EXAMENSARBETE INOM TEKNIKOMRÅDET
BIOTEKNIK
OCH HUVUDOMRÅDET
MEDICINSK TEKNIK,
AVANCERAD NIVÅ, 30 HP
STOCKHOLM, SVERIGE 2021

Design, production and evaluation of cross linked target proteins to an affibody-based carrier framework aimed for affinity protein:antigen structure determination using single particle Cryo-EM

RICHARD BRUNSELL

KTH
SKOLAN FÖR KEMI, BIOTEKNOLOGI OCH HÄLSA

Table of Contents

Abstract

Small proteins are difficult to study at high resolution with single-particle cryo-electron microscopy (cryo-EM). In general, sample properties such as large size (> 80 kDa), symmetry and rigidity are key to utilize this technology. To facilitate structural studies of small proteins as well, using cryo-EM, this project aims to incorporate a photo-inducible cross-link in a large and symmetric scaffold that is amenable for study, and covalently bind small proteins of interest to this scaffold. The scaffold in this project consists of rabbit muscle aldolase (157 kDa in tetrameric state) with an engineered affibody affinity protein (7 kDa) attached to the N-terminus of each aldolase monomer via a rigid helix fusion. The affibody-domain of the scaffold will be cross-linked to small proteins of structural interest, with a focus on a model target consisting of a second affibody with affinity for the affibody displayed on the aldolase scaffold.

Photoconjugation of the affibody Zwt was performed to crosslink both the Fc of IgG and the anti-idiotypic affibody Z963, revealing that a methionine acceptor in the target is preferable but not necessary for UV crosslinking using BPA. Binding of affibodies rigidly displayed on of the scaffold to targets such as affibodies and antibody fragments was demonstrated , using surface plasmon resonance (SPR).

Keywords

Cryo-EM, Structure determination, Protein A, Affibody, Aldolase, Scaffold, Rigid helix fusion, Photoconjugation, UV crosslinking, BPA

Sammanfattning

Att studera små protein vid hög upplösning med enpartikelsrekonstruktion i kryo-elektronmikroskopi (kryo-EM) är utmanande. Generellt så krävs stora (> 80 kDa), symmetriska och stabila protein för att använda sig av kryo-EM. Med målet att möjliggöra strukturbestämning och strukturella studier av små protein, så ska detta projekt föra in en foto-aktiverad korslänk i ett stort och symmetriskt bärarprotein. Bäraren består av aldolas från kaninmuskel (157 kDa som tetramer) med en affibody (7 kDa) kopplad till N-terminalen av varje aldolas-monomer via en rigidt fuserad helix. Affibody-domänen av bärarproteinet kan bilda korslänkar till små protein vars struktur sedan kan studeras. Fokus i projektet är ett modellprotein som består av en annan affibody som binder den affibody i bäraren.

Fotokonjugering av affibodyn Zwt utfördes för att skapa korslänkar till både Fc av IgG, samt den anti-idiotypiska affibodyn Z963, vilket påvisade att en metionin-mottagare i målproteinet är fördelaktigt för UV korslänkning med BPA, men inte ett krav. Affinitet av affibodies i bärarproteinet till målprotein såsom andra affibodies och antikroppsfragment påvisades.

1 Introduction

Single particle cryogenic electron microscopy (cryo-EM) is a powerful and rapidly developing technology for imaging proteins and other macromolecules, which can be used to generate high resolution 3D structures of samples in near natural conditions.

The 2017 Nobel prize in chemistry was awarded to Jacques Dubochet, Joachim Frank, and Richard Henderson for the development of high resolution cryo-EM [1]. A combination of factors such as computational power [2] and new detector hardware for imaging [3] gave rise to the "resolution revolution", i.e. a development where higher and higher resolutions of protein samples have been achieved using electron microscopy.

The current leading techniques for structure determination are X-ray diffraction, NMR, and cryo-EM and all face their own unique bottlenecks. The dominating method of X-ray diffraction requires a crystal of the protein to obtain a high resolution. The process of optimizing the crystallization conditions can be time consuming and requires a lot of sample material [4]. Furthermore, some proteins like membrane bound receptors are not only hard to produce in large quantities, but typically also very difficult to crystallize in solution [5]. NMR is able to solve the structure of a small, spin-labelled protein in solution, and also give insight into their interactions. Like X-ray diffraction, it does however require high concentrations of sample to generate a sufficient signal-to-noise ratio. NMR is also limited to smaller proteins (<40 kDa) [6], which excludes a large portion of the proteins in the cell.

Cryo-EM has an advantage over X-ray crystallography in that no sample crystallization must be performed prior to the imaging. Both X-ray crystallography and NMR require large quantities of sample, while cryo-EM typically does not. The structures that can be directly viewed using cryo-EM are however generally limited to larger proteins. One quarter of cryo-EM entries in EMDB have a resolution of less than 10 Å, with the majority being larger than 500 kDa [7]. This does however not mean that cryo-EM is incapable of studying small proteins. For instance, the 52 kDa protein streptavidin has been solved at 3.2 Å using cryo-EM [8]. However, these types of studies require optimal samples and powerful equipment not available to most institutions.

Cryo-EM structures of a resolution of ~3 Å have been obtained for structures between 80 kDa to 50 MDa using single-particle reconstruction. Single particles of the sample face random directions in the cryo-EM micrograph, allowing the same sample to be viewed from different angles and generating different 2D projections of the same protein. Next, images are reconstructed into a consensus 3D structure [9]. This method is more applicable to symmetric structures, such as multimer complexes, as it effectively multiplies the number of projections from a certain angle, generating more material to refine each image class and build a 3D model.

High resolution cryo-EM structures can usually only be obtained for larger proteins, but many of the proteins of interest are small. In the eukaryotic cell, the majority of proteins are less than 50 kDa [11], which is below the threshold for higher resolution images with cryo-EM. In order to study smaller proteins in cryo-EM, they can be directly or non-covalently attached to a larger scaffold. This way, the entire structure viewed in the single particle can break the size barrier and higher resolution images can be obtained for the smaller associated protein of interest as well. Furthermore, if the scaffold itself is symmetric, this allows for easier imaging [11,13,14,15]. Previous work [10] has shown that an affibody, derived from the B domain of

Protein A from *Staphylococcus aureus*, fused to aldolase (a large, extensively studied, and symmetric protein), is a promising carrier scaffold for such applications. The affibody can bind to the protein of interest, and the aldolase can give the size and symmetry required for high resolution structures. Affibodies are small and stable affinity proteins consisting of three short alpha-helices that through *in vitro* selection can bind to any target, which makes them ideal for capturing the protein of interest. Several affibodies exist as pharmaceutical candidates with affinities towards many target proteins, allowing for modularity in the system.

The design of the linker in an affibody-aldolase fusion protein is important. Single particle reconstruction relies on a structurally homogenous sample where images of different particles can be treated as a single structure from different angles. The linker between the affibody and the aldolase benefits from being rigid, generating as little heterogeneity as possible. The polypeptide link between the carrier scaffold and the affinity protein is here designed as a rigid shared alpha-helix using a procedure described by Youn et al. [27] where a continuous alpha helix extends from the C-terminal helix III of the affibody into the N-terminal helix of aldolase.

The non-covalent bond between the affibody and the protein of interest may be weak and is in such cases expected to make purification of saturated scaffold:protein-of-interest complex difficult. A low saturation of the scaffold would result in a smaller fraction of useful images in the single particle reconstructions contributing to the 3D structure and potentially lowering resolution.

This project aims to evaluate a photoactivated cross-link to covalently associate the scaffold and the protein of interest. The non-canonical amino acid p-benzoyl phenylalanine (BPA) can form covalent cross-links to a proximal residue when exposed to UV light at 365 nm [16,18,19]. The BPA is incorporated into the affibody using an established orthogonal tRNA and aminoacyl-tRNA synthetase system [12]. This unique sidechain was used to photoconjugate the scaffold and the protein of interest and allow for higher saturation after purification of the complex, hopefully assisting in generating high resolution structures using cryo-EM

The B-domain-like affibody Zwt naturally binds to the Fc region of antibodies (50 kDa), which is below lower limit of what traditionally can be studied in cryo-EM. A scaffold with Zwt fused to aldolase (ZwA3) has previously been designed [10]. The photoconjugation of Zwt to Fc has been studied extensively [17,20,21,22] and can be used as a positive control when investigating novel photoconjugations. The anti-idiotypic affibody Z963 [26] also binds Zwt and at a size of 7.5 kDa is smaller than any protein imaged using cryo-EM to date. Z963 bound to the scaffold would therefore be a suitable proof-of-concept for imaging a protein far below the typical size limit in cryo-EM. BPA photoconjugation between Zwt and Z963 has however not been studied previously. Moreover, this investigation could further elucidate whether methionine is required for the cross-linking mechanism, which is debated in literature [18,19,22].

The affibody-aldolase scaffold can itself be used to couple a more general and commonly applied affinity protein such as scFv or Fab antibody fragments from e.g. therapeutic antibody candidates, or in complex with their antigens that alone are too small to be observed in cryo-EM. An affibody binding to the VH3 region of Fab/scFv can be used to link such proteins/protein complexes to the aldolase scaffold to enable electron microscopy studies with

non-covalently bound antigens. The B-domain of protein A (here denoted as Zwt A29G [23,24]), is closely related to the Zwt affibody and has this affinity for VH3.

Previous studies on photoconjugation using BPA have reported a potential dependency on methionine side chain as the target of cross-linking [20,22,25]. Z963 however lacks methionine in its sequence. The related Z964 which also binds Zwt, but with a lower affinity [26], does have a methionine in position 35. The importance of methionine as an acceptor was also investigated by substituting leucine residues in Z963 for methionine based on the position of the methionine in Z964 (*fig 1*). Different variants of the ZwA3 scaffold were investigated. ZwA3 for binding to Fc, Z963 or Z964 and ZwA3 A29G variants for binding to Fab or scFv (*table 1*). Focus in the project was to investigate the interactions of the isolated affibodies before promising findings can be transferred to the assembled aldolase scaffold system.

```
Protein              10        20        30        40        50
                     |         |         |         |         |
Zwt         VDNKFNKEQQNAFYEILHLPNLNEEQRNAFIQSLKDDPSQSANLLAEAKKLNDAQAPK
Z963        --------TQE-SW--FT-----GR-VA---S-LL----------------------
Z963 L34M   --------TQE-SW--FT-----GR-VA---S-ML----------------------
Z963 L35M   --------TQE-SW--FT-----GR-VA---S-LM----------------------
Z964        --------NDA-AW--FS-----GA-TV---R-LM----------------------
```

Figure 1. Sequence of Zwt, Z963 variants and Z964. The sequences are compared to Zwt, where dashed positions are generally not randomized in affibody libraries. Methionine is highlighted in bold.

2 Materials and Methods

2.1 *In silico* design

2.1.1 Positions for BPA incorporation in Zwt

BPA incorporation sites (amber stop codons) in Zwt were selected based on literature to allow photoconjugation to both a wide selection of Fc variants but also to the anti-idiotypic affibody Z963.

Crystal structures were obtained from PDB (5U4Y, 2M5A) and superimposed in PyMOL [30]. The positions for BPA incorporation at the amber stop codon via orthogonal tRNA and aminoacyl-tRNA synthetase were selected based on the proximity to a methionine in potential cross-link targets and the surface accessibility of the substituted residue in Zwt.

2.1.3 Evaluation of scaffold protein design

The scaffold fusion protein design with an affibody attached to aldolase was examined via superimposed structures including bound model targets to the affibody domain (PDB: 1EWD, 5U4Y, 1DEE) and potential steric clashes were evaluated. Residues from one superimposed structure that were close to another structure (3 Å) were flagged as potentially causing steric clashes (*appendix 8.1*). Other scaffold designs were investigated by changing the shared helix fusion between the affibody and aldolase in PyMOL. The affibody Zwt A29G has two binding surfaces, one for Fc and one for VH3 of Fab/scFv. A novel design was created where the second binding surface of the Zwt A29G was exposed (*appendix 8.2*).

2.2 DNA preparation

2.2.1 Site directed mutagenesis

Forward primers for site directed mutagenesis were designed with a 24 nucleotide overlap with the vector on either side of the mutated bases. Reverse primers annealing at a site further downstream were designed to the same length to be used universally for several constructs.

Site directed mutagenesis PCR was performed using a forward primer containing the desired mutation at both 65°C and 70°C annealing temperature in parallel reactions (*appendix 8.3*). The PCR products, representing full plasmid amplifications, were analysed via 2% agarose gel electrophoresis and then transformed to *Escherichia coli* TOP10 for nick repair.

2.2.2 Plasmid prep

Plasmids were replicated *in vivo* in *E. coli* TOP10 and extracted using a plasmid prep kit to create a working stock of plasmids for subsequent mutagenesis and expression. All overnight cultivations were performed in tryptic soy broth supplemented with yeast extract (TSB+Y, 30 g/L tryptic soy broth and 5 g/L yeast extract) and 25 μg/ml kanamycin at 37°C with shaking at 150 RPM.

~150 ng plasmid in 1xKCM (100 mM KCl, 30 mM $CaCl_2$, 50 mM $MgCl_2$) was mixed with an equal volume of freshly thawed *E. coli* TOP10 (cells prepared in-house) and incubated on ice

(20 min). The plasmids were transformed by heat shock (42°C, 40 s), phenotyped in pre-warmed (37°C) TSB+Y (end-over-end, 37°C 1h), and streaked out on antibiotic agar plates (40 g/L blood agar base, 25 µg/mL kanamycin) to grow overnight. The following day, a single colony was picked and grown with shaking overnight. The plasmids were extracted using a QIAprep Spin MiniPrep Kit (QIAGEN, Hilden, Germany), and the concentrations determined via NanoDrop.

Transformants carrying plasmid form site-directed mutagenesis reactions were screened by colony PCR (*appendix 8.4*) to identify clones with correct insert size on a 2% agarose gel, and plasmids prepared as described above were sequenced using Sanger sequencing (Eurofins TubeSeq service, Eurofins, Ebersberg, Germany).

2.3 Production of protein

2.3.1 Protein expression

Sequence verified plasmids were transformed into an *E. coli* expression strain (BL21* (DE3)). The following procedure was common for all constructs that did not contain an amber stop codon for incorporation of BPA.

From an overnight culture, 1 mL was transferred to 200 mL TSB+Y with appropriate antibiotic and grown with shaking to OD600 = 1 (150 RPM, 37°C). The culture was induced with a final concentration of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and harvested the next day.

For constructs aimed for BPA-incorporation, chemically competent cells carrying the plasmid pEVOL were prepared from *E. coli* BL21* DE3 transformed with pEVOL using standard procedures.

Constructs with an (xBPA) mutation were transformed to chemically competent *E. coli* BL21* DE3 containing the plasmid pEVOL. After overnight growth on agar plates (40 g/L blood agar base, 25 µg/mL kanamycin, 25 µg/mL chloramphenicol) the transformants were transferred by scraping of all the colonies from the agar plate (20-40 colonies) to 200 mL TSB+Y chloramphenicol (25 µg/mL) and kanamycin (25 µg/mL) and grown to OD600=1. The induction included, in addition to 1 mM IPTG, also 0.2% arabinose and final concentration of 1 mM BPA (Iris Biotech, Germany), divided in two administrations: one at the time of induction, and one 4 hours later

2.3.2 Purification

After cultivation, the protein purified by immobilized metal affinity chromatography (IMAC) and the buffer exchanged to 1xPBS (150 mM NaCl, 8 mM Na_2HPO_4, 2 mM NaH_2PO_4xH2O).

The cells were harvested by centrifugation (4000 RCF, 10 min, 4°C), the supernatant discarded, and the cell pellet resuspended in 7 ml IMAC wash buffer (1xPBS, 15 mM imidazole, pH 7.5). The cells were mechanically lysed by sonication for 6 minutes (ca. 10 mL sample volume, 1.0 s on: 1.0 s off, microtip, 40% amplitude, Vibracell Sonics, USA). The samples were centrifuged (13000 RPM, 45 min, 4°C) and the cell lysate collected.

IMAC gravity flow columns were packed with 4 mL HisPur cobalt resin slurry (Thermo Scientific, USA) washed with 10 CV MilliQ and equilibrated with 10 column volumes (CV) wash buffer. Cell lysate was applied to the column and washed with 10 CV wash buffer before being eluted with elution buffer (1xPBS, 300 mM imidazole, pH 7.5) in 10 fractions of 500 µL. The protein concentration was estimated using absorbance at 280 nm by an UV/Vis spectrophotometer. The 5 fractions with highest A280 were pooled and the buffer was exchanged to 1xPBS via a gravity flow PD10 desalting column (Sephadex G-25, Cytiva, Uppsala, Sweden).

2.3.3 SDS-PAGE

IMAC elution fractions and buffer-exchanged proteins were verified by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) under reducing conditions for purity and expected size.

3 ug sample, estimated by approximated concentrations above, were mixed with reducing loading dye (50% SDS loading dye, 50 mM TCEP, 40% glycerol) and heated (95°C, 10 min) before running in a pre-cast NuPAGE (Invitrogen, USA) gel (200 V, 35 min, 1x MES running buffer). A LMW standard was used as a protein ladder (Cytiva, Uppsala, Sweden). The SDS-PAGE gel was stained using Coomassie blue.

2.4 Mass Spectrometry

In addition to SDS-PAGE, the sizes of the produced proteins were investigated by matrix assisted laser desorption ionization time of flight (MALDI-TOF) mass spectrometry.

Samples were mixed with equal amounts of MALDI matrix and dried on a MALDI plate. For samples larger than 10 kDa, 5 mg/mL sinapinic acid in 50% acetonitrile with 0.1% TFA was used as matrix, while cyano-4-hydroxycinnamic acid was used for samples smaller than 10 kDa. Mass spectra were obtained using a 4800 MALDI-TOF/TOF analyser (Applied Biosystems, USA).

Selected samples were analysed by LC-MS analysis to further verify purity and size. An Impact II electrospray ionization mass spectrometer (Bruker) was used coupled to an Ultimate 3000 HPLC with a ProSwift RP-4H column (Thermo Scientific).

2.5 CD

Secondary structure content and stability were studied with circular dichroism (CD) using a Chirascan CD spectrometer (Applied Photophysics, UK). Spectra of samples (0.4 mg/mL, 1xPBS) were obtained between 195 and 260 nm (1 nm step size, 1 mm quartz cuvette) before and after heating to study secondary structure content and refolding abilities. The melting point of the samples were determined by heating the sample from ambient 20°C to 95°C at a ramp speed of 5°C/min. A melting curve was obtained by measuring the signal at 221 nm every 1°C increase.

