peds
protein engineering design & selection

Protein Engineering, Design and Selection, 2021, vol. 34, pp. 1–11
doi: 10.1093/protein/gzaa029
Original Article

OXFORD

Original Article

Construction, characterization and crystal structure of a fluorescent single-chain Fv chimera

Nileena Velappan[1,6,†,*], Devin Close[2,6,†], Li-Wei Hung[1,6,†], Leslie Naranjo[1,3], Colin Hemez[4], Natasha DeVore[5], Donna K. McCullough[6], Antonietta M. Lillo[1], Geoffrey S. Waldo[1], and Andrew R.M. Bradbury[3,*]

[1]Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA, [2]ARUP Laboratories, Institute for Clinical and Experimental Pathology, Salt Lake City, UT 84108, USA, [3]Specifica Inc., Santa Fe, NM 87505, USA, [4]Graduate Program in Biophysics, Harvard University, Boston, MA 02115 USA, [5]Chemistry Department, Missouri State University, Springfield, MO 65897, USA, and [6]Microbiology Department, University of Tennessee, Knoxville, TN 37996, USA

*To whom correspondence should be addressed. E-mail: abradbury@specifica.bio (Andrew Bradbury) and nileena@lanl.gov (Nileena Velappan)

[†]Equal contributing first author

Received 24 August 2020; Revised 21 December 2020; Accepted 24 December 2020

Abstract

In vitro display technologies based on phage and yeast have a successful history of selecting single-chain variable fragment (scFv) antibodies against various targets. However, single-chain antibodies are often unstable and poorly expressed in *Escherichia coli*. Here, we explore the feasibility of converting scFv antibodies to an intrinsically fluorescent format by inserting the monomeric, stable fluorescent protein named thermal green, between the light- and heavy-chain variable regions. Our results show that the scTGP format maintains the affinity and specificity of the antibodies, improves expression levels, allows one-step fluorescent assay for detection of binding and is a suitable reagent for epitope binning. We also report the crystal structure of an scTGP construct that recognizes phosphorylated tyrosine on FcεR1 receptor of the allergy pathway.

Key words: crystal structure, epitope binning, improved expression and functionality, intrinsically fluorescent scFv antibodies, soluble scFv antibodies

Introduction

Fluorescently labeled antibodies are extensively used in diagnostic and research methods such as immunofluorescent microscopy and flow cytometry. The latter allows the rapid and quantitative analysis of single cells within a large population, and fluorescent-activated cell sorting (FACS) allows the sorting of individual cells based on their specific fluorescent patterns. Antibodies used in these assays are typically labeled with fluorescent chemical compounds. For example, a succinimidyl ester functional group attached to a fluorophore core may react with primary amines to label the antibody. However, this strategy may reduce or eliminate binding if there are primary amines within the antigen recognition regions (Mccormack *et al.*, 1996). Different antibodies react with different fluorophores at different rates, creating variability in experimental settings. Furthermore, since the labeling is stochastic, some antibodies will have large numbers of fluorophores and some very few. Fluorophore over-conjugation can result in quenching and dimmer antibody molecules than those with less fluorophore conjugation (Haugland, 1995). While standard

© The Author(s) 2021. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/peds/article/doi/10.1093/protein/gzaa029/6135052 by University of Liverpool user on 20 February 2021

techniques have been developed to perform antibody labeling reactions, batch-to-batch variation and labeling can also affect antibody-antigen recognition (Vira *et al.*, 2010).

One of the most effective methods to develop antibodies is to select them from large combinatorial libraries using phage and/or yeast *in vitro* display technologies (Boder and Wittrup, 1997; Ferrara *et al.*, 2012; Marks *et al.*, 1991; Sblattero and Bradbury, 2000). As full-length antibodies cannot be directly displayed easily, Fabs or single-chain variable fragments (scFvs) (Huston *et al.*, 1988), which comprise the antibody heavy-chain variable (VH)- and light-chain variable (VL)-binding domains, can be displayed on phage or yeast fused to surface display proteins. The scFv genes can be constructed by connecting VH and VL regions with an unstructured linker peptide (Huston *et al.*, 1988), or one that comprises a translated DNA recombination signal (Sblattero and Bradbury, 2000). The genes encoding the variable regions are available during the selection process and can be reformatted into many different functional modalities (Bradbury and Pluckthun, 2015).

However, scFvs tend to aggregate, have low expression levels and are unstable during long-term storage (Arndt *et al.*, 1998). To monitor binding, scFvs are typically tagged, then detected using an anti-tag antibody and appropriately labeled secondary antibodies. Several strategies have been used to overcome these limitations (Fig. 1), including cloning as full-length antibodies (Persic *et al.*, 1997) or scFv-Fc fusions (Shu *et al.*, 1993), in which case traditional labeled secondary antibodies can be used. Fusion to enzymatic or fluorescent proteins allows detection without the need for secondary antibodies. The scFv-alkaline phosphatase (AP) fusions (Lindner *et al.*, 1997) are stable and functional, allowing single step detection of bound antibody. Fusion of green fluorescent protein (GFP) or other fluorescent proteins (FPs) to the C terminus of scFv genes yields intrinsically fluorescent scFv. However, scFv-GFP fusions tend to be poorly expressed (Casey *et al.*, 2000). The scFv folds best in oxidizing environments, such as the bacterial periplasm or eukaryotic secretory pathway (Kipriyanov *et al.*, 1997), while most FPs are best expressed in the reducing environment of the cytoplasm. A notable exception is superfolder GFP (Pédelacq *et al.*, 2006). This expression incompatibility has been overcome by fusing a short tag derived from split GFP to the C terminus of the scFv (Ferrara *et al.*, 2011) and carrying out the complementation *in vitro*, fusing GFP to particularly stable scFvs (Nizak *et al.*, 2003), or creating nanobody FP fusions (Rothbauer *et al.*, 2008).

Markiv *et al.* (2011a,b) described an alternative concept to generate fluorescent scFvs in which the FP was placed between the VH and VL as a linker (VH-FP-VL) and could be expressed in the bacterial cytoplasm. The fluorescent proteins used included mRFP1 (derived from DsRed), BFP, mCerulean and mCitrine to produce red, blue, green and yellow antibodies. The alternative orientation, VL-VH, was not attempted. Chain orientation may be important in binding properties, and the strategy described in that publication may not be compatible for scFv libraries constructed in the VL-linker-VH orientation. Markiv *et al.* (2011a,b) use the term REDantibody in their first description of the technology. However, no specific terminology was used in the subsequent paper when additional fluorescent proteins were tested (Markiv *et al.*, 2011a,b). Here we refer to the concept of fluorescent antibodies created using a fluorescent protein as a linker as single-chain fluorescent protein (scFP), with specific formats called out for individual scFPs, e.g. scTGP for the scFPs described here in which the thermal green protein (TGP) (Close *et al.*, 2015) is used as a linker. Figure 1 provides a visual summary of exemplar scFv

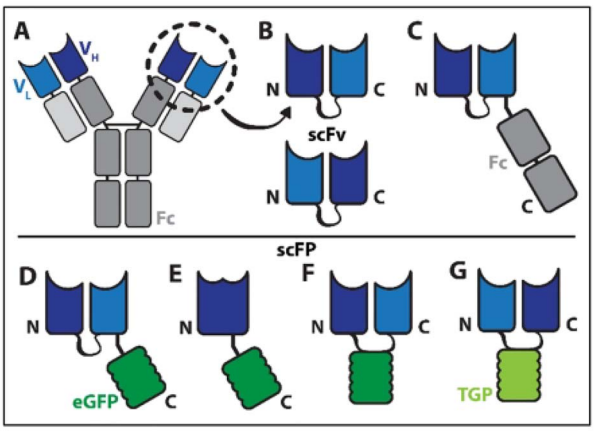


Fig. 1 Schematic diagram depicting examples of antibody formats utilized in research. Panel **A** depicts the traditional IgG and panel **B** depicts the corresponding scFv constructs (in either VH-VL or VL-VH orientation). Examples for scFv fusion constructs are given in panels **C** and **D**. Panel **E** shows a nanobody-FP fusion, VH-FP-VL fusion is given in panel **F** and panel **G** provides the concept for the work described here.

modifications utilized in research in comparison to the scTGP format presented here.

TGP is an extremely stable, highly soluble, non-aggregating green fluorescent protein that was engineered in our lab (Close *et al.*, 2015) to overcome the problems of solubility and aggregation with many traditional GFPs and their variants. TGP was engineered by disrupting crystal lattice contacts and introducing high-entropy glutamate residues to improve crystallization and prevent oligomerization. The monomeric nature of TGP makes it potentially suitable for insertion into scFv genes, substituting the linker originally used to tether the VL and VH. In this study, we construct single-chain fluorescent proteins (scFPs) as VL-TGP-VH. These scFPs are fluorescent, are functional in yeast display technologies and are expressed at high levels in the bacterial cytoplasm. We describe the construction of scFPs and their functionality using multiple assays. We use three scFvs and compare them with the corresponding scTGP formats. The antibodies p1-1 and p13B1 recognize phosphorylated immunoreceptor tyrosine-based activation motif (ITAM) peptides of FcεR1 receptor of the allergy pathway (Velappan *et al.*, 2019) and the anti–tyrosine sulfate antibody recognizes the post-translationally modified tyrosine within proteins or peptides (Kehoe *et al.*, 2006). We report the X-ray crystal structure of scTGP p1-1.

Finally, we use scTGP and IgGs to develop a novel epitope binning method for multimeric proteins. We demonstrate this method using three antibodies against *Yersinia pestis*, the causative agent of plague (Lillo *et al.*, 2011; Lillo *et al.*, 2020). Epitope binning uses competitive pairwise binding assays to sort antibodies into bins recognizing the same or different epitopes. Antibodies binned together often have similar functional activity (e.g. block binding to a cell surface receptor), while those in different bins recognize different epitopes and consequently may have different functional activities. This proof of concept demonstrates the utility of the scTGP format for generating clinically relevant antigen detection tools.

Materials and Methods

Construction of scTGP genes

The scTGP constructs were prepared using inverse PCR and circular polymerase extension cloning (CPEC) assembly. The vector

Downloaded from https://academic.oup.com/peds/article/doi/10.1093/protein/gzaa029/6135052 by University of Liverpool user on 20 February 2021

was amplified using scFv linker–specific primers DCO 111 (TCCTCAAGCGGTACCCAGGTGCA) and DCO 112 (TCTGGAGGGTCGACCATAACTTCG). This primer combination was used to amplify the yeast display vector as well as the bacterial expression plasmids containing the scFv gene(s). The TGP gene was amplified using DCO 109 (TCTGGAGGGTCGACCATAACTTCGGTAATTAAACCGGAAATGAAAATTAAATTGCG) and DCO 110 (TATTCTGGCGGAGGCAGCGGATCCTCAAGCGGTACCCAGGTACA). Q5 polymerase (M4091, New England Biolabs) was used for all amplification reactions according to manufacturer's instructions. Both vector and insert were gel purified prior to CPEC assembly. CPEC assembly was performed by combining equimolar amounts of vector and fluorescent protein insert and reducing the annealing temperature gradually from 70 to 55°C using 10 s intervals. Upon completion of the CPEC assembly, the reaction was purified using a mini-elute column (28004, Qiagen Inc.). The constructs were eluted into 10 µl of 1:2 elution buffer: water and transformed via electroporation to bacterial cells. Yeast display constructs were transformed into Omnimax T1 or TOPO 10 cells (C854003 or C4044003, ThermoFisher Scientific) according to manufacturer's instructions and bacterial expression plasmids were cloned into BL21 DE3 (archival samples LANL) cells. Constructs were verified by DNA sequencing after cloning.

