KTH ROYAL INSTITUTE
OF TECHNOLOGY

Degree Project in Biotechnology

Second Cycle, 30 credits

Selection and Characterisation of Affibody Molecules Intended for Drug Conjugates Targeting Cancer Cells

ELIN HEDBERG

Stockholm, Sweden 2022

Selection and Characterisation of Affibody Molecules Intended for Drug Conjugates Targeting Cancer Cells

Degree Project in Biotechnology, Second Cycle

Elin Hedberg

Supervisor:	Torbjörn Gräslund
Co-supervisor:	Javad Garousi
Examiner:	Yves Hsieh

KTH Royal Institute of Technology, AlbaNova
School of Biotechnology, Department of Protein Science
2022-06-13

Abstract

Affibody molecules are small affinity proteins (6.5 kDa) suggested to substitute monoclonal antibodies in therapeutic applications, e.g., antibody-drug conjugates (ADCs) targeting biomarker proteins expressed on cancer cells. An affibody-drug conjugate (AffiDC) could be used to target these types of overexpressed proteins on cancer cells while offering attractive properties, such as rapid transportation and distribution in the body, as well as efficient tumour penetration. These AffiDCs could be used as a targeted cancer therapy for cancers that are yet to be treatable and curable, like urothelial cancers.

This study suggested the use of ABD-fused affibodies to target a novel cancer protein that has been shown to be overexpressed on cancer cells, including breast, pancreatic and urothelial cancer. Affibody candidates toward this novel target were selected from a recombinant library, of 1×10^{11} variants, that is expressed using *E. coli* cell display system. The final candidates were subsequently biochemically characterized and assessed for affinity for the target. Three affibodies were finally identified and assessed in *in vitro* studies on mammalian cancer cells, revealing two affibodies that appear to bind to the cell lines BT-474 and MCF-7 with K_D ranging 10 to 100 nM.

Keywords: Affibody, AffiDC, affibody-drug conjugate, urothelial cancer, cell display, affibody library.

Sammanfattning

Affibodymolekyler är små affinitetsproteiner (6.5 kDa) som föreslås kunna ersätta monoklonala antikroppar i terapeutiska tillämpningsområden, exempelvis i antikropp-läkemedelskonjugat (ADC:er) som kan navigera sig fram till biomarkörer som är uttryckta på cancerceller. Affibody-läkemedelskonjugat (AffiDC) kan användas för att målsöka just sådana överuttryckta proteiner, samtidigt som de erbjuder goda egenskaper, såsom snabb transportering och spridning i kroppen, och effektiv penetrering genom tumörer. Dessa AffiDC:er skulle kunna användas inom riktad cancerterapi for de cancersjukdomar som fortfarande är i behov av cancerhämmande behandlingar, såsom urotelial cancer.

Den här studien föreslog tillämpning av ABD-kopplade affibodymolekyler för att målsöka ett nytt målprotein som har visats vara överuttryckt i flera olika cancersjukdomar, exempelvis bröst-, pankreas- och urotelial cancer. Affibodykandidater mot målproteinet har valts ur ett rekombinant bibliotek med 1×10^{11} transformanter som uttrycks med hjälp av en så kallad metod med *E. coli* celldisplay där affibodymolekylen visas på cellens ytmembran. De slutliga kandidaterna var sedan identifierade och biokemiskt karaktäriserade i *in vitro*-studie på människocancerceller, som visade att två av kandidaterna verkade binda till cancercellinjerna BT-474 och MCF-7 med K_D omkring 10 till 100 nM.

Nyckelord: Affibody, AffiDC, affibody-läkemedelskonjugat, urotelial cancer, cellvisning, affibodybibliotek, rekombinant bibliotek.

Table of Contents

Abbreviations

ABD	Albumin Binding Domain
ADC	Antibody-Drug Conjugate
AffiDC	Affibody-Drug Conjugate
BIA	Biospecific Interaction Analysis
BLI	Bio-Layer Interferometry
DAR	Drug to Antibody Ratio
FFC	Fluorescence Flow Cytometry
FL	Fluorescence (intensity)
HER2	Human Epidermal Growth Factor Receptor 2
HSA	Human Serum Albumin
IgG	Immunoglobulin G
IMAC	Immobilized Metal-ion Affinity Chromatography
LC-MS	Liquid Chromatography-Mass Spectrometry
mAb	Monoclonal Antibody
MACS	Magnetic-Activated Cell Sorting
MoA	Mechanism of Action
SAPE	Streptavidin-R-phycoerythrin conjugate
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
SPR	Surface Plasmon Resonance
Z	Affibody

1 Introduction

1.1 Targeted cancer therapies

By 2020, cancer diseases caused nearly 10 million deaths worldwide, placing cancer as the second leading cause of death globally. Cancer is a general term for several diseases sharing features such as uncontrolled cell growth and sometimes spread to other tissues in the body (metastasis) which is the primary cause of death in cancers [1,2]. The accumulation of causative mutations results in unique patterns of behaviour, differentiation, and protein expression, which can be exploited when developing novel targeted treatments. Proteins with abnormally elevated expression relating to disease are often suitable biomarker candidates which are used in diagnostics, and may serve as targets for therapeutic agents used in medicine [3]. Targeting cancer-specific proteins is a common strategy in precision medicine, and requires that the expression levels of the target are particularly distinguished between healthy and cancerous cells [4]. Although novel treatments and new candidate targets are constantly in development, many cancer diseases remain very complex to treat and cure, e.g., pancreatic cancer or when the cancer has metastasized [5]. As the conventional treatments, i.e., radical surgery, radiation treatment and chemotherapy, are reaching a therapeutic plateau there is an urgency for novel or complementary treatments [3,6]. Targeted therapies are rising in use with several cancer preventative treatments approved by the FDA and the EMA on the market, like the monoclonal antibody (mAb) trastuzumab (Herceptin®) as a drug targeting human epidermal growth factor 2 (HER2), an upregulated protein in breast cancer. Immunotherapies like this with mAbs are commonly used in targeted therapies thanks to their abilities to navigate and bind to different molecular targets like HER2 [7].

1.1.1 Antibody-drug conjugates

Antibody-drug conjugate (ADC) is an emerging technique that chemically links a toxin to an antibody in order to target a specific protein. The antibody will navigate and bind to the target, while also acting out its immunological role by activating the immune system to kill the cancerous cells. At the same time, the toxin will be delivered to poison the cells, allowing for a double mechanism of action (MoA) to kill the cells [8]. ADCs have been showing promising results with higher tumour specificity and potency in contrast to similar non-targeted cytotoxic drugs, and as of today there are multiple approved ADC drugs available on the market, e.g., Brentuximab vedotin targeting CD30 in Hodgkin's lymphoma and Ado-trastuzumab emtansine targeting HER2 in breast cancer [9].

Urothelial cancer, also called transitional cell carcinoma, is the 10th most common cancer worldwide and is a general term for cancers arising from the urothelial cells lining in the epithelium of the urinary system, like bladder and urethra [10,11]. Although approx. 430,000 patients are diagnosed every year, urothelial cancers have been poorly understood up until the past decade, and causes 170,000 deaths per year worldwide. Urothelial cancers are typically characterized by high expression of numerous surface

proteins, which offers a large range of candidate targets like HER2 and fibroblast growth factor receptors [12,13]. As of today, there are several ADCs available to treat cancers including urothelial cancers, e.g., Sacituzumab govitecan (Trodelvy®) and Enfortumab vedotin (Padcev®). Even so, there are still ongoing clinical trials assessing novel targets and constructs for urothelial cancer [13].

Nonetheless, there are several challenges to therapeutic application of ADCs, one being the risk of cross-reactivity caused by the immunological properties of the antibody [14]. This makes the selection of antibody to conjugate the toxin to crucial. In addition, to reduce the risk of eliciting an immunogenic response, the antibody must be humanized which poses additional requirements, like grafting of the complementarity-determining region [15-17]. The mAb's large size further restricts the application of ADCs to extracellular targets as penetration tissues, which then challenges the delivery of cytotoxic payload to the cells as the ADC must penetrate the tumour mass [14,16]. To facilitate internal delivery of the payload, the ADC can be designed to trigger internalisation which is made possible by the choice of target protein. Still, the degree of successful internalisation may vary, although it may not be required for all ADCs [14]. The drug to antibody ratio (DAR) is another crucial parameter that has been shown to reduce the cytotoxic effect outside the range of 2-4. Theoretically, a higher DAR increases the probability of sufficient toxin delivery but risks a higher frequency of adverse effects which is why DAR traditionally has been maintained between two and four [18]. Although site-specific conjugation of the toxin has allowed for DAR up to 8 in the last decade, the precise control of DAR remains a challenge in practice [16,18,19].

1.1.2 Affibody-drug conjugates

Affibody molecules are engineered scaffold proteins (6.5 kDa) that are much smaller than immunoglobulin G (IgG) molecules (approx. 150 kDa), which are the most frequently used antibodies in ADCs [9,16]. These affibodies consists of mainly one domain of a three-helix bundle with a sequence of 58 amino acids, originating from one of the domains in staphylococcal protein A. Two of these helices make up the binding cleft on targets. By randomizing 13 of the amino acids located in these two helices a library is created from which binders with specificity and affinity towards desired targets can be generated, thus allowing for generation of affibody molecules with novel specificities. This is how recombinantly combinatorial affibody libraries are produced, generating e.g., 1×10^{11} different affibody sequences. These libraries can then be used to select and isolate affibody molecules with good specificity and affinity for different targets, which has been done for more than 40 different disease targets, e.g., HER2, HER3, and TNF-α [20]. Therapeutic applications of affibodies bring many advantages over mAbs, mainly as a result of the much smaller size. As the protein is smaller in space, it is capable of tumour penetration more efficiently than larger proteins, possibly also presenting opportunities of new drug administration routes, i.e., oral, pulmonary, and ocular routes [16,20]. Other aspects include reduced cost as the proteins can be synthesized using bacterial expression, in contrast to antibodies [20].

Substituting antibodies with affibodies in ADCs is an emerging technique to target cancer cells, with promising studies assessing the cytotoxic effect of candidate drugs targeting HER2 [16,21,22]. Instead of using large antibodies, small affibodies are linked to cytotoxic payloads in an affibody-drug conjugate (AffiDC). By introducing cysteines in the affibody sequence, site-specific drug conjugation is possible and allows good control of the DAR [16].

1.2 Outline of Study

The outline of the study presents the main ideas and strategies behind the methods and materials of choice. Due to confidentiality issues, details of the target protein or other components of AffiDC construct are omitted in this report.

1.2.1 Novel cancer target

This study aims to target a novel therapeutic target or biomarker candidate that is expressed on the surface of cancer cells. Several studies have shown that multiple cancers, such as breast, lung, pancreatic and urothelial cancers, show remarkably high expression of this protein. The protein is normally exclusively present in certain tissues that disappear by adulthood, making the protein a suitable candidate target and biomarker in adults. In cancer studies, the protein has also been proposed to contribute to several cancer progression processes, e.g., metastasis, angiogenesis, and tumour relapse.

1.2.2 Affibody construct

To target the proposed target protein, affibody proteins are selected from an affibody library of 10^{11} randomly generated sequences. The affibody sequence (Z) is in fused to a histidine-tag (His_6) and an albumin binding domain (ABD) in a construct (fig 1). The His-tag consists of at least six histidine residues and can bind to divalent ion molecules, like Ni^{2+}, and enables protein capture without affecting the protein's native folding. By including a His-tag in the gene construct, protein purification by immobilised metal ion-affinity chromatography (IMAC) is possible [23], although IMAC was never used in the study.

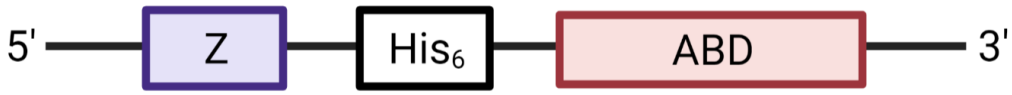


Figure 1. Gene construct of affibody, His-tag and ABD.