2.6 SPR

Non-covalent interactions between proteins were studied by surface plasmon resonance (SPR) using a Biacore 3000 (Cytiva, Uppsala, Sweden). Ligands were diluted in sodium acetate (10 mM NaOAc, pH 4.5) and immobilized onto a CM5 chip (Cytiva, Uppsala, Sweden) using amine coupling (5 µL/min, EDC + NHS activation, ethanolamine-HCl deactivation). 500 response units (RU) Zwt, 2000 RU Z963, 1600 RU scFv, 3400 RU Fc and 6200 RU IgG (Herceptin) were immobilized in separate flow cells of two sensor chips. Analytes were injected at 200 nM and, for selected samples, as a two-fold dilution series between 0.78 nM and 1600 nM in 1xPBST (1xPBS + 0.005% Tween20, pH 7.4), which was also used as the running buffer. The flow during analysis was 30 µL/min, with an analyte association time of 5 minutes and dissociation time of 10 minutes before regeneration by 10 mM HCl.

2.7 Photoconjugation

Photoconjugation was performed by exposing mixtures of specific proteins to UV light. Successful crosslinking was evaluated using reducing SDS-PAGE. Negative controls comprised similar mixtures without UV exposure.

Samples were mixed to a total volume of 50 µL in 1xPBS at 0.3 µg/µL. Cetuximab (Merck, USA) was used as a representative IgG antibody and mixed with an abundance of affibody in a 1:5 molar ratio (calculated by antibody monomers). For photoconjugation of affibody (Zwt) and its anti-idiotypic affibody (Z963 or variants thereof, Z964), purified proteins were mixed in a 1:1 ratio.

The sample mixture was put in a clear flat-bottomed 96 well plate and exposed to UV light at 365 nm using a UVC 500 Ultraviolet Crosslinker (Amersham bioscience, UK) for 2 hours on ice. Molten ice was replaced with new ice every 30 minutes. After UV irradiation, 10 µL of the photoconjugated sample was used for reducing SDS-PAGE.

Table 1. List of proteins produced in the project.

Protein	Size (kDa)	Description
Zwt	7.8	Affibody with natural affinity for Fc of IgG, Z963 and Z964
Zwt L17BPA	7.9	Zwt with BPA for photoconjugation to Z963
Zwt K35BPA	7.9	Zwt with BPA for photoconjugation to Fc of IgG
Zwt A29G	7.8	Zwt with the natural affinity for Fab of IgG of the original Domain B restored
Z963	7.5	Anti-idiotypic affibody against Zwt
Z963 L34M	7.6	Z963 with a Met as an acceptor for BPA photoconjugation
Z963 L45M	7.6	Z963 with a Met as an acceptor for BPA photoconjugation
Z964	7.5	Anti-idiotypic affibody against Zwt with natural Met as an acceptor for BPA photoconjugation
ZwA3	45	Scaffold for cryo-EM with the affibody Zwt attached by rigid helix fusion, for studying Fc of IgG or anti-idiotypic affibodies Z963 or Z964
ZwA3 A29G	45	ZwA3 with the affibody Zwt A29G
ZwA3 A29G -1	45	ZwA3 A29G with a rigid helix fusion that allows study of Fab of IgG
ZwA3 A29G +3	45	ZwA3 A29G with a rigid helix fusion that allows study of Fab of IgG

3 Results

3.1 *In silico* design

3.1.1 Positions for BPA incorporation in Zwt

Before genetic engineering and protein expression, molecular modelling was used to investigate the protein complexes and to determine ideal sites for BPA incorporation.

Positions in Zwt suitable for BPA incorporation for photoconjugation to Fc were determined via literature review and structural evaluation for photoconjugation of PDB: 5U4Y. From a list of potential candidates, positions L17 and K35 were selected.

K35 was selected based on its documented ability to photoconjugate to a large selection of different Fc variants, including human IgG1 [22] which is used in this study. Photoconjugation using BPA may require a methionine in close proximity for the BPA to crosslink [18,19,22]. K35 is in close proximity to a methionine in many different IgG types (*fig 2, A-C*).

L17 was primarily chosen for BPA incorporation due to its proximity to M35 in Z964 (*fig 2, E*). If the binding surface of Z964 is similar to that of Z963, the constructs with methionine substitution in Z963 would also bind to Zwt L17BPA. This position has also displayed photoconjugation to some IgG types such as mouse IgG2A [22] which has M252 in close proximity.

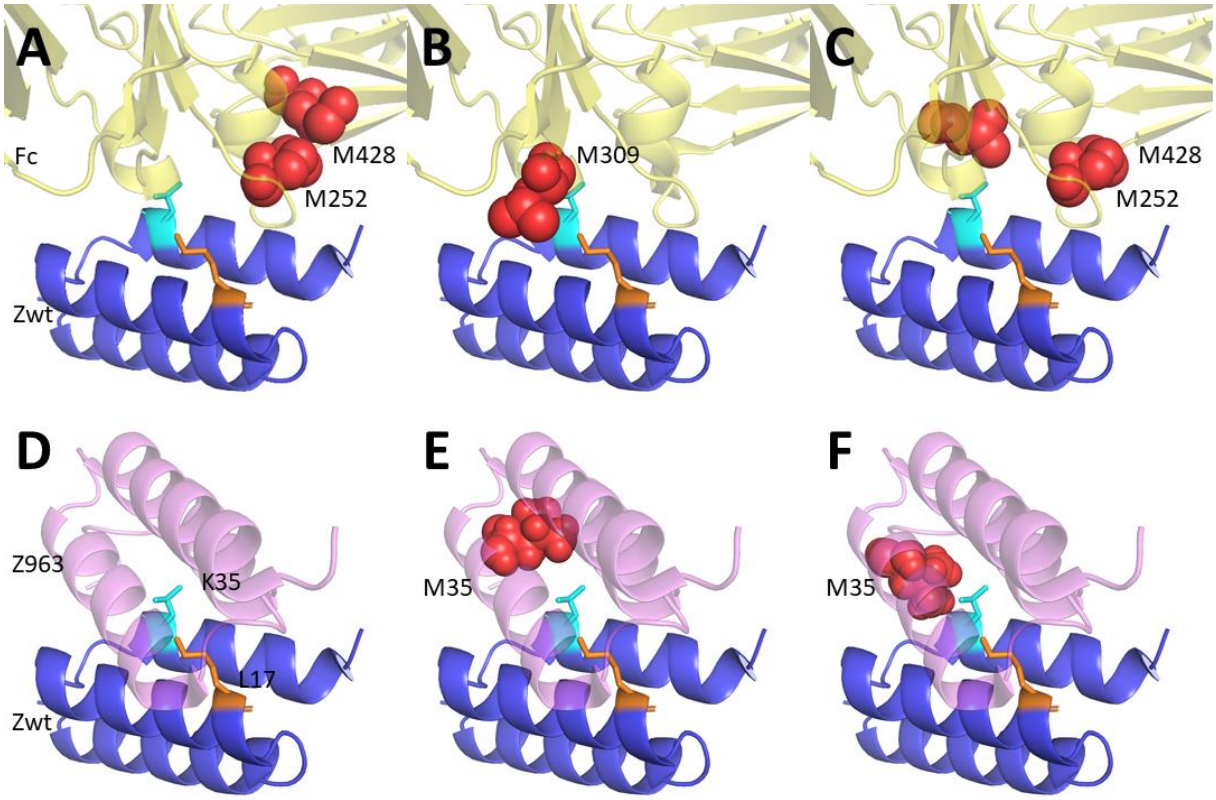


Figure 2. *In silico* inspection of targets to the affibody Zwt (blue). The candidate positions for BPA incorporation in Zwt L17 (cyan) and K35 (orange) are shown. **A:** Human IgG1 (yellow) with M252 and M428 highlighted (red). **B:** Mouse IgG1 M309 highlighted (red). **C:** Mouse IgG2A with M252 and M428 highlighted (red). **D:** Z963 (pink). **E:** Z963 with L34 highlighted (red). **F:** Z963 with L35 highlighted (red). Assembled from PDBs: 5U4Y, 2M5A.

3.1.2 Rigid helix fusion for a more general application

The ZwA3 scaffold has a shared fusion helix between the C-terminal helix III of the Zwt affibody and the N-terminus of aldolase, designed to expose helix I and II on the affibody for binding to targets such as the Fc of IgG or the anti-idiotypic affibody Z963. The binding to VH3 of Fab or scFv affibody fragments for the general application of the scaffold requires binding of helix II and III of the Zwt A29G affibody (*fig 5*). The shared fusion helix design above makes the binding surface in ZwA3 orient itself unfavourably for the intended interactions in this general application. Different scaffold designs are therefore needed for exposing different binding surfaces (*fig 3*).

Two re-designs of ZwA3 aimed to bind VH3 of Fab were made *in silico*: "ZwA3 A29G -1" and "ZwA3 A29G +3", of which the "+3" variant is novel to this study (Figure X, E, F). These have the same starting point of the fusion helix as ZwA3: where the end of the affibody is joined with residue 8 of aldolase. However, the length of the fusion helix between helix III of the affibody and aldolase differ (*appendix, table A1*). Due to the rigid nature of the fusion helix, this turns the angle at which the affibody is attached to the aldolase scaffold (*fig 4*). The "-1" and "+3" variants expose helix II and III rather than I and II, allowing Zwt A29G to accommodate VH3 in the models. These three variants were produced and characterized using SPR, which is described in section 3.5 SPR.

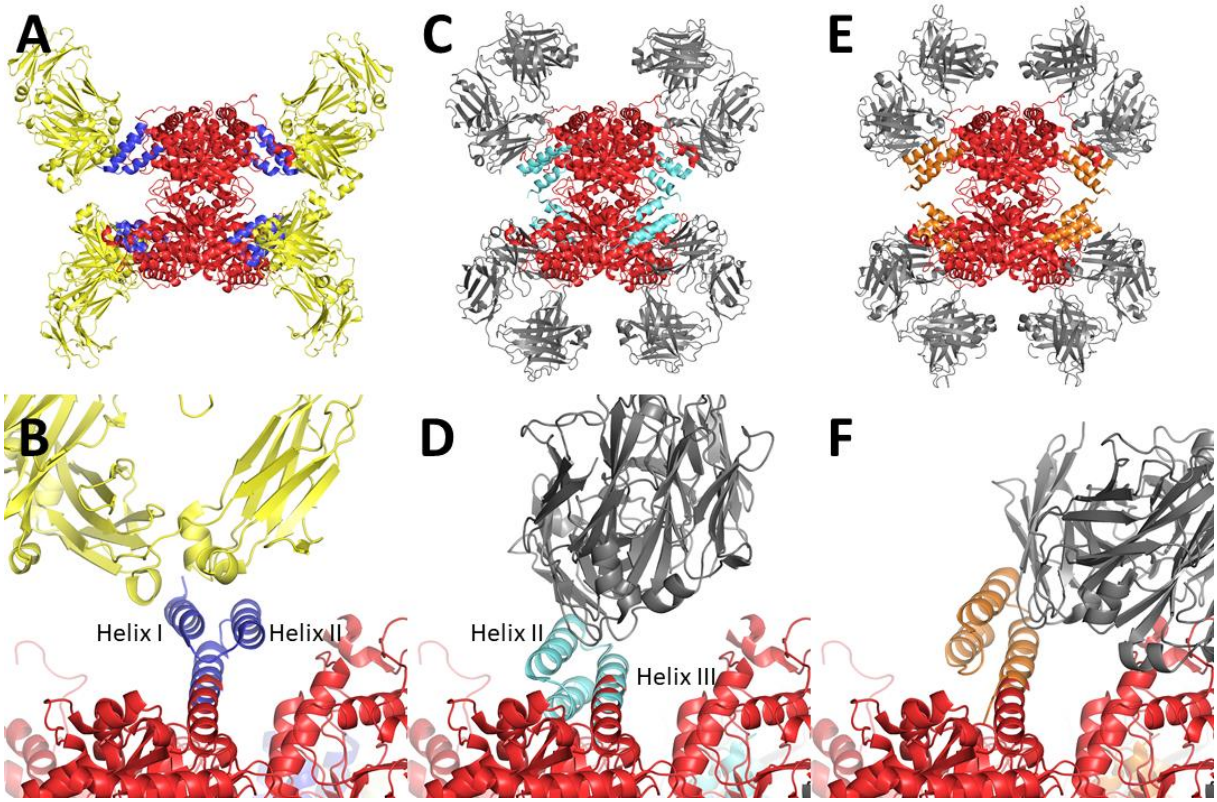


Figure 3. *In silico* designs of the three scaffolds ZwA3, ZwA3 A29G -1 and ZwA3 A29G +3 constructed from crystal structures in PyMOL. The designs differ in the rigid helix fusion from the affibody to aldolase (red). **A:** ZwA3 with the Zwt affibody (blue) bound to Fc of IgG. **B:** closeup of A. **C:** ZwA3 A29G -1 with the Zwt A29G affibody (cyan) bound to Fab of IgG. **D:** closeup of C. **E**: ZwA3 A29G +3 with the Zwt A29G affibody (orange) bound to Fab of IgG. **F:** closeup of E. Assembled from PDBs: 1EWD, 5U4Y, 1DEE.

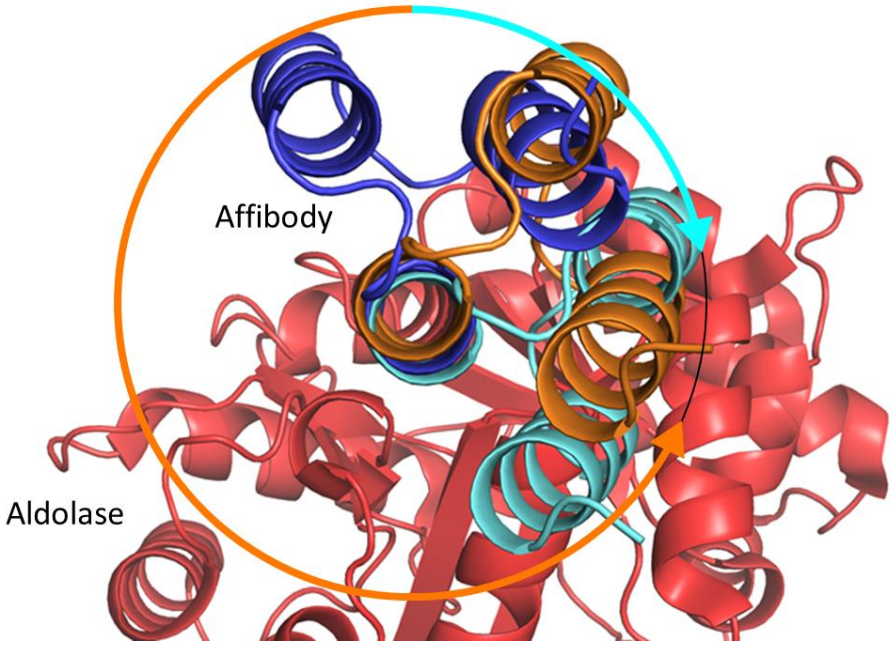


Figure 4. Rotation of the affibody via adjusted amino acid (aa) length of the rigid helix fusion to aldolase (red). With the ZwA3 affibody orientation (blue) as reference, ZwA3 -1 (cyan) is rotated 100° clockwise by the removal of 1 residue in the rigid linker, while the ZwA3 +3 (orange) is rotated 60° clockwise by the addition of three residues in the rigid linker. Assembled from PDBs: 1EWD, 2M5A.

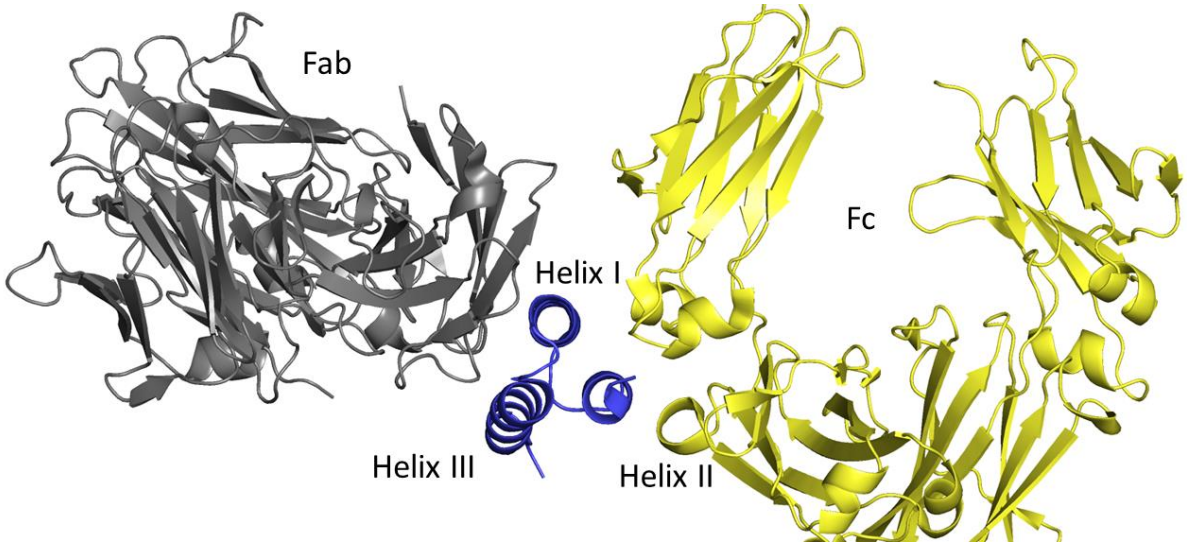


Figure 5. Binding surfaces of Zwt A29G to Fc via helix I and II, and to Fab via helix II and III. The aldolase is in the scaffold design attached to helix III. Assembled from PDBs: 5U4Y, 1DEE.

3.2 Protein expression and purification

Production of protein was successful for most variants with protein yields up to 160 mg protein per litre cultivation (*appendix 8.4*). The induced cultures were harvested and lysed via sonication prior to IMAC purification. The collected IMAC fractions were screened for purity using SDS-PAGE to check for bands corresponding to the expected sizes.

An SDS-PAGE of all the successfully produced proteins can be seen in *figure 6*. For each of the samples, 3 µg as per was loaded onto the gel. The concentrations were estimated by protein absorbance at 280 nm. The intensity of the bands was not the same however, with the Zwt variants weaker than the Z963 variants, Z964 and ZwA3. The other ZwA3 variants also had a lower intensity on the gel.

