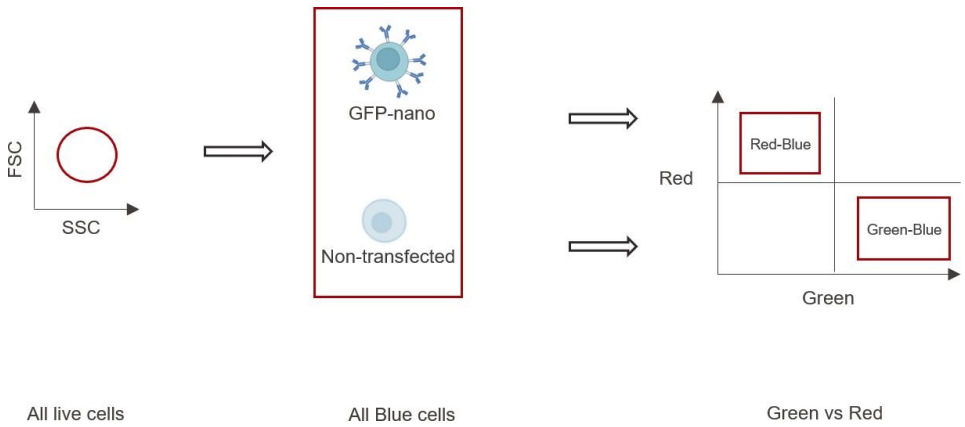


Figure 5: Representation of Gating strategy used in analysing flow cytometer data

Parameter Calculation: -

$$Fold\ Change = \frac{No.of\ doublets\ with\ GFP\ and\ GFPnanobody(Green\&Blue)}{No.of\ doublets\ with\ mCherry\ and\ GFPnanobody\ (Red\&Blue)}$$

2.5 Immunofluorescence

0.5×10^6 transfected cells/well were seeded in a 6 well plate with glass coverslips at the surface. Washes were done with 1× PBS and fixed with 4% paraformaldehyde (Sigma, P6148) prepared in 1× PBS, pH 7.4 at room temperature for 10 minutes, three times in 1× PBS (5 minutes each). Cells were then blocked in 1% bovine serum albumin (BSA) (Sigma, A2153) prepared in 1× PBST (0.1% Triton-X-100 in 1XPBS) for 30 minutes and washed three times with 1× PBS to avoid any unspecific binding of the antibodies. Incubation with primary antibodies was performed in 1% BSA for 90 minutes at RT and with secondary antibodies for 60 minutes at RT, with washes in between using 1× PBST. Primary antibody used was against HA-tag (Invitrogen, 26183) at 1:500 dilution in 0.5% BSA. Secondary antibody was diluted in 1× PBS: Goat anti Mouse-Alexa 568 (Invitrogen, A11031) at 1:1000 dilution. Three washes of 1x PBS were performed. Cells were mounted in Prolong™ Gold Antifade with DAPI (Invitrogen, P36935).

CHAPTER 3 RESULTS

3.1 Expression of candidate proteins at the surface of mammalian cells

For the purposes of our experiments, we wanted three different lines of HEK293T, each one of them displaying a unique protein on its cell surface. To achieve this, three plasmid constructs– pDisplay-GFP, pDisplay-GFPnanobody, and pDisplay-mCherry were transiently transfected into HEK-293T cells using polyethyleneimine, and their surface expression was analyzed through fluorescence microscopy (Fig.6). Since GFP-nanobody is not a fluorescent protein, its expression on the cell membrane was visualized through immunostaining with conjugated antibodies against the HA-tag (Fig.6C). We observed cell surface staining of the nanobody expressing HEK293T, whereas non-transfected HEK293T showed no staining (Fig. S1), confirming the antibody's specificity. As an additional control, both pDisplay-GFP-Nanobody and non-transfected HEK293T cells were stained with only the conjugated secondary antibody, to confirm its specificity to the primary antibody which showed no staining in either case (Data not shown).

Validation of GFP binding specificity with GFPnanobody

The surface of mammalian cell comprises of a diverse set of proteins. It is crucial to ensure that protein of interest is specific to the desired target and not against any other surface protein. As an additional test to validate expression of GFP-nanobody, and to analyze binding specificities, we performed an immunostaining of pDisplay-GFP-nanobody transfected HEK293T cells and stained for surface expression of GFPnanobody using soluble form of the GFP protein. As a control, we stained cells transfected with the empty pDisplay vector (Fig S2.) to validate specificity of GFP to GFP-nanobody and not to any other surface protein.

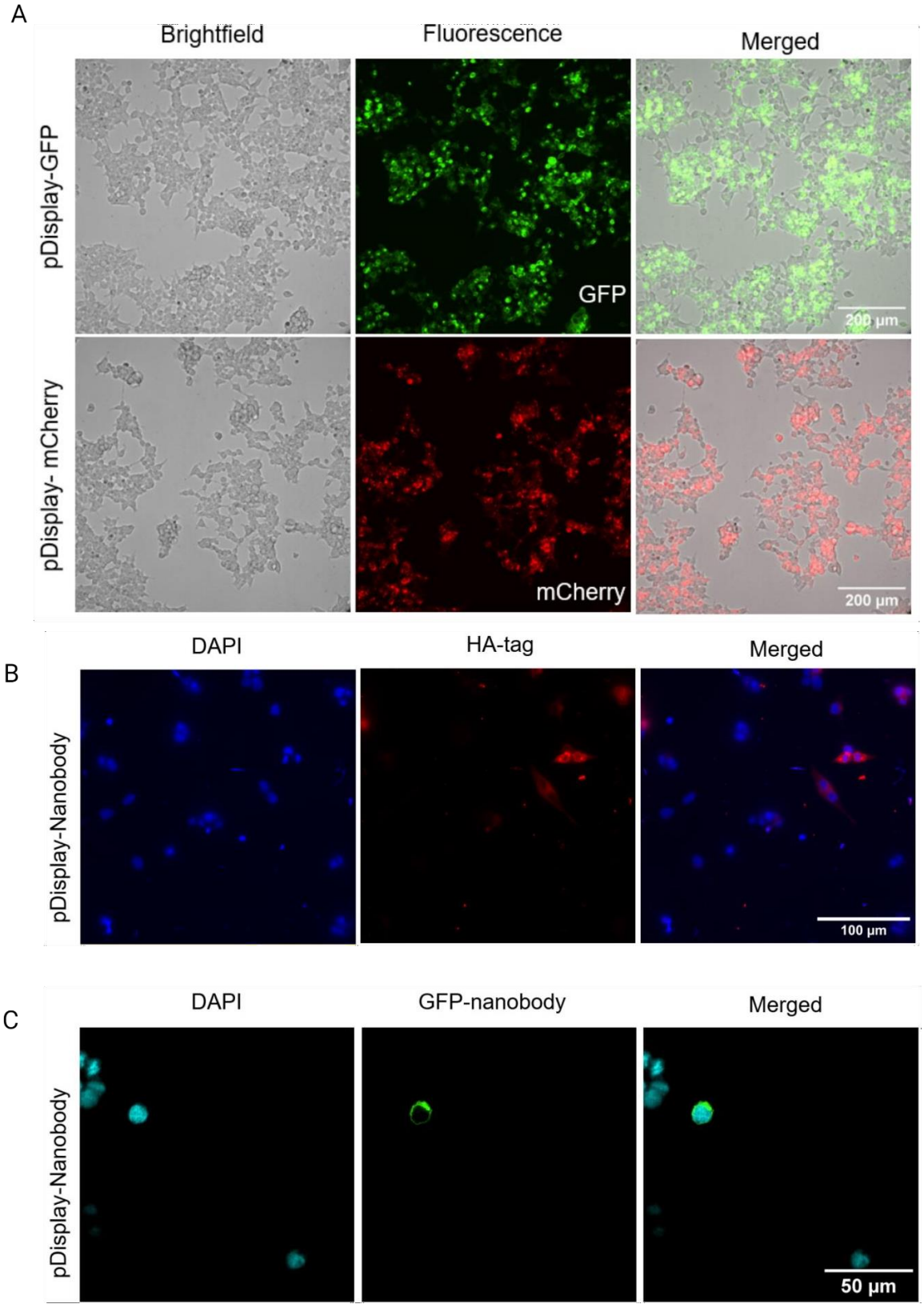

A
Brightfield
Fluorescence
Merged
pDisplay-GFP
GFP
200 µm
pDisplay- mCherry
mCherry
200 µm
B
DAPI
HA-tag
Merged
pDisplay-Nanobody
100 µm
C
DAPI
GFP-nanobody
Merged
pDisplay-Nanobody
50 µm

*Figure 6**: Surface expression of the different pDisplay constructs** (A). Representative images of HEK293T cells transfected with pDisplay-GFP and pDisplay-mCherry constructs. Green-GFP; Red-mCherry. Scale bar = 200um **(B) HA Tag staining** Representative mid-optical sections from confocal microscopy of HEK293T cells immunostained for HA-tag for GFP-nanobody+ cells. Blue-DAPI; Red- HA tagged GFPnanobody. Scale bars = 100uM **(C) Soluble GFP staining**. Representative mid-optical sections from confocal microscopy of pDisplay-GFPnanbody transfected HEK293T cells immunostained for GFP-nanobody expression. Scale Bar = 50uM*

These observations validated surface expression of our constructs and thus implied it could be used for future experiments going forward.