The ABD region acts out two main functions: binding to human serum albumin (HSA) used for protein purification and prolonging half-life in the body. HSA is the most abundant protein in the blood plasma and is known for its long circulation time. This is resulting mainly from its molecular size of 65 kDa, which prevents rapid elimination by kidney filtration. Fusing an ABD to affibodies or other drug constructs has been shown to extend the half-life in the blood plasma as they bind to HSA [16,24,25]. This also allows for protein purification using HSA as a ligand in column chromatography, similar to the His-tag [25].

1.2.3 Methods of selection and screening

The affibody construct was analysed using bacterial cell display in *Escherichia coli* (*E. coli*), allowing for effective expression on the surface of the cell membrane. The combinatorial affibody library is expressed from recombinant pBAD-vectors on the surface of *E. coli* TOP10 cells. The pBAD-vector contains the L-arabinose operon, araBAD, that upon fusion to a gene can act as the sole promoter when binding to L-arabinose, thus inducing protein expression. It also contains components of AIDA transport system to facilitate transportation to the surface (fig 2) [26,27].

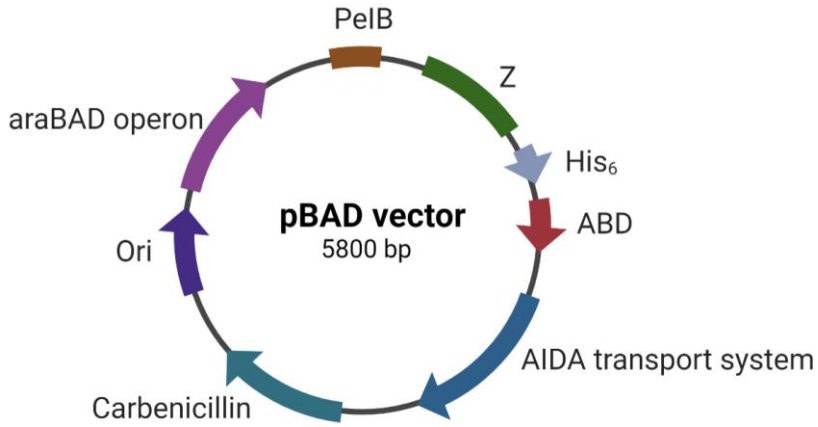


Figure 2. Vector map of pBAD vector and sequences of importance.

Magnetic-activated cell sorting (MACS) was performed for three rounds prior to the study to enrich positive binders in an induced culture of the affibody library. MACS is considered as a useful tool as it offers both negative selection, and protein binding selection (positive), while being quick, cost-effective and can manage large recombinant libraries [26,28]. In short, MACS uses magnetism to capture the particle of interest. In this study, magnetic beads with streptavidin tags are used to bind to the biotinylated target. When the target-bound beads are mixed with affibody-expressing bacterial culture, the target binds to the affibody and these variants can be isolated using a magnet.

In order to screen for stronger binders and candidates for an AffiDC, fluorescence flow cytometry (FFC) was performed using dyes to visualize the binding to target protein. To examine the binding of affibodies to the cell, HSA-binding dye was used. To examine binding of biotinylated target protein, a streptavidin-tagged dye was used.

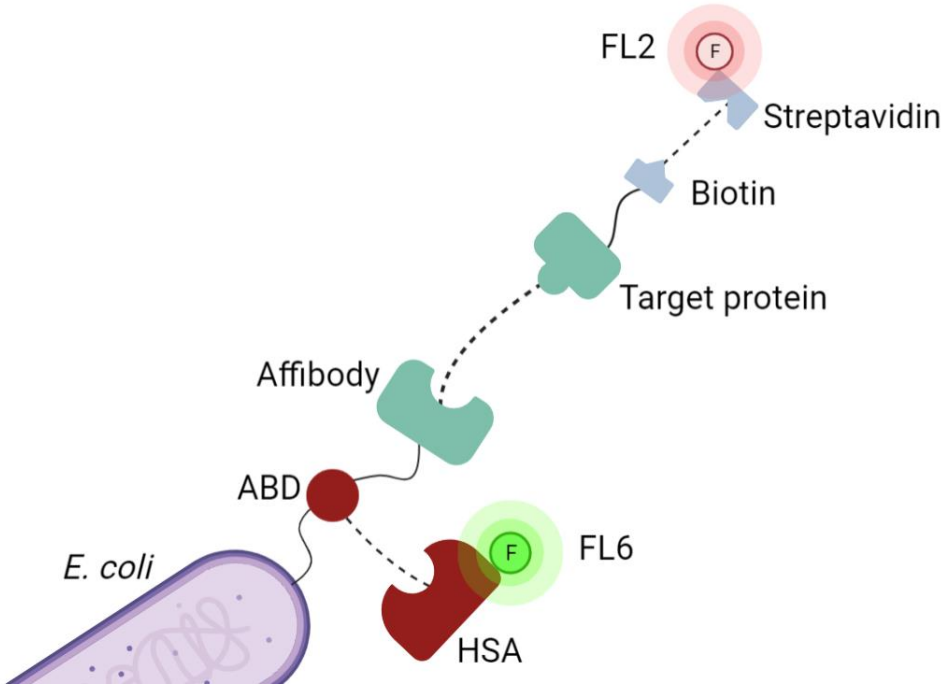


Figure 3. Illustration of the usage of fluorophores to analyse two types of binding; (i) the ABD-fused affibody binds to fluorophore-tagged HSA and gives signal FL6, (ii) the biotinylated target binds to fluorophore-tagged streptavidin and gives signal FL2.

When analysing the binding to soluble, biotinylated target protein, both FL2 and FL6 are used. FL2 gives an indication whether there is affinity toward to target, or the number of bound protein. Though, it is assumed that the affibodies are monovalent. When analysing the binding of soluble affibodies to target expressed on cancer cells, solely FL6 is relevant. If the ABD-fused affibody binds to cancer cells, there will be an increase in FL6.

1.2.4 Methods of characterization

The pBAD-vector is a recombinant plasmid that is mainly intended for surface display of proteins and plasmid extraction. Thus, to produce soluble affibody protein, the gene construct is subcloned into a new host vector pET-26b(+) which is designed to express the protein in the periplasmic space as the periplasm-leading pelB is included in the sequence [29,30]. Finally, competent *E. coli* BL21(DE3) cells are transformed to carry the pET-26b(+) vector. This strain supports protein production in multiple manners, including reduction of the degradation of unfolded protein as a result of deficiency in the Lon protease [31].

Biochemical characterization of the affibody construct involved conventional methods, like SDS-PAGE, DNA-sequencing, and LC-MS, to identify the affibody sequence and determine its molecular size. Affinity analysis was performed to evaluate the constructs' binding to soluble target and target expressed on cancer cells, respectively. This involved methods like FFC, biosensor analysis and bio-layer interferometry (BLI).

For FFC, four cancer cell lines that express the target protein were cultivated to *in vitro* analyse the binding of affibodies to target. For low expression, the cell line A431 was selected. For higher expression, BT-474, MCF-7 and SK-BR-3 were selected, the latter being the highest expressing according to other studies.

The biosensor analysis is based on surface plasmon resonance that quantifies the binding and release of an analyte molecule toward the target that is immobilized on a flow channel chip, and the kinetics are measured as the difference in reflected light that is shifted upon interaction with the target. BLI is similar to SPR as the kinetics are measured from the interference of reflected light, and uses immobilized target on biosensors that are dipped in wells with the analyte [32]. While SPR is known for rendering more consistent results with higher sensitivity, BLI allows for higher throughput and simpler management and preparation [32,33]. In this study, SPR and BLI will be performed as complements to each other.

2 Materials and Methods

2.1 Screening and Selection

The recombinant affibody library had been enriched by three rounds of positive and negative binding using MACS prior to the study. Output 3 had a concentration of 3.2×10^9 cells/ml. Fluorescent screening was performed in a total of three rounds: two for output 3, and one for the new output 4.

2.1.1 Cell cultivation and protein induction

Roughly 5 µl of enriched output cells cultivated in 100 mL of LB buffer and 100 µl of carbenicillin in 37 °C, 150 rpm, overnight. All cultivations with vector pBad2.2 used LB and carbenicillin with a volume ratio 1000:1. The culture was plated in a dilution series ranging 3.2×10^1 to 10^4 and 40 colonies were randomly picked for further cultivation in a volume of 5 ml in 37 °C, 150 rpm, overnight. A new cultivation of 5 ml with 30 µl of overnight culture was grown until it reached an absorbance of 0.5 using optical density at 600 nm (OD600) and was subsequently induced with 0.6% L-arabinose, followed by a final protein induced cultivation in 25 °C, 150 rpm, overnight.

2.1.2 Selection by MACS

MACS was performed to enrich the previous output 3 by off-target selection followed by positive target binding selection. A ratio of 1:50 beads to output cells was used for the selection, with Dynabeads™ MyOne™ Streptavidin C1 beads (ThermoFischer). 20 µl and 10 µl of 10^8 beads for negative, respectively positive control were washed twice in 200 µl 1xPBSP (PBS, 0.1% Pluronic acid) by magnetic separation. 1 µg of biotinylated target protein was added to the positive control and incubated in room temperature (RT) for 1 h, while no protein was added to the negative control. The coated beads were washed with 1 ml PBSP before addition of cells.

Approx. 6.5×10^9 of induced cells were washed twice with 1 ml PBSP during 10 min and 3 min, respectively, and were collected by centrifugation at 4000xg. For off-target selection, 1 ml of coated beads was added and let incubate on rotation 150 rpm, at RT for 30 min. The off-target binders were captured on a magnet for 1 minute so the flowthrough could be collected and re-applied to the magnet for further separation. For the positive selection, the flowthrough and the cells were then added to 1 ml of coated beads and incubated on rotation, 150 rpm, in RT for 2 h. The beads were then captured by a magnet for 2-3 min, the flowthrough was discarded, and the beads were resuspended in 1 ml of fresh PBSP. The selection was repeated 3 times. The new output was centrifuged at 4000xg, 4 °C, 10 min, and was resuspended in 10 ml of LB. A dilution series was prepared to obtain 1 ml samples with concentrations 6.5×10^2 to 10^5, out of which 100 µl was plated. The cultures and plates were grown in 37 °C, 150 rpm, overnight. 40 random colonies were picked and cultivated in a volume of 5 ml overnight and induced with 20% L-arabinose when OD600=0,8.

2.1.3 Screening by fluorescence flow cytometry

The binding of affibodies and target protein was assessed using flowcytometry with labelling of two different fluorophores. Induced culture of the 40 picked colonies were pooled into 4 sample pools by adding 100 µL of each culture. 200 µL of each pool was washed with 500 µL 1xPBSP at 2000xg, 4 °C, 1 min, and subsequently resuspended with 100 µl PBSP. The pools were incubated with 1 µg of biotinylated target protein on a rotamixer in RT for 30 min, followed by wash with PBSP. The pools were then incubated on ice with 0.2 (v/v)% of the fluorophores 5.4 mg/ml HSA-Alexa Fluor™ 647 (Invitrogen) and 1 mg/ml streptavidin R-phycoerythrin conjugate (SAPE; Invitrogen) respectively, in a total volume of 100 µL with PBSP for 20 min and kept dark. The pools were washed and resuspended with 200 µl of PBSP. The pools were analysed for 20.000 events in a Gallios™ flow cytometer (Beckman Coulter).

The analysis was repeated for all colonies in the pools which showed positive binding, using 20 µl of target protein instead. The final positive colonies were cultivated in 5 ml LB overnight, 150 rpm, 37 °C, to extract the plasmids.

2.1.4 Plasmid extraction

To extract the plasmids from cultures, QIAprep® Spin Miniprep Kit (QIAGEN) was used and followed according to QIAGEN Handbook guidelines. For each plasmid extraction, 1-5 ml of overnight culture was pelleted at 6800xg for 3 min. The pellet was resuspended in 250 µl Buffer P1, and 250 µl of Buffer P2 was added and mixed by inversion 4-6 times. The cell lysis was not allowed to proceed longer than 5 minutes. 350 µl of buffer N3 was added and mixed immediately by inversion, followed by centrifugation at 17,900xg for 10 minutes. The supernatant was isolated and added to a QIAprep® spin column and centrifuged at 17,900xg for 1 min. The column was washed by adding 750 µl of Buffer PE, followed by centrifugation twice at 17,900xg for 1 min. The DNA was eluted by adding 30 µl of sterilized water (StAq) 1 min before centrifugation at 17,900xg for 1 min. The final concentration was measured with NanoDrop® 1000 spectrophotometer (NanoDrop® Instruments). To identify the affibody sequence, 20 µl of plasmids, ranging 50-200 ng/µl, were sent for subsequent sanger sequencing (Tubeseq. service, Eurofins Genomics)

2.1.5 PCR

Conventional PCR was performed to confirm the size of affibody gene insert. The Phusion® High-Fidelity DNA Polymerase (#M0530, New England Biolabs® Inc.) was used and the reaction mix (table 1) was set up in reaction volumes of 50 µl according to NEB's guidelines. The Phusion® DNA Polymerase was added to the mix last.