Except for the bands at the expected sizes of the produced protein, a second band could be seen around 50 kDa for the ZwA3 variants. When the same volume of eluate of Zwt was loaded onto the SDS-PAGE gel, a band of similar intensity could be seen at this size, which was not visible when only 3 µg sample is loaded (*appendix, fig A5*). This band is presumably a native E.coli protein.

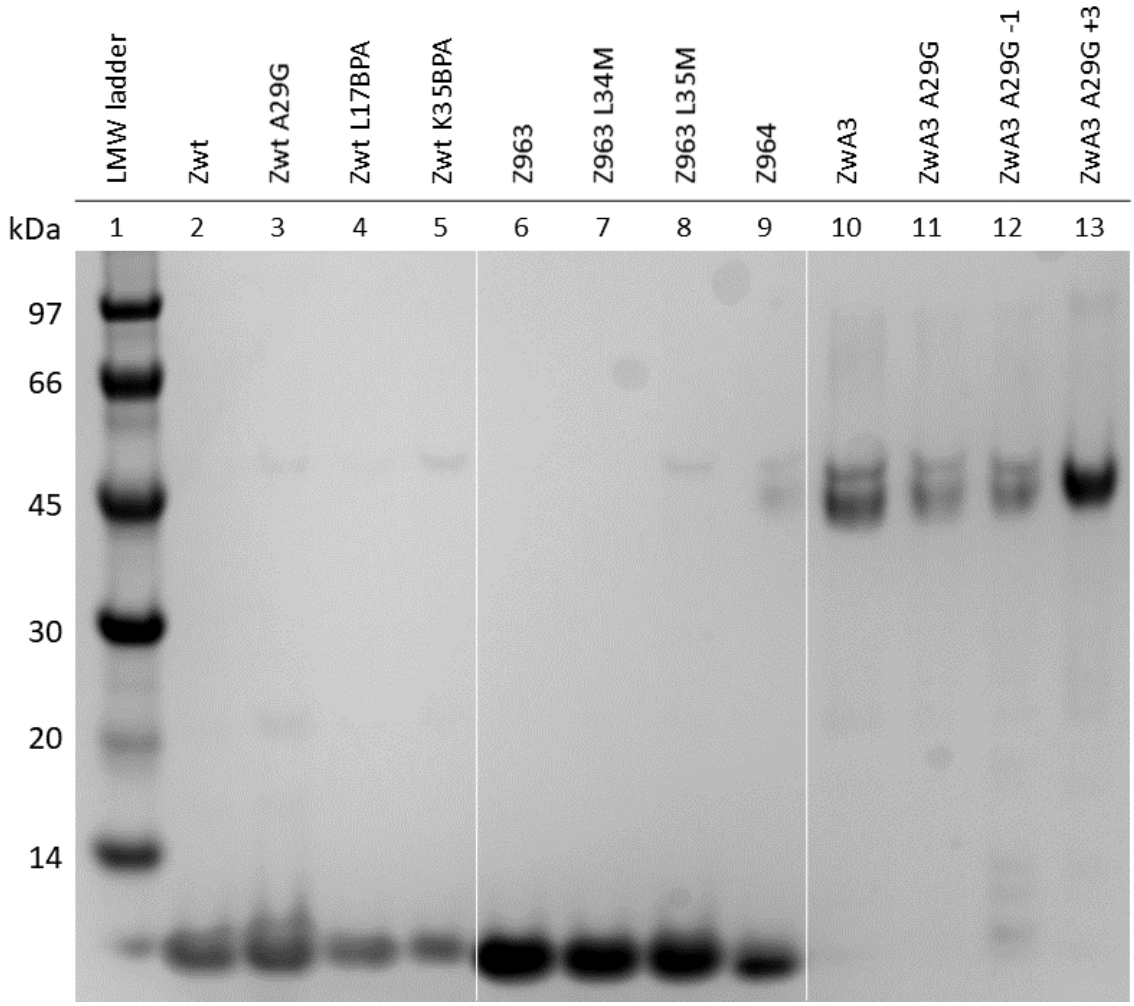


Figure 6. Reducing SDS-PAGE gel of pooled IMAC eluate after buffer change to 1xPBS. 3 µg protein was loaded in each well except for wells 10 and 12 where the maximum volume of 10 µL was loaded due to low concentration.

3.3 Mass Spectrometry

The incorporation of BPA into the Z(xBPA) variants (L17BPA and K35BPA variants of Zwt) was verified via mass spectrometry (*fig 7*). In MALDI-TOF/TOF there was a noticeable shift towards higher mass in the Z(xBPA) samples compared to Zwt, consistent with expectations as BPA is larger than the amino acids it replaces. Both the measured weight and the shifts in weights between the L17BPA or K35BPA variants to Zwt were within 5 Da of their theoretical values. Despite the limited resolution of the MALDI, this is a clear indication that BPA was successfully incorporated at the correct position. This result was confirmed by LC-MS analysis with electrospray ionization (*table 2*).

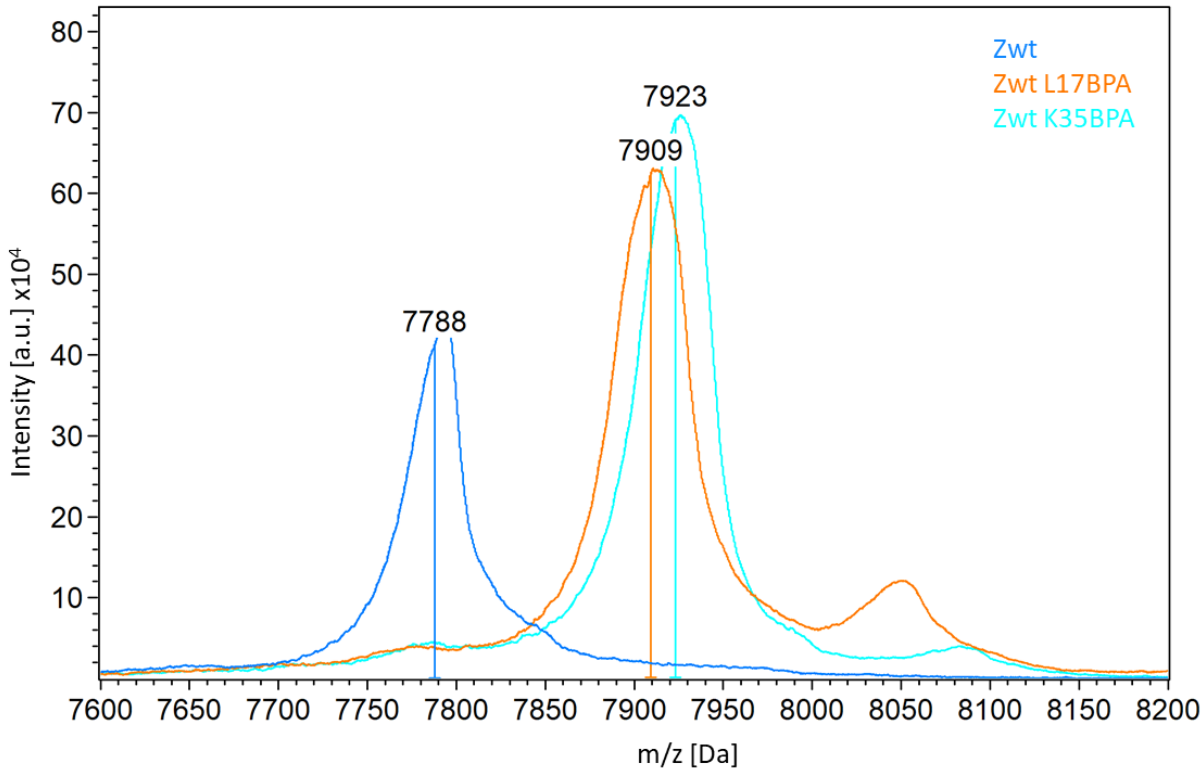


Figure 7. MALDI-TOF/TOF mass spectra of Z(xBPA) variants, focused between 7600 and 8300 Da. Zwt (blue) has a peak at 7790 Da, corresponding to the expected mass. Both Zwt L17BPA (cyan) and Zwt K35BPA (orange) have peaks at higher Da: 7924 respectively 7910, corresponding to their expected masses.

MALDI-TOF/TOF mass spectrometry was also used to evaluate the sizes of affibody variants produced without BPA-incorporation (Zwt, Z963 and Z963 L34M) and ZwA3 variants (ZwA3, ZwA3 A29G, ZwA3 A29G -1 and ZwA3 A29G +3). The measured molecular weight were for these also within a few Da of their expected values (*table 2*).

Table 2. Molecular weights of the produced proteins in Da as measured by MALDI-MS and ESI LC-MS. Samples that were not analysed are marked with N/A. Theoretical molecular weights were estimated using the ExPASy tool ProtParam [31].

Protein	**Theoretical (Da)**	**MALDI (Da)**	**ESI (Da)**
Zwt	7790	7910	N/A
Zwt L17BPA	7928	7924	7927
Zwt K35BPA	7913	7910	7912
Zwt A29G	7749	N/A	N/A
Z963	7537	7538	N/A
Z963 L34M	7555	7556	7555
Z963 L35M	7555	N/A	N/A
Z964	7240	N/A	N/A
ZwA3	45305	45311	N/A
ZwA3 A29G	45291	45269	N/A
ZwA3 A29G -1	45163	45202	N/A
ZwA3 A29G +3	45606	45662	N/A

3.4 Circular Dichroism

To investigate whether the mutations in the affibody variants disrupted the structure or altered stability, the protein secondary structure content was studied using circular dichroism (*fig 8*).

Zwt, Zwt L17BPA and Zwt K35BPA all displayed typical alpha helical spectra. The similarity of the spectra between Zwt and the Z(xBPA) variants indicates that the BPA-carrying proteins have similar secondary structures (*fig 8, A-C*). Variable temperature measurements were performed at 221 nm to obtain melting curves (*fig 8, D*). The thermal stabilities of the Zwt variants were similar to that of the wildtype variant with T_m of ~70-80°C with Zwt K35BPA having the highest T_m at 77°C. Once the samples had cooled down, a second set of spectra were recorded (*fig 8, A-C*). All the Zwt variants had the ability to refold, as the spectra before and after thermal denaturation overlapped.

CD spectra of Z963 and the variants Z963 L34M and Z963 L35M were similar to the Zwt CD spectra with predominantly alpha helical content (*fig 9, A-C*). The spectra returned to their previous shape after heat treatment, although not to the same degree as the Zwt variants with some loss in signal. The melting point of the Z963 variants were lower than Zwt at ~50-55°C, with Z963 (*fig 9, D*).

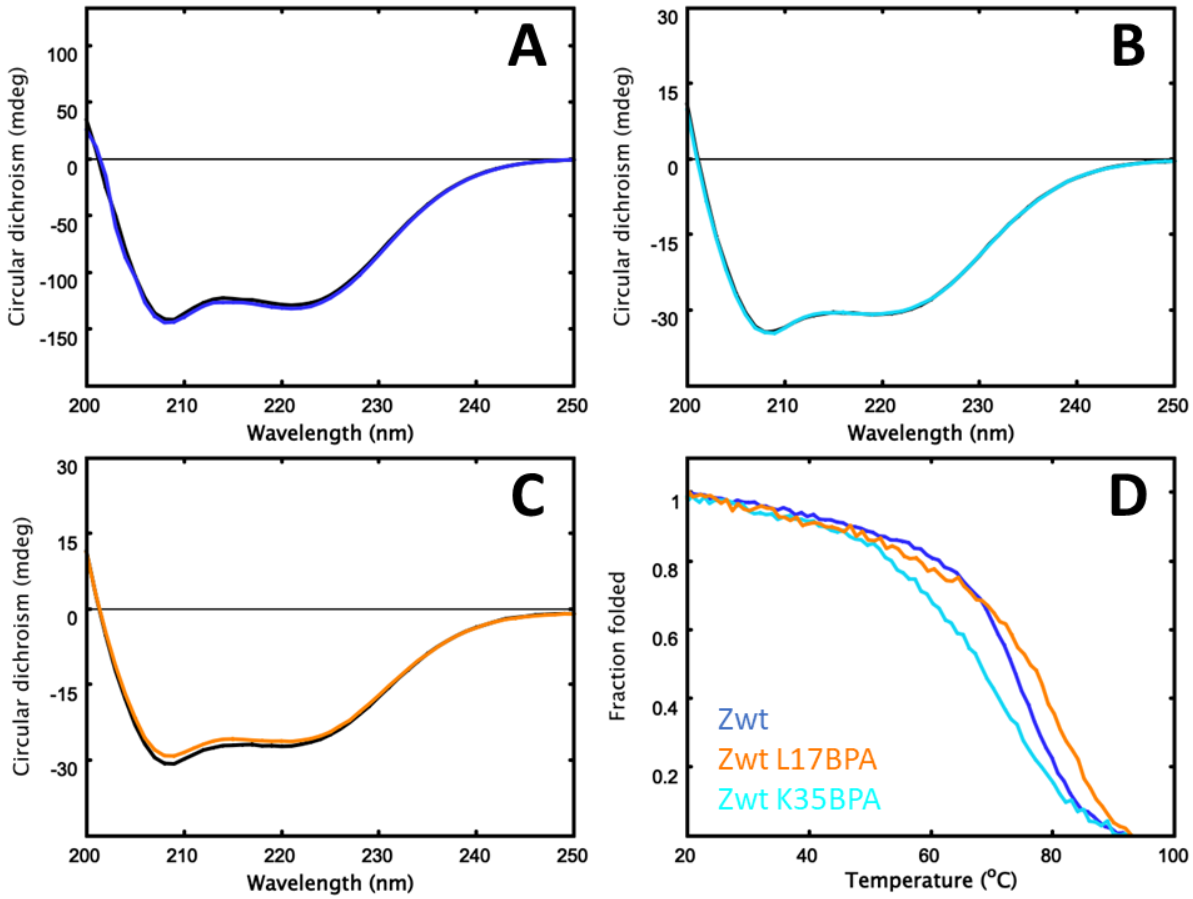


Figure 8. CD spectra of Zwt variants before (black) and after (coloured) heating to 95°C and cooling back down to ambient 20°C. **A**: Zwt. **B**: Zwt L17BPA. **C**: Zwt K35BPA. **D**: Temperature series from 20°C to 95°C with measurements at 221 nm every 1°C increase. The Tm of the Z(xBPA) variants are labelled: Zwt (blue) Tm = 73°C, Zwt L17BPA (cyan) Tm = 68°C and Zwt K35BPA (orange) Tm = 77°C.

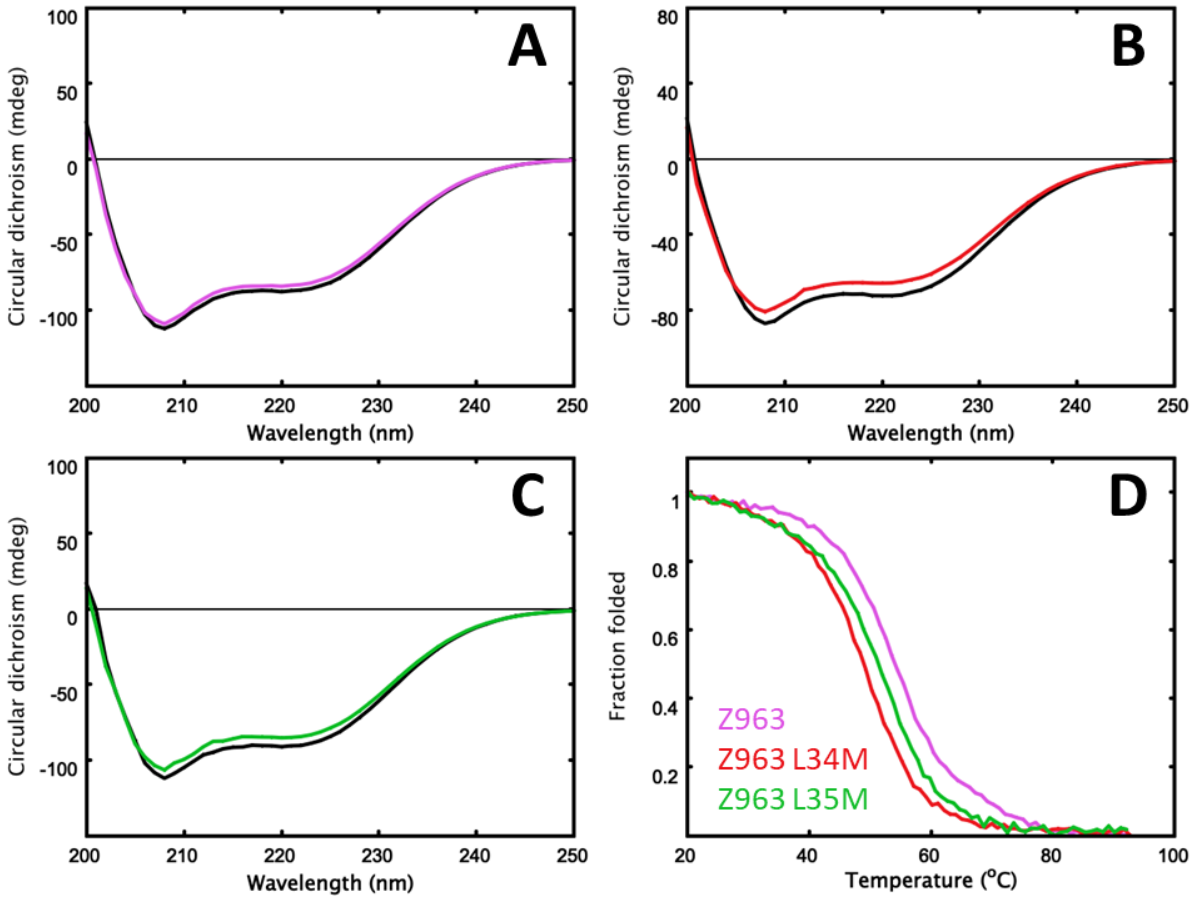


Figure 9. CD spectra of Z963 variants before (black) and after (coloured) heating to 95°C and cooling back down to ambient 20°C. **A**: Z963. **B**: Z963 L34M. **C**: Z963 L35M. **D**: Temperature series from 20°C to 95°C with measurements at 221 nm every 1°C. The Tm of the Z963 variants are labelled: Z963 (purple) Tm = 55°C, Z963 L34M (red) Tm = 49°C and Z963 L35M (green) Tm = 51°C.

3.5 SPR

The binding characteristics of the produced proteins to each other were studied using surface plasmon resonance (SPR).

