Article

Small Bispecific Affinity Proteins for Simultaneous Target Binding and Albumin-Associated Half-Life Extension

Emma von Witting, Sarah Lindbo, Magnus Lundqvist, Marit Möller, Andreas Wisniewski, Sara Kanje, Johan Rockberg, Hanna Tegel, Mikael Åstrand, Mathias Uhlén, and Sophia Hober*

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: Albumin-binding fusion partners are frequently used as a means for the *in vivo* half-life extension of small therapeutic molecules that would normally be cleared very rapidly from circulation. However, in applications where small size is key, fusion to an additional molecule can be disadvantageous. Albumin-derived affinity proteins (ADAPTs) are a new type of scaffold proteins based on one of the albumin-binding domains of streptococcal protein G, with engineered binding specificities against numerous targets. Here, we engineered this scaffold further and showed that this domain, as small as 6 kDa, can harbor two distinct binding surfaces and utilize them to interact with two targets simultaneously. These novel ADAPTs were developed to possess affinity toward both serum albumin as well as another clinically relevant target, thus circumventing the need for an albumin-binding fusion partner. To accomplish this, we designed a phage display library and used it to successfully select for single-domain bispecific binders toward a panel of targets: TNFα, prostate-specific antigen (PSA), C-reactive protein (CRP), renin, angiogenin, myeloid-derived growth factor (MYDGF), and insulin. Apart from successfully identifying bispecific binders for all targets, we also demonstrated the formation of the ternary complex consisting of the ADAPT together with albumin and each of the five targets, TNFα, PSA, angiogenin, MYDGF, and insulin. This simultaneous binding of albumin and other targets presents an opportunity to combine the advantages of small molecules with those of larger ones allowing for lower cost of goods and noninvasive administration routes while still maintaining a sufficient *in vivo* half-life.
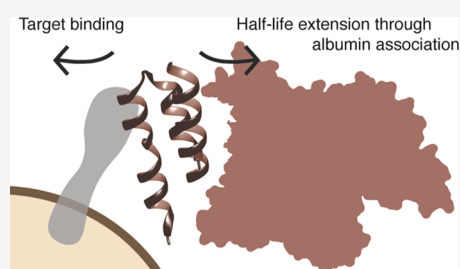
KEYWORDS: protein engineering, phage display, next-generation sequencing, ABD, ADAPT, albumin, half-life extension

INTRODUCTION

Owing to their extraordinary affinity and target specificity, antibodies have been tremendously successful within the pharmaceutical industry. In fact, 6 out of the 10 best-selling drugs on the market today are antibodies,[1] and their combined sales approached $100 billion at the end of 2017.[2] Despite their success, antibodies are not without limitations. Their large and complex structure leads to high production costs and suboptimal pharmacokinetics in many clinical applications. In light of these limitations, and thanks to the development and refinement of protein engineering techniques during the latest decades, other alternatives have emerged. Among these are the small affinity protein domains, collectively known as alternative scaffolds. These scaffolds have several potential benefits as therapeutics and can be engineered to have high affinity and target selectivity. In addition, their small size, which entails both a low cost of goods and an ability to penetrate tissue that is otherwise inaccessible, makes them an attractive alternative to antibodies.

Apart from affecting the production process and the biodistribution, another potential benefit of the small size is the possibility of using alternative administration routes. For large molecules like antibodies, options are limited to intravenous and subcutaneous injection, both being highly inconvenient for the patient and causing a high burden on the healthcare system.[3] For the much smaller alternative scaffolds, nasal or pulmonary delivery could very well serve as a noninvasive alternative, possibly improving patient convenience and compliance. Previous studies have shown that the molecular mass and size of a protein are the most critical factors influencing the bioavailability of drugs administered via the mucosa, and, consequently, keeping the size to a minimum becomes especially important.[3,4]

The small size also causes rapid clearance from the circulation, a feature that can be both an advantage as well as a limitation depending on the application. Fast blood clearance together with high tissue penetration has been essential to achieving high contrast in molecular imaging applications. In fact, many of the alternative scaffolds (e.g., Affibody, DARPin, and ADAPT) have been shown to be very promising as imaging probes, exhibiting extraordinary contrast

from surrounding healthy tissue.[5−7] On the other hand, in most therapeutic applications, rapid clearance is unfavorable since a long residence time is required to keep the dosing low and less frequent. A rapid clearance typically also yields undesired accumulation in the kidneys. To overcome this size-related limitation, several strategies for half-life extension have been developed.[8] One of these strategies is based on fusion to the albumin-binding domain (ABD) of streptococcal protein G. By utilizing the noncovalent association with the patient's own serum albumin, a substantial increase in half-life and decrease in kidney accumulation have been shown for scaffold fusion proteins developed through this strategy.[9−11] The increase of the circulatory half-life is not only attributed to the increase in the size of the complex and hence the ability to avoid filtration through the glomerular barrier but also to the fact that albumin is rescued from lysosomal degradation by association with the neonatal Fc receptor, FcRn.[12] Although frequently used, one obstacle of most half-life extension strategies that exist today is that they counteract some of the benefits associated with the small size. With the aim of developing a new strategy for half-life extension, without altering the size of the molecule, our group has in the past engineered the ABD itself. In this previous work, a novel binding surface was introduced into the ABD, in addition to its native binding to albumin. By randomizing residues in the first and third helix of the ABD, bispecific ABD-derived affinity proteins (ADAPTs) toward the Z_2 domain,[13] TNFα,[14] HER2,[15] and HER3[16] were generated. However, binding both albumin and the desired target simultaneously was not feasible, probably due to steric hindrance. Since albumin concentrations in serum is high ($\sim$600 μM),[17] even a low affinity toward it would most certainly prevent target binding *in vivo* for these molecules. With this in mind, we have in the present study designed a combinatorial library where we have moved the proposed binding surface to helix 1 and 2 of the ABD. Our hypothesis was that the novel library design would increase the possibility of generating minimal proteins capable of binding albumin and other targets simultaneously, thereby combining the advantages of both small and large therapeutic molecules in one single format. To demonstrate the quality and applicability of this new combinatorial library, we performed phage display selections toward a panel of clinically relevant targets of different sizes (TNFα, prostate-specific antigen (PSA), CRP, renin, angiogenin, MYDGF, and insulin). In addition, a standardized workflow including next-generation sequencing and a polymerase chain reaction (PCR)-based method for gene recovery was designed to allow for cost-efficient isolation of target-binding clones. The results from our study showed that the novel binding surface was suitable for target interactions and also enabled simultaneous binding to serum albumin.