Table 1. PCR reaction mixture using NEB Phusion® High-Fidelity DNA polymerase.

Component	Per Reaction (µl)	Final Concentration
5X Phusion HF Buffer	10	1X
10 mM dNTPs	1	200 µM
10 uM Forward Primer	2.5	0.5 µM
10 uM Reverse Primer	2.5	0.5 µM
Template DNA (plasmid)	1	< 250 ng
StAq	32.5	
Phusion® DNA Polymerase	0.5	1.0 units/50 µl PCR

The reaction mixture was prepared in a PCR tray and set for amplification in GeneAmp® PCR System 9700 (Applied Biosystems) with protocol in accordance with NEB's guidelines (table 2).

Table 2. PCR cycling conditions using NEB Phusion® High-Fidelity DNA Polymerase for pBad2.2.

Step	Temp.	Time	Cycles
Initial denaturation	98 °C	30 s	1
Denaturation	98 °C	10 s	35
Primer Annealing	50 °C	30 s	
Extension	72 °C	30 s	
Final extension	72 °C	10 min	1
Hold	4 °C	∞	1

The PCR amplicons were analysed on 1% agarose gels with 6X DNA Loading dye (Thermo Scientific), running at 180 V for 30 minutes. The gels were visualised in UV-light in a gel imaging system Gel Doc™ EZ Imager (Bio Rad).

For PCR purification QIAquick® PCR Purification kit (QIAGEN) was used to remove residuals. 240 µL Buffer PB per each PCR reaction was mixed and placed in a QIAquick column for centrifugation at 13000xg, for 1 min. The column was then washed with 750 µl Buffer PE and centrifuged twice. The amplicons were eluted with 30 µl StAq and analysed on 1% agarose gels under UV-light.

2.2 Subcloning

2.2.1 Competent cell transformation

The new vector plasmid pET-26b(+) was produced by transformation of KCM competent TOP10 *E. coli* cells. The transformation reaction mix (table 3) was heat shocked in 42 °C for 45 s. 200 µL of LB was added and incubated at 37 °C, 150 rpm, for 1 h. To confirm successful transformation, 100 µl of culture was spread of agar plates and grown in 37 °C overnight. The remaining cells were further cultivated in 15 ml LB in 37 °C, 150 rpm, overnight. All cultivations with pET-26b(+) were performed with 1:1000 of kanamycin.

Table 3. Reaction mix for heat-induced transformation using competent cells.

Component	Volume per reaction
DNA insert	2 µl (or 50 ng)
StAq	6 µl
5xKCM	2 µl
Incubate on ice, 5 min	
Thawed competent cells	10 µl
Incubate on ice, 20 min	

2.2.2 Restriction enzyme double digestion

To isolate the gene construct encoding the affibody molecules and insert it into the new plasmid, the purified amplicons were digested by restriction enzymes on sites Xho1 and Nde1. The reaction setup (table 4) was prepared for each of the variants and another one for the new pET-26b(+) plasmid, and the reaction tubes were put for incubation in 37°C for 1 h.

Table 4. Reaction components per reaction for enzymatic double digestion of plasmids.

Component	Z-construct, 50 µl	pET26b$^+$ plasmid, 40 µl
DNA	1 µg	0.8 µg
10X rCutsmart Buffer	5 µl	4 µl
Nde1	1 µl	0.8 µl
Xho1	1 µl	0.8 µl
StAq	To 50 µl	To 40 µl

For a few of the affibody variants and the pET-26b(+) plasmid, the digested mixture was further purified to prevent re-annealing of cut fragments. The cut fragments were run on 1% agarose gels and visualized

using UV-light. The correctly sized band was excised with a scalpel and purified using QIAquick® Gel Extraction Kit (QIAGEN). Assuming 100 mg gel equals approx. 100 µl, 3 volumes of Buffer QG were added to the gel volume, followed by incubation in 50 °C for 10 min. Upon complete gel dissolution, 1 gel volume of 2-propanol and the mix was centrifuged in a QIAquick column at 17000xg for 1 min. 500 µl of Buffer QG was added followed by centrifugation. The column was washed with 750 µl Buffer PE and was let to sit 5 min before centrifugation. The DNA was eluted using 30 µl of StAq and concentration was measured on NanoDrop® 1000 spectrophotometer.

2.2.3 Ligation and transformation

The cut gene constructs, and the digested vector pET-26b(+) were combined by ligation using the molar ratio 7:1 of insert to vector, based on calculations from NEBioCalculator® (New England, BioLabs® Inc.) The ligation reaction mix (table 5) was incubated in 16 °C overnight (or 16 h) and was heat inactivated in 65 °C for 10 min.

Table 5. Reaction mix for ligation of DNA insert and vector plasmid.

Component	**Amount per sample**
Protein gene insert	18.7 ng
Digested pET26b$^+$	48.1 ng
StAq	To 17 µl
Incubate in 50°C for 20 min and let cool	
10X T4 DNA Ligase Buffer	2 µl
T4 DNA Ligase (400U/µl)	1 µl

The ligated plasmid pET-26b(+)-Z-ABD was subsequently transformed to competent *E.coli* TOP10 cells (table 3) and plated on agar plates.

2.2.4 Ensuring affibody integrity

To proceed with affibody sequences with no novel mutations, colony-PCR was performed for 3 different colonies per variant. The master reaction mix (table 6) was prepared in a PCR-plate and one pipette tip of bacterial colony was added into each well. Here, DreamTaq DNA Polymerase (Thermo Scientific) was used. The plate was lightly shaken for a few second, let sit for 5 minutes and the tips were removed before the run.

Table 6. Colony-PCR reaction mixture using DreamTaq DNA polymerase.

Component	Amount	Final Concentration
10 mM dNTPs	5 µl	200 µM
10X DreamTaq Buffer	5 µl	1X
10 uM T7 Forward Primer	1 µl	0.2 µM
10 uM T7 Reverse Primer	1 µl	0.2 µM
DreamTaq DNA Polymerase	0.25 µl	1.25 units/50 µl PCR
StAq	To 50 µl	
Template (cells)	1 colony	

The cycling conditions were designed in accordance with the Thermo Scientific™ guidelines (table 2).

Table 7. PCR cycling conditions using DreamTaq DNA polymerase.

Step	Temp.	Time	Cycles
Initial denaturation	95 °C	3 min	1
Denaturation	95 °C	30 s	35
Primer Annealing	46.2 °C	30 s	
Extension	72 °C	1 min	
Final extension	72 °C	7 min	1
Hold	4 °C	∞	1

The PCR amplicons were run on 1% agarose gels, and *E.coli* clones with amplicons showing a correct size, were cultivated and their plasmids were extracted and subjected to DNA sequencing to confirm the integrity of the genes.

2.3 Protein Production and Purification

2.3.1 Protein production

The ligated plasmids pET-26b(+)-Z-ABD, harbouring the different affibody variants, were subsequently transformed into competent BL21(DE3) cells (table 3), heat shocked in 42 °C for 45 s. 200 µl Tryptic Soy Broth (TSB; Merck) was added and the reaction was incubated and cultivated overnight in TSB. 500 µl of overnight culture was used to inoculate 50 ml of TSB+yeast and induced with 100 µl of 0.5 M isopropyl-β-D-1-thiogalactopyranoside (IPTG) when OD_{600} = 0.8, followed by incubation at 25 °C, 150 rpm, overnight.

The cells were harvested by centrifugation at 4000xg, in 4 °C for 8 minutes and the pellets were then resuspended in 10 ml 1xTST(25 mM Tris, 1 mM EDTA, 0.2 M NaCl, 0.05 % Tween) each. To lysate the cells, the samples were sonicated for 3 min with a pulse of 1.0/1.0 s, at 30% power output using Vibra-Cell™ (Sonics®). To remove cell debris, the sample was centrifuged again in 16000xg, in 4 °C for 20 minutes followed by manual filtration using 0.45 μm filters on syringes.

2.3.2 Protein purification

To purify the proteins, only HSA chromatography columns were used to purify the proteins due to time limitations. For cleaning, 20 ml StAq, 20 ml 1xTST and 20 ml of acetic acid (HAc) pH 2.8 were added separately. 50 ml TST was applied to the column for equilibration before adding 10 ml of protein sample. The column was washed with 50 ml TST and 25 ml of 5 mM ammonium acetate (NH_4Ac), and 10 fractions of protein were be eluted using 1 ml of HAc added at a time. Regeneration of the column was made by adding 20 ml Hac, 20 ml StAq and 20 ml 20% ethanol in TST separately. The protein concentrations in the eluted fractions were measured by A_{280} and the fractions with the highest absorbance values were pooled, lyophilized, and resuspended in 1xPBS.

The purity and molecular size were analysed with SDS-PAGE, using 20 μl of the purified affibody proteins, 3x RED (3x SDS sample buffer, 3% 0.5 M TCEP) and PageRuler Plus Prestained Protein Ladder (Thermo Scientific). The cassette was filled with MES running buffer. The gel was run at 180 V for 40 min, washed with distilled water, stained with Coomassie Brilliant Blue. and brought to a boil. It was washed again and let to de-stain in StAq with gentle shaking, in RT, overnight. The bands in the gel were visualized with visible light.

2.4 Biochemical Characterisation

2.4.1 Purity and molecular weight

The molecular weight and the extinction coefficient at 280 nm were estimated from the amino acid sequence with NEBioCalculator®. Additional measurements of molecular weight were later done by liquid chromatography-mass spectrometry (LC-MS) using an ESI-QTOF mass spectrometer Impact II (Bruker Daltonics), kindly executed by colleagues.

2.4.2 Identification and phylogeny

The ancestral lineage of the variants was analysed by constructing a phylogenetic tree in Geneious Prime (Dotmatics), using the Tamura-Nei model and neighbouring-joining building model with 65% similarity.

2.4.3 Affinity analysis – binding to target

The buffer was exchanged from HAc to 1xPBS by desalting, and further purify, the protein samples with Illustra NAP-5 chromatography columns (Cytiva).

The binding to immobilized target was analysed using SPR instrument Biacore™ T200 (Cytiva) by first executing a trial run of each affibody variant followed by a longer run using a series of concentrations. The instrument was desorbed, primed, and prepared for run, according to the wizard software recommendations. Biotinylated target protein was immobilized to a chip and the protocol was prepared according to the wizard.

The trial run with 400 nM with each variant was performed to determine binding or non-binding variants and for a preliminary K_D-value determination. It was also used to determine the concentration series for the second run. Half of the candidates were further analysed after the trial run. Then, concentrations of 100 nM, 300 nM, 900 nM, and 2700 nM of protein were used, and all samples were filtered in 0.45 um before analysis. The response was set to 500 RU, and the programme was run with PBST (PBS, 0.1% Tween™; Merck), StAq, 10 mM hydrochloric acid (HCl) and 10 mM sodium acetate (NaAc) pH 4. Along with the affibody samples, the rabbit anti-target antibody from Human Protein Atlas (HPA) against the target protein was added as a positive control. The three affibody variants with best affinity were selected for further analysis with screening to cancer cells.

The affinity was also examined using BLI. A tray was prepared for the protocol; probing (200 µl PBS), immobilization (0.2 ug biotinylated protein in 200 µl PBS), washing (200 µl PBS), binding of sample (500 nM of affibody proteins, 50 mM of anti-target antibody, and an affibody control). The tray was run with biosensors on Octet® RED 96 system (fortéBIO) and the results were analysed to determine affinity.