3.5.1 Binding to Zwt

Zwt was immobilized (500 RU) on the SPR chip as a ligand and Z963 and derivatives thereof Z963 L34M, Z963 L35M, and Z964 were injected as analytes at 200 nM. The curves for all the samples were similar, with fast on-rate and strong signal (*fig 10*). No binding could be seen for Zwt.

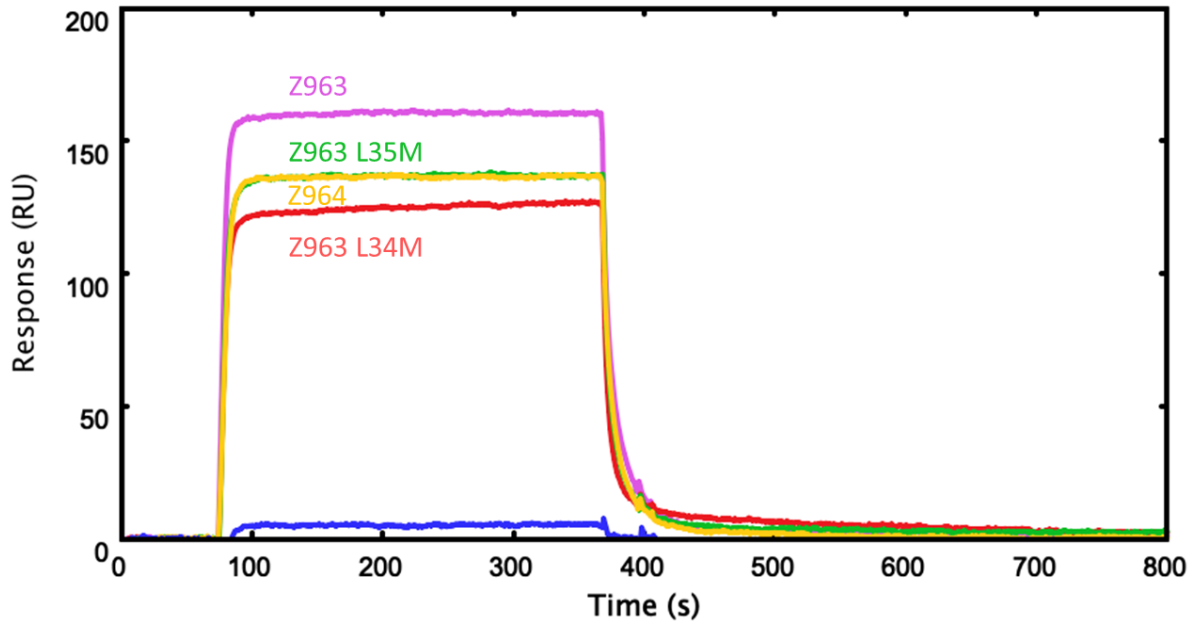


Figure 10. Binding to immobilized Zwt by SPR. Sensogram of Z963 (purple) and variants L34M (red), L35M (green) and Z964 (yellow).

A two-fold dilution series from 0.8 nM to 1.6 µM was used for determining kinetic constant. K_D values for the Z963 variants were estimated (*table 3*) using kinetics non-linear regression model (Langmuir 1:1).

Table 3. K_D values for Z963 variants and Z964 binding to Zwt calculated by non-linear regression using Langmuir 1:1 kinetics.

Protein	**K_D (nM)**
Z963	35
Z963 L34M	62
Z963 L35M	70
Z964	66

3.5.2 Binding to Z963

Z963 was immobilized (2000 RU) on the SPR chip as a ligand. Technical replicates were not consistent in the level of signal, so only qualitative conclusions were drawn from data, and only the first replicate is displayed (*fig 11, 12*)

Zwt and derivatives Zwt L17BPA, Zwt K35BPA, and Zwt A29G were injected as analytes at 200 nM. All the Zwt variants were found to bind to Z963, with Zwt and Zwt A29G displaying the strongest signal (*fig 11*).

ZwA3 and its variants ZwA3 A29G, ZwA3 A29G -1 and ZwA3 A29G +3 were injected as analytes at 200 nM. A strong binding and slow off-rate for all the samples can be seen at a significantly higher signal than the Zwt variants due to their larger size (*fig 12*).

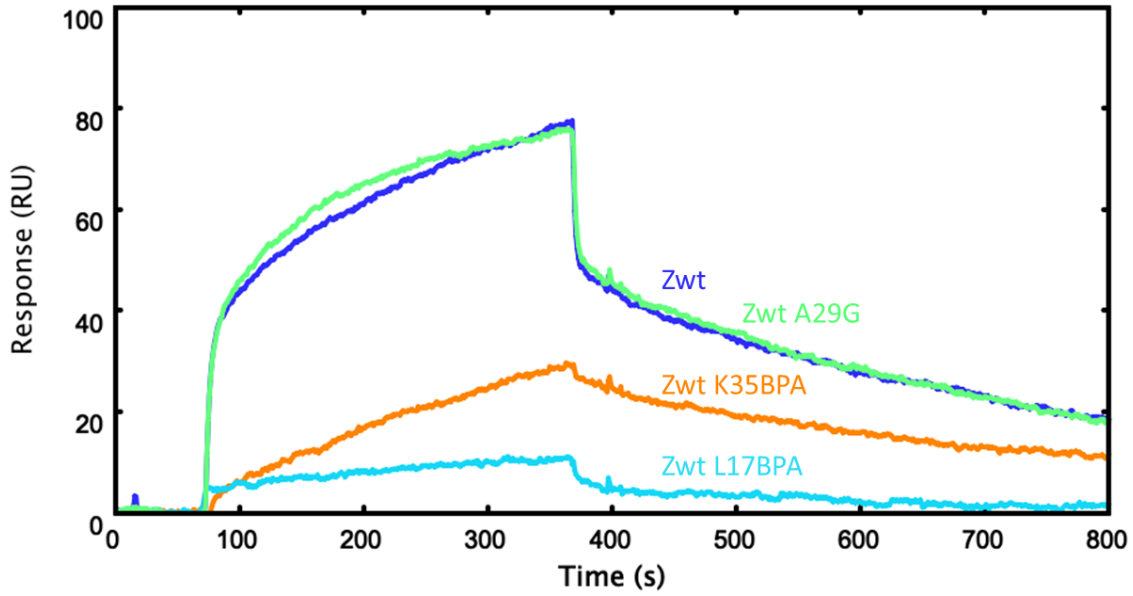


Figure 11. Binding to immobilized Z963 by SPR. Sensogram of Zwt (blue) and Z(xBPA) variants L17BPA (cyan), K35BPA (orange) and Zwt A29G (green).

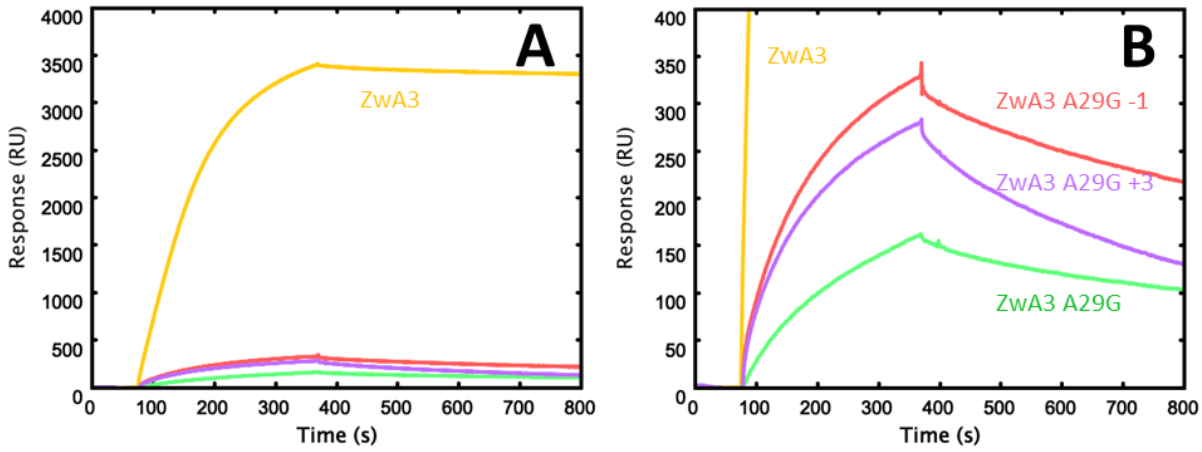


Figure 12. Binding to immobilized Z963 by SPR. Sensogram of ZwA3 (yellow), ZwA3 A29G (green), ZwA3 A29G -1 (red) and ZwA3 A29G +3 (purple). **A:** RU between 0 to 4000 in view. **B:** RU between 0 to 400 in view.

3.5.3 Binding to antibody fragments

Antibody fragments scFv (1600 RU) and Fc (3400 RU), and the IgG antibody Herceptin (6200 RU) were immobilized by amine coupling on the SPR chip as ligands (*appendix 8.8*). 1.25 µM of the ZwA3 variants "ZwA3 A29G -1", "ZwA3 A29G +3", and ZwA3 as well as 2.5 µM of Zwt and Zwt A29G were injected as analytes.

ZwA3 A29G -1 and ZwA3 A29G +3 displayed affinity towards scFv, while ZwA3 did not (*fig 13, A*). Similarly, the Zwt A29G bound to scFv while Zwt did not (*fig 13, B*). Binding of all the analytes could be seen to immobilized Fc (*fig 13, C*) and IgG (*fig 13, D*).

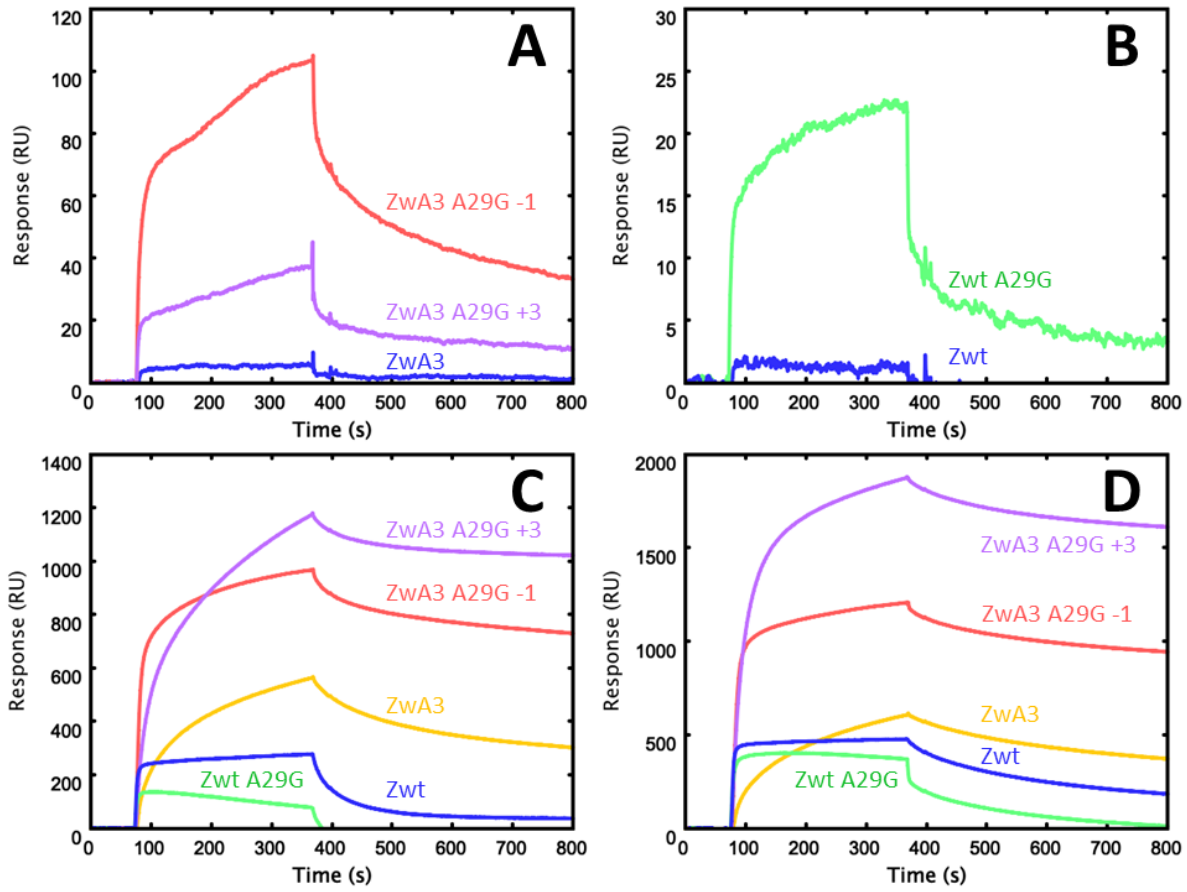


Figure 13. Sensograms of binding to immobilized antibody fragments by SPR. **A:** scFv binding of ZwA3 (blue), ZwA3 A29G -1 (red) and ZwA3 A29G +3 (purple). **B:** scFv binding of Zwt (blue) and Zwt A29G (green). **C:** Fc binding of ZwA3 (yellow), ZwA3 A29G -1 (red), ZwA3A29G +3 (purple), Zwt (blue) and Zwt A29G (green). **D:** IgG binding of ZwA3 (yellow), ZwA3 A29G -1 (red), ZwA3A29G +3 (purple), Zwt (blue) and Zwt A29G (green).

3.6 Photoconjugation

Photoconjugation of the Z(xBPA) variants Zwt L17BPA and Zwt K35BPA to IgG and variants of the anti-idiotypic affibody Z963 were carried out under UV light (365 nm, 2 hours), and analysed via SDS-PAGE and MALDI mass spectrometry.

3.6.1 Z(xBPA) against both IgG and Z963 variants

1000 µM Zwt L17BPA, Zwt K35BPA or Zwt (as a negative control) was mixed with 200 µM of either IgG, Z963 L34M or Z963, resulting in a molar ratio of 5:1 for UV-conjugation. The ability to undergo photoconjugation by Zwt L17BPA and Zwt K35BPA to IgG or Z963 was analysed by a reducing SDS-PAGE gel (*fig 14*). As a size and intensity reference, each protein was also loaded on the SDS-PAGE gel on their own. Successful photoconjugation would be visible on the gel as a new band above the band for the heavy chain of IgG or monomeric affibody, with the combined size of the two proteins. Since the SDS-PAGE was reducing, this band would only show up if the two were covalently bound through photoconjugation.

For the mix of Zwt K35BPA with IgG, a band could be seen above the band for the heavy chain of the antibody. The offset of this band roughly corresponds to the size of an affibody and is likely the crosslinked product. This band was absent from the corresponding Zwt L17BPA sample and Zwt control sample. This difference could also be seen in MALDI mass spectrometry (*appendix 8.10, fig A12*).

A band at roughly twice the size of an affibody was found for Zwt L17BPA mixed with Z963 L34M which was not seen with the Zwt K35BPA variant or the negative control of Zwt mixed with Z963 L34M. This is presumably the hetero-dimer of Zwt L17BPA:Z963 L34M with a covalent crosslink which makes it visible on a reducing SDS-PAGE (*fig 14*). The size observed in MALDI mass spectrometry of a reduced UV sample corresponds only to the heterodimer of Zwt L17BPA:Z963 L34M (*appendix 8.10, fig A13*).

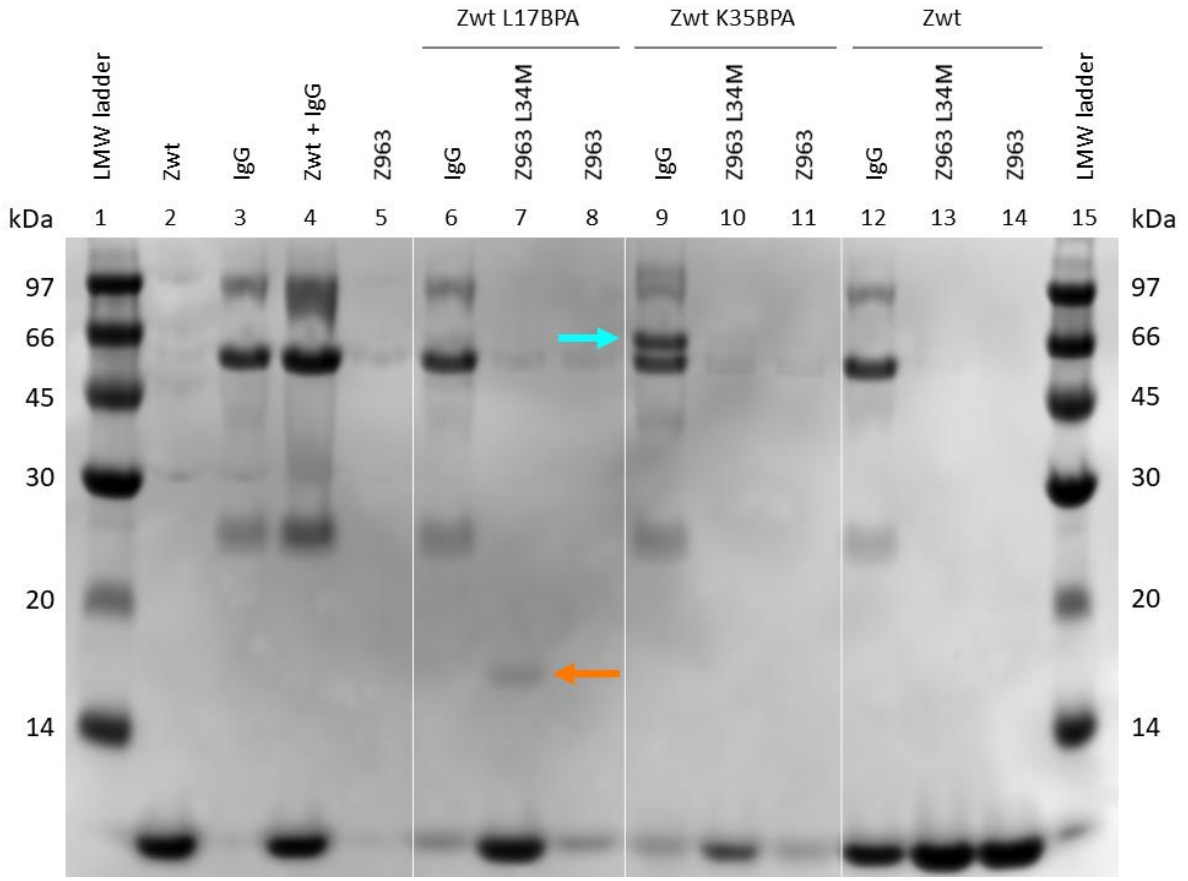


Figure 14. Reducing SDS-PAGE gel of photoconjugated samples. Cyan arrow shows band for IgG cross-linked to Zwt K35BPA. Orange arrow shows band for Z963 L34M cross-linked to Zwt L17BPA. Wells 2-5 are size and intensity reference for the UV exposed protein. Wells 6-14 are UV exposed mixes. **1:** LMW ladder. **2:** Zwt alone. **3**: IgG alone. **4:** Zwt with IgG. **5:** Z963 alone. **6:** Zwt L17BPA with IgG. **7:** Zwt L17BPA with Z963 L34M. **8** Zwt L17BPA with Z963. **9:** Zwt K35BPA with IgG. **10:** Zwt K35BPA with Z963 L34M. **11:** Zwt K35BPA with Z963. **12** Zwt with IgG. **13:** Zwt with Z963 L34M. **14:** Zwt with Z963. **15:** LMW ladder.