The transfection efficiencies for each of the vector constructs was determined via flow cytometry. As seen in Fig.7, the mCherry construct had the highest transfection efficiency with >90% transfected cells, whereas the GFP construct had close to 90% transfected population. Since GFP nanobody is not a fluorescent protein, its surface expression was analyzed through staining with anti-HA antibody for its detection via flow cytometer which showed 82% transfected population.

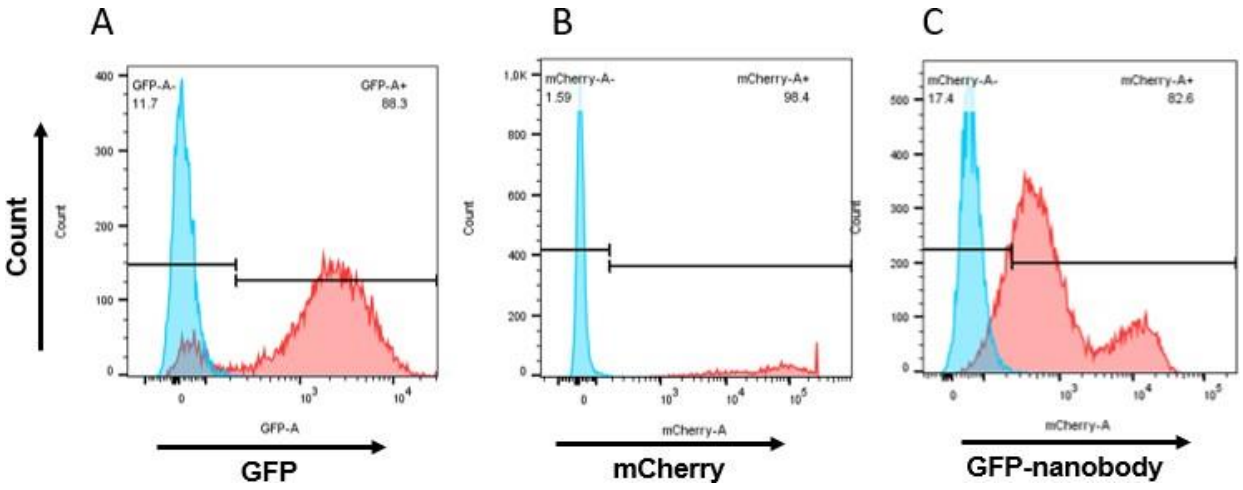


*Figure 7**: Transfection efficiencies** analysed via flow cytometry for HEK293T cells transfected with (A) GFP (B) mCherry and (C) GFP-nanobody. Red-Transfected HEK293T, Blue- Non transfected HEK293T*

3.2 Cell Binding Assay

3.2a Cellular interaction analysis of HEK293T cells

To study the protein-specific interactions and cell-complex formation, we co-incubated the three differently transfected HEK293T populations at equal proportions. GFP displaying cells are specific to GFP nanobody displaying cells, whereas cells displaying mCherry on the surface serve as an internal control that help us quantify cellular complex formation between specific and unspecific interactors. GFP and mCherry signals were gated with respect to homogenous expressing populations of each for every experiment. GFP-Nanobody expressing cells were stained with CellTraceViolet prior to co-incubation for their detection via flow cytometer. A control was set up where GFP-Nanobody transfected cells were replaced with non-transfected cells in the mixture. These were also stained with the same CellTraceViolet dye. The entire gating strategy is described above in materials and methods (Section 2.4). The 2-D dot plot represents the entire population of blue cells. Q1 and Q3 represent the percentage of blue cells associated with mCherry and GFP respectively. The cell-doublets, namely GFP-GFPnanobody complexes, were enriched when compared to GFP-mCherry complexes by 4-fold, indicating a protein-specific enrichment of cell-doublets in our assay system (Fig. 8A-B). We also imaged these cell-doublet events post incubation of the cell mix, and we could visualize them as seen in Fig.8C.The results below were reproducible at least 3 times.

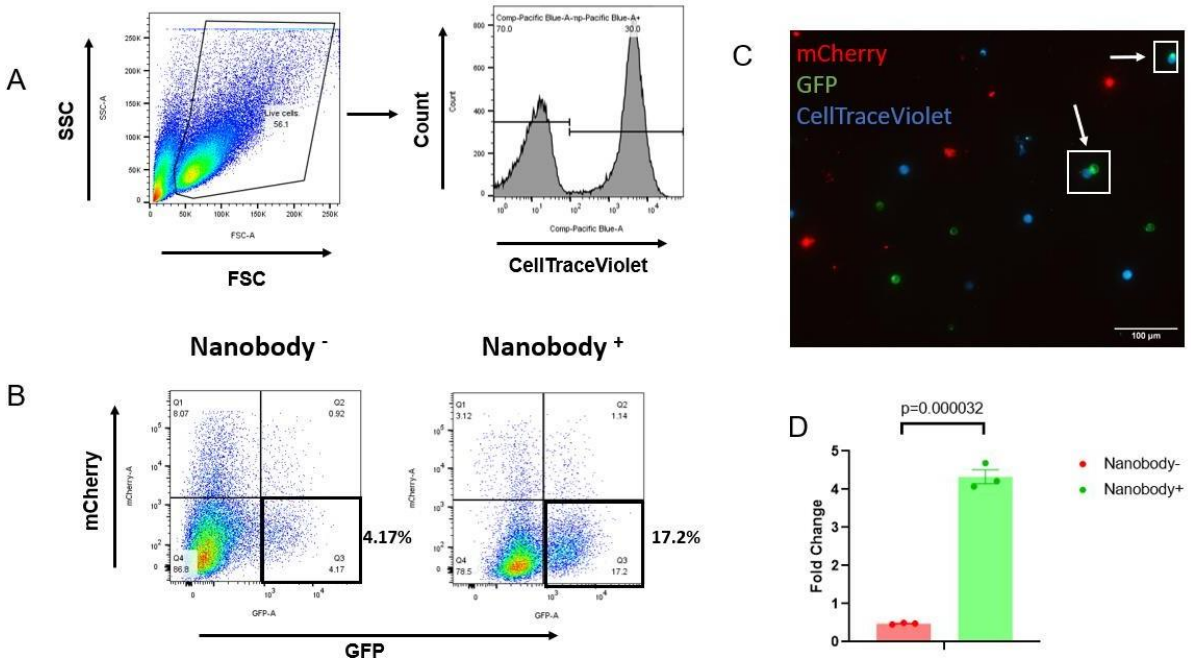


Figure 8: ***Cell-doublet formation is dependent on specific interactions between the cell-surface proteins***

(A) Cells were gated using FSC and SSC to exclude cellular debris, following which only CellTraceViolet positive population was selected in the mixture for further analysis. (B) CellTraceViolet+ populations were plotted on a two-dimensional dot plot - X:Y::GFP:mCherry. The highlighted gate (Q3) shows percentage of GFPnanobody-GFP cell doublets. Q1 refers to GFPnanobody-mCherry doublets. This is only a representative plot of three different technical replicates that were performed for both types (Nanobody and Nanobodynull) of cell mixtures. (C) Fluorescence microscopy images taken post-incubation of the cell mixture containing Nanobody+ cells. 100uL of cells were taken from one of the sample tubes and diluted in 100uL 1xPBS and imaged instantly. The highlighted cells are HEK293TGFP-HEK293TGFPnanobody cell complexes. Scale bar = 100um (D) The data for the ratio of number of Blue-Green (GFPnanobody-GFP) cell doublets to Blue-Red (GFPnanobody-mCherry) cell doublets was plotted on a bar plot for n=3 technical replicates. The error bars indicate +/- SEM and a standard two tailed t-test was performed for comparison.