■ MATERIAL AND METHODS

Construction of Phage Display Library. A phagemid vector (pSE) was constructed based on the phagemid pAffi1.[18] In pSE, the gene fragment encoding helix 3 of protein Z present in pAffi1 was removed and the gene fragment encoding ABDwt was replaced to instead encode the full-length protein Z. Phagemid pSE was ordered from Thermo Fisher Scientific GENEART GmbH (Regensburg, Germany). The restriction enzyme recognition sites for *XhoI* and *NheI*, positioned upstream of Z, were used for subsequent insertion of the library gene cassette.

The first ADAPT library gene was assembled from three separate oligonucleotides: fwdABDlib_h1 covering helix 1 containing eight diversified positions allowing for all amino acids except Pro, Cys, and Gly (Ella Biotech, Martinsried, Germany), revABDlib_h2 covering helix 2 containing three positions diversified through NNK-degeneracy (TAG Copenhagen, Copenhagen, Denmark), and revABDlib_h3 covering helix 3 without any diversified positions (TAG Copenhagen, Copenhagen, Denmark). A second version of the library was assembled by exchanging revABDlib_h2 for revABDlib_h2_v2 (Ella Biotech, Martinsried, Germany), with the same randomized positions but instead diversified to allow for all amino acids except Pro, Cys, and Gly, just like for helix 1. Complete oligo sequences can be found in the Supporting Information Table S1. Further, 800 pmol of each oligo were mixed and assembled by 10 cycles of PCR and amplified for another 10 cycles using external primers. These primers also introduced restriction sites *XhoI* and *NheI* for ligation into the phagemid vector pSE. The ligated phagemid was extracted from a 2% agarose gel and purified using QIAquick gel extraction kit (Qiagen, Hilden, Germany). The product was transformed by electroporation into XL1 blue *Escherichia coli* cells (Lucigen, Middleton, WI) yielding 5×10^9 members.

Production and Biotinylation of Target Protein. All target proteins except insulin were produced in the Human Secretome Project (Stockholm, Sweden) as previously described by Uhlén et al.[19] Human recombinant insulin was ordered from Thermo Scientific (Waltham, MA). For biotinylation of target proteins, EZ-link Sulfo-NHS-LC-Biotin (Thermo Scientific, Waltham, MA) was dissolved in water to a final concentration of 10 mM and a 12-fold molar excess was then added to the target protein solution. The reaction was incubated at room temperature (RT) for 30 min. Nonreacted biotin reagent was removed from the sample by a NAP5 desalting column (GE Healthcare, Stockholm, Sweden).

Phage Display Selections. Four rounds (two rounds for insulin selection) of panning were performed using a biotinylated recombinant target protein (TNFα, CRP, PSA, renin, angiogenin, MYDGF, insulin) in solution, followed by capture on streptavidin-coated magnetic beads. To get rid of nonspecifically binding phages, all tubes and beads used during the selection were washed with PBS-T (phosphate-buffered saline, 0.1% Tween-20) and blocked with PBS-T supplemented with 0.5% gelatin. Furthermore, all panning rounds were preceded by incubation with 0.5 mg of the washed and blocked streptavidin-coated magnetic beads at RT at 150 rpm for 30 min. The resulting supernatant was then used as input in the panning rounds. During panning, the phage supernatants were incubated with 150, 100, 50, and 25 nM target protein for rounds 1−4, respectively. The incubation was allowed to proceed for 2 h under rotation before capture on streptavidin-coated magnetic beads. The beads were washed twice with PBS-T in round 1, four times in round 2, eight times in round 3, and 12 times in round 4. The bound phages were eluted with 50 mM glycine, pH 2.0, at RT for 10 min, followed by neutralization with an equal volume of 10% 1 M Tris-HCl in PBS, pH 8.0. The neutralized eluates were then used to infect at least 100× excess XL1 blue *E. coli* cells compared to the number of eluted phages. After allowing the infection to proceed for 30 min, an equal amount of TSB + YE medium (30 g/L) tryptic soy broth (Merck, Kenilworth, NJ), 5 g/L yeast extract (Merck, Kenilworth, NJ) supplemented with 200 μg/mL ampicillin and 10 μg/mL tetracycline was added to the

cells. Then, 30 min later, a 10× excess M13K07 helper phage, compared to the number of cells used at infection, was added. After an additional 90 min, the cells were pelleted, resuspended in 100 mL of TSB + YE supplemented with 100 μg/mL ampicillin, 25 μg/mL kanamycin, and 0.1 mM IPTG, and finally incubated overnight at 30 °C. The amplified phages were collected by poly(ethylene glycol) (PEG) precipitation and used as input in the next panning round.

Next-Generation Sequencing. Phagemid DNA was extracted from XL1 blue *E. coli* infected with phage eluates from the different panning rounds or the naive phage library using a QIAprep Miniprep kit (Qiagen, Hilden, Germany). A PCR was performed with 50 ng of purified phagemid as a template and 5 pmol of an oligo containing the forward adapter sequence and 5 pmol of an oligo containing the reverse adapter sequence and a sample-specific index sequence. To reduce the risk of bias and errors, the PCR was limited to 15 cycles. The acquired PCR product was extracted from a 2% agarose gel and purified with a QIAquick gel extraction kit. Finally, all samples were pooled in equal amounts to reach a total concentration of 10 nM. Sequencing was done in a flow cell of an Illumina MiSeq v2 instrument and conducted at the National Genomics Infrastructure (Stockholm, Sweden).

Sequencing Data Analysis. FASTQ files were analyzed using in-house software. First, correct sequences were sorted out; a sequence was considered to be correct if two recognition sites, nine bases upstream, and nine bases downstream of the variable region were found. Additionally, the distance between these two sites had to be of correct length, and all bases were required to have a Phred quality score of ≥ 30. Subsequently, the number of each encountered gene variant among the correct sequences was calculated for each of the sequenced samples.

Subcloning, Production, and Purification of Lead ADAPTs. Based on the next-generation sequencing results, 12 different enriched gene variants for each target were selected for further evaluation. The selected genes appeared in high frequency only in the tracks for that specific target protein and were spread over several different sequential clusters. Gene variants for the selections toward TNFα, CRP, PSA, renin, and insulin were synthesized by Thermo Fisher Scientific (Waltham, MA) and amplified using primers with flanking regions containing restriction sites for subcloning into a T7 inducible expression vector. Variants from the selections toward angiogenin and MYDGF were isolated as described in the following section. The final products were sequence-verified using Sanger sequencing by Microsynth Seqlab (Göttingen, Germany). All variants were expressed in *E. coli* BL21*(DE3) cells and extracted by sonication. The resulting lysates were subjected to purification by affinity chromatography using a resin coupled with human serum albumin (HSA) (produced in-house).

Development of a PCR-Based Strategy to Rescue Individual Clones. Although genes from five of the seven selections were initially synthesized as described above, a more cost-efficient strategy based on a standard PCR protocol was also developed and used to rescue clones from the selections toward angiogenin and MYDGF. One unique oligo covering the randomized positions in helix 1 and another covering the randomized positions in helix 2 were used as forward and reverse primers, respectively, to amplify the unique sequence and introduce restriction sites for cloning into a vector. The same T7 inducible expression vector as above was equipped

with the third and nonrandomized helix where we also introduced a *Hind*III restriction site for the assembly of the full-length sequence of the unique ADAPT variant.