Yeast display analysis

Yeast display plasmids (based on pDNL6) (Ferrara *et al.*, 2012) were transformed into *Saccharomyces cerevisiae* EBY100 and cloned by gap repair using the Yeast1 kit (Sigma Inc.). In this vector system, the scFvs are fused to the C terminus of the AGA2 protein of the yeast for surface display and the SV5 peptide tag is used to assess display levels on yeast. The transformed yeast were grown in SD/CAA (5 g/l casamino acids (−ade, −urs, −trp), 20 g/l dextrose, 1.7 g/l yeast nitrogen base, 5.3 g/l ammonium sulfate, 10.19 g/l $Na_2HPO_4.7H_2O$ and 8.58 g/l NaH_2PO_4) agar plates at 30°C for 2 days. Single colonies were further grown in SD/CAA liquid culture at 30°C until the culture reached an OD_{600} of >2. 1/10th volume of this culture was added to SG/R CAA media (same as selective media except substitute the following sugars for dextrose: 20 g/l galactose, 20 g/l raffinose, 1 g/l dextrose) and induced at 20°C as described previously (Ferrara *et al.*, 2012). Yeast inductions were allowed to proceed for 2 days to further improve display of scTGP.

After induction, yeast cells were incubated with phycoerythrin (PE) conjugated anti-SV5 IgG (1 µg/ml) to assess scFv/scTGP display levels. In the case of scTGPs, the display levels were also assessed using the fluorescence at 530 nm when excited using 488 nm laser. The time course experiment reported in Supplementary Fig. S1, available at *PEDS* online, was performed with three different colonies of yeast from each of constructs. For analysis on three consecutive days post-induction, 100 µl of yeast was taken, fluorescence of yeast was analyzed using flow cytometry and average fluorescence along with standard deviation from these replicates was calculated.

Antibody p1-1 (Velappan *et al.*, 2019) recognizes a 25 amino acid peptide that corresponds to the ITAM region of the FcεR1 receptor of mast cells when the first tyrosine of this peptide is phosphorylated, and peptide was named p1 (KVPDDRL**Y**EELHVYSPIYSALEDTR, the phosphorylated tyrosine is bold and underlined). Antibody p13B1 (Velappan *et al.*, 2019) recognizes the same peptide when all three tyrosines are phosphorylated and designated as p123 (KVPDDRL**Y**EELHV**Y**SPI**Y**SALEDTR). The M2 peptide of influenza A was used as non-specific antigen (MSLLTEVETPIRNEWGCRCNDSSD). The yeast (100–200 µl) were first washed with yeast wash buffer (YWB—30 mM Tris pH 8.0 with 0.5% BSA), by centrifugation at 13 000*g*. The binding assays were performed using 100 nM peptide antigens diluted in YWB and incubated for 1 h at room temperature with rotation. Unbound peptide was washed away using YWB. The antigen bound yeast was further stained with anti-SV5 PE (1 µg/ml) and streptavidin labeled with Alexa 633 (S21375, ThermoFisher Scientific, 5 µg/ml) for biotinylated peptide detection. Washes with YWB were conducted to improve specificity of binding. All incubations and washes were performed at room temperature. The yeast display analysis was performed using FACS Aria and/or LSR II cytometers (Becton Dickinson). For each flow cytometry analysis, 10 000 events were collected and average value was used to report mean fluorescence intensity (MFI) for display level and binding signal and the assays were repeated multiple times to ensure accuracy. Experiments were performed in triplicates to measure the K_d values. Affinity measurements were conducted by serial dilution of the biotinylated peptide antigens as described previously (Ferrara *et al.*, 2012; Velappan *et al.*, 2019). K_d determination was performed using GraphPad Prism® software using the equation for one site-specific binding, which uses non-linear regression analysis to fit the curve.

Protein expression in bacteria

All genes were cloned into the vector systems described below, using restriction enzymes BssH II and Nhe I (R0199 and R3131, New England Biolabs). The scFv and scTGP genes were both cloned into pEP, a periplasmic expression vector, while the scTGP constructs were also cloned into pETCK3, a cytoplasmic expression system. The protein expression in both plasmid systems is controlled by the T7 promoter system.

To produce scFvs and scTGPs, bacterial colonies were streaked on LB kan/glu (kanamycin 50 µg/ml and glucose 3%) plates and incubated at 37°C overnight. A second overnight culture was grown in 1–10 ml of LB media containing kanamycin and glucose. Liquid culture was grown at 30°C. The overnight culture volume was set to 1/10th final expression culture volume. The following day, the overnight culture was diluted to half the final culture volume and grown in kan/glu media at 30°C for 2 h. The culture was centrifuged (6000*g* for 6 min) and the pellet was re-suspended in auto-induction (AI) media containing kanamycin (Studier, 2005). AI media allows protein induction without the use of isopropyl β-d-1-thiogalactopyranoside (IPTG). Here, glucose, glycerol and lactose are added to buffered yeast broth growth media. As the *Escherichia coli* cultures grow, they consume glucose first. As the glucose depletes, the bacteria are forced to utilize lactose inducing the T7 promoter and allowing protein production. The culture was incubated at 30°C for 4 h. Subsequently, the temperature was reduced to 18°C and protein expression continued for 2 days (scFvs) or up to 3 days (scTGPs). The cell culture was spun down (10 000*g*, 20 min) and frozen prior to protein purification. Three colonies from each of the constructs (scFv-pEP, scTGP-pEP, scTGP-pET) for antibodies p1-1 and p13B1 were used to analyze functionality of proteins expressed in different bacterial compartments. About 1 ml of expressed culture was spun down and the bacterial pellet was lysed using 1/10 POP culture reagent (71092, EMD Millipore). Incubation time was 30 minutes at room temperature and the insoluble fraction was separated by centrifugation (20 000*g* at 4°C). About 100 µl of cell culture was used to measure the fluorescence of pre-lysed culture and 100 µl of POP

Downloaded from https://academic.oup.com/peds/article/doi/10.1093/protein/gzaa029/6135052 by University of Liverpool user on 20 February 2021

culture supernatant was used to measure fluorescence post-lysis. The cytoplasmically expressed scTGP had higher fluorescence than the periplasmically expressed scTGP. Fluorescence-linked immunosorbent assay (FLISA) was performed with two normalizations: equal volume or equal fluorescence. Here, 100 µl of post-lysis supernatants or volume of supernatants adjusted to provide equal fluorescence was used. Both expression and binding assays were performed in triplicate and average values and standard deviations were calculated. A two-tailed T-test was used to compare the two expression systems.

Protein purification

The bacterial pellets were re-suspended in TNG buffer (50 mM Tris pH 8.0, 150 mM NaCl, 10% v/v glycerol) containing DNase 1 (PI89836, Thermo Fisher Scientific) and EDTA-free protease inhibitor tablet (11873580001, Roche Inc.). Cell lysis was achieved using a cell homogenizer (Avestin Inc.). The lysate was spun at 13 000*g* and the supernatant was then spun at 38 000*g* and filter sterilized to remove cell debris. Protein purification was performed using Ni-NTA agarose beads. The His-tagged proteins were allowed to bind to the beads for 1 h at 4°C with rotation. The beads were washed with 5–10× volume of TNG buffer and TNG buffer containing 20 mM imidazole. Proteins were eluted with TNG buffer containing 250 mM imidazole. The protein was cleaved from the His-tag with TEV protease using 0.5–1 mg of TEV protease and the digestion proceeded in a dialysis cassette in TEV buffer (20 mM Tris [pH 8.0], 100 mM NaCl, 0.5 mM EDTA, 3 mM reduced glutathione) at RT, overnight. The His-tag fragment and any uncleaved protein were subsequently removed by a second pass through fresh Ni-NTA agarose beads. Protein concentration and purity were determined using protein electrophoresis and Coomassie staining, absorbance at 280 nm or using fluorescence intensity. Prior to crystallization, scTGP proteins were further subjected to chromatography on a 320 ml XK 26/60 Sephadex 200 size-exclusion column using an AKTA prime liquid chromatography system (GE healthcare). Fluorescent protein fractions were collected and assessed using gel electrophoresis and Coomassie staining for purity. Pure fractions were combined and concentrated using Amicon centrifugal filters with 30 kDa cutoff. Protein expression for these constructs was conducted multiple times during the course of the project to ensure consistency once the protocol was established. A typical protein expression volume was 1 L followed by FPLC purification.

ELISA and FLISA assays

Neutravidin- (31000, Thermo Fisher Scientific) coated wells (100 µl of 10 µg/mL in PBS) were used in both enzyme-linked immunosorbent assay (ELISA) and FLISA (Velappan *et al*., 2008) to bind biotinylated antigens. Nunc maxisorp plates (12-565-136, ThermoFisher Scientific) and black maxisorp plates (43711, Nunc) were used for FLISA. The peptide antigens were synthesized with biotin at the N terminus and protein antigens were biotinylated using EZ-Link NHS-LC-LC-biotin (21343, Thermo Fisher Scientific) according to manufacturer's protocols. The wells were washed with PBS and blocked with blocking agent containing 2% milk (ELISA) or 2% BSA (FLISA) for 30 mM TRIS pH 8.0 was used as the buffer instead of PBS in binding assays with ITAM peptides. About 1 µg of bio-antigen in 100 µl was added to blocked wells and was allowed to bind to neutravidin for 15 min while shaking. Excess antigen was washed off and equal amounts of scFv and scTGP antibodies were added in blocking buffer diluted to 10% in PBS. Antigen-antibody interactions were allowed to proceed for 1.5 h and excess antibody was washed with PBST (1× PBS, 0.1% tween) three times followed by three washes with 1× PBS. For ELISA, anti-SV5 antibody conjugated to horse radish peroxidase (HRP) was used as secondary reagent. The bound HRP was detected using its substrate TMB (T8665, Sigma Inc.) and the absorbance at 450 nm was read after quenching with 1 M H_2SO_4. For FLISA, fluorescence was directly measured after washing with excitation set at 488 nm and emission set to 530 nm. All binding assays were performed in triplicate and average values and standard deviation were calculated. We repeated the binding assay for p1-1 and p13B1 scFv and scTGP proteins at equimolar concentration (50 nM of antibody). The assay was performed as described above.