3.2b Vortexing of the cell mixture allows breaking unspecific cell-doublets

Vortexing is commonly used in flow cytometry before analysis of samples to help create single cell suspensions. Thus, we wanted to test whether physical disruption would have any significant effect on the doublet status of the cells. Initially, we only tested short bursts of vortexing. We observed that these high-affinity cell-cell interactions could withstand harsh physical disruption. So, we set out to test the effect of vortexing the cells for ~1sec(pre-vortex), 60sec (Vortexed), and 300sec (Vortexed+). We tested longer periods of vortexing and its effects on cell complexes. The sample tube was first analyzed with a short burst of vortex (pre-vortex) to allow for homogenous cell suspension. The same sample tube was then subjected to 60 seconds of constant vortexing, and 300 seconds subsequently, and was analyzed after each step for its effects on cell doublet populations. We observed that with an increase in duration in vortexing, the fold change increased proportionally (Fig.9A-B). Quantitatively, we saw a 35% increase in fold change after 5 mins of harsh

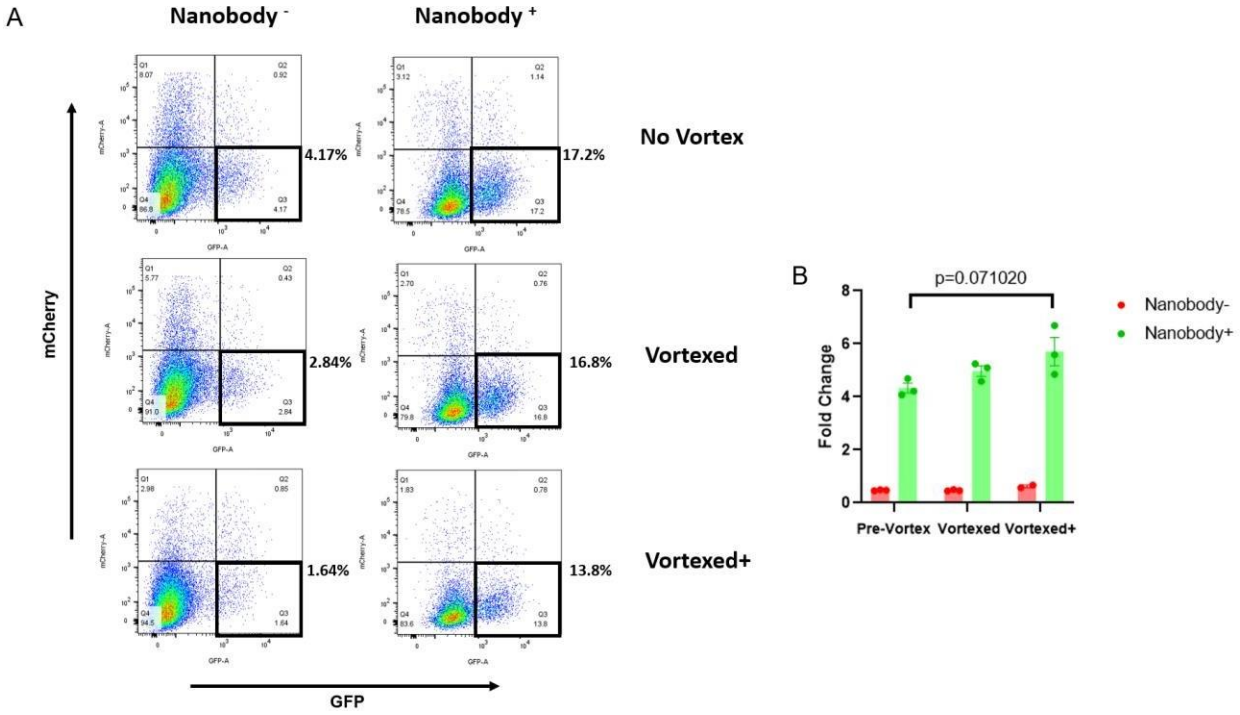


Figure 9: ***Formation of unspecific cell complexes can be reduced by vortexing***
(A) The cell incubation and gating was performed similar to the previous experiment. Each 2-D dot plot is representative of three replicates. The highlighted gate (Q3) shows percentage of GFPnanobody-GFP cell doublets. Q1 refers to GFPnanobody-mCherry doublets. The tubes were subjected to different rounds of vortexing with varying times. (B) The data for the ratio of number of Blue-Green (GFPnanobody-GFP) cell doublets to Blue-Red (GFPnanobody-mCherry) cell doublets was plotted on a bar plot for n=3 technical replicates. The error bars indicate +/- SEM and a standard two tailed t-test was performed for comparison.

vortexing. (Fig9b.) These results were much more pronounced as seen in section3.3d, Fig.11b

3.2c Ethylenediamine tetra-acetic acid (EDTA) does not significantly affect unspecific cell complexes

Another aspect we wanted to test was the addition of a chelating agent in the media and analyze its effect on disrupting non-specific ionic binding between cells. EDTA is a well-known chelating agent that chelates Magnesium and Calcium ions from the medium that are required for integrins to maintain cell adhesion, and is a common cause of cell clumping. We performed the binding assay in two different media- one with 3mM EDTA supplemented and one without any EDTA in it. The rest of the components were kept the same between the two, as outlined in materials and methods above. We observed no significant difference in the ratio of specific binding partners to unspecific partners upon addition of EDTA in the cell culture medium, indicating that it offered no significant benefit. However, we noticed a slight increase in fold change (Fig. 10A-B), which led us to believe that the concentration of EDTA may have been a limiting factor.

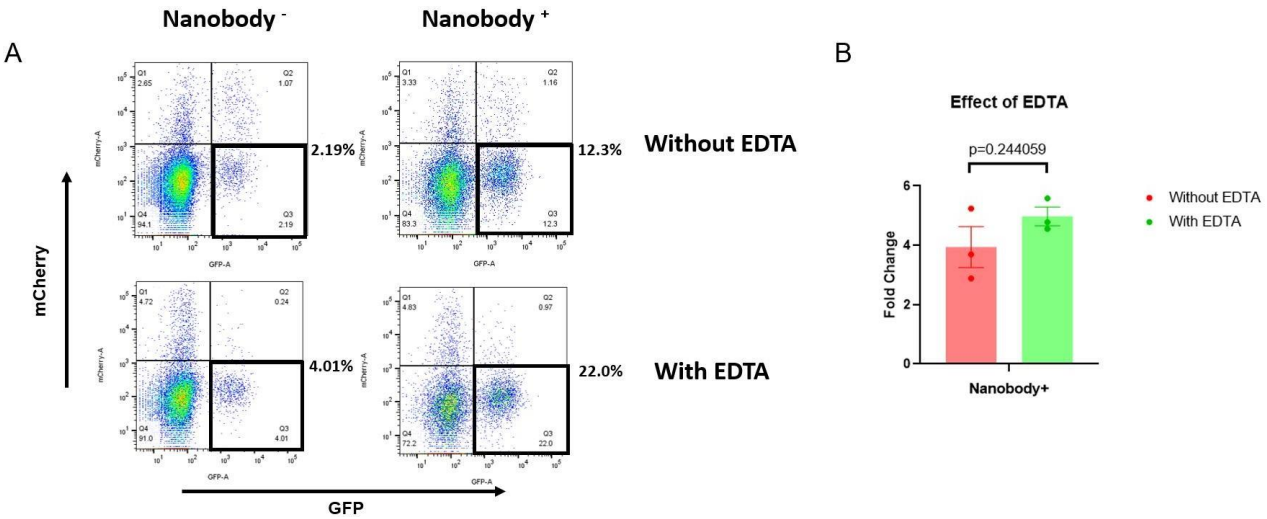


Figure 10: ***Cell-doublet formation is not significantly affected by addition of a chelating agent*** *(A) The cell incubation and gating was performed similar to the previous experiments. Each 2-D dot plot is representative of three replicates. The cells were incubated in two different medias and their respective fold-changes were quantified to compare the effect of addition of EDTA in the mixture. The highlighted gate (Q3) shows percentage of GFPnanobody-GFP cell doublets. Q1 refers to GFPnanobody-mCherry doublets. (B) The data for the ratio of number of Blue-Green (GFPnanobody-GFP) cell doublets to Blue-Red (GFPnanobody-mCherry) cell doublets was plotted on a bar plot for n=3 technical replicates. The error bars indicate +/- SEM and a standard two tailed t-test was performed for comparison*