Circular Dichroism. Secondary structure content, melting temperatures, and ability to refold after denaturation was evaluated in a Chirascan circular dichroism spectrometer (Applied Photophysics, Surrey, U.K.). Samples were diluted in PBS to a concentration of 0.2 mg/mL, and all analyses were performed in a cell with an optical path length of 1 mm. Secondary structure content was evaluated by measuring the ellipticity from 260 to 195 nm at 20 °C. The melting temperatures were determined by monitoring the signal at 221 nm while increasing the temperature from 20 to 100 °C at a rate of 5 °C per min. The refolding capability was assessed by repeating the secondary structure evaluation after the sample had been subjected to heat treatment and cooled to 20 °C.

Evaluation of the Oligomeric State Using Size-Exclusion Chromatography (SEC). The aggregation propensity of the ADAPT variants was evaluated using size-exclusion chromatography (SEC). The experiments were performed on an ÄKTA Pure system equipped with a Superdex 75 5/150 column (GE Healthcare, Stockholm, Sweden). The column was equilibrated with PBS, and elution profiles were acquired by injection of 25 μL of purified protein (with concentrations ranging between 0.5 and 1 mg/mL) at a flow rate of 0.45 mL/min. A calibrant of four different proteins (conalbumin 75 kDa, carbonic anhydrase 29 kDa, ribonuclease A 13.7 kDa, and aprotinin 7.5 kDa) was used for size comparison.

Target-Binding Analysis. The sensor-based binding analysis was performed using a surface plasmon resonance (SPR) Biacore T200 system (GE Healthcare, Stockholm, Sweden). All analyses were conducted using CM-5 sensor chips coated with the target protein, another irrelevant protein, and HSA. One surface was left as the reference. The binding analysis was performed in PBS-T at 25 °C. The ADAPT variants were diluted in PBS-T to 500, 250, 125, 62.5, and 31.25 nM. The samples were injected at a flow rate of 30 μL/min for 120 s, followed by injection of PBS-T for 240 s to monitor the dissociation. Surfaces were regenerated with 10 mM HCl. Kinetic parameters were determined using the Biacore Insight Evaluation software, assuming a 1:1 binding.

The simultaneous binding of ADAPT to albumin and the target protein was evaluated by both a capture assay and a target saturation assay. In the capture assay, 500 nM ADAPT variant was first captured on the HSA surface for 50 s with a flow rate of 10 μL/min, followed by injection of 500 nM target protein. The target protein was injected at a flow rate of 30 μL/min for 120 s, followed by injection of PBS-T for 240 s to monitor the dissociation. Surfaces were regenerated with 10 mM HCl. For the target saturation assay, the ADAPT variants were diluted in PBS-T to 500 nM and incubated with its respective target in 5× excess before injection over a surface immobilized with HSA. The samples were injected at a flow rate of 30 μL/min for 120 s, followed by injection of PBS-T for 240 s to monitor the dissociation. Surfaces were regenerated with 10 mM HCl.

Evaluation of Ternary Complex Formation Using Size-Exclusion Chromatography. For one of the insulin binders (ADAPT$_{Insulin\_02}$), a final characterization step was performed on a SEC column to confirm the formation of the ternary complex. HSA, ADAPT, and insulin were mixed together in a 1:1:1 molar ratio at a total concentration of 50
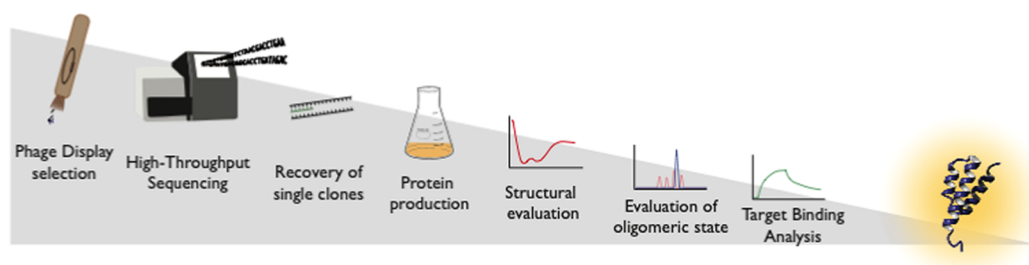
Figure 1. Schematic description of the established workflow.

μM and incubated for 1 h at RT before injection onto the column. The same amounts of each binding partner were also injected separately. The experiments were performed on an ÄKTA Pure system equipped with a Superdex 75 5/150 column. The column was equilibrated with PBS and elution profiles were acquired by injection of 25 μL of the sample at a flow rate of 0.45 mL/min.

■ RESULTS

For the purpose of selecting functional single-domain bispecific binders toward HSA as well as other target proteins, a standard workflow was established and is described in Figure 1. Following phage display selection and next-generation sequencing analysis, single clones were recovered either by synthesis or through a PCR-based method with unique oligos covering the randomized positions. To make sure that all final candidates would be functional in future applications, all protein variants were first evaluated in terms of structure and stability before finally being tested for binding toward the different target proteins as well as to HSA.

Construction of the Phage Display Library. With the aim of allowing for simultaneous binding of both the native ligand HSA as well as other clinically relevant targets, a novel combinatorial library was designed and created. ABD has previously been engineered for this purpose, resulting in the library shown in Figure 2B, with randomized positions spread over the first and the third helix.

The first library was successfully used to develop bispecific binders toward several targets. However, even though they all bound both their target and HSA separately, binding of HSA hindered simultaneous binding to the target in all cases.[13,15] We hypothesized that the quite large and bulky HSA molecule might cause a steric hindrance around the binding surface of helix 1 and 3 and believed that shifting the randomized binding surface to helix 1 and 2 could entail new possibilities (Figure 2A). As we were concerned that mutating positions in helix 2 close to the HSA binding site could affect the binding, the library was based on an affinity-matured ABD known as ABD035.[20]

For the construction of the new library, a forward oligonucleotide covering helix 1 included variations in positions according to Figure 2A,C with an equal distribution of all natural amino acids except proline, cysteine, and glycine, while the reverse oligonucleotide covering helix 2 was based on degenerate codons introducing NNK in positions 22, 23, and 26. The final reverse oligonucleotide covering the third helix contained only constant positions.