Crystallization and structural determination

The scTGP p1-1 was crystallized using the sitting drop vapor diffusion method. Protein (10 mg/ml) and reservoir solutions of 0.1 µl each were mixed and equilibrated against 30 µl reservoir at 298 K using a PHOENIX crystallization robot (Art Robbins Instruments). A set of crystallization reagents consisting of Crystal Screens, PEG/Ion screens (Hampton Research), PACT suite (Qiagen), and JCSG core suites (Qiagen) was used to screen for the propensity of crystallization. Subsequent grid screens to optimize buffer pH, concentrations of salt and precipitants, and additive screens were employed as needed until diffraction-quality crystals were obtained. Crystals of scTGP were harvested with CryoLoops (Hampton Research) and flash-cooled in liquid nitrogen prior to diffraction experiments. Glycerol was added to the mother liquor as cryo-protectants to final concentrations of 5–20% for conditions that were not cryo-ready. X-Ray diffraction experiments were conducted at the 5.0.2 beam line at the Advanced Light Source at the Lawrence Berkeley National Laboratory. The crystal structure of scTGP was determined with the molecular replacement (MR) method using the PHASER program (Mccoy *et al*., 2007) in the PHENIX suite (Liebschner *et al*., 2019). The TGP structure (PDB ID: 4TZA) (Close *et al*., 2015) was used as a search model to locate the TGP moiety of the scTGP. The rest of the protein chain was modeled with the AutoBuild (Terwilliger *et al*., 2008) program in Phenix. Manual model rebuilding and refinement was conducted iteratively until convergence with Coot (Emsley *et al*., 2010) and phenix.refine (Afonine *et al*., 2012), respectively.

Epitope binning using competitive FLISA with scTGPs and IgGs

The anti-F1 scTGPs were constructed by PCR from scFvs (YP 1, 2, 3, 4, 6, 8) (Lillo *et al*., 2011) as described above and cloned into the pETCK3 protein expression vector for expression in BL21 *E.coli*. Protein expression, purification, quantification and binding to specific antigen (biotin-F1V) and control antigen (biotin-myoglobin) were performed with FLISA as described above. FLISA with gradient protein concentration (2×) starting at 20 000 ng in 100 µl volume/well was used to determine the binding curve of scTGP YP2, YP3 and YP8 (these antibodies were determined to be the most functional antibodies in this suite). The corresponding anti-YP IgGs were synthesized at ATUM Inc. and the F1V recognition by these IgGs was evaluated by ELISA using anti-human HRP (97165, Abcam) as the secondary antibody (Lillo *et al*., 2020). The epitope binning assay was performed similarly to FLISA. Here, the biotinylated F1V protein was bound to neutravidin-coated black immunosorp plates (8 wells/sample) and scTGP (5000 ng) was added to all wells. The binding proceeded for 30 minutes without washing while the titrations of the IgGs were performed. IgG concentrations ranging

Downloaded from https://academic.oup.com/peds/article/doi/10.1093/protein/gzaa029/6135052 by University of Liverpool user on 20 February 2021

Table I. Comparison of scFv and scTGP formats

	scFv p1-1	scTGP p1-1	scFv p13B1	scTGP p13B1
Display level (MFI) (SV5-PE)	9111	1280	10 215	5168
Recognition signal for specific target (MFI) (biotinylated target-streptavidin A633)	4696	283	4882	893
Recognition signal for non-specific target M2 (MFI) (biotinylated target-streptavidin A633)	160	127	160	118
Recognition signal for secondary only (no bio peptide) (MFI) (streptavidin A633)	142	115	112	154
K_d values for specific antigen (nM)	2.43 ± 0.22	8.55 ± 2.33	1.40 ± 0.65	0.51 ± 0.56

The table compares amount of protein displayed by the yeast, recognition signal for the specific peptide antigen, non-specific interaction signals and affinity of binding by two antibodies (p1-1 and p13B1) in both scFv and scTGP formats

from 0.75–50 µg were used. One well contained scTGP only. The titrated IgG solutions were added to the wells and incubated for 30–60 min. The unbound molecules were washed using PBST/PBS and the fluorescence of the bound scTGPs was measured using a spectrophotometer with excitation at 488 nm and emission at 530 nm. A decrease in fluorescence or lack thereof was used to determine presence or absence of different epitopes. Duplicate assays were performed to calculate average values and standard deviation was used to determine error bars.

Results

Comparison of scFv and scTGP constructs by yeast display

Three scFvs were used to test TGP in the scFP construct in the VL-VH orientation: p1-1 and p13B1 are two antibodies previously selected against different phosphorylation states of the cytoplasmic domain of the Fc-epsilon receptor (Velappan *et al*., 2019), and αTyrS is an antibody that recognizes the sulfotyrosine post-translational modification, independently of sequence context (Kehoe *et al*., 2006). Figure 2 and Table I together show that p1-1 scTGP and p13B1 scTGP were displayed on the yeast surface and recognized their cognate peptide antigens. Figure 2 shows flow cytometry analysis of protein display and antigen recognition by p1-1 and p13B1 in scFv and scTGP formats. Table I shows mean fluorescent intensity (MFI) values for dot blots and includes comparison with control antigens. The display levels for the scFv or scTGP were measured using labeled anti-SV5 IgG (Hanke *et al*., 1992), an antibody recognizing a peptide tag (SV5) at the C terminus. Data presented in Supplementary Fig. S1, available at *PEDS* online, shows that 2 days of induction is more suitable for scTGP display. The display levels for scTGP formats were reduced by 50–80% compared to the smaller scFv protein format. As expected, the antigen recognition signal was correspondingly reduced. The specific peptide recognition for scTGP p1-1 was about 5% of the recognition signal of the scFv. The pattern was similar for the p13B1 antibody as well and showed nearly 50% reduction in display level and associated drop in signal for antigen recognition. The affinity measurements for scFv and scTGP antibodies showed that despite the reduction in recognition signal the calculated K_d values were similar (within 4-fold) between the formats for a given antigen.

Comparing cytoplasmic and periplasmic scTGP expression systems

The scFvs are commonly expressed in the periplasmic space of the bacterial cell, since this provides the oxidizing environment required (Kipriyanov *et al*., 1997). However, higher levels of recombinant protein expression are usually observed in the *E. coli* cytoplasm. Therefore, we evaluated both compartments for scTGP protein expression using two different protein expression vectors. As shown in Fig. 3A and B, the cytoplasmic expression system (pETCK3) produced a statistically significant higher amounts of fluorescent protein in the soluble fraction after lysis compared to the periplasmic expression system. Furthermore, scTGP constructs expressed in the cytoplasm were more functional than those expressed in the periplasm, as assessed by the specific binding activity of the two formats (Fig. 3C). The recognition signal for specific antigen is more than three times higher than the signal for control antigen. Based on these results, further bacterial expression was carried out in the *E. coli* cytoplasm using the pETCK3 vector system.

Comparing expression levels and functionality of scFv and scTGP constructs

Three scFvs and corresponding scTGPs were used in protein purification and functionality studies. All six protein variants were expressed in 1 l culture volumes, purified using Ni-NTA agarose beads (see methods) in parallel and eluted using identical volumes. Eluates were run on a protein gel to compare relative protein production levels. The Coomassie staining of the protein gel showed scFv and scTGP protein yield was comparable in the case of the p1-1 and p13B1 antibodies. The scTGP format significantly improved soluble yield of the αTyrS antibody (Fig. 4A), consistent with relative total protein quantification (Table II). The relative amount of protein degradation appears to be similar between the scFv and its respective scTGP construct.

The functionality of both antibody formats (scFv and scTGP) was assessed using ELISA against specific and non-specific antigens (Fig. 4B) using equal concentrations of the corresponding scFvs and scTGPs. ITAM peptides p123 and p1 were used as specific antigens for scFv/scTGP p13B1 and p1-1, respectively. Fibrinogen was used as the specific antigen for the anti-tyrosine sulfate antibody (Kehoe *et al*., 2006). The p0 ITAM peptide and all non-cognate targets provide negative controls. The results showed that both the scFv and

Downloaded from https://academic.oup.com/peds/article/doi/10.1093/protein/gzaa029/6135052 by University of Liverpool user on 20 February 2021

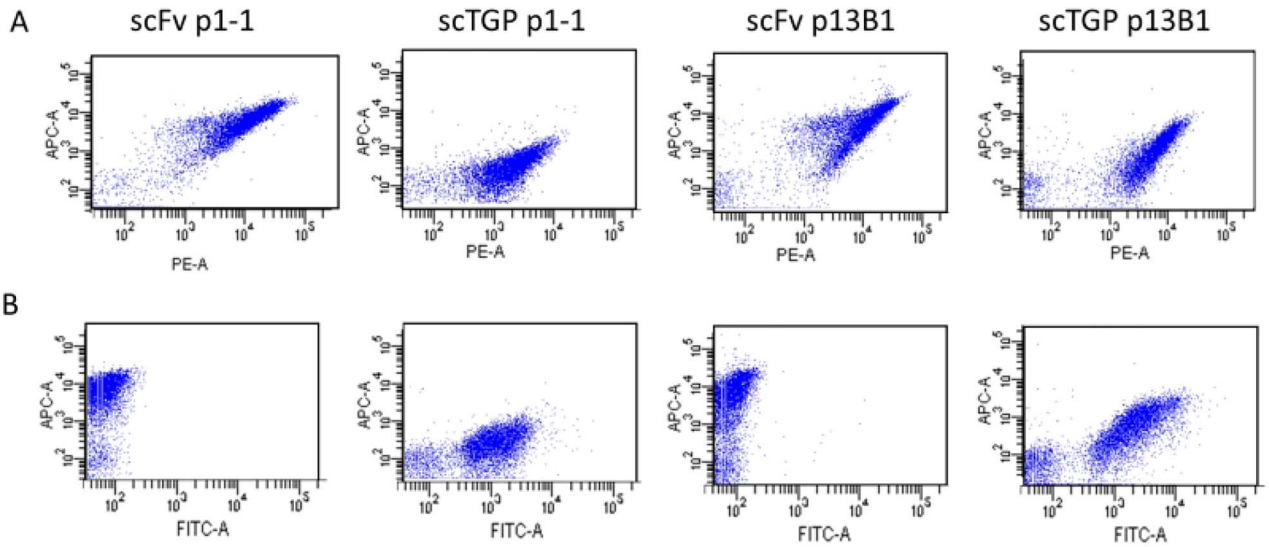


Fig. 2 Yeast display–based analysis of scFv and scTGP constructs for display level and recognition of specific antigen. The scFv and scTGP versions of two antibodies, p1-1 and p13B1, were cloned into the pDNL6 yeast display vector and protein display induced. Protein display levels were assayed using PE conjugated anti-SV5 antibody in panel **A**. In panel **B**, binding of scTGPs to their corresponding targets using the green fluorescence (FITC-A) of the TGP as a detection modality (instead of SV5) is shown in the *x*-axis. The *y*-axis shows the recognition of the biotinylated peptide bound to streptavidin Alexa 633. In the case of the displayed scFvs target binding is shown, but no display as measured by TGP fluorescence.