3.2d Protein-specific cellular complexes can be enriched even at lower abundance

So far, the experiments have been conducted using an equal ratio of GFP and mCherry expressing cells. However, when screening a library of proteins, the abundance of each monoclonal cell type expressing a specific protein on its surface is very low. To mimic these conditions, we spiked the ratio of mCherry expressing cells as compared to the GFP expressing cells at ratios of 1:10 and 1:100 (GFP:mCherry respectively). This would help us analyze whether the assay system allows for rare cell populations to meet and interact with each other. To achieve a 1:10 ratio among competing antigens, we co-incubated 50,000 GFP, 50,000 GPF-nanobody, and 500,000 mCherry transfected cells in the mixture. The population of GFPnanobody expressing cells was also kept at the same ratio as the GFP population to keep both populations of the interacting pair, nanobody and antigen, at lower ratios. We observed a 4-fold change at this ratio after vortexing for 1 minutes as shown in Fig.11B, indicating that even at a lower concentrations the specific binding partners are able to find each other and form cell doublets. The effect of vortexing was much more pronounced at lower proportions than what was observed when cells were present at equal proportion. When present at a 1:100 ratio (5000GFP cells:500,000mCherry cells), we observed a 0.15-fold change, which was more than 20 times higher than the control Nanobodynull cells (Fig11B.). The analysis for the 1:100 sample was done post-vortexing for 1 minutes only since we had established its significance already in achieving a higher fold change. In both cases, we observed a significant enrichment of specific binding partners when compared to the Nanobodynull control. Thus, we show that our system is capable of enriching binding partners even if present at low abundances.

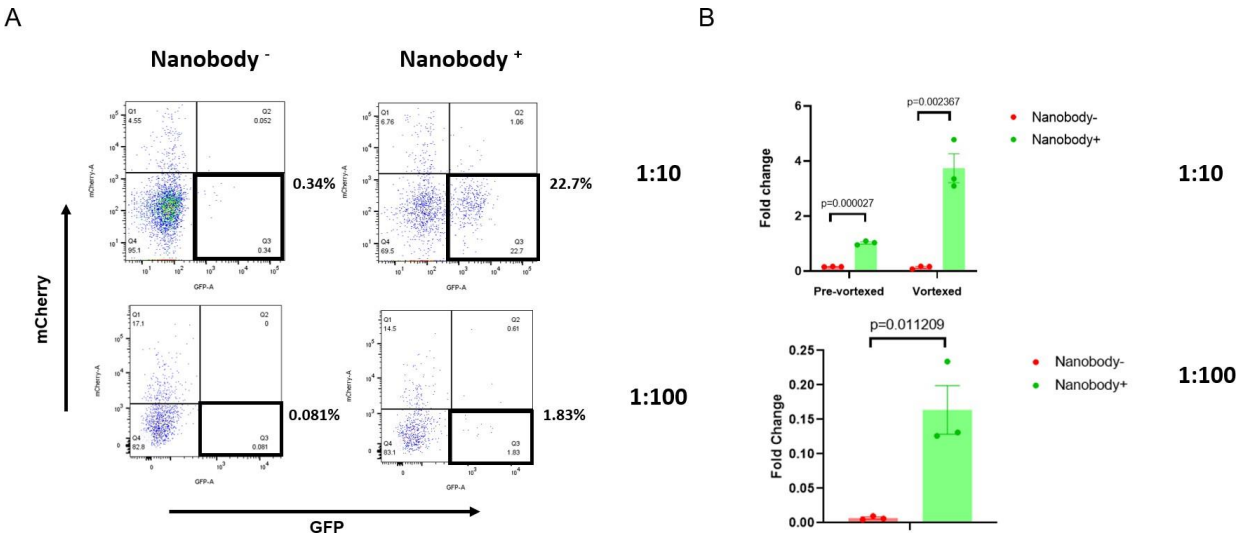


Figure 11: ***Specific cell-doublet formation is enriched even at lower concentrations***
(A) The cell incubation and gating was performed similar to the previous experiments. Each 2-D dot plot is representative of three replicates. The cells were incubated at two different ratios of 1:10 and 1:100 of GFP+:mCherry+ cells respectively. The highlighted gate (Q3) shows percentage of GFPnanobody-GFP cell doublets. Q1 refers to GFPnanobody-mCherry doublets. (B) The data for the ratio of number of Blue-Green (GFPnanobody-GFP) cell doublets to Blue-Red (GFPnanobody-mCherry) cell doublets was plotted on a bar plot for n=3 technical replicates. The error bars indicate +/- SEM and a standard two tailed t-test was performed for comparison

3.2e Incubation at high-speed affects formation of specific cell-doublets

Inspired by results shown above with regards to effect of vortexing on enriching specific cell-doublets, we wanted to test whether incubation at higher speeds would help in disrupting unspecific cell interactions. This would allow for a constant physical stress throughout the incubation period. So, we tested the binding affinity at the maximum possible incubation speed on the thermomixer at 2000RPM. Prior to performing this experiment, we tested the viability of HEK293T cells suspended in these conditions with trypan blue staining and found no change in viability which could potentially affect the final outcome. The higher speed of incubation did not result in a higher enrichment as we could only attain a less than 2-fold change.

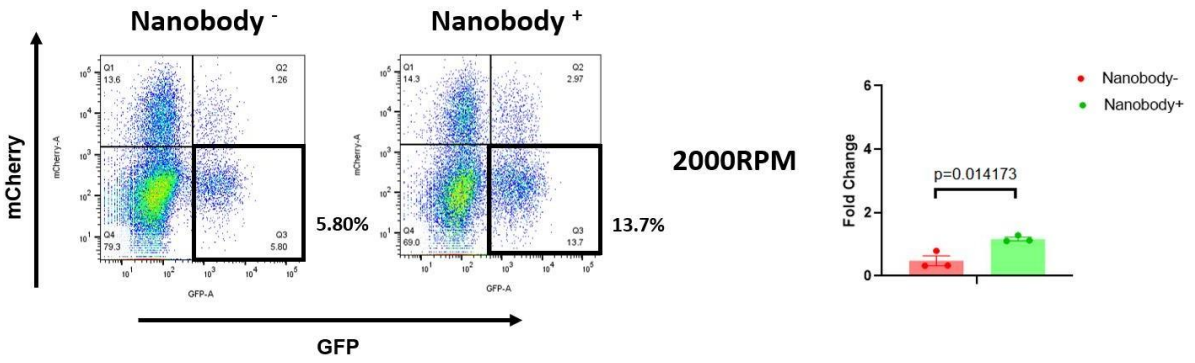


*Figure 12: **Incubation at high speed affects specific cell-doublet formation**.*
(A) The cell incubation and gating was performed similar to the previous experiments. Each 2-D dot plot is representative of three replicates. The cells were incubated at two different ratios of 1:10 and 1:100 of GFP+:mCherry+ cells respectively. The highlighted gate (Q3) shows percentage of GFPnanobody-GFP cell doublets. Q1 refers to GFPnanobody-mCherry doublets. (B) The data for the ratio of number of Blue-Green (GFPnanobody-GFP) cell doublets to Blue-Red (GFPnanobody-mCherry) cell doublets was plotted on a bar plot for n=3 technical replicates. The error bars indicate +/- SEM and a standard two tailed t-test was performed for comparison.

CHAPTER 4 DISCUSSION

In the above study, we cloned three genes of interest- GFP, GFP-nanobody, and mCherry into the pDisplay vector backbone and successfully characterized surface expression of all the genes. Then, we validated GFP's specificity to bind to GFP-nanobody through immunostaining pDisplay-GFPnanobody transfected cells with soluble GFP rather than conjugated antibodies. Next, we quantitatively analyzed HEK293T doublet formation using conventional flow cytometry. Based on the analysis outlined above, we demonstrated that antigen-specific HEK293T interaction is much more frequent (approx. 6 fold) when compared to unspecific interactions in our assay system. Our results further suggest that high-affinity specific interactions between proteins displayed on the cell surface are stable and can withstand physical stress, much more when compared to nonspecific interactions. We also showed that protein specific cellular complexes were enriched even when present at lower abundances (1:10 and 1:100).

