DOI 10.1002/biot.200900075 Biotechnol. J. 2009, 4, 1328–1336

Research Article

scFv-based fluorogen activating proteins and variable domain inhibitors as fluorescent biosensor platforms

Crystal N. Falco[1], Kaitlyn M. Dykstra[1], Bradley P. Yates[1] and Peter B. Berget[1, 2]

[1] Department of Biological Sciences, Carnegie Mellon University, Pittsburgh, PA, USA
[2] Molecular Biosensor and Imaging Center, Carnegie Mellon University, Pittsburgh, PA, USA

Received 26 March 2009
Revised 28 May 2009
Accepted 31 May 2009

Single chain antibodies (scFvs) are engineered proteins composed of IgG variable heavy (V_H) and variable light (V_L) domains tethered together by a flexible peptide linker. We have characterized the individual V_H or V_L domain activities of several scFvs isolated from a yeast surface-display library for their ability to bind environmentally sensitive fluorogenic dyes causing them to fluoresce. For many of the scFvs, both V_H and V_L domains are required for dye binding and fluorescence. The analysis of other scFvs, however, revealed that either the V_H or the V_L domain alone is sufficient to cause the fluorogenic dye activation. Furthermore, the inactive complementary domains in the original scFvs either contribute nothing to, or actually inhibit the activity of these active single domains. We have explored the interactions between active variable domains and inactive complementary domains by extensive variable domain swapping through *in vitro* gene manipulations to create hybrid scFvs. In this study, we demonstrate that significant alteration of the fluorogenic dye activation by the active V_H or V_L domains can occur by partnering with different V_H or V_L complementary domains in the scFv format. Hybrid scFvs can be generated that have fluorogen-activating domains that are completely inhibited by interactions with other domains. Such hybrid scFvs are excellent platforms for the development of several types of genetically encoded, fluorescence-generating biosensors.

Keywords: scFv · Variable domains · Fluorescence · Fluorogenic dye · Yeast surface display

1 Introduction

Single chain antibodies (scFvs) are synthetically constructed proteins that retain the antigen binding capacity of full-length antibodies but are much smaller. Typical IgG molecules have a mass of ~150 kDa, whereas scFvs are only about 25–27 kDa. scFv genes are constructed from gene segments of IgG variable domains of the heavy and light chains (V_H and V_L) connected by a synthetic DNA segment that usually codes for a flexible $(Gly_4Ser)_n$ polypeptide linker. This linker is normally of sufficient length ($n \geq 3$) to allow the translated V_H and V_L domains to associate intramolecularly into the characteristic V_H/V_L domain conformation found at the antigen-binding ends of antibodies. scFvs offer the benefit of being genetically easy to manipulate, and several vector and display library systems have been designed to take advantage of this feature to isolate scFvs that recognize a variety of antigens [1–3].

We previously reported the isolation and characterization of scFvs from a yeast surface-display library that bind to and rotationally constrain several environmentally sensitive organic dye molecules, causing them to fluoresce [4, 5]. We call this phenomenon fluorogenic activation. The fluoro-

Correspondence: Professor Peter B. Berget, Department of Biological Sciences, Molecular Biosensor and Imaging Center, Carnegie Mellon University, 4400 Fifth Avenue, Pittsburgh, PA 15213, USA
E-mail: berget@andrew.cmu.edu
Fax: +1-412-268-7129

Abbreviations: CDR, complementarity determining region; **DIR**, dimethylindole red; **FAP**, fluorogen-activating protein; **MG**, malachite green; **scFv**, single chain variable fragment antibody; **TO**, thiazole orange; **V_H**, heavy chain variable domain; **V_L**, light chain variable domain

genic dyes – thiazole orange (TO), malachite green (MG), and dimethylindole red (DIR) – and their cognate scFv-based fluorogen activating proteins (FAPs) are members of a new class of two-component fluoromodules that have been used to visualize cell-surface and secretory membrane proteins in mammalian cells [4]. In our studies, the DNA sequence of three of the six MG-binding scFv clones predicted the expression of only one domain, either V_H or V_L, in the final protein product. Furthermore, two of the remaining scFvs have V_L chains that are strikingly similar in predicted protein sequence. In addition, five clones that produce scFvs that activate DIR have DNA sequences that suggest that they are full-length scFvs. Interestingly, however, two of them have identical V_L domains and different V_H domains.