Table II. Comparison of scFv and scTGP yield from 1 l of bacterial culture following Ni-NTA purification

Antibody name	scFv (mg)	scTGP (mg)
p1-1	4.3	6.9
p13B1	5.5	6.9
αTyrS	0.7	6.0

scTGP format of the antibodies were functional and recognized the expected specific antigen with comparable recognition signals and the recognition signal for specific antigen is more than three times higher than signal for control antigen. The specificity of the scTGP antibodies were further investigated using a FLISA (Velappan *et al.*, 2008). The results showed that recognition of binding by scTGP can be assayed using fluorescence and all three antibodies showed specific recognition of their target antigen (Fig. 4C). Data presented in Supplementary Fig. S2, available at *PEDS* online, show that at equimolar concentration scFv and scTGP formats of p1-1 and p13B1 provide equivalent binding signals.

Crystallization and structure determination

Previous attempts to crystallize the p1-1 scFv were unsuccessful and we hypothesized that incorporation into the scFP format would promote crystallization. Large (>300 μm), green-colored rod-shaped crystals of scTGP p1-1 were observed in several conditions in the initial screening trials but did not diffract beyond 4 Å. Later, bi-pyramid–shaped crystals of scTGP p1-1 formed with 1.6 M citric acid pH 6.2 as crystallization buffer after 2 months. These crystals diffracted to better than 2.5 Å resolution and belonged to the $P4_12_12$ space group with cell dimensions of $a = b = 130.96$ Å, $c = 122.74$ Å. The complete data set has an R_{merge} of 5.9% for the resolution range of 50–2.5 Å. Cell content analysis gave a Matthews coefficient of 3.17 $Å^3Da^{-1}$ and a solvent content of 61% with one copy of scTGP p1-1 molecule in the asymmetric unit. The final R_{work} and R_{free} values were 16.3% and 21.1%, respectively. The refined structure contains 1 scTGP p1-1, 2 glycerols, and 136 water molecules. Detailed data collection and refinement statistics are listed in Table III. The atomic coordinates and structure factors are available in the RCSB Protein Data Bank under accession code 6WZN.

Figure 5 shows the crystal structure of scTGP p1-1, which revealed a single-chain protein with three distinct domains, the TGP domain (residues 119–335) and the scFv VL (residues 1–112) and VH (residues 346–465) domains. The linker residues 113–118 formed a short α-helix between the scFv VL domain and the N-terminal of TGP. Residues 336–345 linking the C-terminal of TGP to the scFv VH domain have relatively poor electron densities indicating partial disorder. Two glycine residues (338–339) in this linker were not modeled.

The TGP structure was used as the MR search model and as the initial structure of the TGP domain in scTGP p1-1 structural refinement. When superposed together, the Cα root-mean-square deviations (RMSD) between the refined TGP domain (this work) and the four individual protomers in the TGP structure range from 0.24 to 0.41 Å. As expected, several FP structures such as PDB IDs: 4PPL, 5EBJ, 2GX2 have RMSDs within this range when superposed with the TGP domain of scTGP p1-1. This is consistent with the well-conserved sequence of these proteins (>90% identity) in this FP family.

The structures of the VH and VL domains of this scTGP are also well conserved with structures in the Fab structure family such as PDB IDs: 1RZG and 3UX9. The superposed Cα RMSDs are about 0.5 Å between the VL domain of scTGP and similar light-chain structures and are 0.6–0.7 Å between heavy-chain structures. Cysteines 367 and 441 in the VH domain are in close proximity to each other. They did not form a disulfide bond as most of the corresponding cysteine pairs did in similar structures listed in Supplementary Table S1, available at *PEDS* online, except the VH domain of PDB ID: 4NIK (Robin *et al.*,

Downloaded from https://academic.oup.com/peds/article/doi/10.1093/protein/gzaa029/6135052 by University of Liverpool user on 20 February 2021

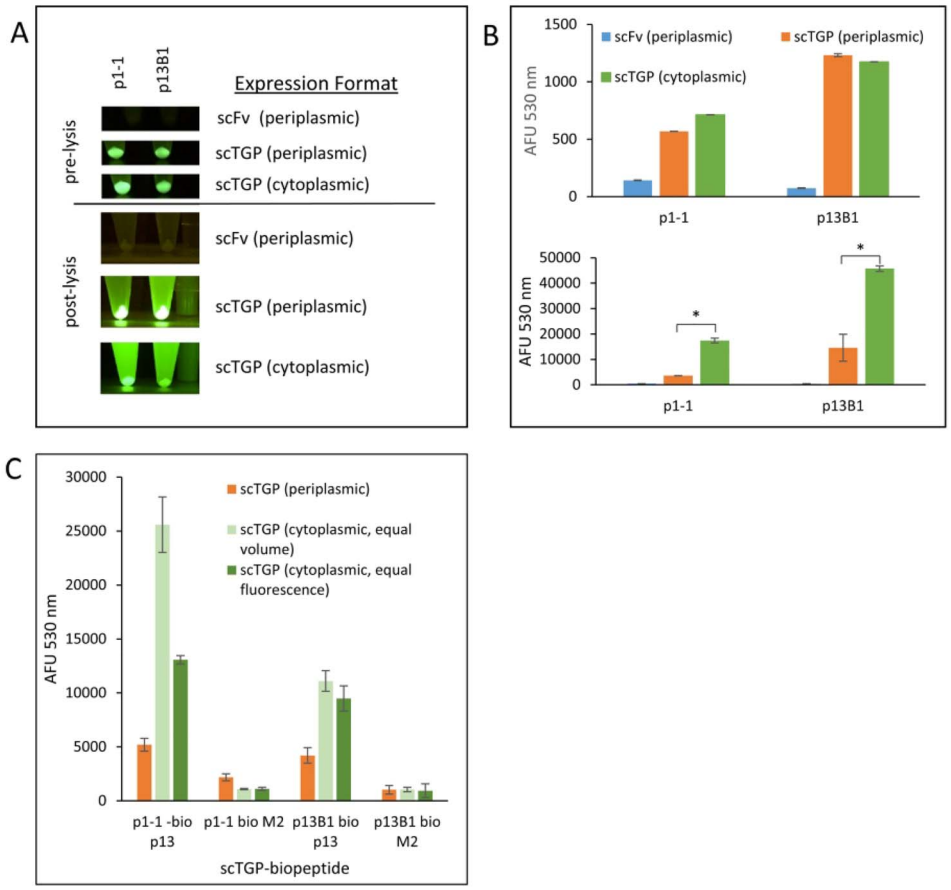


Fig. 3 Assessing protein expression systems for scTGPs. The scFv format of the antibodies is expressed in the periplasmic protein expression vector, pEP. The scTGP (scFP) format of the constructs was expressed in both pEP (periplasmic expression) and pETCK3 (cytoplasmic expression). Panel **A** shows the fluorescence obtained after protein expression in both vector systems using the corresponding scFv constructs as controls. The upper panel **B** shows fluorescence from 100 µl of cell culture after two nights of protein expression (pre-lysis). The lower panel **B** shows fluorescence for equal volume of POP culture supernatant (post-lysis). The error bars were generated from three independent protein expressions. Panel **C** shows the functionality of scTGPs produced in the periplasmic and cytoplasmic compartments. Two scTGPs (p1-1 and p13B1) were analyzed by FLISA for binding to their cognate antigen (ITAM p13 peptide) using M2 peptide of influenza A as negative control. The activity of the scTGPs produced in the cytoplasm (green) was normalized to the activity of scTGP produced in the periplasm (orange) in two different ways: equal volume (light green) or equal fluorescence (dark green). $*$ denotes a p value $\leq$ to 0.05. The error bars represent standard deviation generated from three independent binding assays.

2014). This could be attributed to, in part, the reducing environment provided in protein purification processes.

Epitope binning using scTGPs

We show that scTGPs could be successfully used for epitope binning of antibodies recognizing the recombinant chimeric *Y. pestis* F1V protein (causative agent of plague). Here, the effect of unlabeled IgGs on scTGPs binding was used to determine whether the two antibody formats bind to same or different epitopes, in a one-step assay (Fig. 6A). An scTGP-F1V binding curve (Fig. 6B) was used to estimate the half-saturating concentration of scTGP, optimal for detection of scTGP displacement. Figure 6C and D each show the results of competitive epitope binning assay for anti-F1V antibodies YP2, YP3 and YP8. Binding of scTGP was detected by fluorescence in the presence of non-fluorescent IgGs. Results suggest that antibody YP2 has a different epitope than YP3 and YP8 (Fig. 6C), since no displacement of scTGP YP2 (i.e. lowering of fluorescence) is observed in the presence of IgG YP3 and IgG YP8. As expected, IgG YP2 displaced the scTGP YP2. The data also show that YP3 and YP8 antibodies compete for the same epitope, as displacement of scTGP YP8 was observed in the presence of IgG YP8 and IgG YP3 (Fig. 6D). These results were confirmed by ELISA and flow cytometry experiments (Lillo *et al.*, 2020).