3.6.2 Z(xBPA) to different Z963 variants

The degree of photoconjugation of Zwt L17BPA to variants of Z963 was compared on reducing SDS-PAGE (*fig 15*). Zwt K35BPA mixed with IgG was used as a positive control, and 1000 µM Zwt L17BPA was mixed with either Z963, Z963 L34M, Z963 L35M or Z964 in a 1:1 ratio. Zwt without any BPA was mixed with the same Z963 variants as a negative control.

The band at 17 kDa representing the crosslinked Zwt:Z963 heterodimer is present for all the UV treated samples with Zwt L17BPA mixed with equal molar amounts of any Z963 variant and generally absent in samples that contain Zwt without any BPA (*fig 15*).

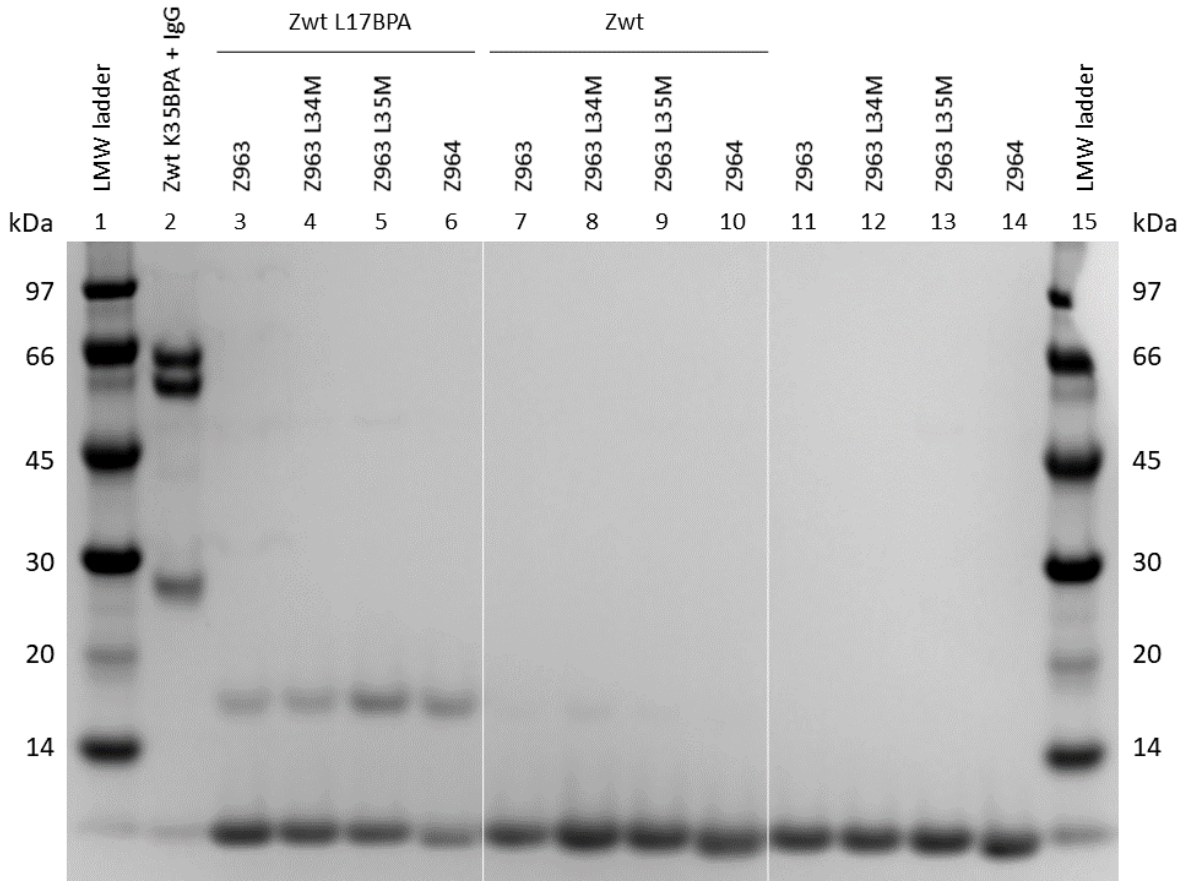


Figure 15. Reducing SDS-PAGE gel of photoconjugated samples. Wells 3-6 are UV exposed mixes with Zwt K35BPA. Wells 7-10 are UV exposed control mixes with Zwt. **1**: LWM ladder. **2:** Zwt K35BPA with IgG. **3**: Zwt L17BPA with Z963. **4:** Zwt L17BPA with Z963 L34M. **5:** Zwt L17BPA with Z963 L35M. **6:** Zwt L17BPA with Z964. **7:** Zwt with Z963. **8:** Zwt with Z963 L34M. **9:** Zwt with Z963 L35M. **10:** Zwt with Z964.

4 Discussion

The two Z(xBPA) variants were selected *in silico* for BPA mutations in the binding surface of Zwt in order to crosslink to its target when exposed to UV light. The Zwt K35BPA variant was intended to accommodate binding to many IgG types, including the antibody Cetuximab used in this study, while the Zwt L17BPA variant was intended for binding to the anti-idiotypic affibody Z964, on account of the methionine acceptor in position 35 of Z964.

The production of Z(xBPA) variants Zwt L17BPA and Zwt K35BPA was successful with high protein yields and sufficient purity. When loading 3 µg of the protein on an SDS-PAGE, only the expected low band for an affibody could be seen, indicating high purity (*fig 6*). Size verification was performed via both MALDI-TOF mass spectrometry, and LC-MS (*table 2*). The reported size corresponded to the expected size, which in the case of BPA incorporation is important. The pEVOL system with orthogonal tRNA and aminoacyl-tRNA synthetase has been reported to substitute BPA with arbitrary amino acids at the amber stop codon when BPA was supplied in insufficient quantities [12]. The reported molecular weight being within 5 Da for MALDI-TOF MS with a relatively homogenous peak, with exception for a small peak ~140 Da above the main peak with no match for known modifications or amino acid substitutions (*fig 7*) and within 1 Da for LS-MS with only a single mass detected confirms that the protein had been produced correctly.

Proteins displaying the correct size in an SDS-PAGE after being reduced and heated, or as ionized particles in a mass spectrum does not guarantee that they are functional. CD was therefore used to verify that they had the correct secondary structure and investigate the thermal stability (*fig 8*). The spectra had the same shape typical of an alpha helical protein, which is the predominant structure of an affibody and therefore expected. The shape of the spectra for Zwt L17BPA and Zwt K35BPA was similar, indicating a correct secondary structure as the BPA mutations are in the binding surface and should not interfere with the overall folding.

Melting curves between 20°C and 95°C revealed that the thermal stability had not been negatively impacted by the BPA mutation. In fact, Zwt K35BPA had a higher melting point than Zwt at 77°C compared to 73°C. Zwt L17BPA had a T_m of 68°C. The melting point does not directly correlate to the stability of the protein but is a good indicator that the BPA mutation did not disrupt the secondary structure. The refolding capacity was also good, with a very minute shift from the original spectrum after heating to 95°C and cooling to ambient 20°C.

With the purity, size and folding verified, a molar abundance of Z(xBPA) variants was mixed with IgG (Cetuximab) and irradiated with UV light (*fig 14*). Zwt K35BPA was found to cross-link to IgG extensively, seen as a second band above the band of the heavy chain on an SDS-PAGE. The Fc region of the chimeric antibody Cetuximab is a human Fcγ1 domain [28] which is present in human IgG1[29] and consistent with literature where Zwt K35BPA was found to photoconjugate to hIgG1 [22]. Using image analysis, the intensity of the photoconjugated band was almost exactly the same intensity as the heavy chain by itself, resulting in a conjugation efficiency of ca 50%. One Zwt K35BPA was crosslinked to one complete IgG on average. The intensity analysis reassuringly showed double the intensity of bands for the heavy chain where photoconjugation did not occur.

Zwt L17BPA did not show any photoconjugation to hIgG1. This was also an expected result as photoconjugation of Zwt L17BPA to IgG has been observed to mouse IgG1 or IgG2A, but not human IgG1 [22].

Several different containers for the photoconjugation were investigated to optimize the conditions (*Appendix 8.9*) now that the positive control was successful. A range of 96-well plates were tested, and it was concluded that the material or purpose for the plate was irrelevant as long as the container was transparent. The most important factor for UV photoconjugation therefore seems to be the amount of light that reaches the sample rather than the material or coating of the plate. The high binding coating on some of the plates could also have been deteriorated by the UV radiation and not been a factor after the first run.

The anti-idiotypic affibody Z963 intended to photoconjugate to the affibody Zwt both in solution as a control and fused to ZwA3 as a proof-of-concept lacks a methionine in its binding surface. Since literature [20,22,25] implies that this is a necessity for BPA photoconjugation, the related Z964 with a methionine in position 35 was also investigated. Very little work had been done on this variant however, and no structure was available. From its similarity in key residues (commonly randomized in affibody libraries) to Z963 (*fig 1*), it was assumed that the binding surface to Zwt was similar and two new Z963 variants were made *in silico*: Z963 L34M and Z963 L35M. Production of Z963 variants Z963 L34M and Z963 L35M was successful, with size and purity verified in the same way as the Zwt variants (*fig* 6).

The Z963 L35M mutation was in the binding surface of the affibody and was assumed to not affect the folding of the protein. The Z963 L34M mutation was however oriented inwards towards the middle of the three-helix bundle. The secondary structure and thermal stability were investigated in CD, where the structure of Z963 and its variants spectra similar to Zwt, with relatively high refolding capacity after heating and cooling to room temperature. The melting point was for all three Z963 variants significantly lower than the Zwt variants, indicating a lower stability for the protein. Z963 L34M did not differ much from the other two variants, and the L34M substitution in the core of the protein did not seem to be detrimental.

Z(xBPA) variants Zwt L17BPA and Zwt K35BPA were mixed with equal molar amounts of Z963 variants, and Z964. The novel photoconjugation of Z(xBPA) variants to an anti-idiotypic affibody was successful for Zwt L17BPA, with all the Z963 variants and Z964 displaying some level of crosslinking while no second band could be seen for Zwt K35BPA (*fig 15*). The highest degree of photoconjugation of Zwt L17BPA was to Z964, which has a naturally occurring methionine at position 35 and Z963 L35M, with the methionine inserted by site directed mutagenesis. Following are the Z963 variants L34M and Z963 with a slightly weaker band at 17 kDa. The slightly higher conjugation efficiency of the variants with M35 was expected as this is the theorized optimal position for cross-linking to Zwt L17BPA. The Z963 L34M variant has the methionine oriented away from the interface surface, but is present, nonetheless. Studies have shown that BPA can move a Met residue out of the way to reach a more optimal residue buried deeper in the target protein [25].

The unaltered Z963 was included as a negative control and not expected to photoconjugate to Zwt L17BPA as it does not contain the necessary methionine. The secondary structure of the protein is disrupted on a reducing SDS-PAGE, eliminating any affinity interactions which means that only crosslinked product can show up at that size. This suggests that BPA

photoconjugation can occur without the need for a methionine acceptor, which has been implied [20,22,25] but not previously shown empirically.

Without a negative control, the possibility of Zwt L17BPA forming a homodimer and crosslinking to each other cannot be ruled out from the SDS-PAGE alone. A reduced UV irradiated sample of Zwt L17BPA and Z963 L34M was examined via MALDI-TOF mass spectrometry (*appendix 8.10*) and a peak identified that could only be the heterodimer of Zwt L17BPA and Z963 L34M. As the sample was reduced, only covalent interactions were observed which supports the claim that no methionine acceptor is required.

The affinity for the affibodies to each other was investigated in SPR. Zwt and Z963 were immobilized as ligands, and the Z963 and Zwt variants injected as analytes. All the Z963 variants and Z964 bound to Zwt with high affinity, with a high signal-to-noise ratio of the data and consistent curves in the dilution series (*fig 10*). Of the anti-idiotypic affibodies, Z963 bound the strongest to immobilized Zwt with a K_D of 35 nM, consistent with the reported affinity of 48 nM [26]. Z963 L35M and Z964 displayed identical binding curves and comparable K_D values (at 70 nM and 66 nM respectively), with Z963 L34M displaying the weakest binding at 62 nM. Reported by Lindborg [26], Z964 had a faster off-rate than Z963 and a weaker binding was expected. This data suggests that M35 could be the main contributor to the difference in affinity between the two affibodies. The Z963 L34M binding was the weakest. No Zwt variant was found to bind to the Zwt surface. While some signal could be seen during sample injection, as soon as PBST was injected, the signal dropped to the baseline, indicating that it was simply particles being close to the surface while not binding.

All the Zwt variants displayed affinity to Z963, however the data was here not as easy to interpret (*fig 11*). The signal-to-noise ratio was significantly lower, with a lower RU signal for all samples. Initially, 500 RU of both Zwt and Z963 were immobilized in separate flow cells, but the Z963 displayed a much lower signal. On a third flow cell, 2000 RU Z963 was immobilized from which the data stems from, however the signal was still relatively low. The fraction of functional Z963 on the surface was presumably much lower, and the surface deteriorated over the SPR run, resulting in technical replicates differing significantly from each other. Only qualitative statements are therefore made for binding on this surface.

All the Z(xBPA) variants bound to the Z963 surface, with Zwt displaying significantly higher affinity than Zwt L17BPA and Zwt K35BPA. This is not surprising, as the large BPA mutation is in the binding surface of the affibody, presumably disrupting some interactions. Both the L17 and K35 positions are randomized in affibody libraries, and the original amino acids could also have a large contribution to the affinity towards Z963.

There were some discrepancies in amplitude of the CD signal among both Zwt variants and Z963 variants (*fig 8, 9*), which could indicate changes in folding, but as CD is a concentration dependent method it is presumably an effect of different concentrations of the samples. SPR is also a method where accurate working concentrations is crucial for determining kinetic constants. An SDS-PAGE of all the proteins produced was made and 3 µg of sample was loaded (*fig 6*). However, when imaging the gel, it became apparent that the intensities of the bands were not all the same. The bands for the Zwt variants were weaker than the Z963 variants (and Z964), resulting in overestimating the actual concentration of Zwt variants and underestimating Z963 variants. The intensity of bands for the ZwA3 variants were closer to the 3 µg, but not consistent among the samples. The volume required was calculated based on A280

measurements and it is possible that this method for concentration determination was not accurate enough or that the extinctions coefficients used to derive the concentrations from A280 was inaccurate. Another possibility however is that the SDS-PAGE gel is misleading, and certain proteins are stained disproportionately to their concentration by Coomassie blue.

Scaffold variants were designed *in silico* and the site directed mutagenesis successfully performed to create the plasmids. ZwA3 would accommodate the binding to Fc of IgG and the anti-idiotypic affibody Z963 with its Zwt affibody domain. ZwA3 A29G, ZwA3 Z29G -1 and ZwA3 Z29G +3 had the Zwt A29G affibody domain to bind to the VH3 of Fab or scFv, however ZwA3 A29G did not expose the correct binding surface. ZwA3 Z29G -1 and ZwA3 Z29G +3 were therefore designed *in silico* to change the orientation of the affibody domain.

The production of scaffolds with canonical amino acids (ZwA3, ZwA3 A29G, ZwA3 Z29G -1 and ZwA3 Z29G +3) gave lower yields than the production of affibodies, but still high enough to provide a working concentration. The production of ZwA3 scaffolds with BPA incorporated in the same positions as the Z(xBPA) variants at L17 and K35 was ultimately unsuccessful. After numerous iterations and troubleshooting, this production was abandoned *in lieu* of time. The analysis therefore focused on the binding of the produced scaffolds, as the photoconjugation of the affibody domain already had been demonstrated.

The affinity of the scaffolds were investigated in SPR. Affinity for immobilized Z963 was seen for all scaffold variants (*fig 12*), with ZwA3 potentially displaying the strongest binding. ZwA3 A29G bound as expected, since affibody orientation was the same as ZwA3, and the A29G mutation is not in its binding surface to Z963. Surprisingly however, both ZwA3 A29G -1 and ZwA3 A29G +3 displayed affinity towards Z963 even though the affibody orientation should have obscured that binding surface. This could mean that the rigid helix has some flexibility that can cause helix I and II of the affibody to be exposed rather than helix II and III.

A scFv, an Fc-fusion and the IgG Herceptin were immobilized as ligands onto another SPR surface, and the scaffold variants injected as analytes (*fig 13)*. The binding of the scaffolds gave a high RU signal, with a slow off-rate due to the large size of the scaffold and the tetrameric structure providing avidity for the binding. The ZwA3 A29G variants (ZwA3 A29G, ZwA3 A29G -1 and ZwA3 A29G +3) all displayed affinity towards the scFv. The fact that ZwA3 A29G still bound to scFv even though the VH3 binding surface is obstructed is consistent with the previous results of the contrary binding to Z963. ZwA3 did however not bind as the A29G mutation is crucial for VH3 affinity. This is further supported by the fact that Zwt A29G bound to the scFv, while Zwt did not. Similar to the affinity of all the Zwt variants to Z963, all of the scaffolds bound to immobilized Fc. This makes interpretation of IgG binding difficult, as the complete antibody contains both an Fc region and VH3, for which the resulting affinity is a combination of the two.

In conclusion, photoconjugation could be observed for Zwt K35BPA to the Fc of hIgG and for Zwt L17BPA to the anti-idiotypic affibodies Z963 and Z964. The ability for BPA to create crosslinks was assisted by the presence of a methionine acceptor in close spatial proximity to the BPA as seen in Z963 L35M and Z964, but it was not a requirement, demonstrated by the misplaced methionine in Z963 L34M, and the lack of a methionine in Z963.

The affinity for the affibodies was measured in SPR, and the mutations in both Z963 and Zwt to accommodate the photoconjugation were found to have some on the binding between them.