HEK293T cells were chosen as it is a well-developed system known for high transfection efficiencies. Prior to HEK293T, attempts were made to perform the above experiments in a suspension cell line, namely Jurkat. A suspension cell line allows cells to stay in a suspension and have no risk of them adhering to a surface. They also do not need to be detached before incubation. However, due to poor and non-reproducible transfection efficiencies, this was later abandoned.
Initially the cell binding assay was performed in a well-plate system. This proved challenging to maintain cells in suspension such that they don't adhere to the surface of the well and constantly mix to be able to interact with new cells. At 37^{o}C the cells would readily attach to the bottom surface of the well plate, which posed a problem as it would not allow cells to be surrounded by a new set of neighbouring cells every few seconds. Thus, the system was shifted to microcentrifuge tubes to disrupt binding to any surface, and it also allowed incubation at 37^{o}C on a thermomixer to maintain a constant homogenous cell suspension in the medium.

As seen in the results, we always observe a baseline level of unspecific doublet formation in all our experiments. This can be explained through two ways- ionic interactions and excessive cell death. Cadherins are transmembrane proteins that mediate cell-cell interactions. The other common cause for cell-clumping is excess DNA and cell debris present in the medium. Due to the sticky nature of this nucleic acid, it causes cells to aggregate and form clumps. Polyethyleneimine is a non-viral cationic polymer commonly used in laboratories for chemical transfections. However,

its positive charges are known to induce necrotic cell death and apoptosis (Yuan & Li, 2017), a phenomenon we observed in our experiments as well. While the addition of DNase I helps alleviate some of the burden caused by excess DNA released by dead cells, it is possible that the load of debris is still significant. This can be avoided by isolating live cells prior to the incubation, or by the usage of stable expressing cell lines.

We hypothesized that addition of EDTA would have a significant impact on the dissociation of non-specific cell doublets. EDTA is commonly used in cell cultures as a chelating agent that binds and sequesters calcium ions, thus preventing interactions between cadherins on the cell which cause the clumping of cells (Lai et al., 2022). EDTA is generally used in a range of 0.5mM to 5mM in cell cultures, where the lower concentration is usually used in suspension cultures and higher concentrations are used for adherent or sticky cell lines. For our purposes we added 3mM EDTA and tested its effect in minimizing non-specific cellular interactions. Contrary to our hypothesis, the addition of EDTA did not have a significant effect on the proportion of cell-doublets. It is possible that we could observe significant effects at higher concentrations of EDTA in the solution, but it is known that EDTA is toxic to in-vitro cultures when present at a higher concentration (Ballal et al., 2009), hence we did not explore this further for the time being. Supplementing a detergent, such as Triton-X-100 or Tween-20, in the incubation media could be a viable strategy to disrupt weak, non-specific interactions among cells. They are widely used at lower concentrations in wash buffers in various methods ranging from immunostaining to phage display.

We observed vortexing to be effective in disrupting unspecific cell-cell interactions to increase the fold-change. In the experiments involving equal ratio of GFP and mCherry (1:1), we observed that vortexing helped increase the overall fold change by close to ~35%, whereas at lower ratios of target antigen (1:10::GFP:mCherry) this effect was even more pronounced with the fold change being 2.65 times higher. It indicates that high-affinity proteins interactions are able to withstand fluid-flow stress as opposed to unspecific interactions.

We quantitatively analysed enrichment of cellular complexes where display cells specific to the protein of interest were present in lower ratios when compared to display cells with an irrelevant protein (1:10 and 1:100 of GFP:mCherry respectively). We observed a greater than 3-fold change when present at 1:10 ratio. Compared to the nanobodynull control it was 30 times higher. For 1:100 ratio, we observe a 24

times higher fold change. It may appear that a less than 1fold-change resulted in no enrichment. However, the current screen only consists of a homogenous positive population of binding proteins. In a real screen consisting of a heterogenous mixture of proteins displayed on the cell surface, each protein-expressing cell exists as a rare population with respect to the rest of the population of cells.

Finally, we observed that antigen specific formation of cell doublets was heavily reduced when cells were incubated at higher speeds. We hypothesize that higher speeds provide an additional external stress in forming the chosen antigen-antibody interaction, while unspecific ionic interactions remain largely unaffected as these are random, not requiring only certain cells to meet and interact, and can occur even during acquisition phase on the flow cytometer. This lowers the overall fold-change.

Antigen-antibody complex formation is known to be affected by various factors, one of which is the affinity between the two. It is known that high affinity reactions Ab/Ag reactions are much more stable when compared to ones with low-affinity (Zehn et al., 2009). Our model system is based on high-affinity interactions using an antibody designed specifically for the protein of interest. Thus, it remains to be tested in the future whether low-affinity reactions on the cell-surface can be observed at similar levels.

Currently, in the field of antibody discovery the methods focus on the screening of an antigen against a library of antibodies, followed by multiple rounds of selection to identify the binding partner with desirable properties (binding affinity or functional effect). While this is beneficial to characterize affinities against a single antigen, limiting to one antigen lowers the final throughput considerably. Yeast biopanning is one such technology that is widely used to screen an antibody library displayed on yeast surface against a target antigen on the surface of mammalian cells. Recent studies on biopanning, such as the one conducted by Yang et al., describe a cell-cell interaction format for selection of a known high-affinity antibody (PcTx1) against a membrane protein (hASIC1a). PcTx1-displaying yeast cells were screened against hASIC1a-overexpressing CHO-K1 cells, and were tested against control yeast cells without PcTx1 expression (Yang et al., 2019). The authors noted a 24-fold increase in the proportion of double-positive clones of interest in the mix with PcTx1^{+} yeast, compared to the control yeast population. We compared our results in enriching double positive clones of interest to these previous studies on yeast biopanning, and we showed similar (and possibly higher) levels of enrichment of GFP-GFPnanobody

double positive cell-cell pairs, indicating the efficiency of the system. We are aware that the comparison is only qualitative since variations in protein binding affinities can skew the results. For a true comparison we would need to enrich for the same set of antibody and antigen on both platforms. However, the results obtained in our study exhibit significant promise, suggesting the possibility of extensive utilization in the future. By leveraging this approach, it may be possible to effectively screen two libraries of display proteins against each other, thereby enabling the identification of novel binding partners for special applications that were not possible before, but are made possible through this technique. For example, a multiplexed approach focused towards screening two libraries could potentially facilitate the discovery of antibodies that selectively bind to only one member of a conserved protein family, thereby substantially augmenting throughput relative to currently established techniques. It is also possible to determine cross-species specificity of target antibodies in a single screen. This can substantially aid pharmacological research that focuses and targets murine and human counterparts of the same antigen in order to develop antibodies in-vivo.

Future Outlook

While an assay condition has been optimized to allow for interactors on the cell membrane to come together, a final readout pertaining to the information of the interacting pairs is currently missing. This can be achieved through the fusion-PCR as mentioned in the introduction. This will allow the fusion of the DNA sequence from two different cell libraries (antigen vs antibody) into one long nucleotide which can then be amplified using a nested PCR reaction. To achieve this, the current plasmid backbone needs to be modified to allow for unique primer binding sites. The upstream and downstream regions of one library needs to be different from the upstream and downstream regions of the second library. To achieve this, a site directed mutagenesis will be performed where the DNA sequence upstream and downstream will be altered to create two different versions of the vector, each one acting as a backbone for one library. The main reason behind this is so that only one set of primers for the fusion PCR can be used for any inserted gene in the vector backbone since they will not be specific to the inserted gene.
Currently the analysis is conducted using a flow cytometer, however going forward the analysis would be implemented on droplet microfluidics to isolate each binding events in a separate reaction vessel and conduct the overlap extension PCR to

retrieve the pairing information. The data presented above indicates that the implemented setup can be utilized for the binding of two known high affinity binders- namely EGFP and GFP-nanobody. While the assay has been tested on high affinity binding proteins, its efficacy on low and medium affinity binding pairs remains to be tested. To this effect, the assay system will be deployed and tested on a range of binding pairs with vastly varying binding affinities. This will allow us to validate the efficacy of the assay system over varying ranges of protein binding affinities, and to alter the conditions if and when necessary, accordingly.