Discussion

A major problem with the use of antibodies in research and diagnostics has been their inconsistency, particularly lot-to-lot variation in polyclonals (Bradbury and Pluckthun, 2015; Schumacher and Seitz, 2016; Slaastad *et al.*, 2011; Weller, 2016). However, monoclonals have also been found to be problematic (Vaezi *et al.*, 2014), in some cases due to the expression of additional chains in hybridomas (Bradbury *et al.*, 2018), reflecting the uncharacterized nature of most commercial antibodies, including monoclonals (Bradbury *et al.*, 2018). In contrast to polyclonal antibodies from animal immunization campaigns, recombinant antibodies are expressed from defined plasmids where the sequences of the VH and VL genes are known. Recombinant antibodies can be obtained from hybridomas by cloning the VH and VL genes (Bradbury *et al.*, 2018; Ruberti *et al.*, 1994),

Downloaded from https://academic.oup.com/peds/article/doi/10.1093/protein/gzaa029/6135052 by University of Liverpool user on 20 February 2021

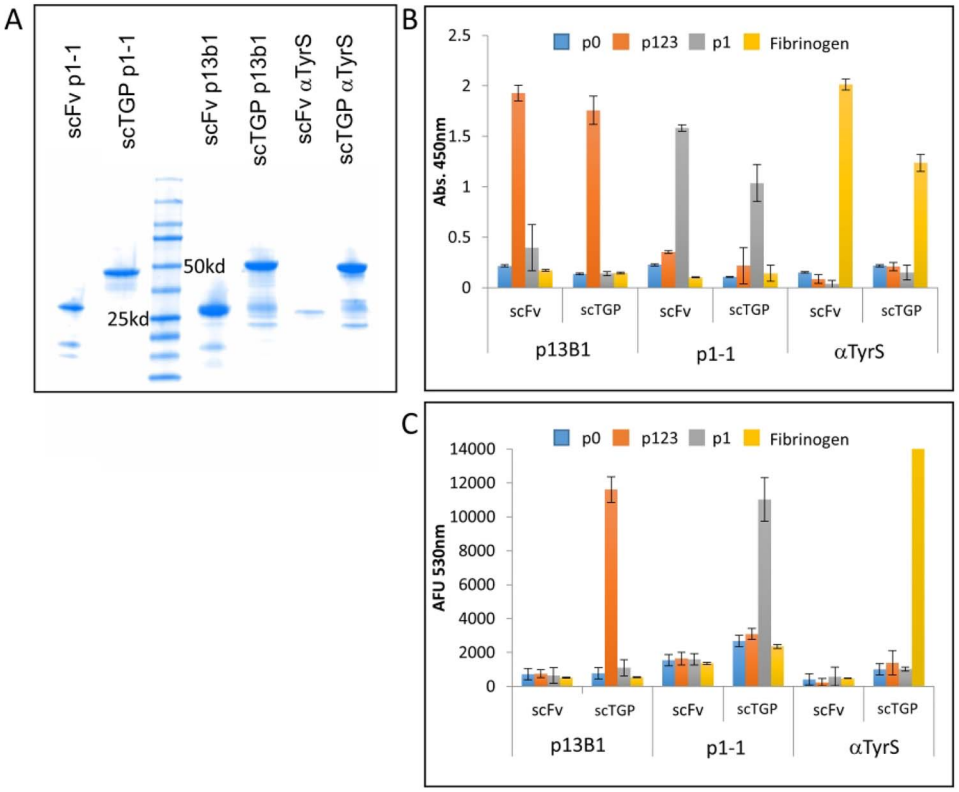


Fig. 4 Protein purification and functionality assays. Panel **A** shows scFv and scTGP proteins purified using Nickel NTA 6× HIS tag from a 1-l culture volume. The scFvs run slightly above 25 kDa, while scTGP run at ~50 kDa. Panel **B** shows a comparison of the functionality of scFv and scTGP by ELISA. Panel **C** shows that only scTGPs are functional in FLISA assays. Data for three different antibodies (p13B1, p1-1, αTyrS) with binding profiles against four different antigens (one specific, and three non-specific—ITAM peptides p0, p123, p1 and fibrinogen) are shown. The ELISA/FLISA assays were performed in triplicate and error bars refer to standard deviation.

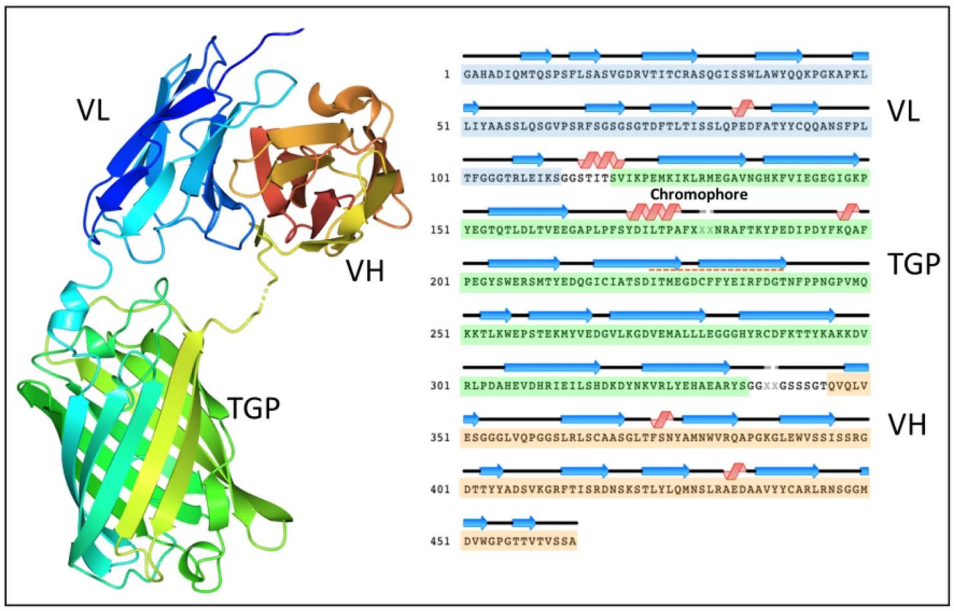


Fig. 5 Crystal structure of scTGP p1-1 (left). The sequence and secondary structure assignments of the structure (right). Residues 338–339 in the linker between TGP and VH are disordered.

as well by selecting from large combinatorial antibody libraries displayed on phage and/or yeast (Bradbury *et al.*, 2011; Marks *et al.*, 1991). Finally, scFv can be converted into full IgGs identical to animal-derived antibodies. A significant advantage to using display technologies to select antibodies is that the genes encoding the VH and VL genes are obtained concurrently with selection of the

Downloaded from https://academic.oup.com/peds/article/doi/10.1093/protein/gzaa029/6135052 by University of Liverpool user on 20 February 2021

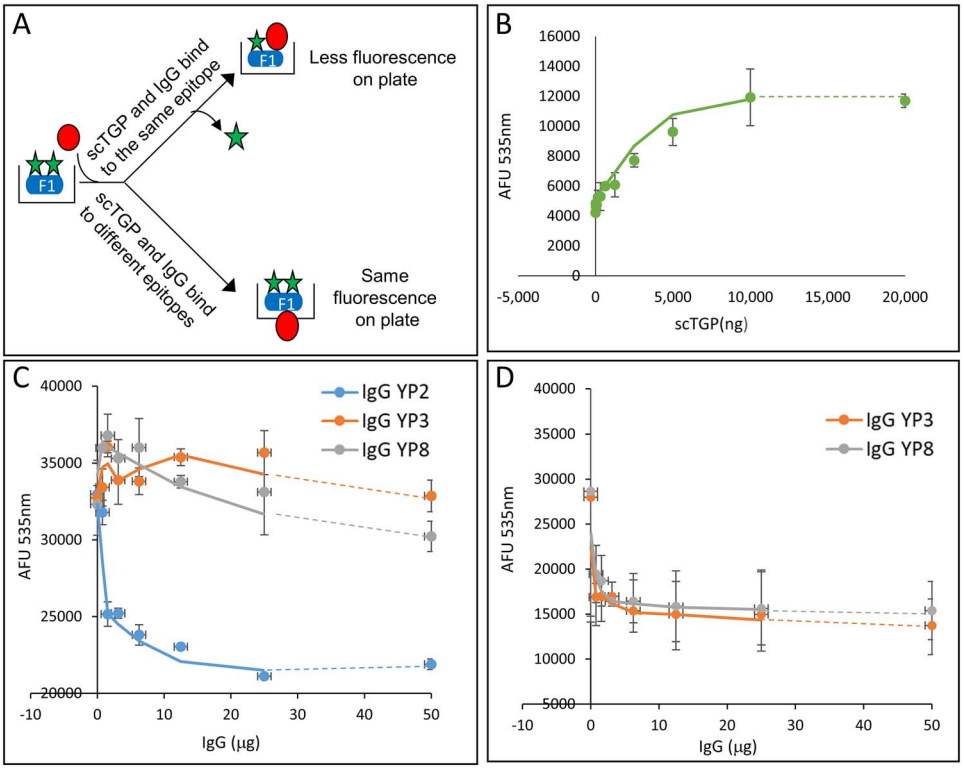


Fig. 6 Epitope mapping using scTGPs. The rationale behind epitope binning using scTGPs and IgGs is depicted in panel **A**. Here, the scTGP (green star) is allowed to bind to a dimeric protein. Upon addition of unlabeled IgG (red oval) scTGP is either displaced (lowering the amount of fluorescence) or remains the same (unchanged fluorescence). These two scenarios indicate IgG competition for the same epitope or IgG recognition of a different epitope, respectively. Panel **B** shows an example of a scTGP-F1V binding curve that was used to determine suitable non-saturating amount of scTGP needed for the epitope binning assay (in this case 5000 ng). Panel C and D show scTGP binding (fluorescence) in the presence of increasing concentrations of IgG (0–50 µg). The dots represent average fluorescence values with corresponding standard deviations (error bars). Panel **C** shows the displacement of the scTGP 2 by IgGs YP2 and lack of displacement by YP3 and YP8. Panel **D** shows the displacement of scTGP 8 by IgGs YP3 and YP8. For panels B, C, D two per moving average trend line is given as solid line and the line is extended to last value using dotted line.

antibody specificity, making the generation of potentially useful derivatives relatively straightforward. Apart from reformatting as full-length IgGs (Persic *et al.*, 1997), VH and VL genes can be transformed into dimerizing (De Kruif and Logtenberg, 1996) or multimerizing (Dübel *et al.*, 1995; Hudson and Kortt, 1999) antibody fragments, as well as novel functional entities generated by fusion to enzymes, tags (Cloutier *et al.*, 2000) or fluorescent proteins (Casey *et al.*, 2000). While antibody fragments selected from display libraries have been reformatted as full-length antibodies for catalog sale (Kehoe *et al.*, 2006), the potential utility of other formats as commercial therapeutics has been limited.

One particularly useful format enabled by the work described here is recombinant intrinsically fluorescent antibodies, which have the potential to be widely used in flow cytometry and immunofluorescence. These could replace chemically labeled fluorescent antibodies, which suffer from a number of problems including variability in labeling efficiency, inactivation and quenching (Haugland, 1995; Vira *et al.*, 2010). In this study, we build on previous efforts (Markiv *et al.*, 2011a,b) aimed at ameliorating these problems by showing that: (i) the VL-FP-VH orientation can also be used (in addition to the original VH-FP-VL orientation); (ii) additional scFvs can be effectively modified in this way; (iii) suitable FPs are not limited to those described by Markiv *et al.* but also include TGP, an extremely stable monomeric fluorescent protein (Close *et al.*, 2015); (iv) scFPs can be effectively displayed on yeast, where they function similarly to their corresponding scFv counterparts, opening the possibility to selecting variants with improved properties either directly from naïve, or preselected, libraries; and (v) functional expression levels of scFPs in the bacterial cytoplasm are relatively high and can rescue some scFvs otherwise expressed at low levels in the periplasm. This was particularly true of the anti-TyrS scFv, which has been challenging to use in the past due to poor expression levels and a propensity to aggregate. The expression level of the scFP version was dramatically higher than the scFv version, increasing by nearly 10-fold (6 vs 0.7 mg/ml). While we have previously demonstrated the value of TGP as a fluorescent protein (Close *et al.*, 2015), we have not compared the properties of scFPs made using TGP with others such as sfGFP or those used by Markiv *et al.* (2011a,b) (monomeric red fluorescent protein; blue, cerulean and citrine fluorescent proteins). The ability to crystallize one scTGP construct, as well as express the others within the normal *E. coli* cytoplasm (i.e. without the need for specialized strains with thioredoxin reductase and glutathione reductase mutations), suggests increased stability of the scTGP construct. However, further experiments to confirm this increased stability will be required.