2.4.4 Affinity analysis – binding to cancer cells

The cell lines, A-431, BT-474, MCF-7 and SK-BR-3 (ATCC®) were grown in RPMI medium 1690 (Gibco) with 10% FBS-HI and 1% PEST, in 37 °C incubation chamber until a high concentration was reached, e.g. 500 000 cells/ml. The cells were detached by aspirating the medium, adding 1 ml TryplE™ Express Enzyme (Thermo Scientific) and let incubate for 10-15 min, depending on the cells' morphology and tendency to adhere to each other. PBST was added to reach a volume of 10 ml, and approx. 500,000 cells/sample was obtained for each cell line. The cells were centrifuged at 4000xg for 6 min and washed with 500 µl of PBSP (PBS, 0.1% FBS-HI).

The three selected affibody variants and the anti-target antibody were tested for each of the cell lines. 300 nM affibody protein was stained with 0.2 (v/v)% 5.4 mg/ml HSA-Alexa Fluor™ 647 (Invitrogen) and the anti-target antibody was stained with 2 mg/ml Alexa Fluor™ goat anti-rabbit IgG (H+L) (Invitrogen). For each run, meaning each combination of cell line and sample), the cell pellets were resuspended in 100 µl stained sample and 400 µl PBS, and thereafter analysed for 20.000 events in Gallios™ flow cytometer (Beckman Coulter).

3 Results

3.1 Screening and selection

3.1.1 Selection by MACS

The recombinant *E. coli* library had been enriched in three rounds of MACS previously, and a fourth round was performed to select variants again. Titration of the MACS output 4 estimated a concentration of 3.2×10^7 cells/ml, a 10-fold enrichment in comparison to output 3 of 3.2×10^8 cells/ml. The cells were then dyed with fluorophore tagged SAPE (FL2) and HSA (FL6), and the fluorescence was visualized in a flow cytometer (fig.4). The results implies that the fraction of stronger FL2 intensity is larger for the latter output (fig.4b).

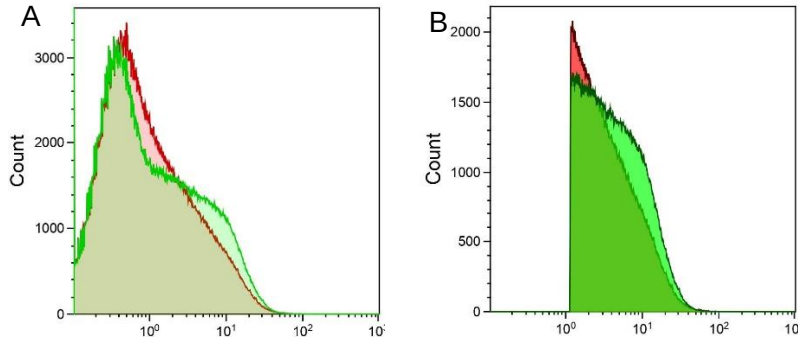


Figure 4. Comparison of MACS enrichment output 3 (red) and 4 (green) in FL2 for (a) ungated events and (b) gated events.

3.1.2 Selection by fluorescence flow cytometry

To assess the binding of particular variants rather than analysing a whole library, the outputs from MACS were used to screen 40 randomly picked variants using FFC. Two runs were made for output 3, and one for output 4. For each of the flow cytometry runs, 4 out of 40 variants were selected for further analysis resulting in a total of 12 affibodies, labelled variant 1-12 (fig.5). All variants show a shift in FL2 intensity, indicating that there is target protein bound to the variants expressed on cells.

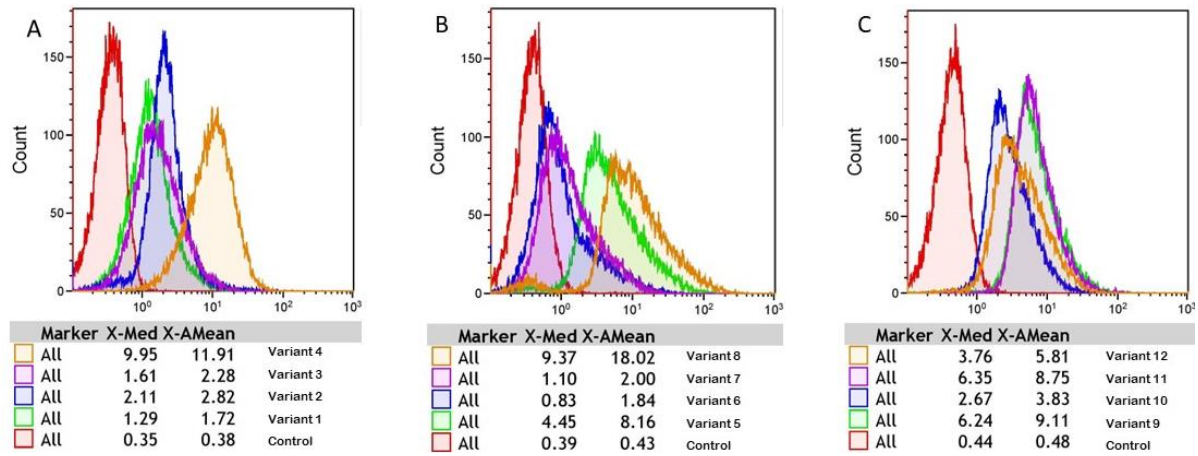


Figure 5 (a-c). Flow cytometry results with count plotted against FL2 intensity of four variants and one control picked from (A) MACS output 3, variants 1-4, (B) MACS4 output 4, variants 5-8, and (C) MACS output 3, variants 9-12.

3.2 Subcloning

Multiple attempts of sub-cloning of the affibody variants were required in order to successfully obtain correct plasmids. The first digested plasmids that were not purified caused some variants' ligation to be unsuccessful, and required thus additional steps of gel extraction to remove other fragments (fig.6).

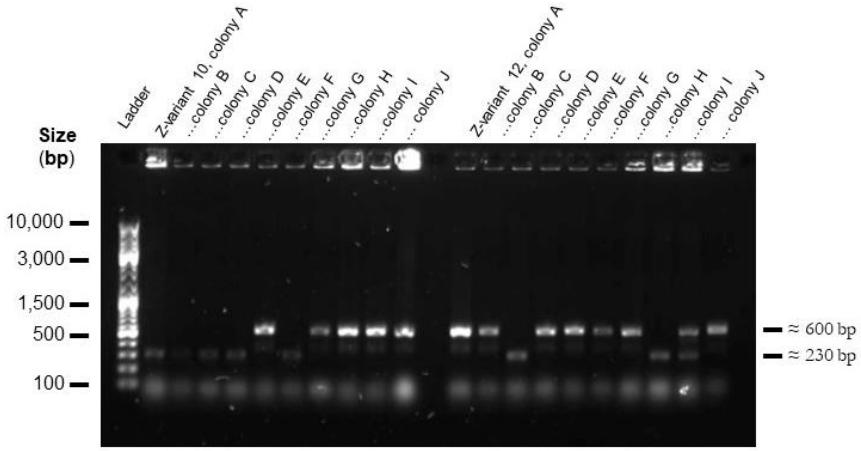


Figure 6. Colony-PCR results of 10 colonies (A-J) screened for variant 10 and 11 each, running with DreamTaq under UV-light. The larger bands correspond to full and successful ligation of the new host vector, anticipated to approx. 580 bp. Bands with size approx. 230 bp are expected to be empty vectors.

3.3 Protein production and purification

Purification and measurements of protein concentration showed that three of the variants 9, 11 and 12 yielded remarkably low protein production, even after repeated production attempts. Based on the estimated molecular weights (table 7), the highest obtained elution concentrations ranged between 0.70 to 1.70 mg/ml, while ranging around 0.03 to 0.07 mg/ml for the low expressing variants. The SDS-PAGE shows thick bands at approx. 14 kDa, and while there are additional bands on the gel, the proteins were deemed to be containing the protein(fig.7). Expression for all variants except colonies 9, 11 and 12 appears to be high, causing overloaded sample in the gel. In this case, further purification would increase the reliability of the other characterisation measurements. IMAC would have been used to purify the proteins further if time allowed.

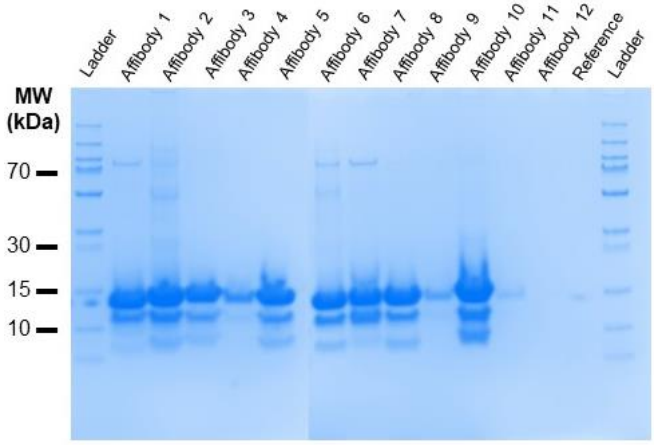


Figure 7. SDS-PAGE results visualized on gel under UV-light.

3.4 Biochemical analysis

3.4.1 Purity and molecular size

The estimated (NEBioCalculator®) and the measured (LC-MS) molecular weights were similar (table 7), although the purity of the proteins fluctuated as multiple compounds were detected in the LC-MS plots (appendix fig.A4). The results from the SDS-PAGE (fig.7), LC-MS and the estimation generated similar results of roughly 13.5 kDa for the whole ABD-fused affibody construct.

Table 7. Estimated extinction coefficient, molecular weight and measured molecular weight.

	Estimated		Measured
Variant	ε_{280nm} [$M^{-1}cm^{-1}$]	M [kDa]	M [kDa]
1	17,780	13.54	13.67*
2	15,220	13.45	13.45*
3	13,370	13.60	13.59
4	19,630	13.55	13.55
5	20,910	13.54	13.54
6	17,780	13.48	13.47*
7	17,780	13.36	13.36*
8	19,630	13.49	13.49*
9	19,630	13.59	13.48*
10	20,910	13.56	13.56
11	26,600	13.60	13.36*
12	22,219	13.55	13.36*

*two or more compounds detected in LC-MS

3.4.2 Identification and phylogeny

Identification and sequencing of the affibody genes shows that all appear to be healthy, and does not contain any stop codon. Instead of the 13 randomized amino acids, 15 positions turned out to be differing. The affibody sequences cannot be revealed due to confidentiality issues.

Analysing the genes' lineage in a phylogenetic tree shows that there are three early branching points and four main clusters, A-D (fig.8). An obvious observation is the fact that cluster A, with variant 4 and 9, is early distinguished from the other variants. Cluster C, with the variants 2, 5 and 7 shows an earlier branching point, although the variants are not very closely related according to the branch lengths.

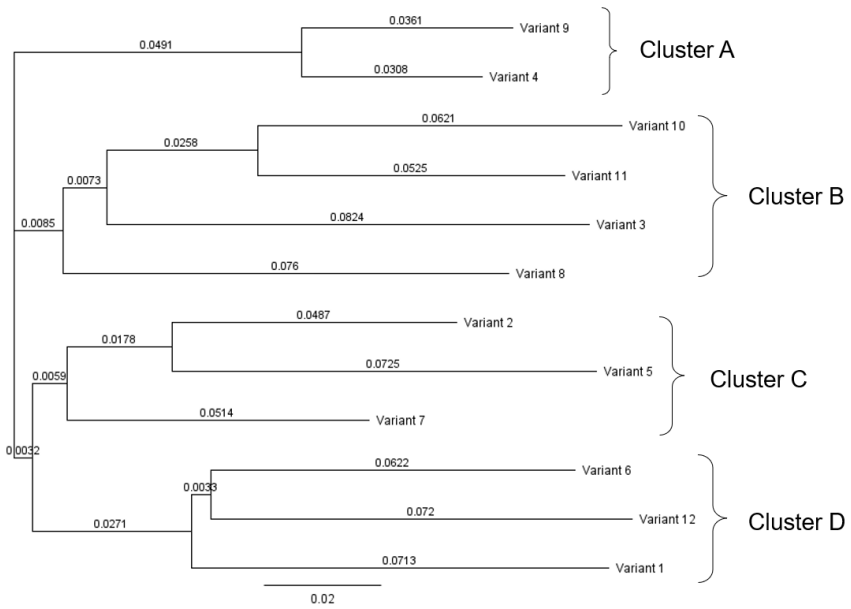


Figure 8. Phylogenetic tree of the genetic lineage descendance of the 12 affibody variants with four clusters (A-D).