The randomized positions in helix 1 and 2 were chosen based on the position of their side chains, which were suggested to be located at a sufficient distance from the HSA binding surface, as shown in the crystal structure described by
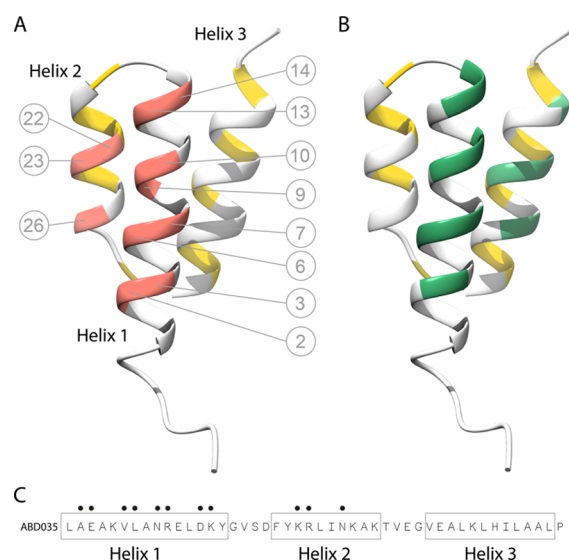

Figure 2. (A) Design of the new library with the randomized positions shown in red and the binding toward HSA shown in yellow. (B) Previously published library design with the randomized positions shown in green, which was used to select for bispecific binders that turned out to be sterically hindered from binding HSA and other targets simultaneously. (C) Sequence of the parental molecule ABD035 and the randomized positions marked with black dots.

Lejon et al.[21] In addition, the library design was also based on previous knowledge on residues that are not directly involved in the HSA binding as well as positions that had previously been successfully substituted.[22] Residues that were highly conserved in sequence alignments among homologous domains, and therefore thought to be of importance, were avoided. After assembly of the complete gene encoding ADAPT variants in an amount corresponding to 10^{14} individuals, the library was transformed into XL1 blue E. coli cells with total diversity of 5×10^9. The final naive library was sequenced with Illumina MiSeq, yielding more than 500 000 reads and showing a variation and an amino acid distribution consistent with the design parameters (Supporting Information Figure S1).

Phage Display Selection and Sequence Evaluation. Four rounds of panning were performed against four of the five biotinylated target proteins: TNFα (19.5 kDa), CRP (25.7 kDa), PSA (29.5 kDa), and renin (44.9 kDa). The selection toward insulin (5.8 kDa) was already sequenced after two panning rounds, since data from the other selections already indicated enrichment in this round. Next-generation sequencing with barcoded adapter sequences enabled evaluation of sequence enrichment throughout the selections. The first four selections demonstrated enrichment of specific sequences as
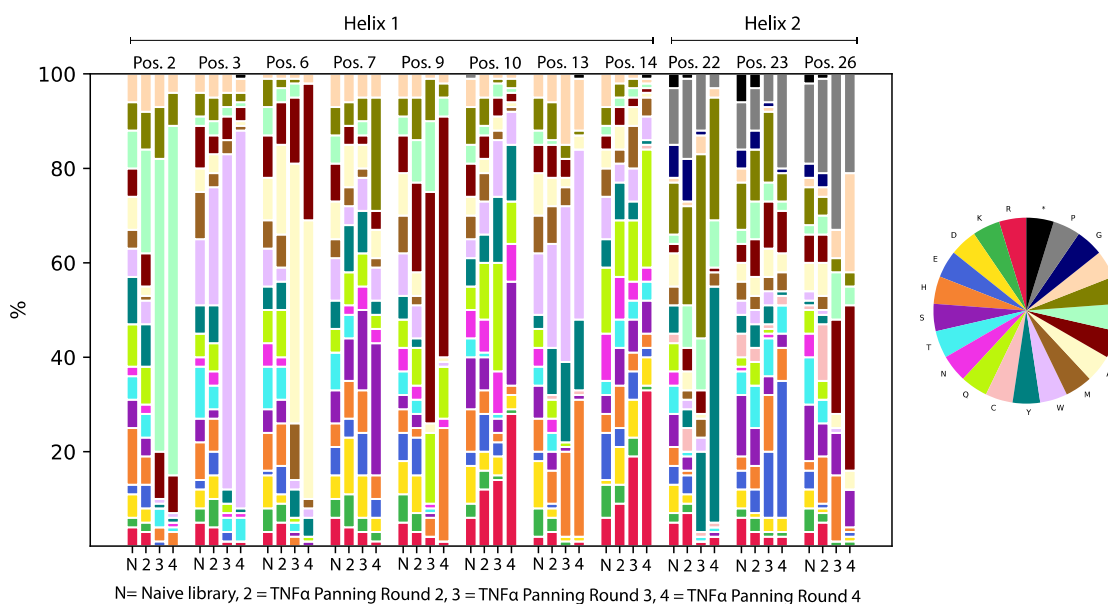
Figure 3. Amino acid distribution of each variable position of the 100 most enriched variants of the different panning rounds in the selection toward TNFα.

well as a strong consensus sequence after the fourth panning round, and enrichment was present also after the second panning round of the insulin selection, although, as expected, not as distinct. The selection toward TNFα serves as an example in Figure 3, where the amino acid distribution of the 100 most enriched sequences in the different panning rounds is shown. In this figure, it can be seen that the even distribution of amino acids in the naive library changes toward certain preferred amino acids during the panning. In the same figure, it can also be seen that the amino acid proline is quite prominent in positions 23 and 26 of helix 2. Proline, known for its rigid structure and helix-breaking properties, together with the flexible glycine and disulfide-forming cysteine were excluded from the design of helix 1. We hypothesized that these amino acids would not be as much of a problem in helix 2, but on comparing the different selection results, we saw an unexpected enrichment of one or more of these three amino acids for all selections. Based on these results, a new library was designed, with an even distribution of the 17 remaining amino acids in both helix 1 and 2. To evaluate the new version of the library, two additional selections were performed using the target proteins angiogenin (16.8 kDa) and MYDGF (19 kDa). These selections resulted in enrichment of relevant sequences not containing cysteine, proline, or glycine, as expected from the library design.

By comparing the amino acid enrichment in each position for the different selections (Supporting Information Figures S2−S7), it seems as if the localization of the binding surface is somewhat different for the different target proteins. This is not very surprising since the target proteins are different in their molecular properties and one could assume that different amino acid positions might be variously accessible to the different target proteins.

Another interesting finding derived from the sequencing data was that in all selections regardless of the target, position 6 showed a strong preference for hydrophobic amino acids (A, I, L, V, and W) and a large portion of the sequences from the TNFα, PSA, CRP, and angiogenin selections even contained valine, which is present also in the parental ABD035. These

results suggest that it is possible to improve and redesign the library to be even more focused.