The introduction of methionine in Z963 lowered the affinity for Zwt, but only to the same degree as Z964. The introduction of BPA in the binding interface of Zwt did decrease the affinity for Z963, but not to the point of non-binding. Scaffolds were found to bind both Z963 and Fc regardless of the orientation of the affibody domain, indicating some flexibility in the rigid helix fusion, or the aldolase itself. The affinity for VH3 in a scFv was completely dependent on the presence of the A29G mutation in the affibody domain but retained the Z963/Fc binding.

While not produced and studied in cryo-EM, the scaffold ZwA3 was demonstrated to bind to Z963, and the Zwt affibody domain was photoconjugated to Z963 with and without methionine. The assembled proof-of-concept scenario where a small protein, here Z963 at 7.5 kDa, is studied in cryo-EM with the help of a covalent crosslink between the protein of interest and the scaffold is therefore feasible. The scaffolds ZwA3 A29G -1 and ZwA3 A29G +3, but also ZwA3 A29G were found to bind VH3 which is present in both scFv and Fab antibody fragments. The assembled general application scenario where a protein traditionally too small for cryo-EM such as the pharmaceutically interesting scFv or Fab fragments is studied in cryo-EM using the scaffold is feasible. The photoconjugation of the affibody domain may not require a methionine in the antibody fragment, making it a generally applicable method.

5 Future perspectives

This project suggests that the proof-of-concept scenario where a small protein is attached to the scaffold and the general application scenario where an antibody fragment instead is bound to the scaffold and then studied in cryo-EM are feasible. The photoconjugation of the scaffold and its target can allow for higher saturation of the affibody domains, resulting in better cryo-EM resolution. The affibodies can be fused to different scaffold proteins than aldolase to provide other properties such as thermal stability or different symmetry. The flexibility of the rigid helix in ZwA3 could be solved by using a protein with a terminal helix that interacts more with the rest of the protein. A different scaffold with the terminal helix in the C-terminus can allow for helix I of the affibody to construct the shared fusion helix, and potentially avoiding issues with obstructed binding surfaces.

The structural determination of affibody targets can help pharmaceutical development of affibody drugs as their interaction can be studied with high resolution in cryo-EM. With the general application scenario, this can also be used to characterize Fab or scFv antibody fragments and aid in their development as drugs or biotechnological tools. This allows in addition to studying the antibody fragment in cryo-EM, also the interaction between it and its ligand. Not only can any protein be attached to the scaffold as long as there is an affibody for it, but any protein with an antibody targeting it can in a two-step method be attached to the scaffold. The scaffold can be photoconjugated with an scFv or Fab fragment, which can thanks to the covalent bond be purified to create an affinity scaffold without any cloning. This can then in turn bind to the protein of interest and be studied in cryo-EM, creating a platform for rapid structure determination where the scaffold can be mixed with the antibody fragment, irradiated with UV and mixed with the ligand for cryo-EM.

6 Acknowledgements

This project would not have been possible without my supervisor Johan Nilvebrandt, thank you for guiding me and supporting me through the project. Thank you, Kim Anh Giang for always being available for questions and helping me in the lab. To Per-Åke, thank you for your input throughout the project and your creative solutions to unexpected problems. Thank you Ábel Nagy for your expertise and help regarding BPA and for letting me borrow some of your stock. I would also like to thank everyone at Floor 3 of AlbaNova for welcoming me to the floor and providing help when needed. Finally, I would like to thank my co-master's thesis student Hanna Lundahl for all the help in the lab and letting me tag along for your adventures with cryo-EM.

7 References

1. The Nobel Prize in Chemistry 2017 [Internet]. NobelPrize.org. 2017 [cited 31 May 2021]. Available from: https://www.nobelprize.org/prizes/chemistry/2017/8660-pressmeddelande-nobelpriset-i-kemi-2017

2. Hanske J, Sadian Y, Müller C. The cryo-EM resolution revolution and transcription complexes. Current Opinion in Structural Biology. 2018;52:8-15.

3. Kuhlbrandt W. The Resolution Revolution. Science. 2014;343(6178):1443-1444.

4. Wang H, Wang J. How cryo-electron microscopy and X-ray crystallography complement each other. Protein Science. 2016;26(1):32-39.

5. Kermani A. A guide to membrane protein X-ray crystallography. The FEBS Journal. 2020;.

6. The Rise of Cryo-Em Among Structural Characterization Methods: NMR, X-Ray Crystallography [Internet]. Kbdna.com. 2021 [cited 31 May 2021]. Available from: https://kbdna.com/publishinglab/struc-bio

7. Milne J, Borgnia M, Bartesaghi A, Tran E, Earl L, Schauder D et al. Cryo-electron microscopy - a primer for the non-microscopist. FEBS Journal. 2012;280(1):28-45.

8. Fan X, Wang J, Zhang X, Yang Z, Zhang J, Zhao L et al. Single particle cryo-EM reconstruction of 52 kDa streptavidin at 3.2 Angstrom resolution. Nature Communications. 2019;10(1).

9. Wang H. Cryo-electron microscopy for structural biology: current status and future perspectives. Science China Life Sciences. 2015;58(8):750-756.

10. Bergström O. Design, production and evaluation of rigid helix fusion between affibody molecules and framework proteins aimed for Affibody:antigen structure determination using single-particlecryo-EM [Internet] [Dissertation]. 2020. (TRITA-CBH-GRU). Available from: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-278582

11. Yeates T, Agdanowski M, Liu Y. Development of imaging scaffolds for cryo-electron microscopy. Current Opinion in Structural Biology. 2020;60:142-149.

12. Young T, Ahmad I, Yin J, Schultz P. An Enhanced System for Unnatural Amino Acid Mutagenesis in E. coli. Journal of Molecular Biology. 2010;395(2):361-374.

13. Liu Y, Gonen S, Gonen T, Yeates T. Near-atomic cryo-EM imaging of a small protein displayed on a designed scaffolding system. Proceedings of the National Academy of Sciences. 2018;115(13):3362-3367.

14. Yao Q, Weaver S, Mock J, Jensen G. Fusion of DARPin to Aldolase Enables Visualization of Small Protein by Cryo-EM. Structure. 2019;27(7):1148-1155.e3.

15. Coscia F, Estrozi L, Hans F, Malet H, Noirclerc-Savoye M, Schoehn G et al. Fusion to a homo-oligomeric scaffold allows cryo-EM analysis of a small protein. Scientific Reports. 2016;6(1).

16. Leisle L, Valiyaveetil F, Mehl R, Ahern C. Incorporation of Non-Canonical Amino Acids. Advances in Experimental Medicine and Biology. 2015;:119-151.

17. Konrad A, Eriksson Karlström A, Hober S. Covalent Immunoglobulin Labeling through a Photoactivable Synthetic Z Domain. Bioconjugate Chemistry. 2011;22(12):2395-2403.

18. Wittelsberger A, Thomas B, Mierke D, Rosenblatt M. Methionine acts as a "magnet" in photoaffinity crosslinking experiments. FEBS Letters. 2006;580(7):1872-1876.

19. Wittelsberger A, Mierke D, Rosenblatt M. Mapping Ligand–receptor Interfaces: Approaching the Resolution Limit of Benzophenone-based Photoaffinity Scanning. Chemical Biology & Drug Design. 2008;71(4):380-383.

20. Perols A, Karlström A. Site-Specific Photoconjugation of Antibodies Using Chemically Synthesized IgG-Binding Domains. Bioconjugate Chemistry. 2014;25(3):481-488.

21. Stiller C, Aghelpasand H, Frick T, Westerlund K, Ahmadian A, Karlström A. Fast and Efficient Fc-Specific Photoaffinity Labeling To Produce Antibody–DNA Conjugates. Bioconjugate Chemistry. 2019;30(11):2790-2798.

22. Hui J, Tsourkas A. Optimization of Photoactive Protein Z for Fast and Efficient Site-Specific Conjugation of Native IgG. Bioconjugate Chemistry. 2014;25(9):1709-1719.

23. Nilsson B, Moks T, Jansson B, Abrahmsén L, Elmblad A, Holmgren E et al. A synthetic IgG-binding domain based on staphylococcal protein A. "Protein Engineering, Design and Selection." 1987;1(2):107-113.

24. Jansson B, Uhlén M, Nygren P. All individual domains of staphylococcal protein A show Fab binding. FEMS Immunology & Medical Microbiology. 2006;20(1):69-78.

25. Vance N, Zacharias N, Ultsch M, Li G, Fourie A, Liu P et al. Development, Optimization, and Structural Characterization of an Efficient Peptide-Based Photoaffinity Cross-Linking Reaction for Generation of Homogeneous Conjugates from Wild-Type Antibodies. Bioconjugate Chemistry. 2018;30(1):148-160.

26. Lindborg M, Dubnovitsky A, Olesen K, Bjorkman T, Abrahmsen L, Feldwisch J et al. High-affinity binding to staphylococcal protein A by an engineered dimeric Affibody molecule. Protein Engineering Design and Selection. 2013;26(10):635-644.

27. Youn S, Kwon N, Lee J, Kim J, Choi J, Lee H et al. Construction of novel repeat proteins with rigid and predictable structures using a shared helix method. Scientific Reports. 2017;7(1).

28. Cetuximab (Erbitux) [Internet]. Ki.se. 2021 [cited 31 May 2021]. Available from: https://ki.se/en/cns/cetuximab-erbitux

29. Soleimanpour S, Hassannia T, Motiee M, Amini A, Rezaee S. Fcγ1 fragment of IgG1 as a powerful affinity tag in recombinant Fc-fusion proteins: immunological, biochemical and therapeutic properties. Critical Reviews in Biotechnology. 2016;37(3):371-392.

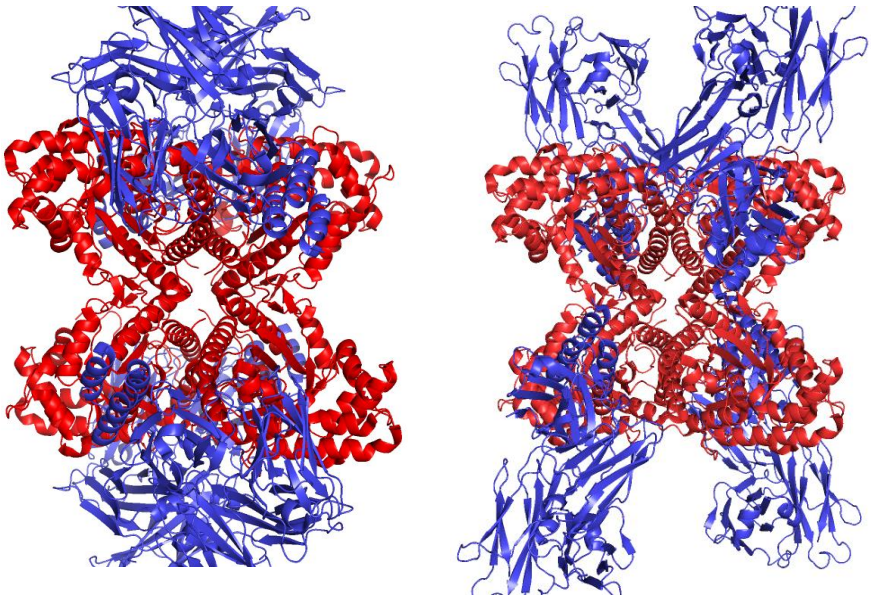
30. PyMOL | pymol.org [Internet]. Pymol.org. 2021 [cited 31 May 2021]. Available from: https://pymol.org/2/

31. ExPASy - ProtParam tool [Internet]. Web.expasy.org. 2021 [cited 31 May 2021]. Available from: https://web.expasy.org/protparam/

8 Appendix

8.1 *In silico* evaluation of scaffold protein design

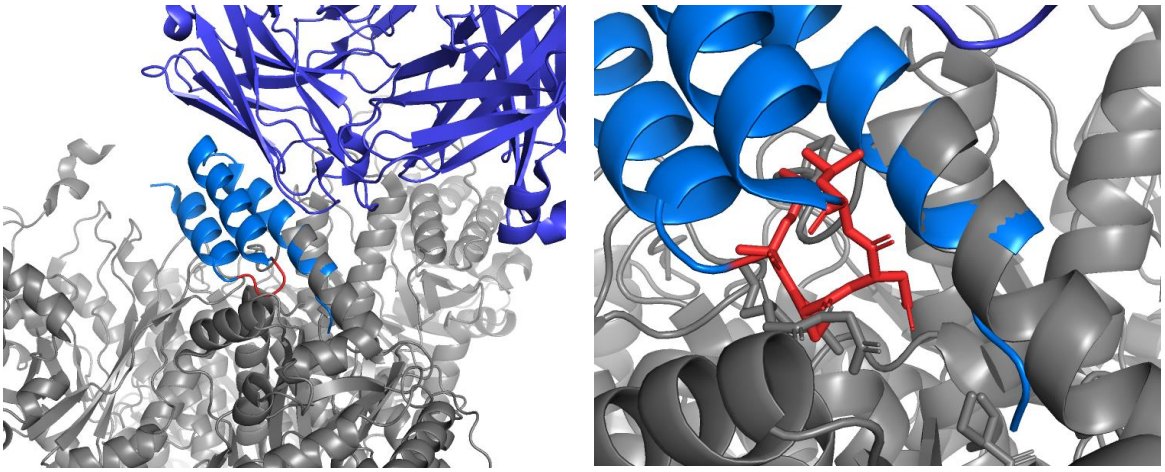
Structures from PDB (1EWD, 5U4Y, 1DEE) were assembled in PyMOL. Using the ZwA3 A29G design to attach Fab fragments to the scaffold uses helix II and III as the binding surface on the Zwt A29G domain, rather than helix I and II, creating massive steric clashes (*fig A1, A*). The ZwA3 A29G -1 design did not have any large overlaps of the models.



A **B**

Figure A1. Superimposition of the scaffold (red) and Fab fragments (blue). **A:** ZwA3 A29G with Fab. **B:** ZwA3 A29G -1 with Fab.

Selecting residues in the affibody withing 3Å of the scaffold reveals minor steric clashes between the loop connecting helix I and helix II and the aldolase molecule for the ZwA3 A29G -1. The offenders in the affibody are highlighted in red, and all the residues close to each other are shown as sticks rather than cartoon (*fig A2*).



A **B**

Figure A2. Superimposition of the ZwA3 A29G -1 scaffold and the bound Fab fragment (blue). Aldolase (grey) has some minor overlap with the attached affibody (light blue). **A:** overview. **B:** closeup of the steric clash.

8.2 *In silico* fusion helix design

An extension to accommodate Fab without clashes can be based on ZwA3 A29G -1 and extend the helix ~1 turn (*fig A3*). The following sequences are displayed in *table A1*.

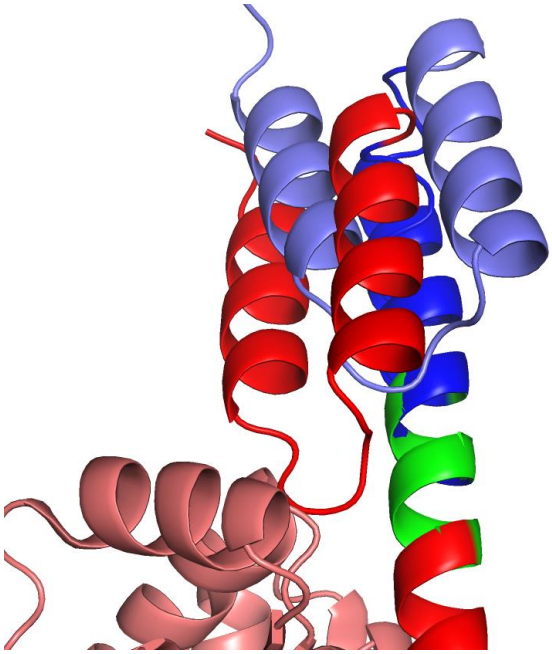


Figure A3. ZwA3 A29G -1 (red) and ZwA3 A29G +3 (blue) attached to the aldolase scaffold (salmon).

Table A1. Helix fusions between Zwt (blue) and aldolase (red). The overlapping helix region is highlighted in green, with the colours corresponding to the origin of the residues.

Name	Fusion helix sequence (5' – 3')
ZwA3	LAEAKK[green]ELSDIA
ZwA3 -1	LAEAKK[green]ELSDIA
ZwA3 +3	LAEAKK[green]ELSDIA

8.3 Site directed mutagenesis

Primers for site directed mutagenesis were designed with a 24 nt overlap either side of the mutation (*table A2*). The codons use can be found in *table A5*. PCR reactions were made (*table A3*) and mutagenesis was performed in a thermocycler (*table A4*) with an annealing temperature of 65°C and 70°C.

Table A2. Primer used for site directed mutagenesis.

Primer	Construct	Sequence (5' – 3')
L17x Zwt Fwd	Zwt L17BPA, ZwA3 L17BPA	CAACAGAACGCGTTCTATGAGATC**TAG**CA TTTACCTAACTTAAACGAAGAA
K35x Zwt Fwd	Zwt K35BPA, ZwA3 K35BPA	CGAAACGCCTTCATCCAAAGTTTA**TAG**GA TGACCCAAGCCAAAGCGCTAAC
A29G Zwt Fwd	Zwt A29G, ZwA3 A29G	AACTTAAACGAAGAACAACGAAAC**GGC**TT CATCCAAAGTTTAAAAGATGAC
L34M Z963 Fwd	Z963 L34M	CAAGTTGCTGCCTTCATCTCTTCT**ATG**TT GGATGACCCAAGCCAAAGCGCT
L35M Z963 Fwd	Z963 L35M	GTTGCTGCCTTCATCTCTTCTTTA**ATG**GA TGACCCAAGCCAAAGCGCTAAC
Helix_Ext_+3 ZwA3 Fwd	ZwA3 A29G +3	CTAGCAGAAGCTAAAAAGCTAAAT[redacted] GAACTGTCTGATATTG CGCATCGT
Helix_Ext_-1 ZwA3 Fwd	ZwA3 A29G -1	CTAGCAGAAGCTAAAAAGCTAAAT[redacted] GAACTGTCTGATATTGCGCATCGT
Zwt-His6 / Z963-His6 Rev	All affibodies	TGGTAGTATTTCGGCGCCTGAGCATCATT TAGCTTTTTAGCTTCTGCTAGC
Zwt-Aldolase Rev	All ZwA3	TCTCTTCAGTATTTTCTGTCCCAATAGAC TGCAGACGTTTGGCGATGCTAC

Table A3. Primer used for site directed mutagenesis.