Functional V_H or V_L domains are of particular interest because they are half the size of scFvs (12–15 kDa) yet they bind their target antigens with high affinity and specificity. Furthermore, active domains (V_H or V_L) and their cognate fluorogenic dyes are fluoromodules that are half the size of typical fluorescent proteins. Because we have already demonstrated the existence of FAPs comprised of only of V_H or V_L domains that activate MG [4] and because of the structural similarities among domains of other scFvs in our collection, we decided to explore whether any V_H or V_L domains of these scFv FAPs are functional in isolation. In this paper we report the molecular dissection of scFvs into individually expressed V_H and V_L domains. We observed that some scFvs require both V_H and V_L domains for detectible activity, whereas others only need one domain. Molecular hybrid scFvs consisting of active and inactive V_H and V_L domains show remarkable modulation of the activity of the single active domains.

2 Materials and methods

2.1 FAP nomenclature

FAP names consist of three components: (i) the scFv chain configuration, with H designating the heavy variable (V_H) region and L designating the light variable (V_L) region; (ii) a unique numerical identifier designating the parent isolate and its affinity maturation lineage in the format '*parent#.1stgeneration#.2ndgeneration#*', etc. An entry of '*0*' describes the case where the '*parent*' is a population of different isolates whose DNA may be recombined; and (iii) the dye used to isolate the FAP [4]. Thus L5.1-MG indicates the first affinity-matured variant of the MG-activating, V_L-only clone 5.

2.2 Genetic construction of plasmids

2.2.1 Separation of variable domains

Single V domain reduction plasmids were constructed using restriction sites within the pPNL6 vector and the $(Gly_4Ser)_3$ linker. V_H-only plasmids were constructed by subcloning the V_H domain-coding restriction fragments into empty pPNL6 vector [3] as follows. After cleavage with the restriction enzymes *Nhe*I and *Bam*HI (see Fig. 1), V_H domain-coding fragments from two-domain scFvs were separated from the rest of the plasmid DNA by agarose gel electrophoresis and purified with a Qiagen Gel Extraction Kit (Qiagen, Valencia, CA, USA). Empty pPNL6 vector was cut using the same pair of enzymes to remove the *Nhe*I/*Bam*HI stuffer fragment, and the backbone was purified in a similar manner to prepare for ligation. V_H domains were then ligated into this pPNL6 vector backbone following the suggested protocol in the NEB Quick Ligation Kit (New England Biolabs, Ipswich, MA, USA). A special vector pPNL6(HL1-TO1 V_L) was constructed for cloning V_L domains. pPNL6 carrying the two-domain scFv gene HL1-TO1 was digested with *Bmt*I (an isoschizomer of *Nhe*I) and *Bam*HI. The DNA ends were treated with T4 DNA polymerase (according to the NEB protocol for "polishing" DNA ends). DNA was purified from this reaction using a Qiagen PCR Cleanup Kit. DNA molecules were circularized by ligation at a total DNA concentration <1 μg/mL. This series of enzymatic treatments deletes the HL1-TO1 V_H domain-coding DNA while retaining the V_L domain, restores the *Bam*HI site, and preserves the reading frame between the Aga2 gene and the remaining V_L domain. V_L-only plasmids were constructed by gel-purifying the V_L domain-coding restriction fragments from all other two-domain plasmids after cleavage with *Bam*HI and *Not*I. The pPNL6(HL1-TO1 V_L) vector was cut using the same pair of enzymes to remove the HL1-TO1 V_L stuffer fragment, and the plasmid backbone was gel-purified to prepare for ligation. V_L domains were ligated into this vector backbone following the suggested protocol in the NEB Quick Ligation Kit.