3.4.3 Affinity analysis – binding to target

The trial experiment to analyse affinity revealed that all affibody proteins seemed to be able to bind to the target protein. at concentrations of 400 nM, resulting in bulky binding curve. The visually most promising candidates, variants 1, 2, 3, 4, 5 and 7, were selected to run in the second biosensor analysis. As the parameter fitting is poor due to the "curve bulkiness", the trial results are deemed non-reliable.

The second SPR run also show fully saturated curves for all concentrations (ranging from 100 to 2700 nM) and the parameter fitting remains poor (fig.9). The affibody appears to both bind and dissociate from the target quickly. The variants with strongest affinity, that is the lowest K_D, was selected for further analysis.

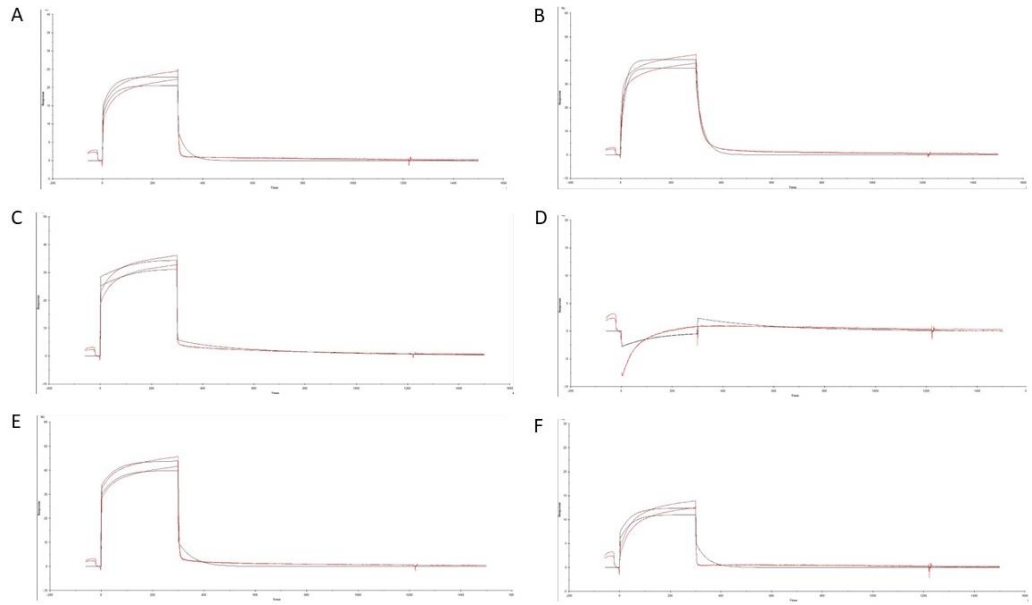


Figure 9 (a-f). Biosensor binding curves for the trial run of the variants; (a) 1, (b) 2, (b) 3, (d) 4, (e) 5 and (f) 7.

K_D for variants 2, 3 and 5 was measured to 12.6 nM, 26.5 nM, and 91.3 nM respectively, reaching response ranging up to 50 RU (fig.10) while other variants gave a weaker response (fig.10 b, c, e). Variant 1 gave similar response levels but shows very rapid dissociation from target (fig.10 a). Variant 4 shows the slowest dissociation out of all variants, but reaches very low response levels, i.e., 7 RU (fig.10 d). Thus, variants 2, 3 and 5 was selected to continue to *in vitro* analysis (fig.10 b, c, e).

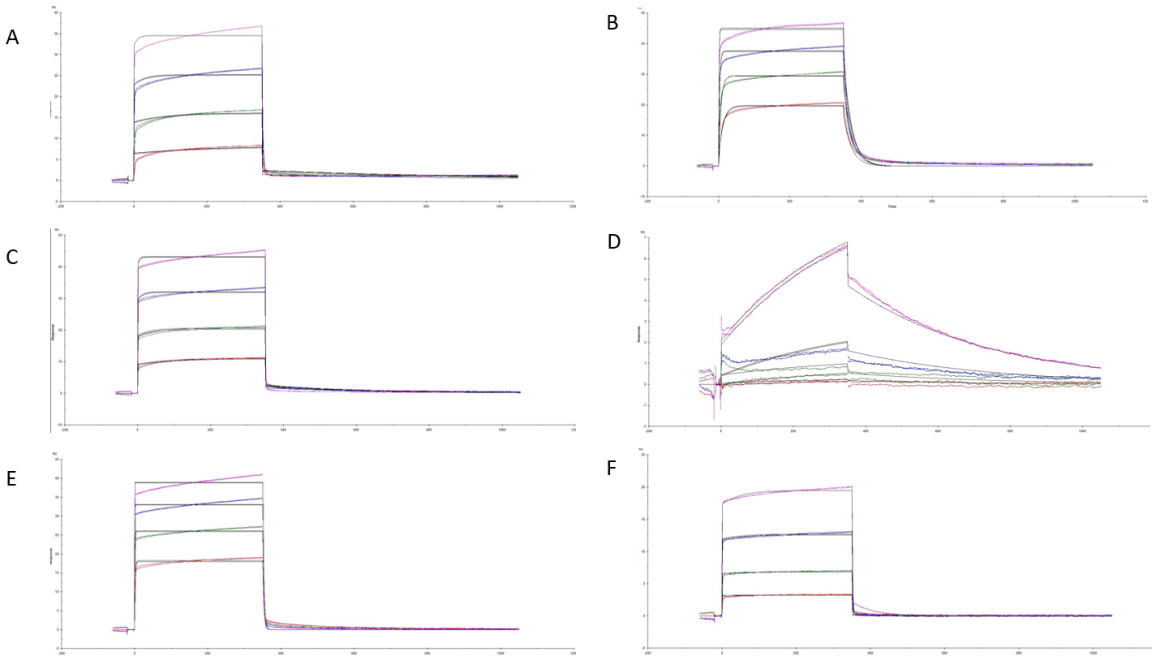


Figure 10 (a-f). Biosensor binding curves for the second of the six selected variants; (a) 1, (b) 2, (b) 3, (d) 4, (e) 5 and (f) 7, each with concentrations of 100 nM (red), 300 nM (green), 900 nM (blue), and 2700 nM (magenta).

The amount of human mAb was possibly miscalculated and the binding curve shows only straight lines, thus making fitting unsuitable.

The BLI analysis was thus run for variants 1-5, 7, an affibody control and the anti-target antibody. The binding curves obtained were less "bulky" than the SPR analysis and the fitting appears to be better for each of the variants, ranging in K_D from 10.1 to 84.2 nM and reaching similar response levels (fig.11). The antibody does not appear to be binding at all to the target as expected, as it immediately started washing out and reached negative response (fig.11 g). The affibody control that was intended to be a negative binder appears instead to be a positive binder as there appears to be binding and dissociation, although reaching lower response level than the variants did (fig.11 h).

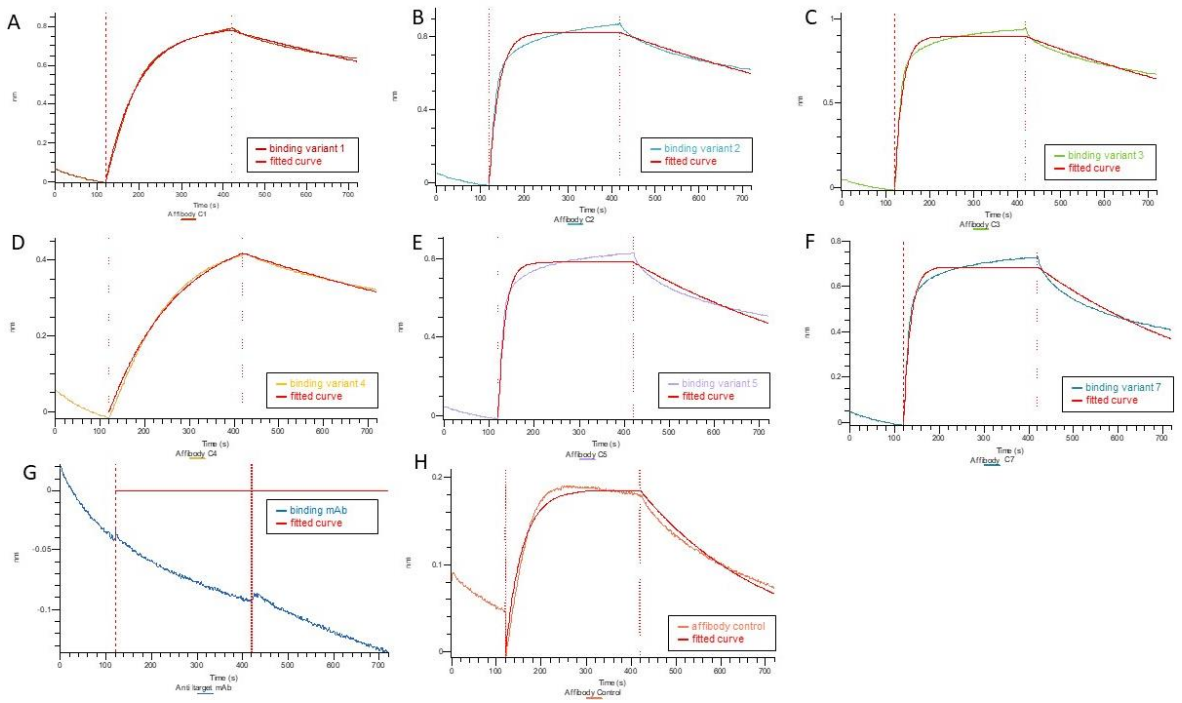


Figure 11 (a-h). BLI binding curves, with the measured interference [nm] plotted against time, for the second of the six selected variants in 500 nM for variants (a) 1, (b) 2, (c) 3, (d) 4, (e) 5, (f) 7, and for (g) 50 nM antibody, and (h) 500 nM affibody control.

For the SPR run, 382 out of 500 RU was bound. For the BLI analysis, the equilibrium observed binding signal R (R_{eq}) was used to calculate the ratio of R_{eq}/R_{max}.

Comparing the kinetic parameters from SPR and BLI, variant 2, 3 and 5 shows highest affinity toward the target, despite the difference in K_D (table 8). According to the BLI results, these variants also reach high response close to R_{max}. As the anti-target antibody did not bind to the target, the kinetic parameters could not be fitted or calculated.

Table 8. Parameters obtained from SPR and BLI analysis for variants 1-5, 7 and the anti-target mAb.

Variant	SPR	BLI			
	K_D [M]	K_D [M]	R_{max} [nm]	R_{eq} [nm]	R_{eq}/R_{max} [%]
1	4.10×10^{-8}	3.09×10^{-8}	0.84	0.79	94.2
2	1.26×10^{-8}	1.24×10^{-8}	0.84	0.82	97.6
3	4.30×10^{-8}	1.01×10^{-8}	0.91	0.90	98.0
4	2.65×10^{-5}	8.42×10^{-8}	0.57	0.49	85.6
5	9.13×10^{-8}	1.35×10^{-8}	0.80	0.78	97.4
7	2.87×10^{-5}	1.68×10^{-8}	0.71	0.69	96.7
mAb	N/A	1.00×10^{-7}	0	0	N/A

3.4.4 Affinity analysis – binding to cancer cells

The *in vitro* screening of variants revealed that out of the four cell lines, BT-474 and MCF-7 gave rise to the most apparent shift in FL6 (HSA) intensity in comparison to the control (fig.11 b, c). Overall, affibody variant 3 appears to be very similar in fluorescence to the control as the graphs overlap for all cell lines, while variants 2 and 5 shows shift for both cell line BT-474 and MCF-7. In coincidence with previous results, the anti-target antibody appears to bind weakly or off-site to the target, as there is no FL6 shift, and the fluorescence is lower than the affibodies (appendix fig.A5).