References

Ballal, N. V., Kundabala, M., Bhat, S., Rao, N., & Rao, B. S. S. (2009). A comparative in vitro evaluation of cytotoxic effects of EDTA and maleic acid: Root canal irrigants. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, *108*(4), 633–638. https://doi.org/10.1016/j.tripleo.2009.05.039

Beerli, R. R., Bauer, M., Buser, R. B., Gwerder, M., Muntwiler, S., Maurer, P., Saudan, P., & Bachmann, M. F. (2008). Isolation of human monoclonal antibodies by mammalian cell display. *Proceedings of the National Academy of Sciences*, *105*(38), 14336–14341. https://doi.org/10.1073/pnas.0805942105

Boder, E. T., Raeeszadeh-Sarmazdeh, M., & Price, J. V. (2012). Engineering antibodies by yeast display. *Archives of Biochemistry and Biophysics*, *526*(2), 99–106. https://doi.org/10.1016/j.abb.2012.03.009

Boder, E. T., & Wittrup, K. D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nature Biotechnology*, *15*(6), 553–557. https://doi.org/10.1038/nbt0697-553

Bowers, P. M., Horlick, R. A., Kehry, M. R., Neben, T. Y., Tomlinson, G. L., Altobell, L., Zhang, X., Macomber, J. L., Krapf, I. P., Wu, B. F., McConnell, A. D., Chau, B., Berkebile, A. D., Hare, E., Verdino, P., & King, D. J. (2014). Mammalian cell display for the discovery and optimization of antibody therapeutics. *Methods*, *65*(1), 44–56. https://doi.org/10.1016/j.ymeth.2013.06.010

Cherf, G. M., & Cochran, J. R. (2015). Applications of Yeast Surface Display for Protein Engineering. *Methods in Molecular Biology (Clifton, N.J.)*, *1319*, 155–175. https://doi.org/10.1007/978-1-4939-2748-7_8

De Las Rivas, J., & Fontanillo, C. (2010). Protein–Protein Interactions Essentials: Key Concepts to Building and Analyzing Interactome Networks. *PLoS Computational Biology*, *6*(6), e1000807. https://doi.org/10.1371/journal.pcbi.1000807

Di Noia, J. M., & Neuberger, M. S. (2007). Molecular Mechanisms of Antibody Somatic Hypermutation. *Annual Review of Biochemistry*, *76*(1), 1–22. https://doi.org/10.1146/annurev.biochem.76.061705.090740

Farooq, Q. ul A., Shaukat, Z., Aiman, S., & Li, C.-H. (2021). Protein-protein interactions: Methods, databases, and applications in virus-host study. *World Journal of Virology*, *10*(6), 288–300. https://doi.org/10.5501/wjv.v10.i6.288

Hartley, O. (2002). The use of phage display in the study of receptors and their ligands. *Journal of Receptor and Signal Transduction Research*, *22*(1–4), 373–392. https://doi.org/10.1081/rrs-120014608

Hoogenboom, H. R., Lutgerink, J. T., Pelsers, M. M. A. L., Rousch, M. J. M. M., Coote, J., van Neer, N., de Bruïne, A., van Nieuwenhoven, F. A., Glatz, J. F. C., & Arends, J.-W. (1999). Selection-dominant and nonaccessible epitopes on cell-surface receptors revealed by cell-panning with a large phage antibody library. *European Journal of Biochemistry*, *260*(3), 774–784. https://doi.org/10.1046/j.1432-1327.1999.00214.x

Hutchings, C. J., Colussi, P., & Clark, T. G. (2019). Ion channels as therapeutic antibody targets. *MAbs*, *11*(2), 265–296. https://doi.org/10.1080/19420862.2018.1548232

Hutchings, C. J., Koglin, M., & Marshall, F. H. (2010). Therapeutic antibodies directed at G protein-coupled receptors. *MAbs*, *2*(6), 594–606. https://doi.org/10.4161/mabs.2.6.13420

Jaroszewicz, W., Morcinek-Orłowska, J., Pierzynowska, K., Gaffke, L., & Węgrzyn, G. (2022). Phage display and other peptide display technologies. *FEMS Microbiology Reviews*, *46*(2), fuab052. https://doi.org/10.1093/femsre/fuab052

Lai, T.-Y., Cao, J., Ou-Yang, P., Tsai, C.-Y., Lin, C.-W., Chen, C.-C., Tsai, M.-K., & Lee, C.-Y. (2022). Different methods of detaching adherent cells and their effects on the cell surface expression of Fas receptor and Fas ligand. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-09605-y

Lu, R.-M., Hwang, Y.-C., Liu, I.-J., Lee, C.-C., Tsai, H.-Z., Li, H.-J., & Wu, H.-C. (2020). Development of therapeutic antibodies for the treatment of diseases. *Journal of Biomedical Science*, *27*(1), 1. https://doi.org/10.1186/s12929-019-0592-z

Parthiban, K., Perera, R. L., Sattar, M., Huang, Y., Mayle, S., Masters, E., Griffiths, D., Surade, S., Leah, R., Dyson, M. R., & McCafferty, J. (2019). A comprehensive search of functional sequence space using large mammalian display libraries created by gene editing. *MAbs*, *11*(5), 884–898. https://doi.org/10.1080/19420862.2019.1618673

Phizicky, E. M., & Fields, S. (1995). Protein-protein interactions: Methods for detection and analysis. *Microbiological Reviews*, *59*(1), 94–123. https://doi.org/10.1128/mr.59.1.94-123.1995

Pleiner, T., Bates, M., Trakhanov, S., Lee, C.-T., Schliep, J. E., Chug, H., Böhning, M., Stark, H., Urlaub, H., & Görlich, D. (2015). Nanobodies: Site-specific labeling for super-resolution imaging, rapid epitope-mapping and native

protein complex isolation. *ELife*, *4*, e11349. https://doi.org/10.7554/eLife.11349

Schwimmer, L. J., Huang, B., Giang, H., Cotter, R. L., Chemla-Vogel, D. S., Dy, F. V., Tam, E. M., Zhang, F., Toy, P., Bohmann, D. J., Watson, S. R., Beaber, J. W., Reddy, N., Kuan, H.-F., Bedinger, D. H., & Rondon, I. J. (2013). Discovery of diverse and functional antibodies from large human repertoire antibody libraries. *Journal of Immunological Methods*, *391*(1), 60–71. https://doi.org/10.1016/j.jim.2013.02.010

Sellmann, C., Pekar, L., Bauer, C., Ciesielski, E., Krah, S., Becker, S., Toleikis, L., Kügler, J., Frenzel, A., Valldorf, B., Hust, M., & Zielonka, S. (2020). A One-Step Process for the Construction of Phage Display scFv and VHH Libraries. *Molecular Biotechnology*, *62*(4), 228–239. https://doi.org/10.1007/s12033-020-00236-0

Sergeeva, A., Kolonin, M. G., Molldrem, J. J., Pasqualini, R., & Arap, W. (2006). Display technologies: Application for the discovery of drug and gene delivery agents. *Advanced Drug Delivery Reviews*, *58*(15), 1622–1654. https://doi.org/10.1016/j.addr.2006.09.018

Smith, G. P. (1985). Filamentous Fusion Phage: Novel Expression Vectors That Display Cloned Antigens on the Virion Surface. *Science*, *228*(4705), 1315–1317. https://doi.org/10.1126/science.4001944

Smith, G. P. (2019). Phage Display: Simple Evolution in a Petri Dish (Nobel Lecture). *Angewandte Chemie International Edition*, *58*(41), 14428–14437. https://doi.org/10.1002/anie.201908308

Smothers, J. F., Henikoff, S., & Carter, P. (2002). Tech.Sight. Phage display. Affinity selection from biological libraries. *Science (New York, N.Y.)*, *298*(5593), 621–622. https://doi.org/10.1126/science.298.5593.621

Tillotson, B. J., Cho, Y. K., & Shusta, E. V. (2013). Cells and cell lysates: A direct approach for engineering antibodies against membrane proteins using yeast surface display. *Methods*, *60*(1), 27–37. https://doi.org/10.1016/j.ymeth.2012.03.010

Valldorf, B., Hinz, S. C., Russo, G., Pekar, L., Mohr, L., Klemm, J., Doerner, A., Krah, S., Hust, M., & Zielonka, S. (2022). Antibody display technologies: Selecting the cream of the crop. *Biological Chemistry*, *403*(5–6), 455–477. https://doi.org/10.1515/hsz-2020-0377