2.2.2 Variable domain shuffling

Full-length, two-domain plasmids were digested with appropriate restriction enzymes to remove the variable domain-coding DNA of interest, which was then replaced with a new variable domain segment. All scFv genes were in the pPNL6 vector backbones; thus, for swapping V_H domains *Nhe*I/*Bam*HI restriction digests were performed to remove and replace V_H domains. *Bam*HI/*Not*I restriction digests were performed to remove and re-

place V_L domains. Plasmid backbones and domain fragments were purified by gel electrophoresis as described above. Purified V_H and V_L domains and plasmid backbones from different starting plasmids were combined and joined by DNA ligation to form new full-length hybrid genes.

2.2.3 Transformation, plasmid recovery and validation
Plasmid DNA was transformed into chemically competent TOP10 or Mach1 *E. coli* (Invitrogen, Carlsbad, CA, USA) or electroporation competent DH5α *E. coli* (Bioline, Boston, MA, USA). Plasmid DNA was extracted from *E. coli* using a miniprep kit (Qiagen). scFv genes in all plasmids were re-sequenced to confirm that they contained the correct fragments and were in frame with the C-terminal c-myc epitope (EQKLISEEDL).

2.3 Yeast surface-display analysis

2.3.1 Growth conditions and induction
Selective growth media (SD+CAA) and induction media (SGR+CAA) for yeast carrying the pPNL6 surface-display plasmid have been previously described [3]. pPNL6 plasmids containing different scFv gene constructs were transformed into EBY100 yeast by EZ Yeast Transformation Kit (BIO 101, La Jolla, CA, USA). The EBY100 yeast transformants were grown for 48 h at 30°C in SD+CAA selective growth media. When the cultures reached an OD at 600 nm of >1.0 (>2×10^7 cells/mL), the yeast cells were harvested by centrifugation and resuspended in SGR+CAA selective induction medium at a concentration of 2×10^7 cells/mL. These induction cultures were incubated with shaking for 72 h at 20°C. For each transformant, an additional culture was maintained in selective growth media as an uninduced control.

2.3.2 Flow cytometry
For each induced and uninduced sample, 10^6 cells were washed twice in wash buffer containing 1× PBS, 2 mM EDTA, 0.1% Pluronic F-127 (a bifunctional block copolymer surfactant that prevents dye binding to plastic) and resuspended in 100 µL wash buffer containing 2 µg mouse anti-c-myc mAb (clone 9E10, Roche Applied Science, Indianapolis, IN, USA). Following a 1-h incubation on ice, the cells are washed twice in wash buffer and resuspended in 100 µL wash buffer containing 0.8 µg appropriately labeled goat anti-mouse secondary antibody. Alexa-fluor 647-conjugated antibodies were used for TO dye activating FAPs. Alexa-fluor 488-conjugated antibodies were used for MG and DIR activating FAPs. All Alexa-fluor antibodies are from Invitrogen. After a 1-h incubation on ice, the cells were washed twice and resuspended in 500 µL wash buffer to which fluorogenic dye was added to a concentration ten times the measured cell-surface K_d for the particular FAP. A parallel set of samples was treated with a 1 µM final concentration of propidium iodide, a vital dye used as an indicator of cell viability. Cells were held on ice until cytometric analysis. FAP activation of the dye was assessed by flow cytometry on a Becton Dickinson FACSVantage SE cytometer. Fluorescence was excited using the 488 nm laser for either the TO dye or the Alexa-fluor 488-conjugated antibodies and fluorescence signals from either of these molecules were measured at 530 nm. MG dye, DIR dye or Alexa-fluor 647-conjugated antibodies were excited using the 635 laser and fluorescence signals from any of these molecules were measured at 685 nm. A ratio of fluorescence signal from the appropriate dye and c-myc channels was calculated to determine signal per scFv molecule compare one scFv-expressing cell sample to another.