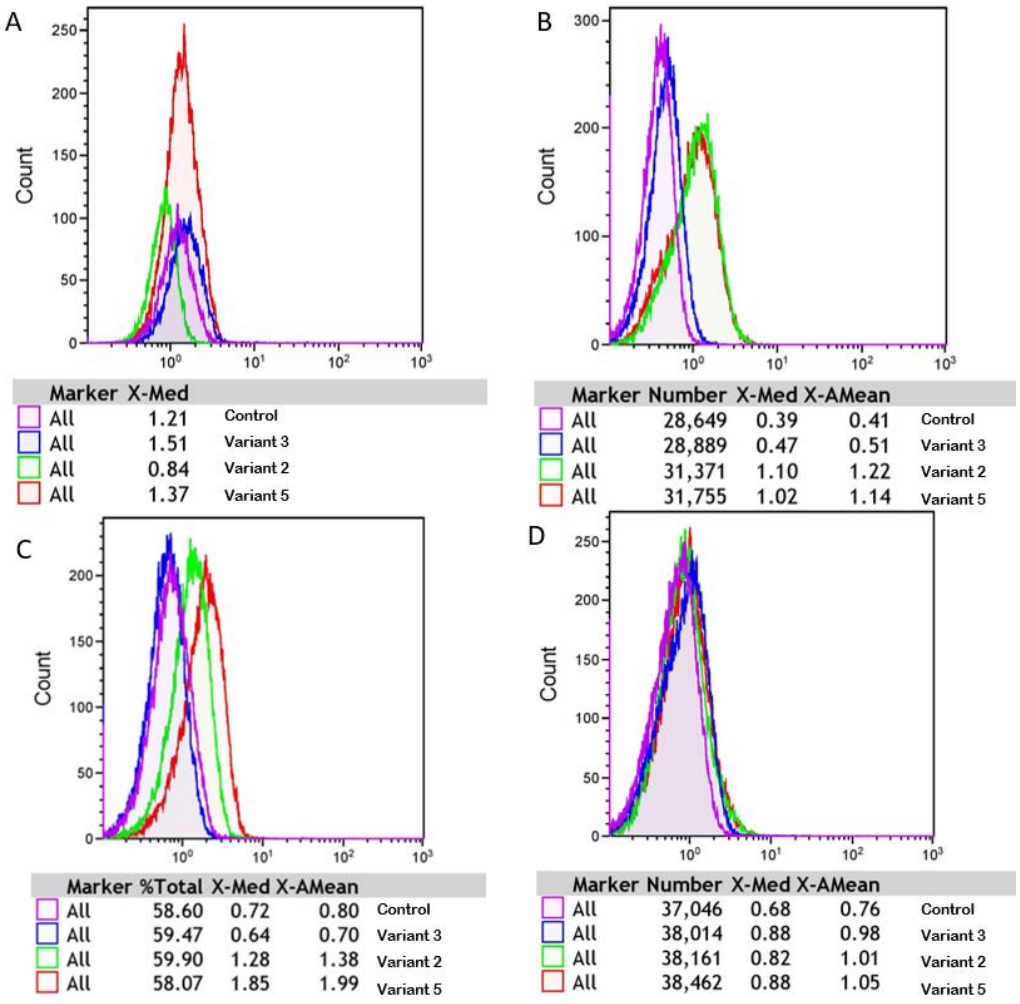


Figure 12 (a-d). FFC results for *in vitro* studies for the variants 2, 3 and 5, to screen the cell lines (a) SK-BR-3, (b) BT-474, (c) MCF-7, and (d) A-431.

4 Discussion

This study aimed to select and characterise ABD-fused affibodies to target a novel candidate biomarker and target protein. From the enriched affibody library, selection and screening isolated a total of 40 variants that were screened toward the target protein, out of which 12 variants were further analysed for biochemical characterisation and *in vitro* study to evaluate the affinity to cell-expressed target protein.

4.1 Results

The fourth MACS round successfully enriched the library an additional 10-fold and showed a shift in FL2 (SAPE) intensity, indicating that the library now contains a larger fraction of variants with stronger affinity to the target protein, in comparison to the previous outputs (fig.4). As MACS is limited to off-target binding, it is unable to remove the weak variants from the output, which could be an explanation as to why there still appears to be a lot of variant with very weak fluorescence or background noise. Additionally, although MACS is supposed to remove variants binding to tags like ABD and His_6, there is still a risk of non-specific binding variants remaining the enriched library.

Followingly, the affibodies to proceed with were picked based on the flow cytometry scatter plots, by picking the variants with highest shift in FL2 as this theoretically indicates that the affibodies bind strongly to the target or that the affibodies binds the target bivalently. Based on the FL2 shift the variants 4, 5, 7 and 12 appears to be the best candidates (fig.5) Although, as there still is a considerable risk of weak or non-specific variants being selected, it would have been more preferable to account to this risk by proceeding with a higher number of variant for screening.

Upon identification and comparison of the affibody sequences, some similarities among the variants could be detected in the sequence of the two helices forming the cleft, although it is unclear to what extent these similarities are useful for affinity or specificity. The genetic lineage analysis of the variants showed that there are four clusters, and interestingly, variant 2, 5, and 7 all belong to the same cluster C (fig.8). Even here, some common amino acids in certain positions could be found, especially for variant 2 and 7 where five of the randomized amino acids in helix 1 proved to be identical.

The affinity presents as rather vague as the SPR and BLI analyses do not agree visually. The SPR exhibits very bulky curves with very rapid binding and dissociation from the target for almost all of the variants, which is undesirable as this could mean that the AffiDC would not be bound long enough to trigger internalisation into the cell. Overall, the affinity is measured to less than 100 nM for all variants, which is decent yet not fully satisfying. Looking in detail, variant 2 shows one of the slower dissociations rates, which still is quicker than desired. Interestingly, variant 4 showed one of the largest intensity shifts in FFC and has slower on-and off-rates, but does not appear to reach full saturation in SPR, even for

concentrations up to 2700 nM. This would mean that large amounts of affibody would be required to elicit a response, ultimately making it a seemingly unfit candidate for therapeutic applications.

Unlike the SPR results, the BLI analysis showed both slower binding and dissociation, thus rendering curves that could be fitted to parameters more efficiently and accurately. Still, the affinity for each variant proved to be rather similar, all ranging under 100 nM. For variant 2, the difference in K_D is only 2 nM. These differences that occur are likely to be caused by differences in methods or fitting. Still, both analyses imply that variants 2, 3 and 5 are the most promising.

The screening with cell lines was run with the three final variants that exhibited affinity to cell lines BT-474 and MCF-7, meaning variants 2, 3 and 5. Surprisingly, no binding was seen in SK-BR-3 which makes it believable that the expression of target was reduced, possibly as a result of long-running cultivations or poor quality due to other external factors. Since the binding of affibodies are of interest, the fluorescence shift of the fluorophore-bound HSA is measured. This FL6 (HSA) intensity shift is not as high as anticipated and this would presumably be a result of the fast dissociation from target, as shown from SPR and BLI. Again, the antibody does not appear to be binding to target since higher intensity is expected. Thus, it fails to fill its function as positive control, so another type of positive control could have been used to ensure reliability of the results. Additionally, studies have shown that the antibody in question *does* bind efficiently to the target, leaving doubts whether the correct antibody was used for the study.

Upon comparison of all variants, it can be concluded that variants 2 and 5 exhibit the strongest binding toward the target. These variants are also members of the same cluster, which is interesting as that could offer a sense of target binding properties. On the other hand, these sequences are rather different and only two identical amino acids in helix 1 could be found. Even though variants 2 and 5 are the best-binding affibody candidates out of the 12 screened ones, they are still lacking in performance as they in several analyses appear to dissociate very quickly. Before proceeding further with these affibodies, additional affinity measurements should be run and repeated.

To conclude, the screening and selection were successful in finding and identifying affibodies toward the target protein. This supports the fact that the methods and the *E. coli* expression display systems can successfully be applied to find new affibodies from recombinantly produced libraries, which is in line with previous studies of affibody cell display. Furthermore, this also provides support for the idea to target this particular protein as a novel therapeutic agent, in spite of the varying degree of affinity. Although a larger number of affibodies would be required to determine any affinity-beneficial patterns or regions in the amino acid sequences, this study provides with interesting preliminary results for this novel cancer therapeutic target.

4.2 Future perspectives

Finally, it would be interesting to perform an extension of this study in order to screen more variants and many aspects could be repeated, changed, or added. To be more time-efficient, a method that allows for higher throughput of screened variants would be preferred, like fluorescence activated-cell sorting (FACS). Other runs of SPR or BLI could be run initially to decide whether the variants bind the target or not, without any quantitative analysis. Still, managing large numbers of unique samples bring other practical challenges. As the quality *in vitro* study was low, due to faulty antibody and cell lines, repetition would provide with stronger support to the results. One could also attempt to dimerize the affibody as this has been suggested to improve the stability [34]. Continuing beyond the scope of the study, if certain motifs or sequences in the affibodies helices were to be identified, next generation DNA sequencing could allow for a first filtering of uninteresting variants. And certainly, if a satisfying affibody with suitable properties were to be found, an AffiDC would be designed and evaluated as a possible cancer treatment.

Acknowledgement

I am truly honoured and forever thankful to my main supervisor Torbjörn Gräslund for offering me the opportunity to write my master's thesis in his group. These weeks have been the most fun I have had during my studies, and I feel incredibly privileged to have partaken in the frontline of cancer therapy studies. Thank you for your wisdom, expertise, and teaching – you have taught me things in minutes that took hours previously! 😊 I wish you and the affibodies all the best and good luck!

Next, I would like to thank my co-supervisor Javad Garousi. I don't really know where to begin, I am so, so thankful for all your time that you have allocated to guide me during these weeks. I could not have had a better co-supervisor. I feel so spoiled to have you on my shoulder as a wisdom guardian *every day*, to answer my stupid questions or explain the same thing over and over again. Thank you for your encouragement and guidance, and most importantly – thank you for reminding me to take coffee breaks!

A big thanks to the rest of the group; Haozhong Ding, Jie Zhang, and Wen Yin. Many thanks to Haozhong for lending us your knowledge whenever we needed it (also, your latest paper has been very useful to me in my thesis writing, so thank you for writing a great publication!). Thanks to Jie for lending us materials and going out of your way to guide me throughout these weeks, even when the sonicator was uncooperative! Thank you to Wen for inspiring me during your thesis defence, and for sharing my column purification pains.

To all of the staff on plan 3, thank you for making my time here pleasant and enjoyable. Special thanks to Jonas Persson for being the allt-i-allo hero that we all apparently *really* need. Thanks to Linnea Hjelm, Andreas Wisniewski, Moira Ek, Savvina Gkouma and Emma Larsson for arranging introductions and being incredibly helpful. Thanks to Aman Mebrahtu for letting us use your equipment for the Octet, it was really useful in my study! Thanks to Maria Geis for saving my samples and buffers, when I was confusedly occupied with other things. And thanks to Gabriella Jensen for kindly running the MS for us! Lastly, many thanks to all the people who have cheered on us and offered their help and knowledge, you made the experience the best.

Of course, nothing would have been emotionally possible without my master thesis colleagues; Tasnim Idris, Barathram Swaminathan, Arman Swaich, Kristel Jaiu, Ebba Bergström, Oscar Friberg, Ayleen Burt, Sebastian Augustinsson and Matilda Fresk. I wish you best of luck in the future!

My most heartfelt thanks goes to my family and friends - thank you for encouragement and support! And finally, many thanks to you who read my thesis. I hope you found something useful or interesting!

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin. 2020;70(1):7-30.

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209-49.

3. Zugazagoitia J, Guedes C, Ponce S, Ferrer I, Molina-Pinelo S, Paz-Ares L. Current Challenges in Cancer Treatment. Clinical Therapeutics. 2016;38(7):1551-66.

4. Stoletov K, Beatty PH, Lewis JD. Novel therapeutic targets for cancer metastasis. Expert Rev Anticancer Ther. 2020;20(2):97-109.

5. Torphy RJ, Fujiwara Y, Schulick RD. Pancreatic cancer treatment: better, but a long way to go. Surg Today. 2020;50(10):1117-25.

6. Gotwals P, Cameron S, Cipolletta D, Cremasco V, Crystal A, Hewes B, et al. Prospects for combining targeted and conventional cancer therapy with immunotherapy. Nat Rev Cancer. 2017;17(5):286-301.