Yang, Z., Wan, Y., Tao, P., Qiang, M., Dong, X., Lin, C.-W., Yang, G., Zheng, T., & Lerner, R. A. (2019). A cell–cell interaction format for selection of high-affinity antibodies to membrane proteins. *Proceedings of the National Academy of Sciences*, *116*(30), 14971–14978. https://doi.org/10.1073/pnas.1908571116

Yuan, W., & Li, H. (2017). Chapter 14—Polymer-based nanocarriers for therapeutic nucleic acids delivery. In E. Andronescu & A. M. Grumezescu (Eds.), *Nanostructures for Drug Delivery* (pp. 445–460). Elsevier. https://doi.org/10.1016/B978-0-323-46143-6.00014-2

Zehn, D., Lee, S. Y., & Bevan, M. J. (2009). Complete but curtailed T-cell response to very low-affinity antigen. *Nature*, *458*(7235), Article 7235. https://doi.org/10.1038/nature07657

Appendix

Primer Name	Primer Sequence
EGFP_FP	GATCTCCCGGGATCCGCGGCTGCAGGTCGACagtgagcaagggcgaggagctgttcaccggg
EGFP_RP	CAGATCCTCTTCTGAGATGAGTTTTTGTTCGTCGACcttgtacagctcgtccatgccgagagtgat
Nano_FP	GATCTCCCGGGATCCGCGGCTGCAGGTCGACagtacagctggttgaaagcggcggtgcgctg
Nano_RP	CAGATCCTCTTCTGAGATGAGTTTTTGTTCGTCGACgctgctaactgtaacctgtgtgccctggcc
mCherry_FP	GATCTCCCGGGATCCGCGGCTGCAGGTCGACagtgagcaagggcgaggaggataacatggcc
mCherry_RP	CAGATCCTCTTCTGAGATGAGTTTTTGTTCGTCGACcttgtacagctcgtccatgccgccggtggag

Table S1- Primer sequences *used for eGFP, GFP-nanobody, and mCherry for gibson assembly into the pDisplay vector*

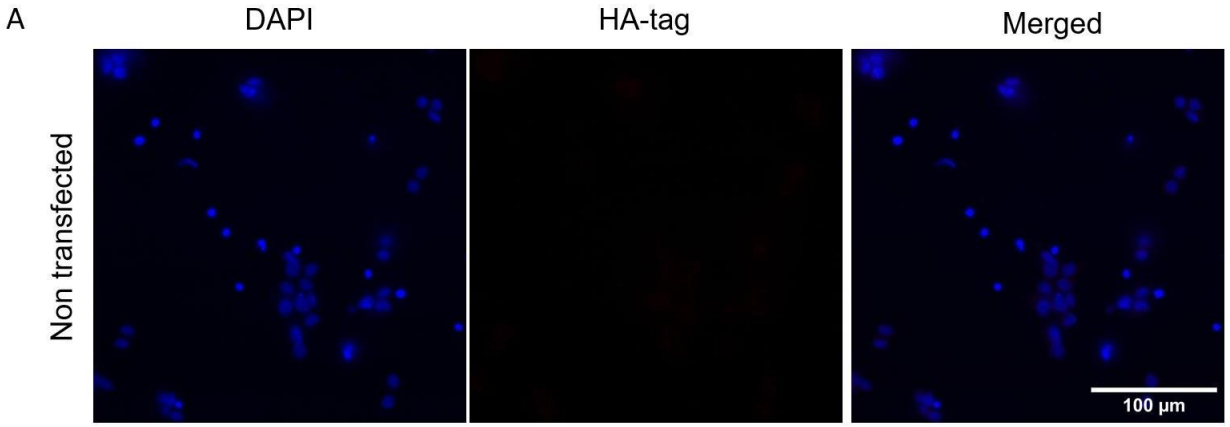


*Figure S1: **HA Tag staining control**. Representative mid-optical sections from confocal microscopy of HEK293T cells immunostained for HA-tag for Non transfected. Blue-DAPI; Red-HA tag. Scale bars = 100uM*

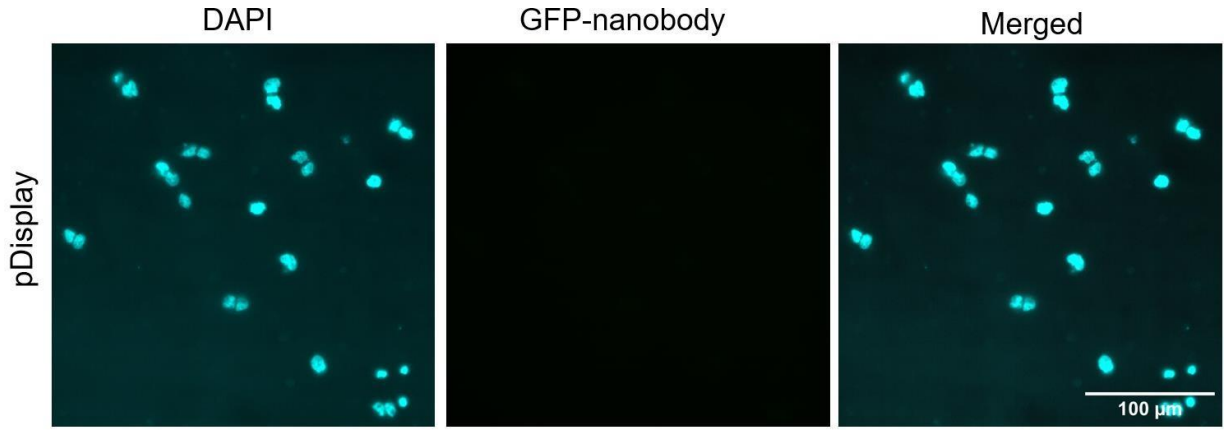


*Figure S2: **Soluble GFP staining control**. Representative mid-optical sections from confocal microscopy of pDisplay transfected HEK293T cells immunostained for GFP-nanobody expression. Scale Bar =100uM*

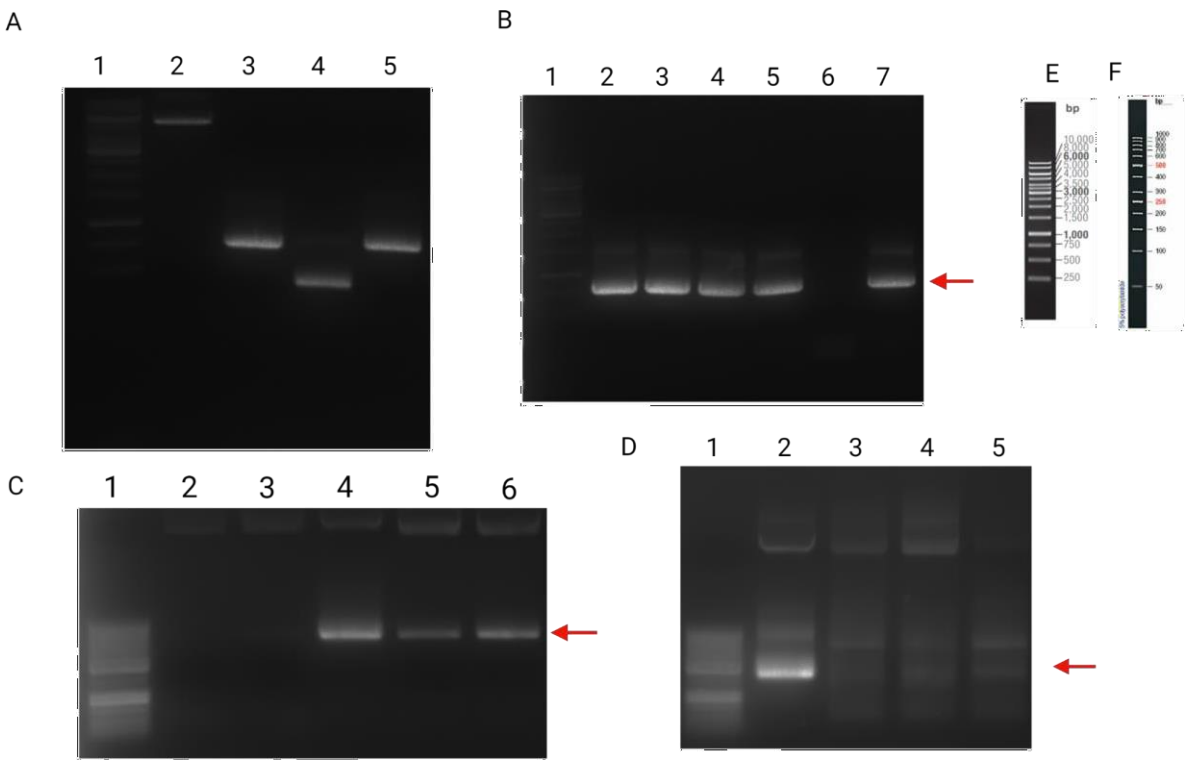
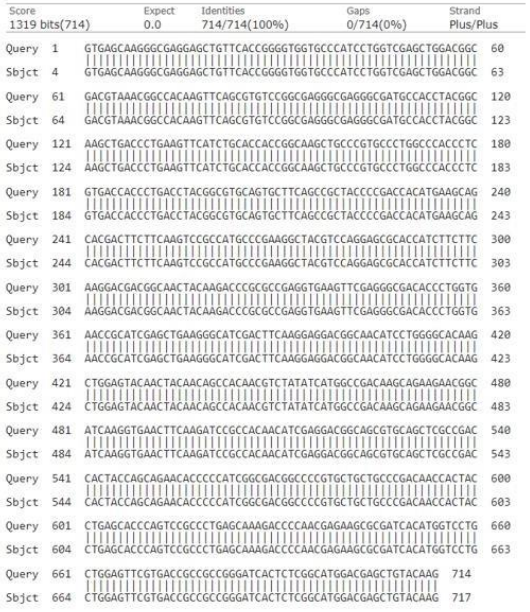