2.4 FAP and scFv hybrid protein purification

The scFv genes were isolated from pPNL6 clones and tailed with *Sfi*I restriction enzyme sites by anchored PCR. The forward primer for amplifying and *Sfi*I-tailing the H6-MG gene was 5′-GGCCCAGCCGGCCATGGCGCAGGTGCAGCTGCAGGAGTGC-3′; the reverse primer for amplifying and *Sfi*I-tailing the H6-MG gene is 5′-GGCCCCCGAGGCCTCGGAGACAGTGACCAGGGTACC-3′. The forward primer for amplifying and *Sfi*I-tailing the two-domain hybrid containing the V_H of H6-MG and the V_L of HL1-TO1 was the same as above. The reverse primer for amplifying and *Sfi*I-tailing the two-domain hybrid containing the V_H of H6-MG and the V_L of HL1-TO1 was 5′-GGCCCCCGAGGCCCCTAGGACGGTGAGCTTGGTCC-3′. PCR products were TOPO-cloned (Invitrogen) and sequenced to verify faithful amplification. *Sfi*I fragments were gel-purified after *Sfi*I digestion and ligated into *Sfi*I-digested pAK400 [6] between a pelB leader sequence and His_6 tag. pAK400 in an *E. coli* periplasmic secretion vector for high-level expression of scFvs under the control of a wild-type lac promoter and IPTG-inducible promoter. *E. coli* transformed with the pAK400 plasmids were grown in LB broth to late log phase and induced with 1 mM IPTG in fresh media for 5 h at 25°C. Periplasmic proteins were isolated by osmotic shock [7] and dialyzed into 10 mM Tris, 500 mM NaCl buffer, pH 8.0. Periplasmic, secreted scFv proteins were purified *via* the C-terminal His_6-tag by Nickel-NTA chromatography (Qiagen). The periplasmic protein extract from 1 L

culture was incubated with 0.5 mL settled volume of Nickel-NTA resin for 1 h. The column was poured with this resin and the eluate applied to the column a second time. The column was washed with buffer containing 20 mM imidazole, 10 mM Tris, 100 mM sodium phosphate, 300 mM NaCl, and pH 8.0. His_6-tagged protein was eluted in the same buffer containing 250 mM imidazole. All purification steps and storage of proteins were performed at 4°C. All fractions were analyzed by 15% SDS-PAGE gel electrophoresis to monitor purification. This procedure generates scFv proteins that are homogeneous.

2.5 Soluble protein activity assay and dye titration analysis

All fluorometric analyses, activity assays and the dye titration analyses were performed on a TECAN Safire2 plate reader in black, 96-well, flat-bottom microtiter plates. Purified protein (500 ng H6-MG, or 1 µg H6-MG/HL1-TO1 hybrid) was mixed with MG-11P-NH_2 dye in wash buffer in a final volume of 100 µL. Dye concentrations varied from 0 to 20 µM in a threefold serial dilution series. MG dye samples were excited at 625 nm and emission was detected at 660 nm.

3 Results

3.1 Genetic dissection of two-domain scFvs and functional analysis by surface expression

In the Pacific Northwest National Laboratory yeast (*S. cerevisiae*) surface-display library [2, 3] scFv genes are cloned in the pPNL6 plasmid where they are expressed as fusion proteins between the N-terminal hemagglutinin epitope (YPYDVPDYA)-tagged AGA2p protein and a C-terminal c-myc epitope tag (Fig. 1a). The V_H and V_L gene segments are separated by a synthetic DNA segment that codes for a flexible 15-amino acid linker made up of three repeats of the sequence Gly_4Ser.

scFv clones from our collection of TO-, MG- and DIR-binding FAPs were reduced to their V_H-only or V_L-only plasmids by DNA manipulation (as described in Materials and methods). The modified scFv genes were expressed from the galactose-inducible pPNL6 plasmid to generate surface-displayed fusion proteins tagged with both hemagglutinin and c-myc epitopes (Figs. 1b and c). Flow cytometry of induced, plasmid-carrying yeast cells was used to measure both the amount of scFv (c-myc epitope) expressed and the amount of fluorogenic dye activation when these surface-displayed proteins bound their cognate dye molecules.