From the sequencing data, certain promising ADAPT variants were chosen for further characterization. Twelve variants from each selection toward TNFα, CRP, PSA, renin, MYDGF, and angiogenin were synthesized and cloned into an expression vector for production and purification. The variants were chosen based on their relative frequency as well as them being spread over several sequential clusters (dendrograms of the sequences can be seen in the Supporting Information Figures S8−S13). For the insulin selection, an alternative strategy was used where the variants were already selected after two rounds of panning, encouraged by the fact that Figure 3 implies that enrichment is already sufficient at that early stage. Table 1 shows the top-ranked clones from the MiSeq data derived from the second panning round in the selection toward insulin. Four of these variants showed high sequence similarity and were therefore selected for further characterization.

A few of the other top-ranked clones in the insulin selection could also be found in selections toward other targets and were

Table 1. Top-Ranked Clones after Two Rounds of Selection toward Insulin[a]

rank	sequence (variable positions)	# sequences	% of total
1	**VI..AY..QY..AF..SD..K**	**148**	**0.114**
2	**AI..AY..QY..SF..RE..V**	**91**	**0.070**
3	NA..WV..LN..YD..SR..I	88	0.068
4	**AV..AY..KY..AF..NE..L**	**70**	**0.054**
5	IW..AS..VR..HR..YE..V	55	0.042
6	EQ..WA..MW..ST..IL..T	32	0.025
7	YR..LW..WV..KK..MA..F	26	0.020
8	FL..WA..NY..AR..LY..A	24	0.018
9	IW..VL..HS..WQ..YP..F	22	0.017
10	**QV..AY..KY..AF..MS..Q**	**22**	**0.017**
11	NA..WH..VN..YY..AR..L	12	0.009
12	LW..SV..RM..AD..LC..K	10	0.008

[a]Clones selected for further characterization are marked in bold.

Table 2. Results from the Characterization of the Target-Binding ADAPT Variants, Describing Melting Temperatures, Ability to Refold, Ability to Bind Simultaneously with HSA, and Determined Equilibrium Dissociation Constants

ADAPT variant	T_m (°C)	refold after heat treatment	simultaneous binding to HSA	K_D [M] target	K_D [M] HSA
TNFα_01	59	yes	yes	3.6×10^{-9}	9.5×10^{-9}
CRP_05	58	yes	no	5.2×10^{-8}	1.3×10^{-8}
CRP_27	67	yes	no	6.2×10^{-8}	1.5×10^{-8}
CRP_112	59	yes	no	1.1×10^{-6}	7.7×10^{-9}
CRP_244	68	yes	no	1.9×10^{-6}	2.3×10^{-8}
PSA_05	52	no	yes	3.0×10^{-7}	3.3×10^{-9}
Renin_01	63	yes	no	7.6×10^{-8}	2.3×10^{-8}
Renin_40	52	yes	no	4.1×10^{-7}	5.5×10^{-9}
Angiogenin_02	46	yes	yes	5.4×10^{-9}	7.3×10^{-10}
Angiogenin_06	43	yes	yes	1.0×10^{-11}	5.9×10^{-10}
Angiogenin_13	38	yes	yes	2.8×10^{-11}	2.5×10^{-9}
Angiogenin_16	42	yes	yes	2.4×10^{-9}	6.0×10^{-10}
MYDGF_07	61	yes	yes	2.3×10^{-7}	4.7×10^{-9}
MYDGF_08	56	no	yes	1.8×10^{-7}	1.2×10^{-8}
Insulin_01	31	no	yes	9.5×10^{-7}	3.9×10^{-9}
Insulin_02	44	yes	yes	6.7×10^{-7}	1.5×10^{-9}

therefore disregarded. This is likely a result of the high background still visible due to the moderate enrichment in panning round two. Possible explanations for these sequences being present in several selections could be, for example, sequencing errors, target-unspecific enrichment, or contamination during selection. Regardless of the reason, comparing results from several selections toward different targets has proven to be a useful strategy to exclude non-target-specific sequences when analyzing early panning rounds.

Characterization of Structure and Stability of Lead ADAPT Variants. All ADAPT variants that were produced in reasonable yields (see Supporting Information Table S2) were evaluated with regard to their thermal stability and secondary structure content by circular dichroism. In total, 39 candidates (of the 76 analyzed variants) demonstrated the expected α helical content with two characteristic dips at around 222 and 208 nm (Supporting Information Figure S14) and were chosen for further characterization. The thermal stability of the variants varied with melting temperatures ranging from 34 to 67 °C (Table 2), which means that some selected variants had increased stability while others had decreased stability compared to the parental molecule ABD035 that had a reported T_m of 58 °C.[20] Several variants also had the ability to refold after heat denaturation (Table 2, Supporting Information Figure S14), which indicates that they can withstand harsh treatments both during production and in future downstream processes. Those variants that demonstrated high α helical content were further assessed for their tendency to form unwanted oligomers. This was done by analysis on a size-exclusion chromatography column. Twenty-two variants (of the 39 analyzed variants) demonstrated a main peak at the correct elution time compared to the standard and were regarded as being in a monomeric state (Supporting Information Figure S15). Variants with peaks corresponding to larger complexes or degradation products were omitted from later characterizations.

Evaluation of Target Binding, Affinity, and Simultaneous Binding to HSA. Variants with sufficient α helical structure as well as monomeric elution profiles on the SEC column were further analyzed for binding toward their target protein as well as to HSA. The binding toward the target used in the selection was measured by a multicycle analysis using

SPR, showing that several binders have affinities in the micromolar to nanomolar range. ADAPT variants were considered binders only if they demonstrated binding exclusively for the target molecule and not to other irrelevant target proteins (see Supporting Information Table S2 for information on what targets were used for cross-validation of each binder). The analysis revealed 16 ADAPT variants that bound their target protein in a specific manner (the variants and their characteristics are summarized in Table 2). After target binding had been established, the variants were also evaluated for their ability to simultaneously bind to HSA. This was done both through an SPR capture assay with immobilized HSA and a target saturation assay. Figure 4 shows a
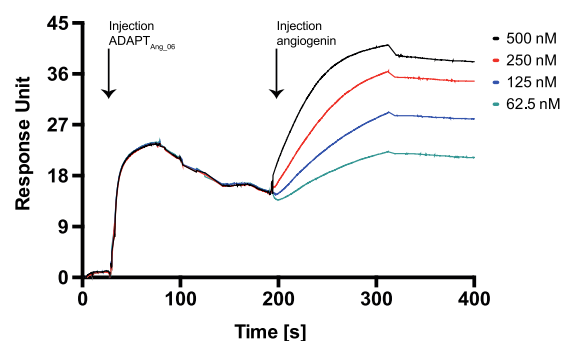

Figure 4. SPR sensorgram of ADAPT$_{Angiogenin06}$ analyzed for binding toward four different concentrations of angiogenin after being captured on a sensor chip immobilized with HSA. Arrows indicate the time of injection for ADAPT$_{Angiogenin06}$ and angiogenin.

representative example of ADAPT$_{Angiogenin06}$, binding to its target angiogenin in a concentration-dependent manner at the same time as it is captured on the chip through its binding to HSA (Supporting Information Figure S16 shows sensorgrams for the other variants in the capture assay). All variants also demonstrated binding following target saturation (Supporting Information Figure S17), with an increased response due to the increased molecular weight of the complex. This was true for all variants except the one targeting TNFα where the target seems to interfere somehow with the HSA binding, resulting in a lowered response. In total, 10 of the characterized binders

from the selections toward TNFα, PSA, angiogenin, MYDGF, and insulin could bind both HSA and its target protein simultaneously.