Figure S3: ***Quality control of construct cloning by Gel Electrophoresis.***
(A) Represents (from left to right)- 1- 1kb Ladder, 2-pDisplay vector linearized with SalI, 3-PCR product of eGFP, 4- PCR product of GFP-nanobody, 5-PCR product of mCherry. (B) PCR products of the inserted gene eGFP in pDisplay vector from DNA that was extracted from bacterial plate post transformation, to confirm the colonies that contained the desired plasmid insertion. 1- 1Kb ladder (2-7) Represent different colonies from which plasmid was extracted (C) PCR product from bacterial colonies for mCherry. 1-50bp Ladder, (2-6) Represent the different colonies. (D) PCR product from bacterial colonies for GFP-nanobody insertion. 1-50bp ladder, (2-5) Represent the different colonies. (E-F) Represent 1kb and 50bp DNA ladders respectively. Red arrows in all the images indicate expected band size for each of the inserted genes

The positive clones were sent for sequencing to confirm insertion of our genes of interest in-frame with the leader sequence and the PDGFR- transmembrane domain. The colonies that showed a positive band were sent for sequencing to confirm correct insertion of gene and below are the sequence alignment results to confirm error-proof cloning.

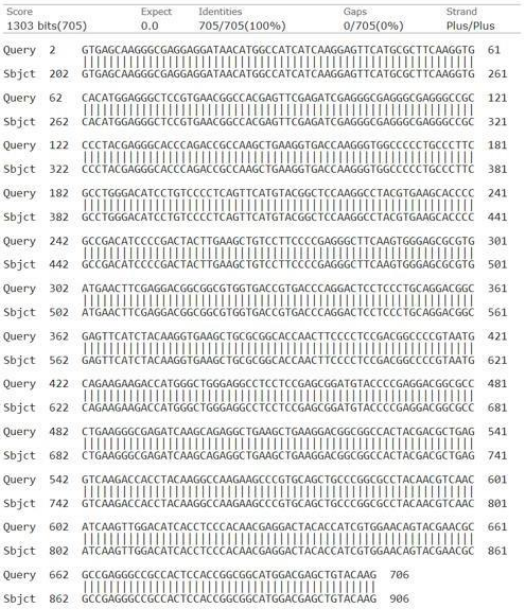
A



Score	Expect	Identities	Gaps	Strand
1319 bits(714)	0.0	714/714(100%)	0/714(0%)	Plus/Plus

```
Query  1    GTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4    GTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGC  63

Query  61   GACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  64   GACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGC  123

Query  121  AAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  124  AAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTC  183

Query  181  GTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAG  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  184  GTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAG  243

Query  241  CACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTC  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  244  CACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTC  303

Query  301  AAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  304  AAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTG  363

Query  361  AACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAG  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  364  AACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAG  423

Query  421  CTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGC  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  424  CTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGC  483

Query  481  ATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGAC  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  484  ATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGAC  543

Query  541  CACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTAC  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  544  CACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTAC  603

Query  601  CTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTG  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  604  CTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTG  663

Query  661  CTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG  714
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  664  CTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG  717
```

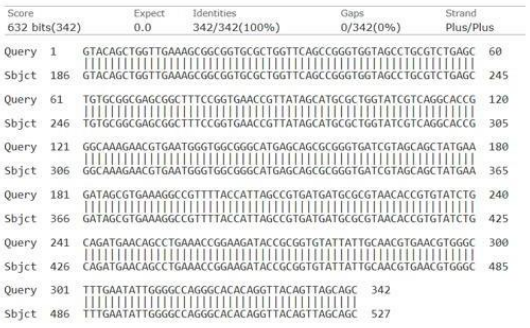
B



Score	Expect	Identities	Gaps	Strand
1303 bits(705)	0.0	705/705(100%)	0/705(0%)	Plus/Plus

```
Query  2    GTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTG  61
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  202  GTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTG  261

Query  62   CACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGC  121
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  262  CACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGC  321

Query  122  CCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTC  181
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  322  CCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTC  381

Query  182  GCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCC  241
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  382  GCCTGGGACATCCTGTCCCCTCAGTTCATGTACGGCTCCAAGGCCTACGTGAAGCACCCC  441

Query  242  GCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTG  301
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  442  GCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTG  501

Query  302  ATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGC  361
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  502  ATGAACTTCGAGGACGGCGGCGTGGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGC  561

Query  362  GAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATG  421
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  562  GAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCCGTAATG  621

Query  422  CAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCC  481
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  622  CAGAAGAAGACCATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCC  681

Query  482  CTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAG  541
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  682  CTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGAG  741

Query  542  GTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAAC  601
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  742  GTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAAC  801

Query  602  ATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGC  661
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  802  ATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAACGC  861

Query  662  GCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAG  706
            ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  862  GCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAG  906
```

C



Score	Expect	Identities	Gaps	Strand
632 bits(342)	0.0	342/342(100%)	0/342(0%)	Plus/Plus

```
Query  1    GTACAGCTGGTTGAAAGCGGCGGTGCGCTGGTTCAGCCGGGTGGTAGCCTGCGTCTGAGC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  186  GTACAGCTGGTTGAAAGCGGCGGTGCGCTGGTTCAGCCGGGTGGTAGCCTGCGTCTGAGC  245

Query  61   TGTGCGGCGAGCGGCTTTCCGGTGAACCGTTATAGCATGCGCTGGTATCGTCAGGCACCG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  246  TGTGCGGCGAGCGGCTTTCCGGTGAACCGTTATAGCATGCGCTGGTATCGTCAGGCACCG  305

Query  121  GGCAAAGAACGTGAATGGGTGGCGGGCATGAGCAGCGCGGGTGATCGTAGCAGCTATGAA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  306  GGCAAAGAACGTGAATGGGTGGCGGGCATGAGCAGCGCGGGTGATCGTAGCAGCTATGAA  365

Query  181  GATAGCGTGAAAGGCCGTTTTACCATTAGCCGTGATGATGCGCGTAACACCGTGTATCTG  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  366  GATAGCGTGAAAGGCCGTTTTACCATTAGCCGTGATGATGCGCGTAACACCGTGTATCTG  425

Query  241  CAGATGAACAGCCTGAAACCGGAAGATACCGCGGTGTATTATTGCAACGTGAACGTGGGC  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  426  CAGATGAACAGCCTGAAACCGGAAGATACCGCGGTGTATTATTGCAACGTGAACGTGGGC  485

Query  301  TTTGAATATTGGGGCCAGGGCACACAGGTTACAGTTAGCAGC  342
            |||||||||||||||||||||||||||||||||||||||||
Sbjct  486  TTTGAATATTGGGGCCAGGGCACACAGGTTACAGTTAGCAGC  527
```

*Figure S4: **Sequence alignment** results to confirm correct insertion of the gene (A) GFP (B) mCherry (C) GFPnanobody*