As shown in Fig. 2, cytometric analysis of both the full-length scFv and individual variable domain plasmids gave results that fell into two distinct groups. For one group, represented here by yeast displaying the HL1-TO1 proteins, the full-length (V_H+V_L) scFv shows typical fluorescence activation of the dye compared to uninduced cells (significant numbers of cells detected at higher fluorescence). However, neither of the individual V_H or V_L domain proteins could activate the dye (no difference compared to uninduced cells). The other group, represented here by yeast displaying the HL7-MG proteins, also shows typical fluorescence activation of the dye by the full-length scFv compared to uninduced cells. As was the case for the HL1-TO1 protein, the V_H domain of HL7-MG was not capable of activating the dye. However, cells displaying the V_L domain of HL7-MG show as much (if not more) activity compared to the full-length scFv protein. Thus, the molecular dissection of the V_H and V_L domains of HL7-MG reveals that the V_L domain is ca-

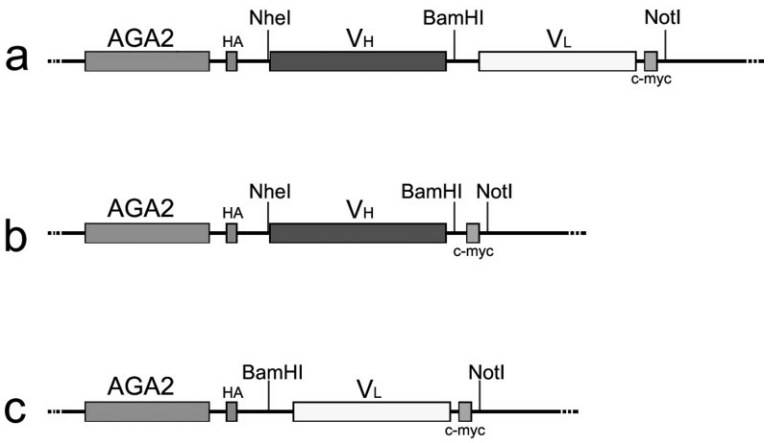

Figure 1. (a) The general structure of scFv clones in the pPNL6 surface display library. The *Nhe*I, *Bam*HI and *Not*I sites are usually unique to the plasmid with the occasional exception of *Nhe*I and *Bam*HI sites appearing in the V_H or V_L coding segments. (b, c) The structure of in-frame deletions of the V_L or V_H domains, respectively.

Biotechnol. J. 2009, 4, 1328–1336

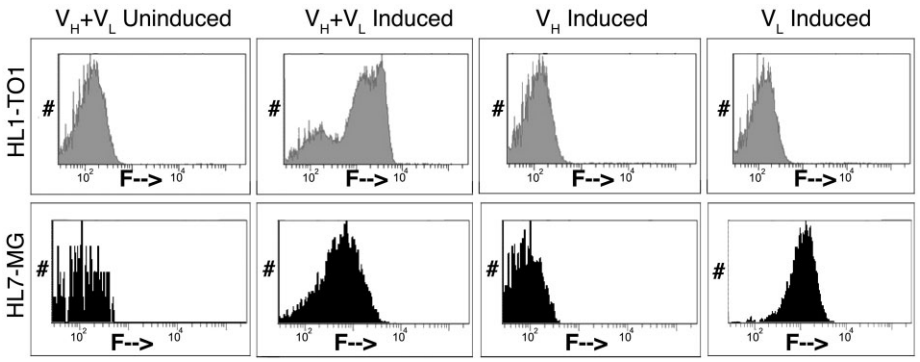

Figure 2. Qualitative cytometric analysis of the modified scFv gene products displayed on the surface of yeast. Full length (V_H+V_L), V_H domain only and V_L domain only scFvs genes were induced for surface display by galactose induction (Induced) and compared to uninduced samples. The panels represent histograms of numbers of cells (#) with a particular fluorescence intensity (F) caused by the activation of the cognate dye for the scFv (note the logarithmic scale). The top panel represents the data from a TO-activating scFv (HL1-TO1). The bottom panel represents the data from a MG-activating scFv (HL7-MG).

pable of binding and fluorogenically activating the dye.