Evaluation of Ternary Complex Formation Using Size-Exclusion Chromatography. As a final characterization, the ability to form the ternary complex consisting of one of the ADAPT variants (ADAPT$_{Insulin\_02}$) together with albumin as well as the target was evaluated. The chromatogram is shown in the Supporting Information Figure S18. A small shift in the elution profile can be seen when the ternary complex has been allowed to form, compared to both the injection of the binary complex as well as of the individual binding partners. The signal overlap is attributed to the limited resolution of the SEC column. Further strengthening the claim of complex formation, no free ADAPT$_{Insulin\_02}$ and only a small amount of free insulin can be detected when injected as a complex. These data confirm what we could see in the SPR results that the ADAPTs are not only bispecific but they can also bind both albumin and its target at the same time.

DISCUSSION

Small size is an attractive feature of therapeutic proteins since it allows for high tissue penetration and noninvasive administration routes, as well as easy and cost-efficient production. However, advantages often come with a price. These small proteins typically have a limited *in vivo* half-life leading to substantial kidney uptake as well as lower therapeutic efficacy. A well-known strategy to overcome these limitations is to utilize the long half-life of human serum albumin either by direct fusion to albumin itself or to different albumin-binding moieties. Previous studies have shown that fusing therapeutic molecules to an ABD derived from streptococcal protein G leads to a substantial decrease in kidney uptake and considerably longer half-lives.[10,23]

In this particular study, we have utilized the albumin-binding properties of ABD, but instead of fusing it to a therapeutic molecule, we have introduced an additional binding site in the molecule itself and hence obtained small single-domain bispecific binders with potentially improved pharmacokinetic properties. To do this, we designed a combinatorial phage display library that was used to select molecules with the aim of binding both HSA and other clinically relevant targets at the same time.

We have previously confirmed that it is possible to introduce dual binding specificity into this scaffold.[13−16] However, in these studies, binding of one of the partners somehow prevented simultaneous binding of the other one. We hypothesized that this might be due to steric hindrance and therefore designed our library with a novel binding surface on the opposite side of the molecule (see Figure 1). This proved to be a successful strategy as it resulted in several bispecific molecules exhibiting simultaneous binding.

The binders were identified through the use of phage display in combination with Illumina high-throughput sequencing to gain a broader perspective on the enrichment process. The phage display technique has allowed for incredible discoveries and led to a great paradigm shift within the field of antibody development, but it is not without limitations. Performing repetitive panning rounds, each followed by an amplification step in bacterial cells, inevitably introduces bias. Furthermore, the high enrichment of one or a few sequences leads to a loss of sequence diversity and hence loss of some of the relevant variants that simply do not propagate as fast as others. Next-

generation sequencing is therefore particularly suitable for the application of phage-displayed ADAPTs since one 150 bp read of the Illumina MiSeq covers the entire length of the randomized sequence. Also, by pooling together many individually barcoded samples in one single analysis, we have acquired hundreds of thousands of sequences from each of many different selections and panning rounds at a low cost. Another advantage of next-generation sequencing is of course the depth of the data. Looking at the frequency at which the selected binders appear in their respective phage pools (Supporting Information Table S3), it can be seen that not all of these sequences have a fair chance of being discovered using traditional Sanger sequencing methods. In addition, by sequencing earlier panning rounds, we could see that the binding variants were already sufficiently enriched to be discovered after round 2 (Supporting Information Figures S2−S7). To test this in a real setting, we made a selection toward insulin and picked four sequenced clones that appeared already after two panning rounds, which indeed resulted in two functional and simultaneous binders. This confirms previously reported findings where positive clones have been identified without the need for tedious repetitive panning rounds.[24,25] This strategy, in particular, would not be possible without next-generation sequencing since the variants binding to insulin were only present in less than 0.1% of the total phage pool.

Despite the advantages of next-generation sequencing, a major limitation of this workflow is the recovery of individual clones from the phage pool after identification. During this project, we have mainly worked with synthesized gene fragments, which is not a very cost-efficient alternative if one wants to screen many variants. To overcome this, we later also developed a PCR-based strategy where unique but shorter oligos were used to rescue individual clones. Taken together, the fact that positive clones can be discovered already after two panning rounds as well as recovered with a straightforward PCR-based method leads to significant time and cost savings in the phage display process. It also removes some obstacles associated with performing many parallel selections in a multiplex and high-throughput manner.

Following characterization of the enriched variants, it could be noted that several bispecific binders were identified in all seven different selections. The measured affinities were in the nanomolar to micromolar range, and with angiogenin being the only exception, they are a result of relatively fast on-rates but also fast off-rates (Supporting Information Table S3). Surprisingly, the binders that were isolated in the selection toward angiogenin demonstrated much slower off-rates. For the remaining molecules to be of major therapeutic value, affinity maturation is needed to, most importantly, improve the off-rates. We believe that more stringent selection conditions could be one way to achieve this. We also think that the insight gained into the amino acid preferences in each position will help guide us in the construction of new and even more focused maturation libraries. Furthermore, since the rationale behind choosing these ADAPT variants was that they appeared in different sequential clusters, another strategy could be to look deeper into the clusters where binding variants were identified.

Interestingly, even though bispecific binders were identified in all selections regardless of target protein, not all resulted in binders that could bind both targets simultaneously. The SPR capture assay revealed simultaneous binders for the targets TNFα, PSA, angiogenin, MYDGF, and insulin, which was

confirmed also in the target saturation assay. In the latter, the binder targeting TNFα deviates from the rest in that the binding response seems to be influenced to some extent by the binding of the target. A possible explanation for this is the fact that TNFα exists both in a monomeric and a trimeric form,[26] where the trimeric form might cause sterical hindrance in the SPR setup.

For the remaining targets, CRP and renin, simultaneous binding was not observed. Looking at the sizes of these targets, it is not evident that the molecular weight would influence the ability of simultaneous binding, even though it seems plausible that larger target proteins have a higher risk of causing sterical hindrance. With the increasing knowledge gained from these and future selections, we are confident that it is possible to create a more focused library that is even more successful in selecting not only numerous bispecific but also simultaneous binders.