TGP is a monomeric fluorescent protein amenable to crystallization. These features provide unique and valuable characteristics to scTGP molecules for their use in protein chemistry. In addition to utility in downstream assays such as immunofluorescence and flow cytometry, the intrinsic fluorescence of the molecules allows straightforward assessment of expression levels, monitoring of protein

Downloaded from https://academic.oup.com/peds/article/doi/10.1093/protein/gzaa029/6135052 by University of Liverpool user on 20 February 2021

Table III. Data collection and refinement statistics for scTGP p1-1

Data collection statistics	
Wavelength (Å)	1.0
Resolution range (Å)	50–2.50 (2.50–2.54)
Space group	$P4_12_12$
Unit-cell parameters (Å)	$a = b = 103.96$, $c = 122.74$
Total reflections	3 33 231
Unique reflections	24 030 (1177)
Completeness (%)	100.0 (100.0)
Multiplicity	13.9 (13.7)
R_{merge}[a] (%)	5.9 (100.0)
$R_{p.i.m}$[b]	0.019 (0.467)
$R_{meas}/R_{r.i.m}$[b]	0.062 (1.000)
Wilson B factor	34.51
Average $I/\sigma(I)$	40.3 (1.7)
Refinement statistics	
Resolution (Å)	50.00–2.50 (2.59–2.50)
Completeness (%)	91.85 (75.20)
Working reflections	21 973 (1771)
Test reflections	1839 (150)
R_{work}[c] (%)	0.1630 (0.2343)
R_{free}[c] (%)	0.2111 (0.2711)
No. of waters	136
No. of other solvent molecules	2 GOL
R.M.S. deviations from ideal geometry	
Bonds (Å)	0.006
Angles (°)	0.84
Average B factors (Å^2)	43.74
Ramachandran analysis by MOLPROBITY	
Residues in favored regions (%)	98.90
Residues in allowed regions (%)	1.10

Values in parentheses represent the outermost resolution shell.

[a] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - <I(hkl)>|/\Sigma_{hkl}\Sigma_i I_i(hkl)$ where $I_i(hkl)$ is the intensity of the ith observation and $<I(hkl)>$ is the mean intensity of the reflections. The values are for unmerged Friedel pairs.

[b] $R_{p.i.m.}$ (precision-indicating R_{merge}) and $R_{meas}/R_{r.i.m.}$ (redundancy-independent R_{merge}).

[c] The crystallographic R factor $R = \Sigma_{hkl}|\,|F_{obs}| - |F_{cal}|\,|/\Sigma_{hkl}\,|F_{obs}|$; $R_{free} = \Sigma_{hkl}\,|\,|F_{obs}| - |F_{cal}|\,|/\Sigma_{hkl}\,|F_{obs}|$ where all reflections belong to a test set of randomly selected data.

purification steps, and antigen recognition. The one step FLISA utilized here saves time and improves the utility of antibodies produced in this format. We also describe the use of the scTGP proteins for epitope binning, a useful and easy technique if the antigen is multimeric that capitalizes on the monomeric and fluorescent properties of the scTGP constructs to assess and develop non-competitive sets of antigen detection proteins. IgGs in general have higher affinity for the antigen than the corresponding scFvs, since IgGs have two copies of VH-VL and scFPs have one. Therefore, eventual competition between the two antibody formats is strong and easy to detect. Our ongoing work includes development of an LFA-based detection system for plague using antibodies YP2 and YP8, identified as orthogonal binder pairs using the scTGP displacement assay described here (Lillo *et al.*, 2020).

A critical result obtained in this study was the observation that the affinity values for the scFv and scTGP formats of the antibodies were very similar, providing an additional indication that the addition of fluorescent protein added utility without decreasing antibody functionality. However, the display levels of scTGP on the yeast surface were significantly lower compared to the display of the scFv and required additional induction time (~48 vs 12–16 h for scFvs) to increase the fluorescence of yeast displaying scTGPs by ~20%, as shown in Supplementary Fig. S1, available at *PEDS* online. This extended time likely reflects the need for additional time to fold the larger scFP molecules, or mature the fluorophore, and may vary depending upon the scFv sequence. Therefore, the development of high-throughput methods to select scFvs suitable for conversion to scTGP would be highly advantageous. Interactions between the folding of the FP and the VH, VL domains are difficult to predict. For example, yields may be improved by other robustly folding fluorescent proteins such as superfolder GFP (Pédelacq *et al.*, 2006), which was specifically selected to avoid interfering with the folding of fused domains. The availability of several, alternative, robustly folded fluorescent proteins to serve as linkers should extend the applicability of the approach described (Velappan *et al.*, 2020).

Attempts to obtain crystal structures of scTGP complexed with 3 versions of the 17mer antigen peptides were unsuccessful. The 17mer peptide sequence was LYEELHVYSPIYSALED and peptides containing the first tyrosine phosphorylated (p1-17), the second tyrosine phosphorylated (p2-17), and both first and second tyrosines phosphorylated p12-17 were selected for crystallization experiments based on ELISA, FLISA and flow cytometry studies with peptides of varying lengths. Over 800 screening conditions were attempted to obtain co-crystals of scTGP and these peptides but resulted in no crystal hits. Soaking apo-scTGP crystals in peptide-containing mother liquors shattered and dissolved crystals almost instantaneously, and crystal pieces harvested after flash-soaking in these solutions did not diffract. These observations suggest that strong interactions exist between scTGP and its antigen peptides. Such interactions are likely to induce substantial conformational changes between the scFv domains causing deformation of the crystal lattice of the apo-scTGP crystal. The bound peptides could also interfere with crystal formation in the case of co-crystallization trials. Systematic peptide construct screening may be required to obtain the co-crystals of the scTGP and its antigen peptides.

One of the major problems faced by scFvs identified using display technologies is the quality, quantity and stability of purified scFv protein available for downstream characterization. *E. coli* periplasmic expression produces limited quantities of scFv, which can lead to poorly purified, aggregated and unstable protein (Arbabi-Ghahroudi *et al.*, 2005; Arndt *et al.*, 1998), which is why scFvs are not widely used as research reagents. N- and C-terminal FPs require additional flexible linkers that are prone to proteolytic cleavage resulting in potential decoupling of binding from fluorescence. scFv constructs use a flexible linker to connect the VL and VH regions of the scFv, while in naturally occurring IgGs, the VH/VL interaction is stabilized by CH1 and CL. It is possible that the use of a fluorescent protein as connector between the VL and VH domains of a scFv may promote greater stability by mimicking the Fab architecture of natural IgGs, while also providing the added biotechnological functionality of intrinsic fluorescence. The additional stability provided in this orientation relative to N- and C-terminal constructs may promote crystallization through minimization of interdomain flexibility. The crystal structure reported here provides some evidence to support this concept, as the scFv version of the scFP was previously refractory to crystallization, and within the structure, the TGP is found in close proximity to the lower scFv interface. In conclusion, we report that the scTGP format holds promise for the development of robust and readily expressible antibody fragments with a wide variety of uses in research and clinical settings.

Downloaded from https://academic.oup.com/peds/article/doi/10.1093/protein/gzaa029/6135052 by University of Liverpool user on 20 February 2021

Supplementary data

Supplementary data are available at *PEDS* online.

Acknowledgments

Authors would like to acknowledge Dr Thomas Terwilliger for his guidance. We would like to thank the staff at the beam line 5.0.2 managed by the Berkeley Center for Structural Biology (BCSB) at the Advanced Light Source (ALS) for technical support. The Berkeley Center for Structural Biology is supported in part by the Howard Hughes Medical Institute. The ALS is a Department of Energy Office of Science User Facility under Contract No. DE-AC02-05CH11231.

Funding

This work was supported by National Institutes of Health grant (P50GM085273); Foundation for the National Institutes of Health (P50GM085273); Los Alamos National Laboratory's Laboratory Directed Research & Development grant (20180005DR, 20160054DR). Research was also supported by the DOE Office of Science through the National Virtual Biotechnology Laboratory, a consortium of DOE national laboratories focused on response to COVID-19, with funding provided by the Coronavirus CARES Act (#KP160101).

Conflict of Interest

No potential conflicts of interest were disclosed.

Author Contributions

Concept was developed by D.C. D.C., N.V., L.H., C.H., A.M.L., G.S.W., A.B. contributed toward manuscript preparation and revision. D.C., N.V., L.H., L.N., C.H., N.D., D.K.M., A.M.L. contributed toward experimental results and the work was performed under the direction of A.B.

References

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H. and Adams, P.D. (2012) *Acta Crystallogr. D Biol. Crystallogr.*, **68**, 352–367.

Arbabi-Ghahroudi, M., Tanha, J. and MacKenzie, R. (2005) *Cancer Metastasis Rev.*, **24**, 501–519.

Arndt, K.M., Muller, K.M. and Pluckthun, A. (1998) *Biochemistry*, **37**, 12918–12926.

Boder, E.T. and Wittrup, K.D. (1997) *Nat. Biotechnol.*, **15**, 553–557.

Bradbury, A.R.M. and Pluckthun, A. (2015) *Protein Eng. Des. Sel.*, **28**, 303–305.

Bradbury, A.R.M., Sidhu, S., Dübel, S. and McCafferty, J. (2011) *Nat. Biotechnol.*, **29**, 245–254.

Bradbury, A.R.M., Trinklein, N.D. *et al.* (2018) *MAbs*, **10**, 1–19.

Casey, J.L., Coley, A.M., Tilley, L.M. and Foley, M. (2000) *Protein Eng.*, **13**, 445–452.

Close, D.W., Paul, C.D., Langan, P.S. *et al.* (2015) *Proteins*, **83**, 1225–1237.

Cloutier, S.M., Couty, S., Terskikh, A., Marguerat, L., Crivelli, V., Pugnières, M., Mani, J.C., Leisinger, H.J., Mach, J.P. and Deperthes, D. (2000) *Mol. Immunol.*, **37**, 1067–1077.

de Kruif, J. and Logtenberg, T. (1996) *J. Biol. Chem.*, **271**, 7630–7634.

Dübel, S., Breitling, F., Kontermann, R., Schmidt, T., Skerra, A. and Little, M. (1995) *J. Immunol. Methods*, **178**, 201–209.

Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.

Ferrara, F., Listwan, P., Waldo, G.S. and Bradbury, A.R.M. (2011) *PLoS One*, **6**, e25727.