7. Maximiano S, Magalhaes P, Guerreiro MP, Morgado M. Trastuzumab in the Treatment of Breast Cancer. BioDrugs. 2016;30(2):75-86.

8. Tsao LC, Force J, Hartman ZC. Mechanisms of Therapeutic Antitumor Monoclonal Antibodies. Cancer Res. 2021;81(18):4641-51.

9. Thomas A, Teicher BA, Hassan R. Antibody-drug conjugates for cancer therapy. Lancet Oncol. 2016;17(6):e254-e62.

10. Miyazaki J, Nishiyama H. Epidemiology of urothelial carcinoma. Int J Urol. 2017;24(10):730-4.

11. Sheng X, Yan X, Wang L, Shi Y, Yao X, Luo H, et al. Open-label, Multicenter, Phase II Study of RC48-ADC, a HER2-Targeting Antibody-Drug Conjugate, in Patients with Locally Advanced or Metastatic Urothelial Carcinoma. Clin Cancer Res. 2021;27(1):43-51.

12. Patel VG, Oh WK, Galsky MD. Treatment of muscle-invasive and advanced bladder cancer in 2020. CA Cancer J Clin. 2020;70(5):404-23.

13. Vlachostergios PJ, Jakubowski CD, Niaz MJ, Lee A, Thomas C, Hackett AL, et al. Antibody-Drug Conjugates in Bladder Cancer. Bladder Cancer. 2018;4(3):247-59.

14. Abdollahpour-Alitappeh M, Lotfinia M, Gharibi T, Mardaneh J, Farhadihosseinabadi B, Larki P, et al. Antibody-drug conjugates (ADCs) for cancer therapy: Strategies, challenges, and successes. J Cell Physiol. 2019;234(5):5628-42.

15. Castelli MS, McGonigle P, Hornby PJ. The pharmacology and therapeutic applications of monoclonal antibodies. Pharmacol Res Perspect. 2019;7(6):e00535.

16. Ding H, Xu T, Zhang J, Tolmachev V, Oroujeni M, Orlova A, et al. Affibody-Derived Drug Conjugates Targeting HER2: Effect of Drug Load on Cytotoxicity and Biodistribution. Pharmaceutics. 2021;13(3).

17. Nagano K, Tsutsumi Y. Phage Display Technology as a Powerful Platform for Antibody Drug Discovery. Viruses. 2021;13(2).

18. Tumey LN. An Overview of the Current ADC Discovery Landscape. Methods Mol Biol. 2020;2078:1-22.

19. Panowski S, Bhakta S, Raab H, Polakis P, Junutula JR. Site-specific antibody drug conjugates for cancer therapy. MAbs. 2014;6(1):34-45.

20. Stahl S, Graslund T, Eriksson Karlstrom A, Frejd FY, Nygren PA, Lofblom J. Affibody Molecules in Biotechnological and Medical Applications. Trends Biotechnol. 2017;35(8):691-712.

21. Altai M, Liu H, Ding H, Mitran B, Edqvist PH, Tolmachev V, et al. Affibody-derived drug conjugates: Potent cytotoxic molecules for treatment of HER2 over-expressing tumors. J Control Release. 2018;288:84-95.

22. Xu T, Ding H, Vorobyeva A, Oroujeni M, Orlova A, Tolmachev V, et al. Drug Conjugates Based on a Monovalent Affibody Targeting Vector Can Efficiently Eradicate HER2 Positive Human Tumors in an Experimental Mouse Model. Cancers (Basel). 2020;13(1).

23. Spriestersbach A, Kubicek J, Schafer F, Block H, Maertens B. Purification of His-Tagged Proteins. Methods Enzymol. 2015;559:1-15.

24. Hoppmann S, Miao Z, Liu S, Liu H, Ren G, Bao A, et al. Radiolabeled affibody-albumin bioconjugates for HER2-positive cancer targeting. Bioconjug Chem. 2011;22(3):413-21.

25. Malm M, Bass T, Gudmundsdotter L, Lord M, Frejd FY, Stahl S, et al. Engineering of a bispecific affibody molecule towards HER2 and HER3 by addition of an albumin-binding domain allows for affinity purification and in vivo half-life extension. Biotechnol J. 2014;9(9):1215-22.

26. Fleetwood F, Andersson KG, Stahl S, Lofblom J. An engineered autotransporter-based surface expression vector enables efficient display of Affibody molecules on OmpT-negative E. coli as well as protease-mediated secretion in OmpT-positive strains. Microb Cell Fact. 2014;13:179.

27. Guzman LM, Belin D, Carson MJ, Beckwith J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol. 1995;177(14):4121-30.

28. Rashid Z, Shokri F, Abbasi A, Khoobi M, Zarnani AH. Surface modification and bioconjugation of anti-CD4 monoclonal antibody to magnetic nanoparticles as a highly efficient affinity adsorbent for positive selection of peripheral blood T CD4+ lymphocytes. Int J Biol Macromol. 2020;161:729-37.

29. Deb A, Johnson WA, Kline AP, Scott BJ, Meador LR, Srinivas D, et al. Bacterial expression, correct membrane targeting and functional folding of the HIV-1 membrane protein Vpu using a periplasmic signal peptide. PLoS One. 2017;12(2):e0172529.

30. Du F, Liu YQ, Xu YS, Li ZJ, Wang YZ, Zhang ZX, et al. Regulating the T7 RNA polymerase expression in E. coli BL21 (DE3) to provide more host options for recombinant protein production. Microb Cell Fact. 2021;20(1):189.

31. Ratelade J, Miot MC, Johnson E, Betton JM, Mazodier P, Benaroudj N. Production of recombinant proteins in the lon-deficient BL21(DE3) strain of Escherichia coli in the absence of the DnaK chaperone. Appl Environ Microbiol. 2009;75(11):3803-7.

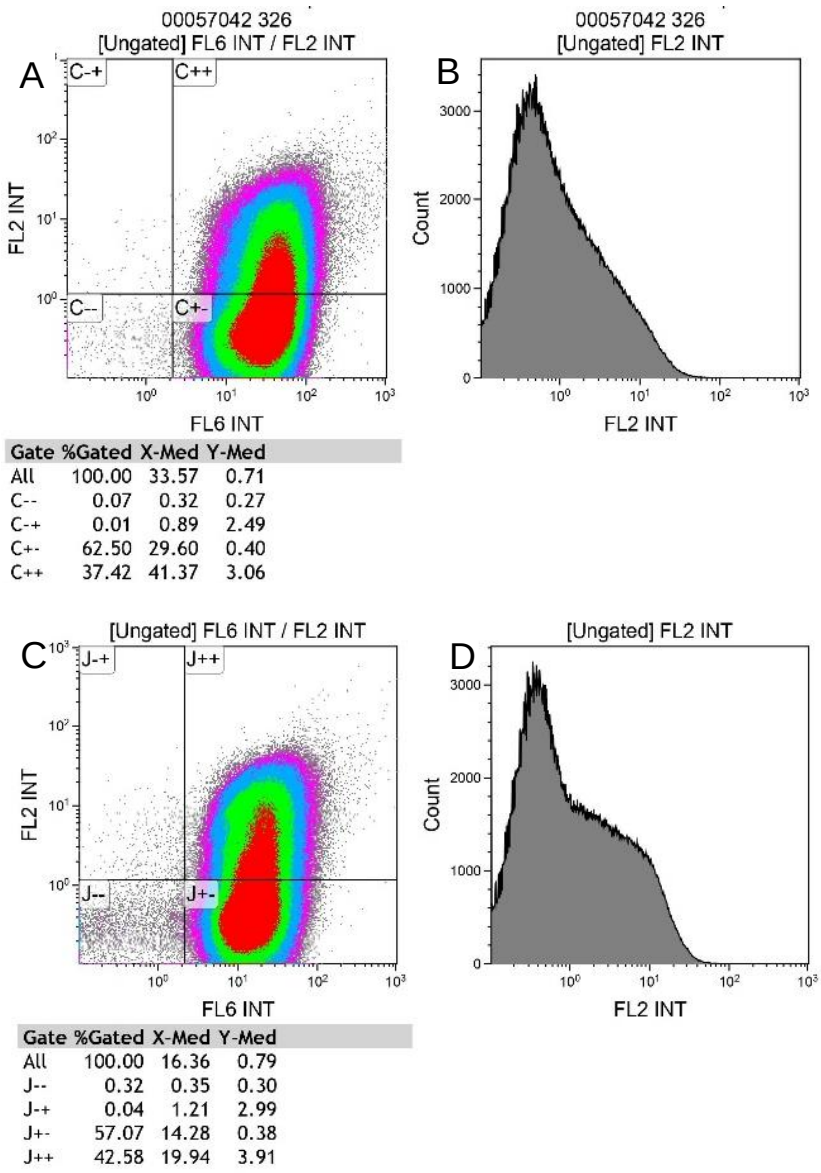
32. Murali S, Rustandi RR, Zheng X, Payne A, Shang L. Applications of Surface Plasmon Resonance and Biolayer Interferometry for Virus-Ligand Binding. Viruses. 2022;14(4).

33. Yang D, Singh A, Wu H, Kroe-Barrett R. Comparison of biosensor platforms in the evaluation of high affinity antibody-antigen binding kinetics. Anal Biochem. 2016;508:78-96.

34. Westerlund K, Myrhammar A, Tano H, Gestin M, Karlstrom AE. Stability Enhancement of a Dimeric HER2-Specific Affibody Molecule through Sortase A-Catalyzed Head-to-Tail Cyclization. Molecules. 2021;26(10).

Appendix

A1. MACS enrichment



Gate	%Gated	X-Med	Y-Med
All	100.00	33.57	0.71
C--	0.07	0.32	0.27
C-+	0.01	0.89	2.49
C+-	62.50	29.60	0.40
C++	37.42	41.37	3.06

Gate	%Gated	X-Med	Y-Med
All	100.00	16.36	0.79
J--	0.32	0.35	0.30
J-+	0.04	1.21	2.99
J+-	57.07	14.28	0.38
J++	42.58	19.94	3.91

Figure A1. FFC results of enrichment analysis, FL2 = SAPE corresponding to binding to target, FL6 = HSA-Alexa Fluor™ 647, corresponding to binding of analyte. MACS output 3 events plotted (a) with FL2 and FL6, and (b) event count and FL2. MACS output 4 events plotted with (c) FL2 and FL6, and (d) event count and FL2.