As we expected, all variants bind HSA weaker than the parental ABD035. However, the affinities are still in the low nanomolar range, and considering the very high HSA concentration in serum, these affinities are well above what is needed for efficient half-life extension. Indeed, previous studies have shown that the half-life extension is only weakly influenced by the affinity, even with affinities in the micromolar range.[27]

It remains to be evaluated how these minimized scaffolds behave in an *in vivo* setting compared to other small molecules simply fused to the ABD. Earlier studies where we have fused two ADAPTs together,[28] as well as other studies on the Affibody scaffold,[29] suggest that even a quite small size increase from ∼6 to ∼12 kDa leads to decreased uptake in tumor tissue despite increased affinities. Other studies show a linear relationship between smaller size and increasing tissue distribution also for many other tissues.[30] Association with albumin *in vivo* of course increases the size of the complex substantially, which could potentially limit the transport benefits of the small protein domains. However, since this is a noncovalent interaction, the possibilities of tunability are almost unlimited. In addition, in tumor-targeting applications, albumin probably also contributes with other accumulative properties thanks to the enhanced permeability and retention (EPR) effect.[31] Due to the many different interactions, the biodistribution *in vivo* is hard to predict a priori and future studies are necessary to answer these questions.

CONCLUSIONS

In this study, we have designed a new combinatorial library with the aim of developing bispecific ADAPTs toward various new targets in addition to serum albumin. Through the use of phage display and next-generation sequencing, we have identified and evaluated bispecific binders toward seven different targets and demonstrated simultaneous binding to albumin for five of these. These single-domain bispecific binders offer an opportunity of combining the advantages of small molecules, such as good tissue penetration and noninvasive administration routes, with those of much larger molecules that have a considerably longer *in vivo* half-life.

The results from this study confirm that a small 6 kDa domain can indeed accommodate two separate binding surfaces and hence bind to albumin without compromising either its size or the binding to its intended target. Furthermore, we have also developed an efficient workflow combining phage display with next-generation sequencing to

allow for the time- and cost-efficient selection of these binders toward many more targets in the future.

ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.molpharmaceut.0c00975.

> Sequencing data and experimental data of yield; α-helical structure; monomeric state; and binding of the evaluated protein variants (PDF)

AUTHOR INFORMATION

Corresponding Author

Sophia Hober − *Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm 10691, Sweden;* ⓞ orcid.org/0000-0003-0605-8417; Email: sophia@kth.se

Authors

Emma von Witting − *Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm 10691, Sweden;* ⓞ orcid.org/0000-0002-4695-7858

Sarah Lindbo − *Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm 10691, Sweden*

Magnus Lundqvist − *Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm 10691, Sweden*

Marit Möller − *Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm 10691, Sweden*

Andreas Wisniewski − *Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm 10691, Sweden*

Sara Kanje − *Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm 10691, Sweden;* ⓞ orcid.org/0000-0002-4751-2519

Johan Rockberg − *Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm 10691, Sweden*

Hanna Tegel − *Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm 10691, Sweden*

Mikael Åstrand − *Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm 10691, Sweden*

Mathias Uhlén − *Protein Science, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, Stockholm 10691, Sweden*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.molpharmaceut.0c00975

Author Contributions

E.v.W. and S.L. contributed equally to this work. All authors have given approval to the final version of the manuscript.

■ ABBREVIATIONS

ABD, albumin-binding domain; ADAPT, ABD-derived affinity protein; HSA, human serum albumin

■ REFERENCES

(1) Urquhart, L. Top drugs and companies by sales in 2017. *Nature* **2018**, *17*, No. 232.

(2) Grilo, A. L.; Mantalaris, A. The Increasingly Human and Profitable Monoclonal Antibody Market. *Trends Biotechnol.* **2019**, *37*, 9−16.

(3) Bajracharya, R.; Song, J. G.; Back, S. Y.; Han, H.-K. Recent Advancements in Non-Invasive Formulations for Protein Drug Delivery. *Comput. Struct. Biotechnol. J.* **2019**, *17*, 1290−1308.

(4) Donovan, M. D.; Huang, Y. Large molecule and particulate uptake in the nasal cavity: the effect of size on nasal absorption. *Adv. Drug Delivery Rev.* **1998**, *29*, 147−155.

(5) Orlova, A.; Wållberg, H.; Stone-Elander, S.; Tolmaehev, V. On the selection of a tracer for PET imaging of HER2-expressing tumors: Direct comparison of 124I-labeled affibody molecule and trastuzumab in a murine xenograft model. *J. Nucl. Med.* **2009**, *50*, 417−425.

(6) Goldstein, R.; Sosabowski, J.; Livanos, M.; Leyton, J.; Vigor, K.; Bhavsar, G.; Nagy-Davidescu, G.; Rashid, M.; Miranda, E.; Yeung, J.; Tolner, B.; Plückthun, A.; Mather, S.; Meyer, T.; Chester, K. Development of the designed ankyrin repeat protein (DARPin) G3 for HER2 molecular imaging. *Eur. J. Nucl. Med. Mol. Imaging* **2015**, *42*, 288−301.

(7) Garousi, J.; Lindbo, S.; Nilvebrant, J.; Åstrand, M.; Buijs, J.; Sandström, M.; Honarvar, H.; Orlova, A.; Tolmachev, V.; Hober, S. ADAPT, a Novel Scaffold Protein-Based Probe for Radionuclide Imaging of Molecular Targets That Are Expressed in Disseminated Cancers. *Cancer Res.* **2015**, *75*, 4364−4371.

(8) Kontermann, R. E. Strategies for extended serum half-life of protein therapeutics. *Curr. Opin. Biotechnol.* **2011**, *22*, 868−876.

(9) Tolmachev, V.; Orlova, A.; Pehrson, R.; Galli, J.; Baastrup, B.; Andersson, K.; Sandström, M.; Rosik, D.; Carlsson, J.; Lundqvist, H.; Wennborg, A.; Nilsson, F. Y. Radionuclide Therapy of HER2-Positive Microxenografts Using a Lu-Labeled HER2-Specific Affibody Molecule. *Cancer Res.* **2007**, *67*, 2773−2782.

(10) Orlova, A.; Jonsson, A.; Rosik, D.; Lundqvist, H.; Lindborg, M.; Abrahmsen, L.; Ekblad, C.; Frejd, F. Y.; Tolmachev, V. Site-specific radiometal labeling and improved biodistribution using ABY-027, a novel HER2-targeting affibody molecule-albumin-binding domain fusion protein. *J. Nucl. Med.* **2013**, *54*, 961−968.

(11) Liu, H.; Lindbo, S.; Ding, H.; Altai, M.; Garousi, J.; Orlova, A.; Tolmachev, V.; Hober, S.; Gräslund, T. Potent and specific fusion toxins consisting of a HER2-binding, ABD-derived affinity protein, fused to truncated versions of Pseudomonas exotoxin A. *Int. J. Oncol.* **2019**, *55*, 309−319.