Ferrara, F., Naranjo, L.A., Kumar, S., Gaiotto, T., Mukundan, H., Swanson, B. and Bradbury, A.R.M. (2012) *PLoS One*, **7**, e49535.

Hanke, T., Szawlowski, P. and Randall, R.E. (1992) *J. Gen. Virol.*, **73**, 653–660.

Haugland, R.P. (1995) *Monoclonal Antibody Protocols*. Humana Press, Totowa, NJ, pp. 205–222.

Hudson, P.J. and Kortt, A.A. (1999) *J. Immunol. Methods*, **231**, 177–189.

Huston, J.S., Levinson, D., Mudgett-Hunter, M., Tai, M.S., Novotny, J., Margolies, M.N., Ridge, R.J., Bruccoleri, R.E., Haber, E. and Crea, R. (1988) *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 5879–5883.

Kehoe, J.W., Velappan, N., Walbolt, M. *et al.* (2006) *Mol. Cell. Proteomics*, **5**, 2350–2363.

Kipriyanov, S.M., Moldenhauer, G. and Little, M. (1997) *J. Immunol. Methods*, **200**, 69–77.

Liebschner, D., Afonine, P.V., Baker, M.L. *et al.* (2019) *Acta Crystallogr. D Struct. Biol.*, **75**, 861–877.

Lillo, A.M., Ayriss, J.E., Shou, Y., Graves, S.W., Bradbury, A.R. and Pavlik, P. (2011) *PLoS One*, **6**, e27756.

Lillo, A.M., Velappan, N., Kelliher, J.M. *et al.* (2020) Development of anti-yersinia pestis human antibodies with features required for diagnostic and therapeutic applications. *Immunotargets Ther.*, **9**, 299–316.

Lindner, P., Bauer, K., Krebber, A., Nieba, L., Kremmer, E., Krebber, C., Honegger, A., Klinger, B., Mocikat, R. and Plückthun, A. (1997) *Biotechniques*, **22**, 140–149.

Markiv, A., Anani, B., Durvasula, R.V. and Kang, A.S. (2011a) *J. Immunol. Methods*, **364**, 40–49.

Markiv, A., Beatson, R., Burchell, J., Durvasula, R.V. and Kang, A.S. (2011b) *BMC Biotechnol.*, **11**, 117.

Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., McCafferty, J., Griffiths, A.D. and Winter, G. (1991) *J. Mol. Biol.*, **222**, 581–597.

McCormack, T., O'Keeffe, G., Mac Craith, B. and O'Kennedy, R. (1996) *Anal. Lett.*, **29**, 953–968.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) *J. Appl. Cryst.*, **40**, 658–674.

Nizak, C., Monier, S., del Nery, E., Moutel, S., Goud, B. and Perez, F. (2003) *Science*, **300**, 984–987.

Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T.C. and Waldo, G.S. (2006) *Nat. Biotechnol.*, **24**, 79–88.

Persic, L., Roberts, A., Wilton, J., Cattaneo, A., Bradbury, A. and Hoogenboom, H.R. (1997) *Gene*, **187**, 9–18.

Robin, G., Sato, Y., Desplancq, D., Rochel, N., Weiss, E. and Martineau, P. (2014) *J. Mol. Biol.*, **426**, 3729–3743.

Rothbauer, U., Zolghadr, K., Muyldermans, S., Schepers, A., Cardoso, M.C. and Leonhardt, H. (2008) *Mol. Cell. Proteomics*, **7**, 282–289.

Ruberti, F., Cattaneo, A. and Bradbury, A. (1994) *J. Immunol. Methods*, **173**, 33–39.

Sblattero, D. and Bradbury, A. (2000) *Nat. Biotechnol.*, **18**, 75–80.

Schumacher, S. and Seitz, H. (2016) Quality control of antibodies for assay development. *N. Biotechnol.*, **33**, 544–550. ISSN 1871-6784.

Shu, L., Qi, C.F., Schlom, J. and Kashmiri, S.V. (1993) *Proc. Natl. Acad. Sci. U.S.A.*, **90**, 7995–7999.

Slaastad, H., Wu, W., Goullart, L., Kanderova, V., Tjønnfjord, G., Stuchly, J., Kalina, T., Holm, A. and Lund-Johansen, F. (2011) *Proteomics*, **11**, 4578–4582.

Studier, F.W. (2005) *Protein Expr. Purif.*, **41**, 207–234.

Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H., Hung, L.W., Read, R.J. and Adams, P.D. (2008) *Acta Crystallogr. D Biol. Crystallogr.*, **64**, 61–69.

Vaezi, A.E., Bepler, G., Bhagwat, N.R. *et al.* (2014) *Cancer*, **120**, 1898–1907.

Velappan, N., Clements, J., Kiss, C., Valero-Aracama, R., Pavlik, P. and Bradbury, A.R.M. (2008) *J. Immunol. Methods*, **336**, 135–141.

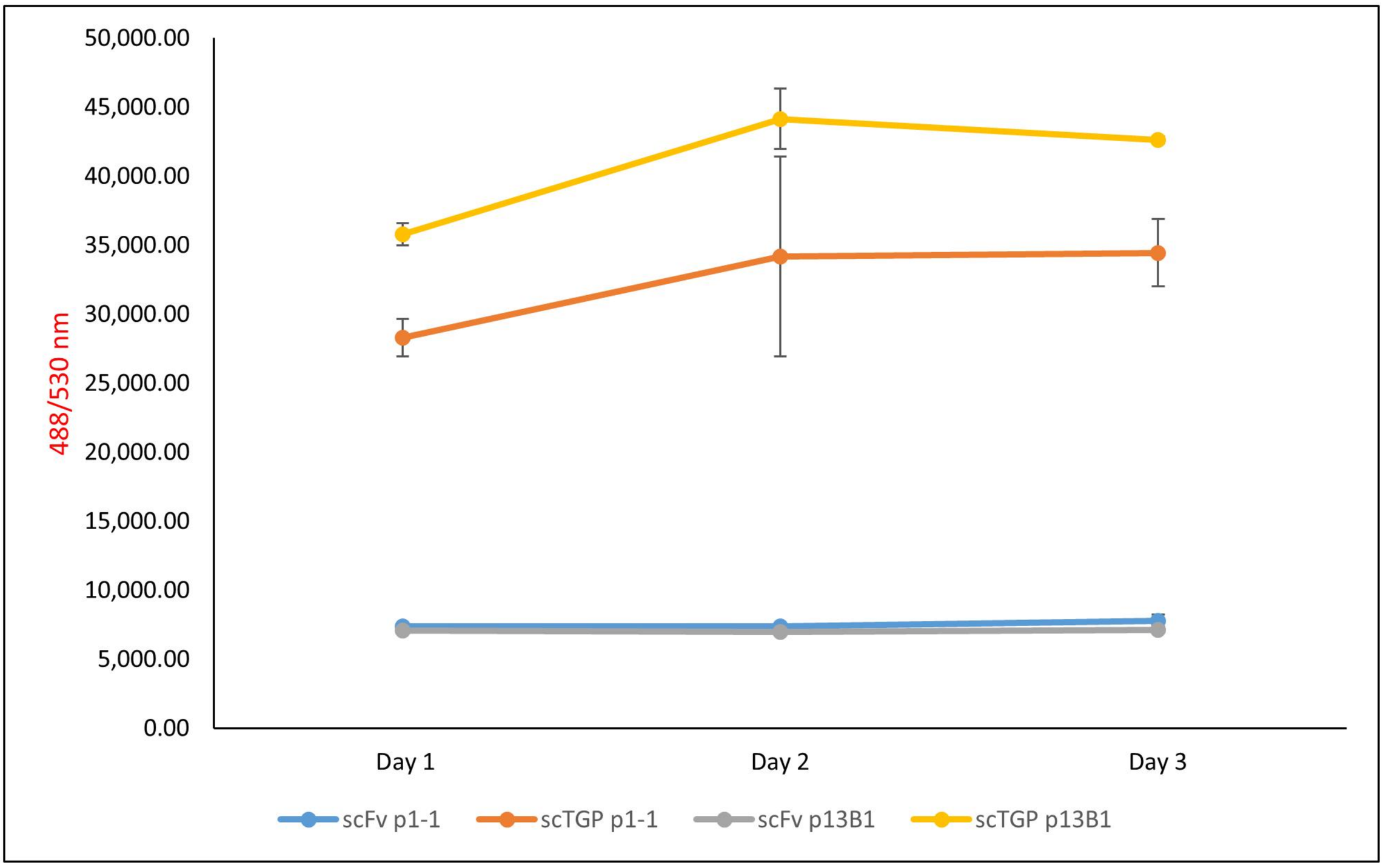
Velappan, N., Mahajan, A., Naranjo, L. *et al.* (2019) Selection and characterization of FceRI phospho-ITAM specific antibodies. *MAbs*, **11**, 1206–1218.