Component	Volume (µL)
Nuclease free water	14.35
5x Phusion buffer	5
dNTPs (2 mM of each)	2.5
DMSO	0.75
Fwd primer	0.5
Rev primer	0.5
Template plasmid	1
Phusion polymerase	0.4
Total	**25 µL**

Table A4. Thermocycler scheme for site directed mutagenesis.

Temp	95°C	95°C	65/70°C	72°C	72°C	4°C
Time	0:30	0:30	1:00	5:00	7:00	hold
Purpose	Initial denaturation	Denaturation	Annealing	Elongation	Final elongation	
Repeats	1x		30x		1x	1x

Table A5**:** ELLA codons used in primer design

Lys	Asn	Thr	Ile	Met	Gln	His	Pro	Arg	Leu	Glu	Asp	Ala	Gly	Val	Tyr	Ser	Cys	Trp	Phe
K	N	T	I	M	Q	H	P	R	L	E	D	A	G	V	Y	S	C	W	F
AAA	AAC	ACC	ATC	ATG	CAG	CAT	CCA	CGT	CTG	GAA	GAT	GCA	GGT	GTT	TAC	TCT	TGC	TGG	TTC

8.4 Colony PCR screen

Colonies were picked and screened for an insert of the expected size. While the mutagenesis did not introduce changes large enough to be visible on an agarose gel, a PCR screen was still performed.

Table A6. Primer used for colony PCR screen

Primer	**Sequence (5' – 3')**
LaMa27 Fwd	ATCCCGCGAAATTAATACGACTCAC
LaMa14 Rev	ATGCTAGTTATTGCTCAGCGGTGG

Table A7. Primer used for site directed mutagenesis.

Component	**Volume (µL)**
dNTP	12.5
10x DreamTaq buffer	2.5
LaMa27 (5 pmol/µL)*	0.5
LaMa14 (5 pmol/µL)*	0.5
DreamTaq DNA polymerase (5 U/µl)	0.15
MQ	7.35
Colony water	1.5
Total	**25 µL**

Table A8. Thermocycler scheme for colony PCR screen.

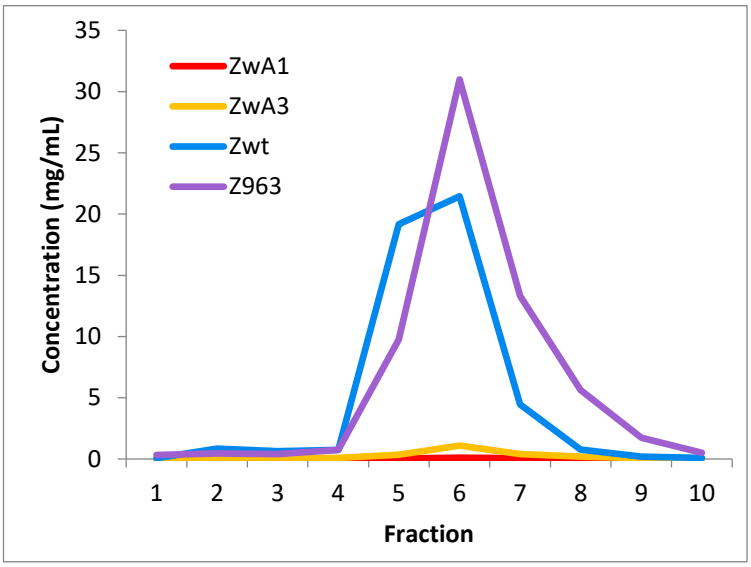
Temp	94°C	96°C	55°C	72°C	72°C	4°C
Time	5:00	0:40	0:40	1:40	10:00	hold
Purpose	Initial denaturation	Denaturation	Annealing	Elongation	Final elongation	
Repeats	1x		35x		1x	1x

8.4 Example production workflow

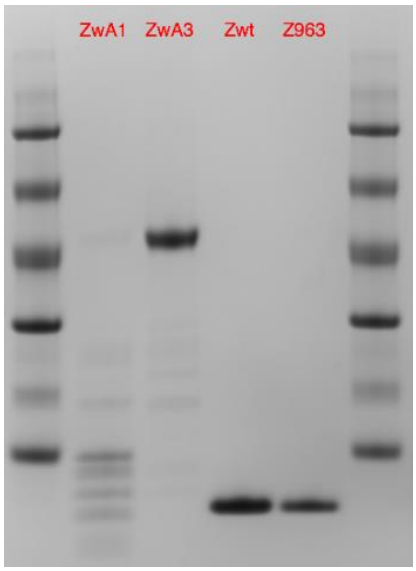
Plasmids for Zwt, Zspa963, ZwA1 and ZwA3 were successfully transformed into *E. coli* BL21* expression cells, cultivated, and induced to produce protein. Fractions from IMAC elution were collected and the protein concentration estimated by absorbance at 280 nm (*fig A4, A*). By plotting these concentrations against the fractions, an elution profile was constructed allowing the fractions with the highest concentration to be pooled. The shape of the elution profile corresponds with what is expected from a column elution, with a void volume of 4 fractions (2 mL).

The two affibodies Zwt and Zspa963 were in much higher concentration with a total yield of 120 respectively 160 mg protein produced per litre cultivation, compared to the 1.3 mg/L and 6.0 mg/L of ZwA1 and ZwA3.

The five fractions with the highest concentrations were pooled and the buffer changed to 1xPBS via a PD10 column, after which the purity was validated via SDS-PAGE. However, ZwA1 did not yield protein in sufficient amounts or purity and was therefore not studied further (*fig A4, B*).



A



B

***Figure A4*. A:** Elution profile from the concentrations of fractions after IMAC purification. **B:** Reducing SDS-PAGE of pooled IMAC eluate. 3 µg protein was loaded in each well when possible. In the well for ZwA1, the maximum volume 10 µL was loaded. For ZwA3, Zwt and Z963, a band can be seen at the expected size, however for ZwA1 there are several smaller bands but nothing at the expected size.

8.5 Protein expression and purification

Protein were expressed and purified via IMAC. The A280 of each IMAC fraction was measured, and the 5 highest used for buffer change via PD10 columns. A selection of the A280 values from the IMAC fractions were converted into concentrations, and the theoretical yield of the cultivation calculated (*table A9*). The concentration was measured again after buffer change to PD10 columns, which was later used as the working concentration.

For constructs with a lower yield, a second band could be seen at 45 kDa. When loading the same volume onto a reducing SDS-PAGE, this band was also visible for constructs with a high concentration (*fig A5*), meaning that when 3 µg was loaded the band was present, but too weak to be visible. The second band is presumably a contamination from a native protein in *E. coli* with affinity for the IMAC column.

Table A9. Theoretical yields in mg protein per litre culture.

Protein	Yield (mg/L)
Zwt	121
Zwt L17BPA	205
Zwt K35BPA	50
Zwt A29G	N/A
Z963	N/A
Z963 L34M	120
Z963 L35M	N/A
Z964	160
ZwA3	6
ZwA3 A29G	N/A
ZwA3 A29G -1	N/A
ZwA3 A29G +3	N/A

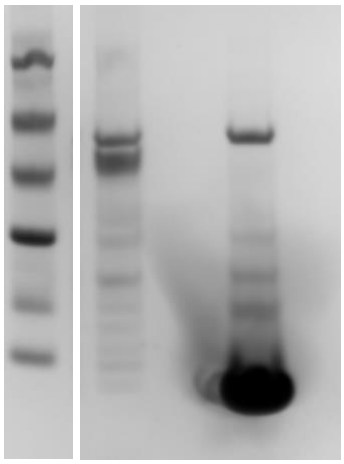


Figure A5. Reducing SDS-PAGE gel of ZwA3 and Zwt. 10 µL sample was loaded in each well.

8.6 MALDI-TOF MS

Successful mutations to create the ZwA3 variants ZwA3 A29G, and the subsequent mutation of this to the ZwA3 A29G -1 and +3 variants were verified using MALDI mass spectrometry (*fig A6*). ZwA3 was used as a reference which allowed mass differences created by the mutations to be visible even if the m/z of the MALDI was inaccurate. The peaks are broad but generally have the expected shifts: the ZwA3 A29G mutation has a small shift towards lower mass since glycine is slightly smaller than alanine. The -1 and +3 mutations have shifts in both the right direction and approximate size for the change in length of the protein.

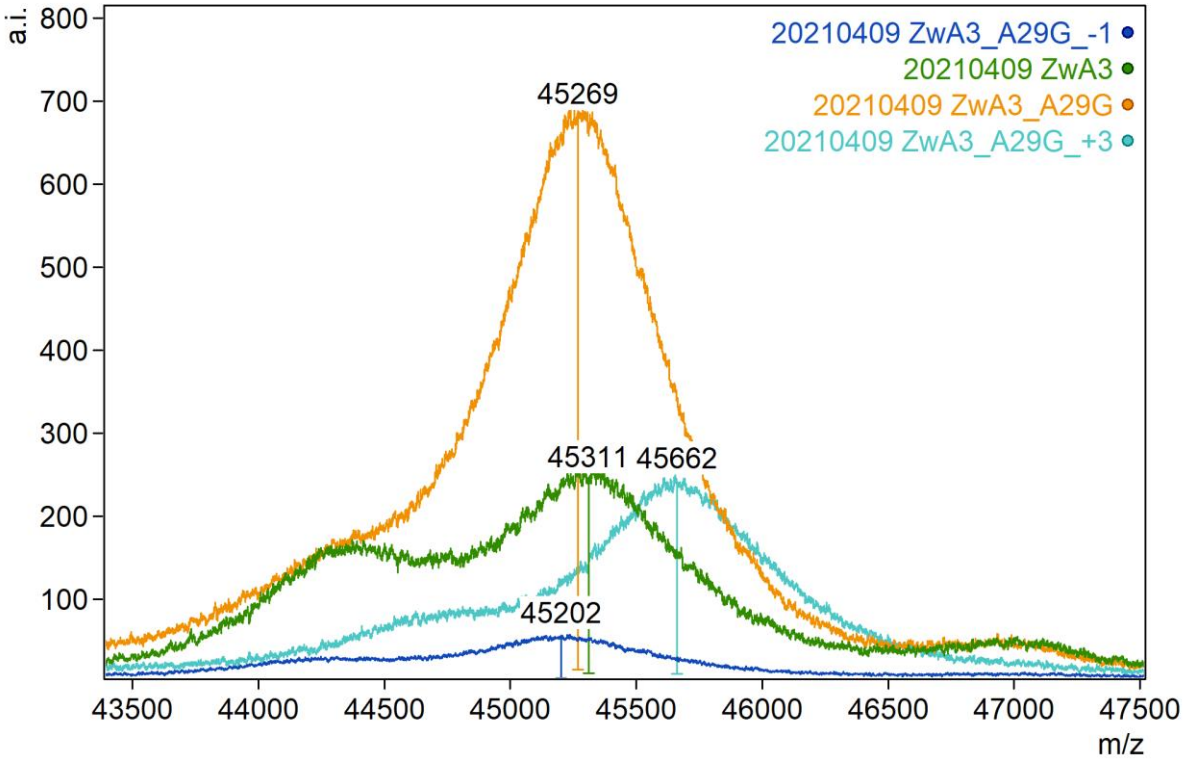


Figure A6. Mass spectrum of ZwA3 variants, focused between 43000 and 48000 m/z. The labelled peaks correspond to approximately the expected size for the produced protein.

The m/z and the shifts in m/z from ZwA3 of the mutants are within 50 Da of the expected values (*table 2*).

8.7 ESI-LC-MS

Selected samples were analysed by LC-MS, with the purpose of detecting UV cross-linked product. However, due to limitations with the LC-MS system, this peak was not analysed. The part of the sample with monomer protein was however analysed (*fig A7*). The theoretical size is reported in *table 2.*

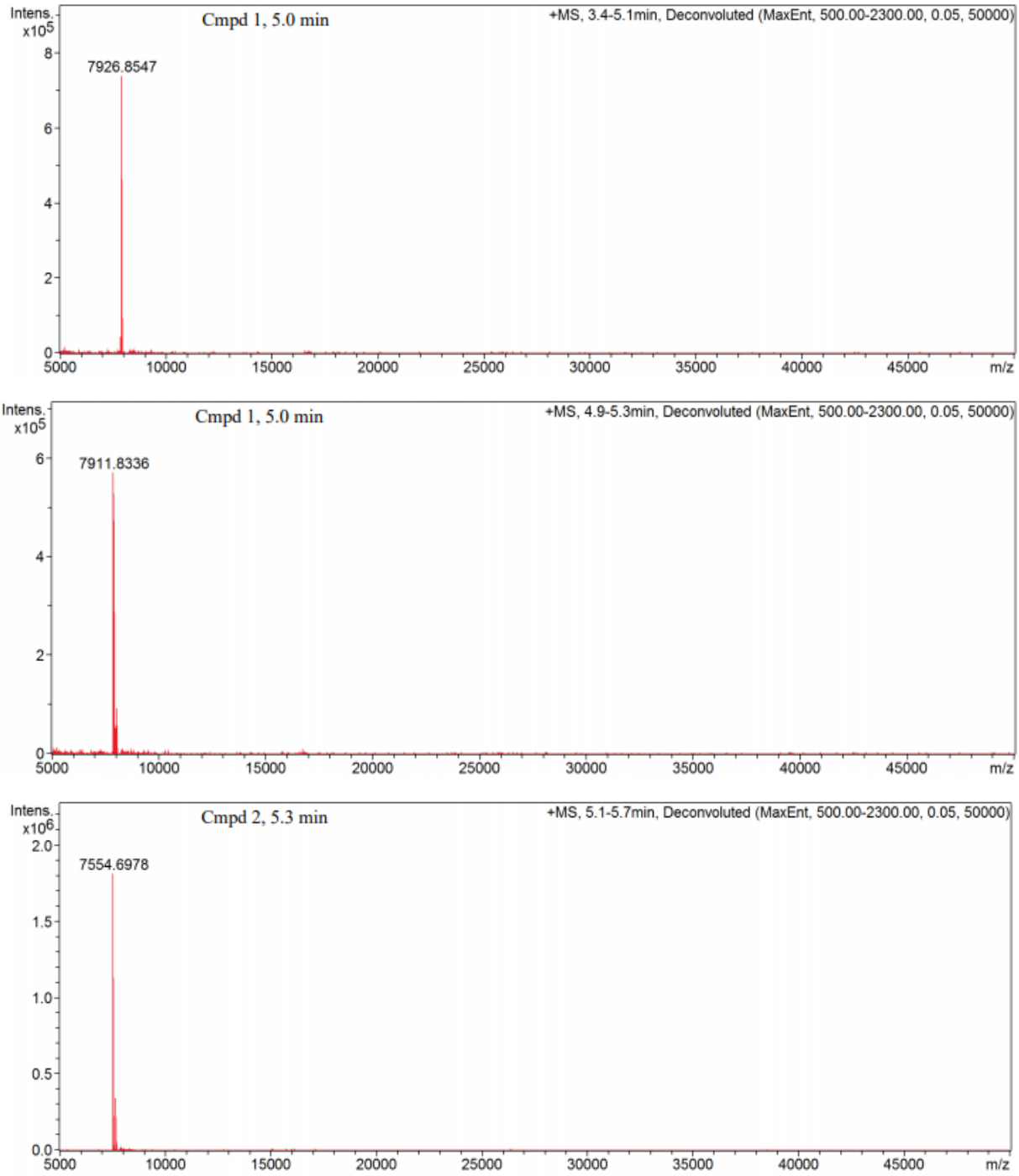


Figure A7. Deconvoluted mass spectra of Zwt L17BPA, Zwt K35BPA and Z963 L34M from LC-MS.

8.8 SPR

Immobilization of ligand onto the CM5 chip can be seen in *figure A8* for scFv, Fab and IgG. First negative peak is to see how much the sample is attracted to the surface. First square positive peak is activation of the surface. A series of injections of the ligand are then performed until the desired RU is reached, after which the second positive square signal is the deactivation of the surface.

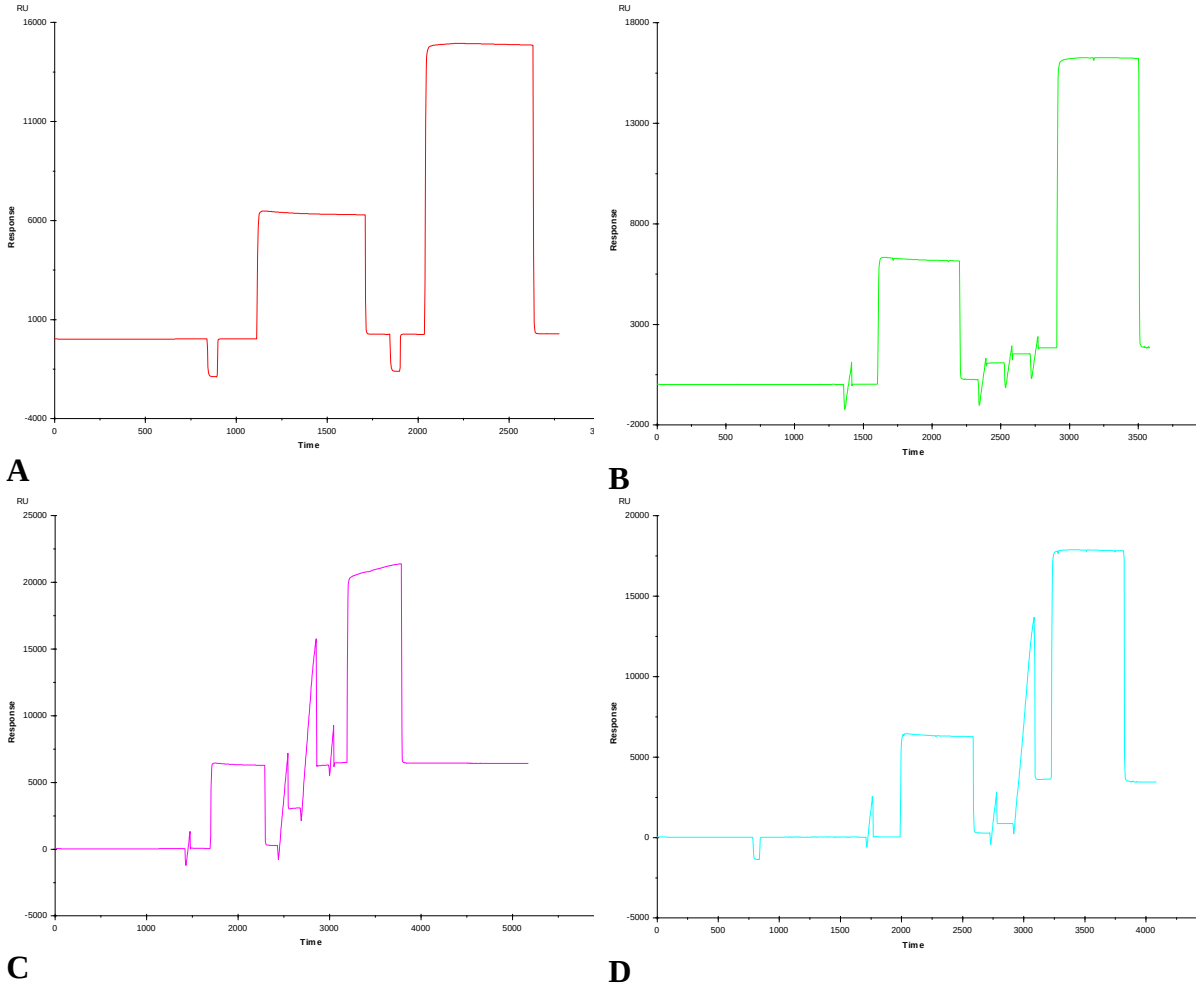


Figure A8. Immobilization of ligand. **A:** negative control. **B:** scFv. **C:** Fab. **D:** IgG.