(12) Sand, K. M. K.; Bern, M.; Nilsen, J.; Noordzij, H. T.; Sandlie, I.; Andersen, J. T. Unraveling the Interaction between FcRn and

Albumin: Opportunities for Design of Albumin-Based Therapeutics. *Front. Immunol.* **2015**, *5*, 1−21.

(13) Alm, T.; Yderland, L.; Nilvebrant, J.; Halldin, A.; Hober, S. A small bispecific protein selected for orthogonal affinity purification. *Biotechnol. J.* **2010**, *5*, 605−617.

(14) Nilvebrant, J.; Alm, T.; Hober, S.; Löfblom, J. Engineering Bispecificity into a Single Albumin-Binding Domain. *PLoS One* **2011**, *6*, No. e25791.

(15) Nilvebrant, J.; Åstrand, M.; Georgieva-Kotseva, M.; Björnmalm, M.; Löfblom, J.; Hober, S. Engineering of Bispecific Affinity Proteins with High Affinity for ERBB2 and Adaptable Binding to Albumin. *PLoS One* **2014**, *9*, No. e103094.

(16) Nilvebrant, J.; Åstrand, M.; Löfblom, J.; Hober, S. Development and characterization of small bispecific albumin-binding domains with high affinity for ErbB3. *Cell. Mol. Life Sci.* **2013**, *70*, 3973−3985.

(17) Zorzi, A.; Linciano, S.; Angelini, A. Non-covalent albumin-binding ligands for extending the circulating half-life of small biotherapeutics. *MedChemComm* **2019**, *10*, 1068.

(18) Grönwall, C.; Jonsson, A.; Lindström, S.; Gunneriusson, E.; Ståhl, S.; Herne, N. Selection and characterization of Affibody ligands binding to Alzheimer amyloid β peptides. *J. Biotechnol.* **2007**, *128*, 162−183.

(19) Uhlén, M.; Karlsson, M. J.; Hober, A.; Svensson, A. S.; Scheffel, J.; Kotol, D.; Zhong, W.; Tebani, A.; Strandberg, L.; Edfors, F.; Sjöstedt, E.; Mulder, J.; Mardinoglu, A.; Berling, A.; Ekblad, S.; Dannemeyer, M.; Kanje, S.; Rockberg, J.; Lundqvist, M.; Malm, M.; Volk, A. L.; Nilsson, P.; Månberg, A.; Dodig-Crnkovic, T.; Pin, E.; Zwahlen, M.; Oksvold, P.; von Feilitzen, K.; Häussler, R. S.; Hong, M. G.; Lindskog, C.; Ponten, F.; Katona, B.; Vuu, J.; Lindström, E.; Nielsen, J.; Robinson, J.; Ayoglu, B.; Mahdessian, D.; Sullivan, D.; Thul, P.; Danielsson, F.; Stadler, C.; Lundberg, E.; Bergström, G.; Gummesson, A.; Voldborg, B. G.; Tegel, H.; Hober, S.; Forsström, B.; Schwenk, J. M.; Fagerberg, L.; Sivertsson, Å. The human secretome. *Sci. Signal.* **2019**, *12*, No. eaaz0274.

(20) Jonsson, A.; Dogan, J.; Herne, N.; Abrahmsén, L.; Nygren, P. Å. Engineering of a femtomolar affinity binding protein to human serum albumin. *Protein Eng., Des. Sel.* **2008**, *21*, 515−527.

(21) Lejon, S.; Frick, I.-M.; Björck, L.; Wikström, M.; Svensson, S. Crystal Structure and Biological Implications of a Bacterial Albumin Binding Module in Complex with Human Serum Albumin. *J. Biol. Chem.* **2004**, *279*, 42924−42928.

(22) Johansson, M. U.; Frick, I.-M.; Nilsson, H.; Kraulis, P. J.; Hober, S.; Jonasson, P.; Linhult, M.; Nygren, P.-Å.; Uhlén, M.; Björck, L.; Drakenberg, T.; Forsén, S.; Wikström, M. Structure, Specificity, and Mode of Interaction for Bacterial Albumin-binding Modules. *J. Biol. Chem.* **2002**, *277*, 8114−8120.

(23) Andersen, J. T.; Pehrson, R.; Tolmachev, V.; Daba, M. B.; Abrahmsén, L.; Ekblad, C. Extending half-life by indirect targeting of the neonatal Fc receptor (FcRn) using a minimal albumin binding domain. *J. Biol. Chem.* **2011**, *286*, 5234−5241.

(24) Christiansen, A.; et al. High-throughput sequencing phage display enables the identification of patient-specific epitope motifs in serum. *Sci. Rep.* **2015**, *5*, No. 12913.

(25) 't Hoen, P. A. C.; Jirka, S. M. G.; Ten Broeke, B. R.; Schultes, E. A.; Aguilera, B.; Pang, K. H.; Heemskerk, H.; Aartsma-Rus, A.; Van Ommen, G. J.; Den Dunnen, J. T. Phage display screening without repetitive selection rounds. *Anal. Biochem.* **2012**, *421*, 622−631.

(26) Smith, R. A.; Baglioni, C. The active form of tumor necrosis factor is a trimer. *J. Biol. Chem.* **1987**, *262*, 6951−6954.

(27) Hopp, J.; Hornig, N.; Zettlitz, K. A.; Schwarz, A.; Fuß, N.; Mü, D.; Kontermann, R. E. The effects of affinity and valency of an albumin-binding domain (ABD) on the half-life of a single-chain diabody-ABD fusion protein. *Protein Eng., Des. Sel.* **2010**, *23*, 827−834.

(28) Garousi, J.; Lindbo, S.; Borin, J.; von Witting, E.; Vorobyeva, A.; Oroujeni, M.; Mitran, B.; Orlova, A.; Buijs, J.; Tolmachev, V.; Hober, S. Comparative evaluation of dimeric and monomeric forms of ADAPT scaffold protein for targeting of HER2-expressing tumours. *Eur. J. Pharm. Biopharm.* **2019**, *134*, 37−48.

(29) Tolmachev, V.; Friedman, M.; Sandström, M.; Eriksson, T. L. J.; Rosik, D.; Hodik, M.; Ståhl, S.; Frejd, F. Y.; Orlova, A. Affibody molecules for epidermal growth factor receptor targeting in vivo: Aspects of dimerization and labeling chemistry. *J. Nucl. Med.* **2009**, *50*, 274−283.

(30) Li, Z.; Krippendorff, B. F.; Sharma, S.; Walz, A. C.; Lavé, T.; Shah, D. K. Influence of molecular size on tissue distribution of antibody fragments. *MAbs* **2016**, *8*, 113−119.

(31) Maeda, H. Tumor-Selective Delivery of Macromolecular Drugs via the EPR Effect: Background and Future Prospects. *Bioconjug. Chem.* **2010**, *21*, 797−802.