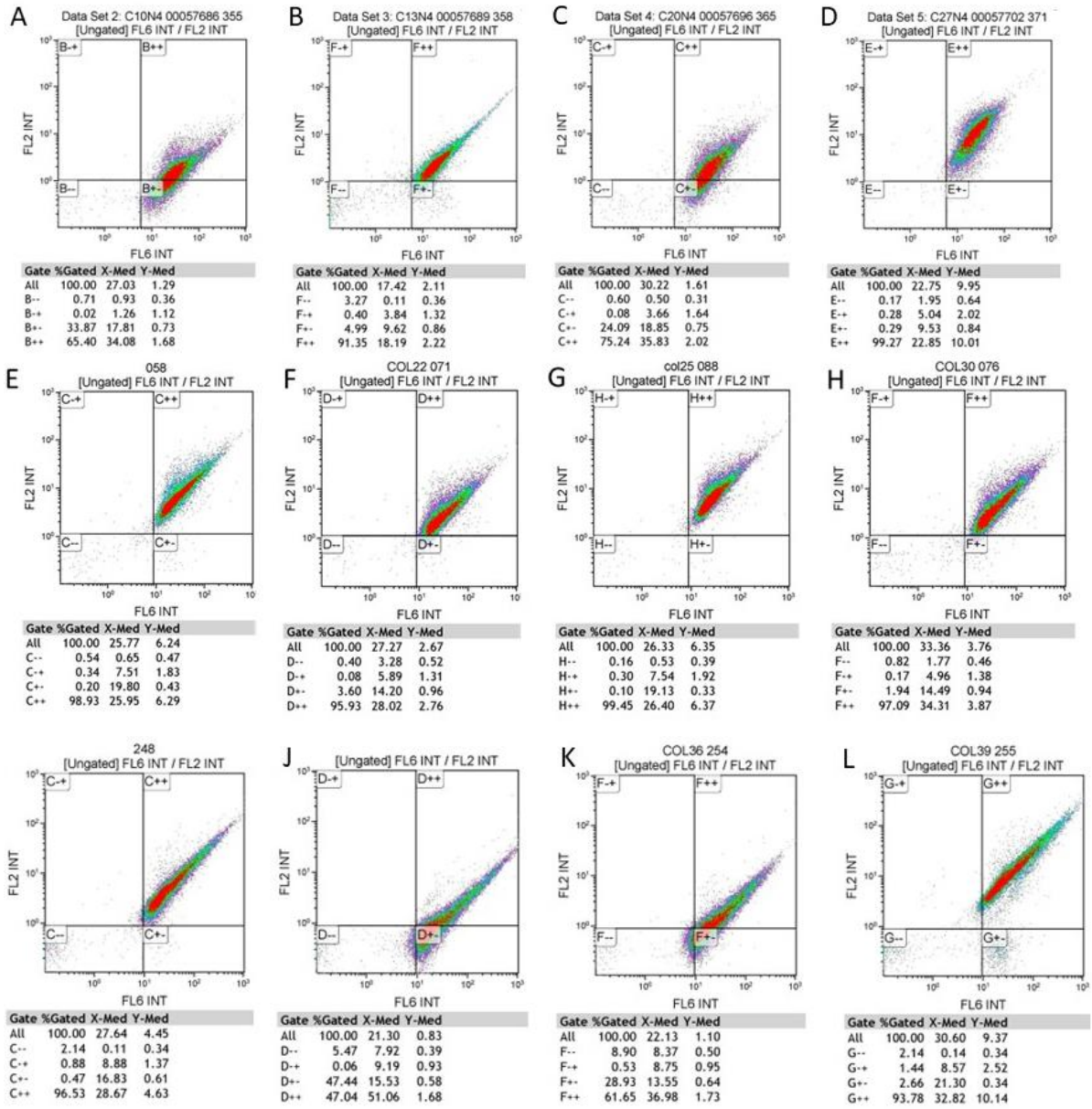
A2. pBAD2 PCR-primers

Forward primer: 5'-ATA TGG CAT ATG GTG GAT AAC AAA TTC-3'

Reverse primer: 5'-GTA AAT GTC GAG CTA CGG CAG TGC-3'

Figure A2. Custom primers for used for conventional PCR on pBAD-vector.

A3. Screen by flow cytometry



Gate	%Gated	X-Med	Y-Med
All	100.00	27.03	1.29
B--	0.71	0.93	0.36
B-+	0.02	1.26	1.12
B+-	33.87	17.81	0.73
B++	65.40	34.08	1.68

Gate	%Gated	X-Med	Y-Med
All	100.00	17.42	2.11
F--	3.27	0.11	0.36
F-+	0.40	3.84	1.32
F+-	4.99	9.62	0.86
F++	91.35	18.19	2.22

Gate	%Gated	X-Med	Y-Med
All	100.00	30.22	1.61
C--	0.60	0.50	0.31
C-+	0.08	3.66	1.64
C+-	24.09	18.85	0.75
C++	75.24	35.83	2.02

Gate	%Gated	X-Med	Y-Med
All	100.00	22.75	9.95
E--	0.17	1.95	0.64
E-+	0.28	5.04	2.02
E+-	0.29	9.53	0.84
E++	99.27	22.85	10.01

Gate	%Gated	X-Med	Y-Med
All	100.00	25.77	6.24
C--	0.54	0.65	0.47
C-+	0.34	7.51	1.83
C+-	0.20	19.80	0.43
C++	98.93	25.95	6.29

Gate	%Gated	X-Med	Y-Med
All	100.00	27.27	2.67
D--	0.40	3.28	0.52
D-+	0.08	5.89	1.31
D+-	3.60	14.20	0.96
D++	95.93	28.02	2.76

Gate	%Gated	X-Med	Y-Med
All	100.00	26.33	6.35
H--	0.16	0.53	0.39
H-+	0.30	7.54	1.92
H+-	0.10	19.13	0.33
H++	99.45	26.40	6.37

Gate	%Gated	X-Med	Y-Med
All	100.00	33.36	3.76
F--	0.82	1.77	0.46
F-+	0.17	4.96	1.38
F+-	1.94	14.49	0.94
F++	97.09	34.31	3.87

Gate	%Gated	X-Med	Y-Med
All	100.00	27.64	4.45
C--	2.14	0.11	0.34
C-+	0.88	8.88	1.37
C+-	0.47	16.83	0.61
C++	96.53	28.67	4.63

Gate	%Gated	X-Med	Y-Med
All	100.00	21.30	0.83
D--	5.47	7.92	0.39
D-+	0.06	9.19	0.93
D+-	47.44	15.53	0.58
D++	47.04	51.06	1.68

Gate	%Gated	X-Med	Y-Med
All	100.00	22.13	1.10
F--	8.90	8.37	0.50
F-+	0.53	8.75	0.95
F+-	28.93	13.55	0.64
F++	61.65	36.98	1.73

Gate	%Gated	X-Med	Y-Med
All	100.00	30.60	9.37
G--	2.14	0.14	0.34
G-+	1.44	8.57	2.52
G+-	2.66	21.30	0.34
G++	93.78	32.82	10.14

Figure A3 (a-l). FFC scatter plot result with FL2 intensity plotted against FL6 intensity, for all 12 selected variants from MACS output 3 (col 1-4, and col 9-12) and MACS output 4 (variant 5-8).

A4. LC-MS

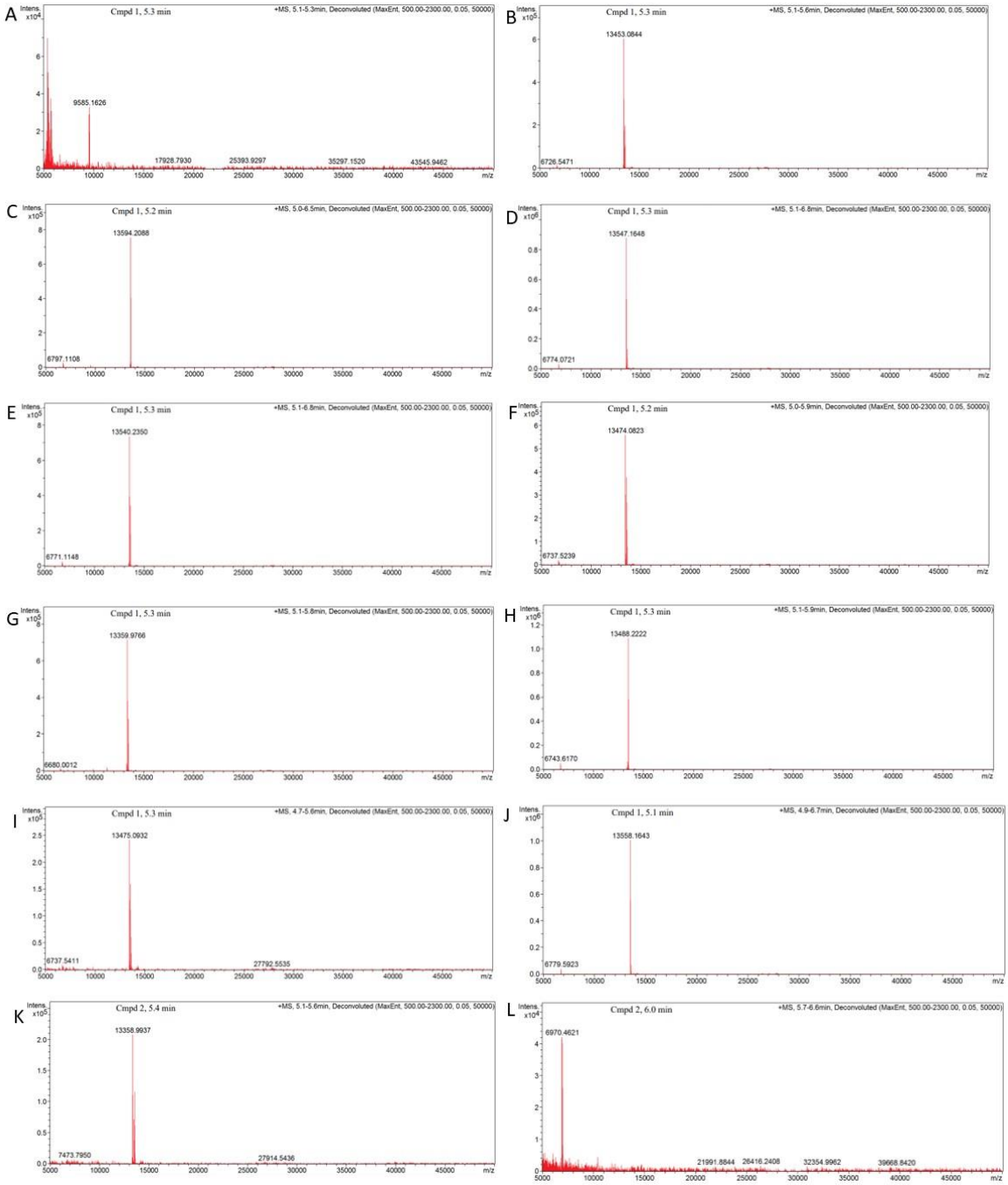


Figure A4 (a-l). LC-MS mass spectrum for all variants 1-12 (a-l) with relative intensity plotted against mass-to-charge (m/z) ratio.

A5. *In vitro* analysis with antibody

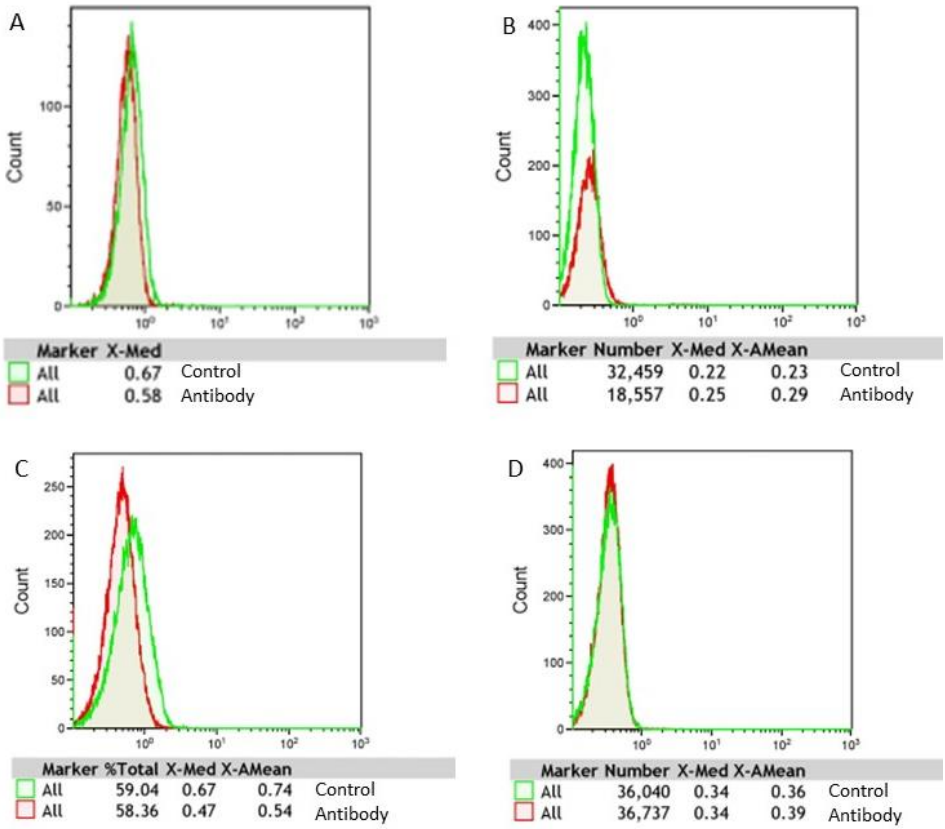


Figure A5 (a-d). The fluorescence results with count plotted against FL6 (HSA) for the anti-target antibody (red) and a control (green), for the cell lines (a) SK-BR-3, (b) BT-474, (c) MCF-7, and (d) A-431.