The binding of two-fold dilution series of Z963 variants and Z964 to 500 RU Zwt immobilized to the CM5 chip can be seen in *figure A9.* Concentrations varied between 1.6 µM and 0.8 nM. From these dilutions, kinetic constants such as K_D values seen in *table 3* were calculated using Langmuir 1:1 kinetics. No curve fit was made for the other samples, as they did not display characteristic curves for Langmuir 1:1 kinetics. Some deviation can be seen in **B**, where biphasic behaviour was displayed for two of the dilutions. These were not included in the curve fit, and **C** where the point after the signal becomes abnormal was not used.

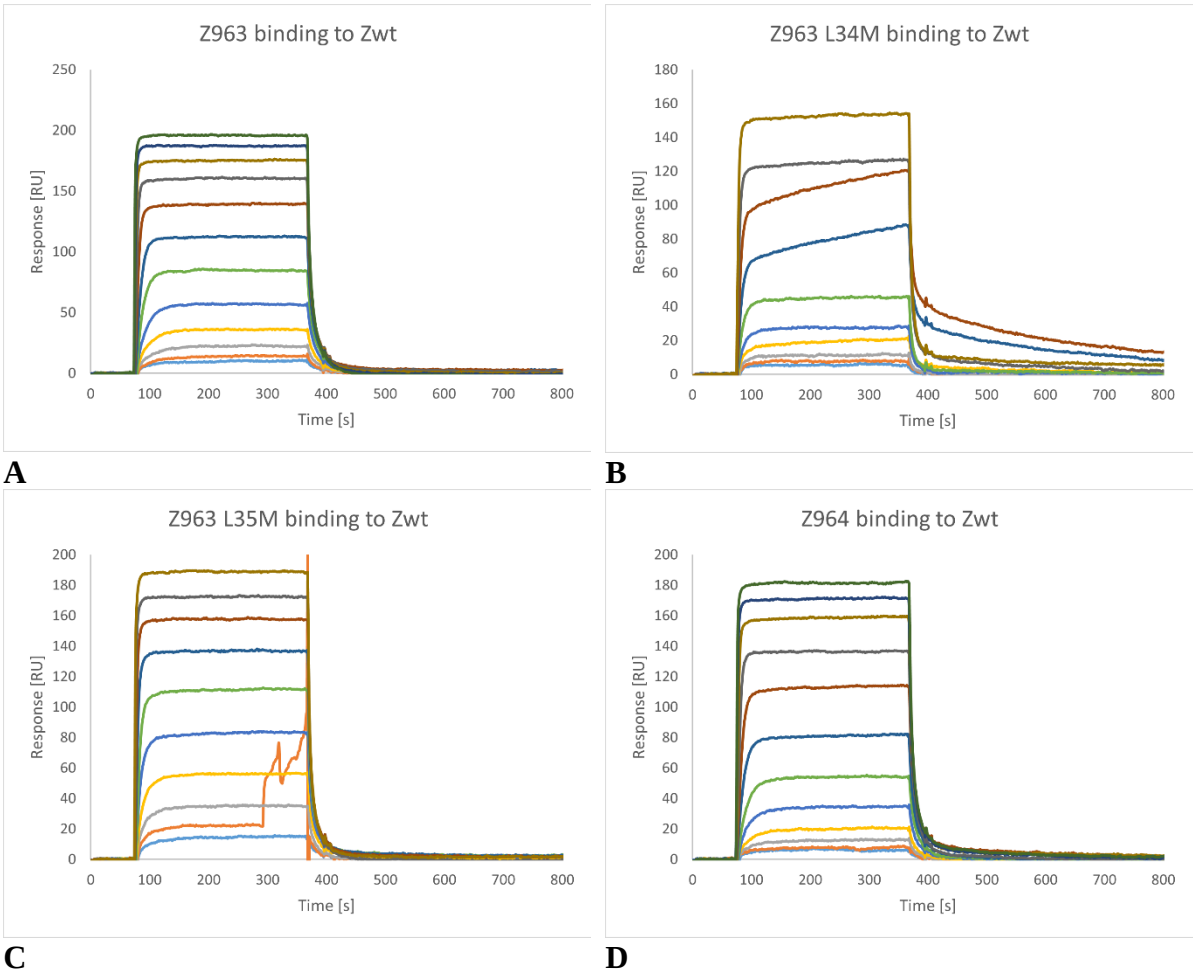


Figure A9. Dilution series of anti-Zwt variants to immobilized Zwt. **A**. 0.8 nM to 1.6 µM Z963. **B**. 0.8 nM to 400 nM Z963 L34M. **C**. 3.2 nM to 1.6 µM Z963 L35M. **D**. 0.8 nM to 1.6 µM Z964.

8.9 Photoconjugation SDS-PAGE

Other factors that could affect the degree of photoconjugation were investigated, such as the container in which the crosslinking takes place. The photoconjugation protocol was carried out in different material plates and run on a reducing SDS-PAGE (*fig A10*). The material of the container did not matter as much as anticipated, but rather the colour. Transparent tubes and plates yielded a ~50% crosslinking efficiency, while black opaque containers were at a much lower efficiency, regardless of its original purpose.

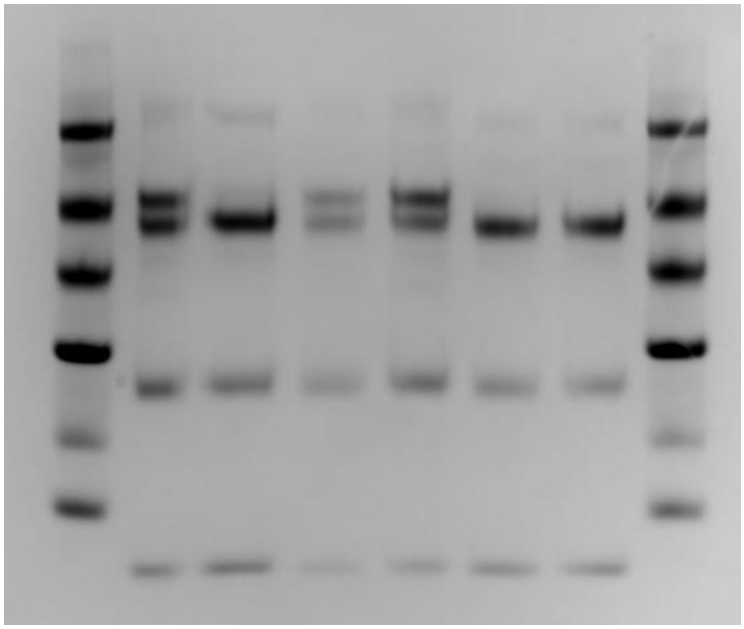


Figure A10. Reducing SDS-PAGE of UV conjugated samples in different containers. **1.** LMW ladder **2.** Transparent flatbottomed high binding 96-well ELISA plate **3.** Black opaque untreated flatbottomed 96-well plate **4.** One singular transparent PCR tube **5.** Transparent 96-well PCR plate **6.** Black opaque 96-well PCR plate **7.** Negative control **8.** LMW ladder

8.10 MALDI-TOF MS on photoconjugated sample

The photoconjugated reaction mix was reduced and analysed in MALDI mass spectrophotometry (*fig A11*). A large peak can be seen at the expected size for the Zwt K35BPA affibody, with dimers or fractionally charged peaks at multiples of the affibody size. The peak at 23 kDa is larger than the previous peak, due to the overlapping sizes of the affibody "tri-mer" and the light chain of the antibody.

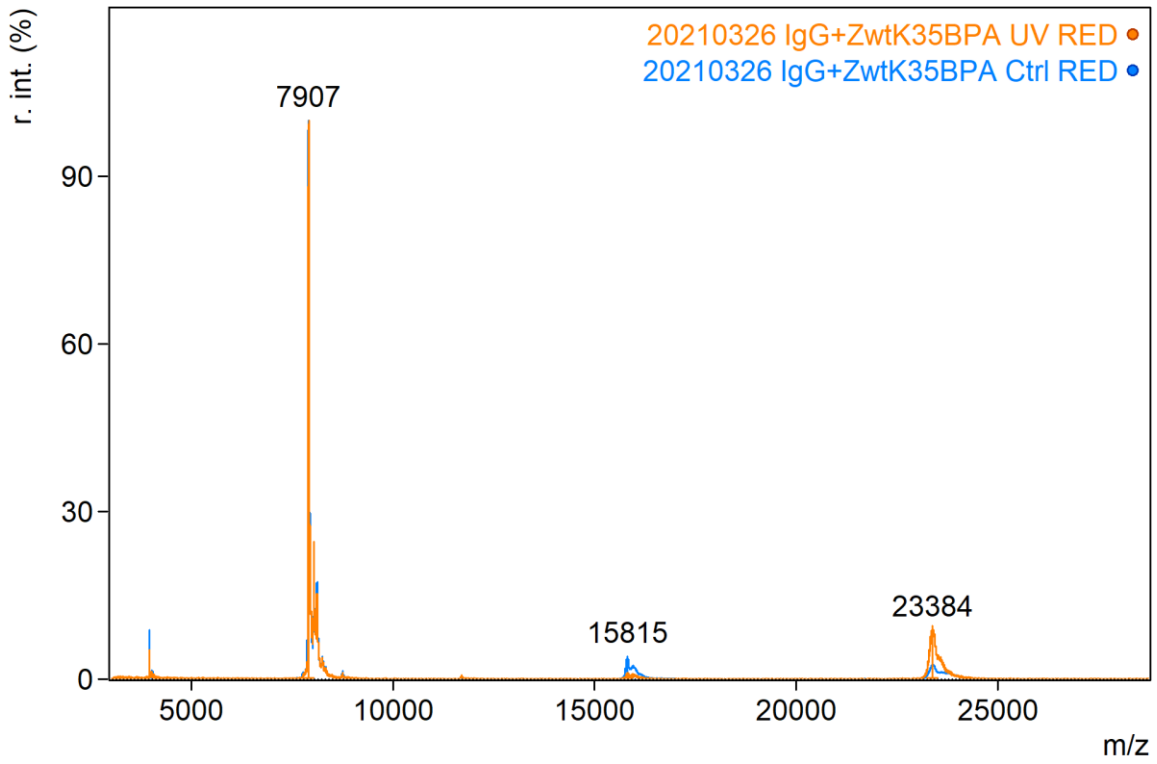


Figure A11. Mass spectra of Z(xBPA) variants mixed with IgG between 3000 and 28000 m/z with (orange) and without (blue) UV treatment. One major peak can be seen at 8 kDa, and two smaller at 16 kDa and 23 kDa.

The repeating pattern of affibody multimers continues, and as the peaks increase in mass, they decrease in relative intensity. The last peak with a difference of ~8 kDa from the previous can be seen at 55 kDa (*fig A12*). There is however a large peak slightly below at 53 kDa which corresponds to the heavy chain of IgG. At 61 kDa, a peak can be seen which is only 5 kDa from the previous affibody multimer at 55 kDa, ruling out the possibility that it is an octamer. This peak is approximately 8 kDa from the heavy chain peak, matching the mass of the Zwt K35BPA crosslinked to the Fc of the IgG. The sample had been reduced by TCEP, and the Zwt K35BPA:Fc peak is absent from the control sample that has not been exposed to UV, confirming that the interaction is indeed covalent.

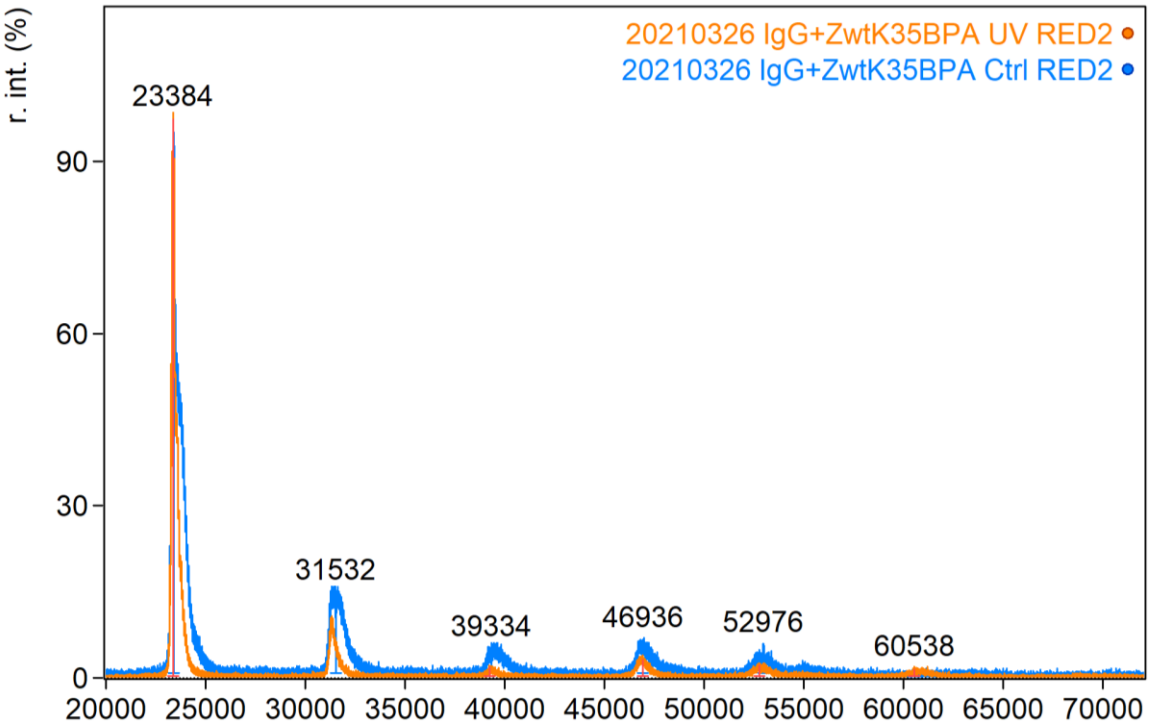


Figure A12. Mass spectra of Z(xBPA) variants mixed with IgG between 20000 and 70000 m/z with (orange) and without (blue) UV treatment. A repeating pattern of diminishing peaks can be seen every 8 kDa starting at 23 kDa. A peak at 53 kDa can be seen that is not part of this pattern, and another at 61 kDa, which is only present for the UV treated sample.

For Zwt L17BPA photoconjugated to Z963 L34M, MALDI was also performed. Around the base peaks of 7555 Da for Z963 L34M and 7928 for Zwt L17BPA, there is a third peak at 7694, which presumably is the 2+ charged heterodimer of Z963:Zwt. The peak is less prominent in the non-UV treated sample.

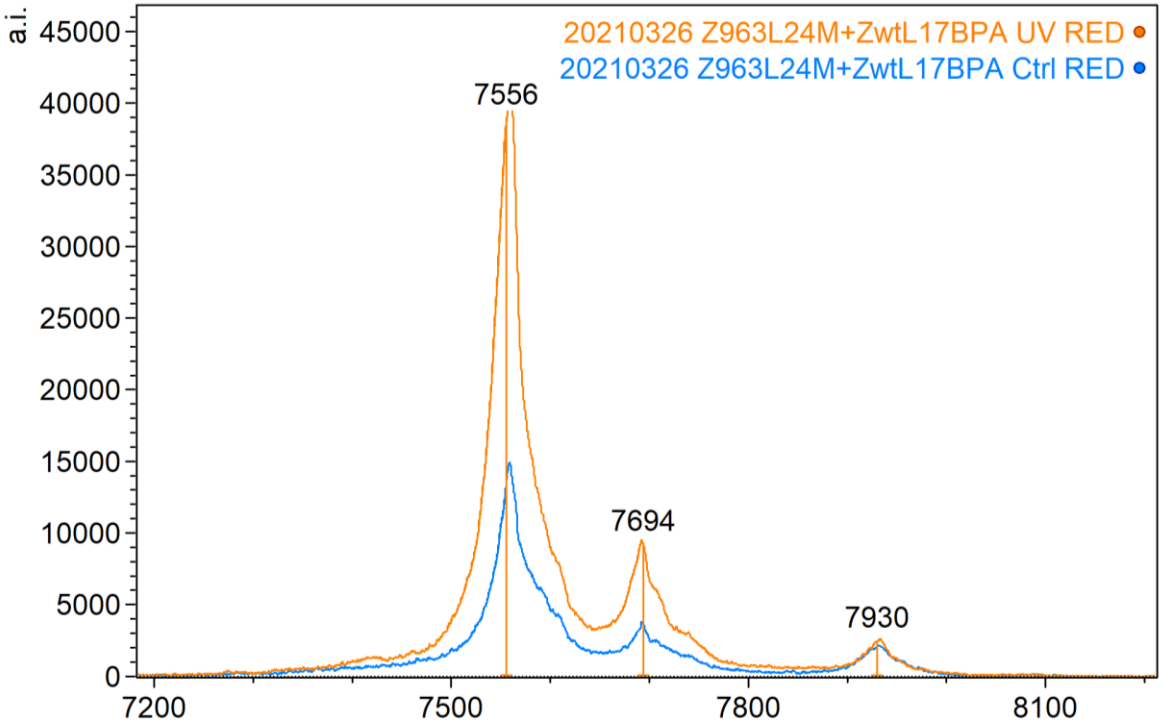


Figure A13. Mass spectra of Zwt L17BPA variants mixed with Z963 L34M with (orange) and without (blue) UV treatment.