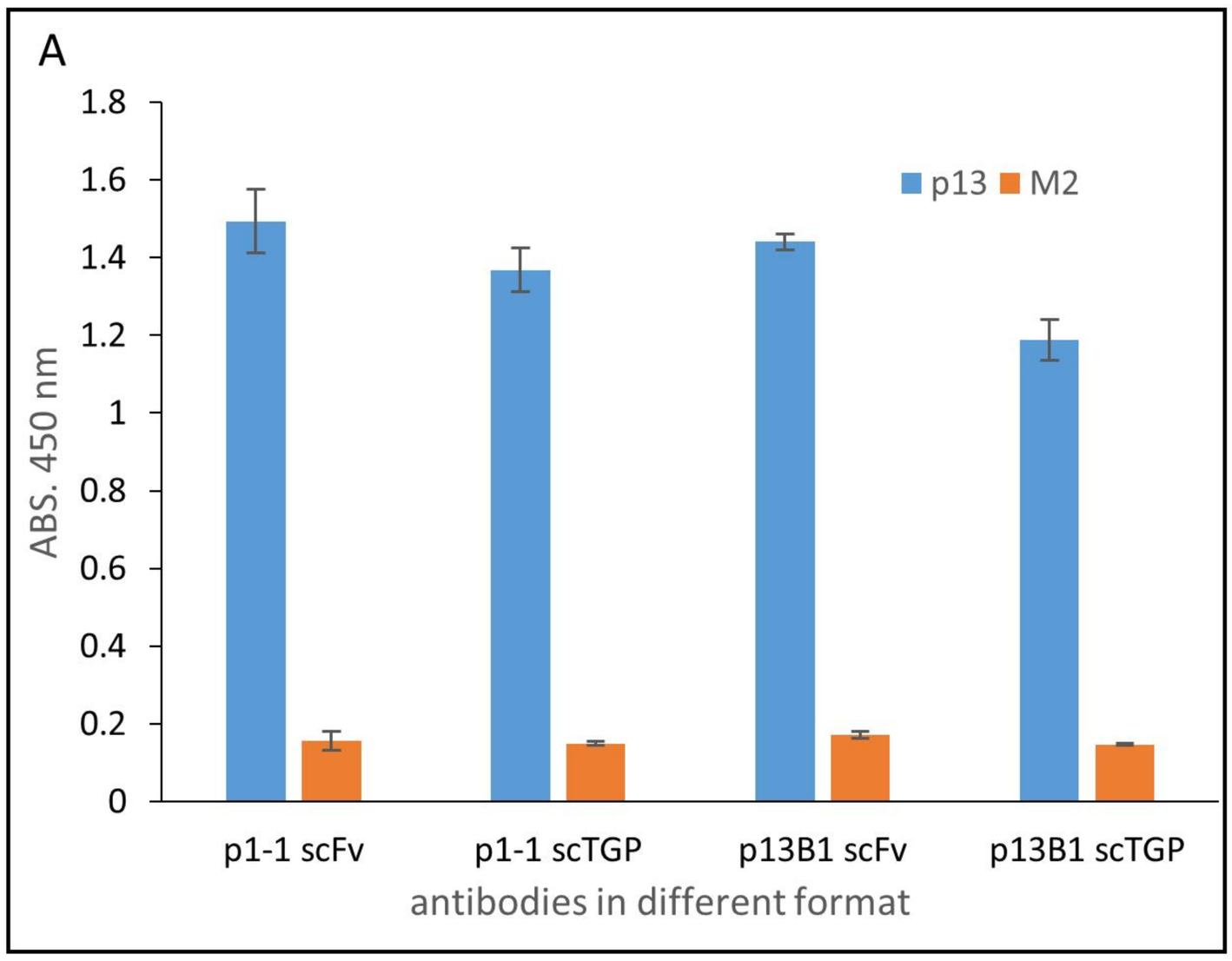
Velappan, N., Micheva-Viteva, S., Adikari S.H., Waldo G.S., Lillo A.M., Bradbury A.R.M. (2020) Selection and verification of antibodies against the cytoplasmic domain of M2 of influenza, a transmembrane protein. *MAbs*, **12**, 1843754.

Vira, S., Mekhedov, E., Humphrey, G. and Blank, P.S. (2010) *Anal. Biochem.*, **402**, 146–150.

Weller, M.G. (2016) *Anal. Chem. Insights*, **11**, 21–27.

Downloaded from https://academic.oup.com/peds/article/doi/10.1093/protein/gzaa029/6135052 by University of Liverpool user on 20 February 2021


50,000.00
45,000.00
40,000.00
35,000.00
30,000.00
25,000.00
20,000.00
15,000.00
10,000.00
5,000.00
0.00
488/530 nm
Day 1
Day 2
Day 3
scFv p1-1
scTGP p1-1
scFv p13B1
scTGP p13B1


A
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
ABS. 450 nm
p13
M2
p1-1 scFv
p1-1 scTGP
p13B1 scFv
p13B1 scTGP
antibodies in different format

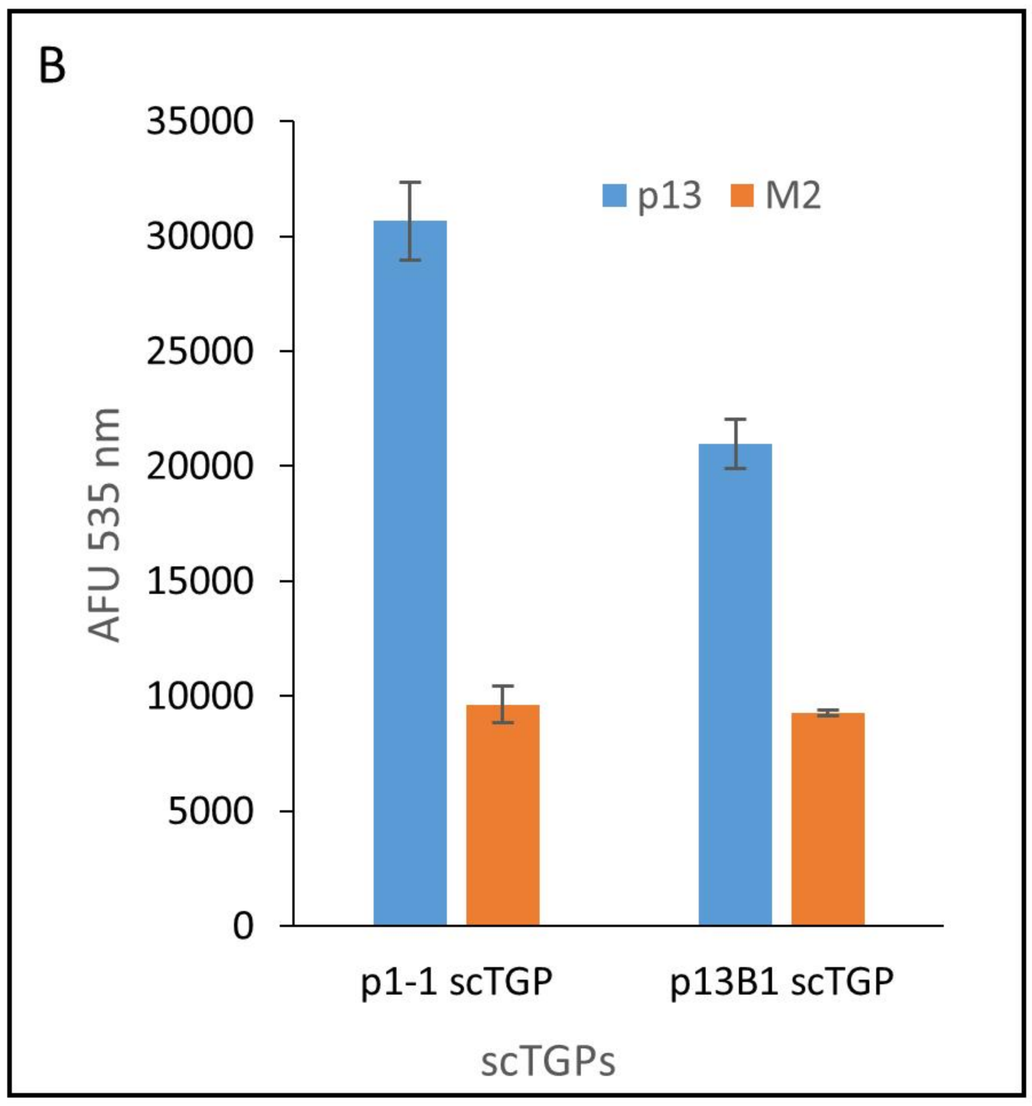

B
35000
30000
25000
20000
15000
10000
5000
0
AFU 535 nm
p13
M2
p1-1 scTGP
p13B1 scTGP
scTGPs

Table S1. Structure used in comparison

A. TGP domain

PDB ID	Chain	RMSD	Aligned Residues
4TZA	C	0.244	212
4PPL	A	0.278	213
4TZG	A	0.281	213
5EBJ	A	0.286	212
4PPK	A	0.294	213
2GX2	A	0.334	212
4IZN	A	0.335	213
4DXM	A	0.337	213
4HQ8	A	0.339	213
6D39	A	0.341	212
4HQC	B	0.342	213
5HZU	A	0.346	212
4HQC	D	0.348	213
4DXI	A	0.348	213
4HQC	C	0.349	213
4HQC	A	0.352	213
4DXP	A	0.352	213
6CIU	A	0.354	213
4DXN	A	0.355	212
4JC2	A	0.356	213
2Z1O	A	0.357	212
5EB6	A	0.357	212
4DXO	A	0.358	212
2IE2	A	0.361	212
2Z6X	A	0.382	213
2VZX	A	0.399	213
5EJU	A	0.404	211
2GX0	A	0.405	213
4TZA	A	0.411	213
4UTS	A	0.433	212
3ADF	A	0.436	210
4PPJ	A	0.438	213
5DTL	A	0.446	213
4R6B	A	0.452	213
4P76	A	0.454	213
6MXW	A	0.461	213
5EXB	A	0.485	213
1MOV	A	0.545	213
1MOU	A	0.552	213
2P4M	A	0.563	213

B. VL domain

PDB ID	Chain	RMSD	Aligned Residues
5GS3	H	0.328	107
5VYF	L	0.435	108
5GGV	L	0.438	108
1DEE	A	0.447	108
6CNR	L	0.453	108
4L1H	A	0.455	107
5CZX	D	0.472	108
1RZG	B	0.479	108
6BTJ	B	0.481	108
1HEZ	A	0.486	108
5I8C	B	0.507	108
5FV1	L	0.515	108
6EHY	A	0.517	107
4XAK	L	0.521	108
5OD8	B	0.523	108
4PY7	J	0.525	108
1T3F	A	0.536	108
6EHX	B	0.543	108
1TJG	L	0.543	108
4KMT	L	0.554	108
2CMR	L	0.570	108
5D1Z	A	0.585	108
4YK4	Y	0.601	107
1BJ1	J	0.601	107
5D1Q	A	0.603	108
6DKJ	B	0.605	108
5SX5	K	0.607	107
6APB	L	0.610	107
1B6D	A	0.616	107
1CZ8	L	0.622	107
1RZ7	L	0.624	106
6BFT	B	0.641	107
3JWD	L	0.646	107
3HMW	L	0.648	107
3JWO	L	0.650	107
1U8K	A	0.658	107
6I04	L	0.661	107
4M6O	L	0.670	107
4NRX	B	0.690	107
6OGX	D	0.703	108

C. VH domain

PDB ID	Chain	RMSD	Aligned Residues
4UU9	H	0.466	114
5GS3	A	0.468	114
4G5Z	H	0.506	117
6QF9	A	0.532	117
6CYF	D	0.543	117
3UX9	B	0.544	117
5BVJ	B	0.558	117
5I1A	B	0.583	118
6GHG	A	0.586	115
6FS0	H	0.59	118
1T2J	A	0.594	109
6J6Y	B	0.6	118
6QBC	H	0.605	118
5D70	H	0.613	115
5DR5	H	0.632	113
3SO3	C	0.633	118
4Y5Y	A	0.65	116
6NMS	B	0.653	118
5TSJ	1	0.66	115
6II8	J	0.681	109
3TNN	H	0.682	117
6E63	H	0.684	117
5MYK	B	0.688	117
4QCI	B	0.694	118
5CUS	H	0.695	117
3QOS	H	0.706	118
6PE7	H	0.706	115
4YBQ	D	0.707	117
5TZU	H	0.713	117
4P59	H	0.722	117
6JC2	H	0.723	118
5TRU	H	0.727	118
3KDM	H	0.728	117
2V7N	B	0.73	116
5HHV	H	0.746	117
6B0A	H	0.748	118
2VXS	H	0.755	117
5VEB	A	0.768	118
5BK5	B	0.769	110
5X8M	B	0.78	118