To quantify the activities of isolated V_H or V_L domains, the amount of protein expressed on the surface of the yeast cells was determined by fluorescent immunolabeling of the c-myc epitope fused to the C-terminus of each protein molecule. The population average fluorescence intensity of that c-myc signal was used to normalize the population average dye fluorescence signal to provide a specific activity of the surface-displayed protein. In the bar chart in Fig. 3, the activity of each dissected domain construct is shown as a percentage of the activity of its parent scFv clone.

The quantitative data presented in Fig. 3 demonstrate that five of the scFvs require both V_H and V_L domains for fluorogenic dye activation, *i.e.*, individual V_H or V_L domain proteins retain only a few percent of the activity of the parent scFv protein. The data also allowed us to identify four scFvs for which fluorogenic dye binding and activation can be attributed to only one domain. Expression of the V_L domain contained within HL7-MG, HL9-MG, HL-K10-DIR, or HL-M8-DIR scFv was sufficient to activate the fluorescence of the cognate dye. Sequence analysis of the individual domains of HL7-MG and HL9-MG (not shown) revealed 92% sequence identity at the protein level for the V_L domains, whereas the V_H domains shared only ~75% sequence identity. The V_L domains of HL-K10-DIR and HL-M8-DIR were 100% identical, whereas the V_H domains shared only ~46% sequence identity. Thus, the sequence similarities of the V_L domains in these different scFvs and the lack of similarity of their partner V_H domains are mirrored in their activity.

3.2 Construction and analysis of genetic hybrid scFvs

The apparent increase in activity of the isolated V_L domains of HL7-MG and HL9-MG compared to their V_H/V_L parent FAPs (Fig. 3) suggested that there might be some interference caused by the association of the non-functional V_H domains with the active V_L domains. This observation motivated us to create a set of shuffled V_H/V_L hybrids fusing various active single-domain FAPs with inactive

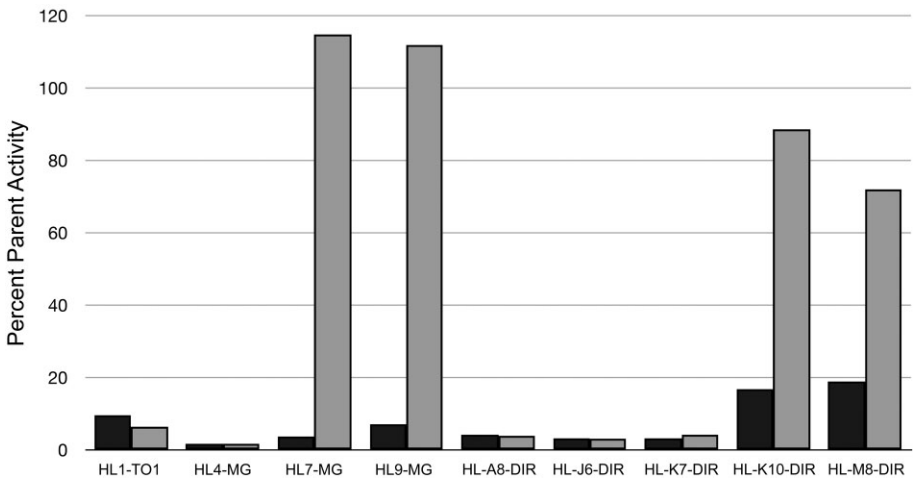

Figure 3. Quantitative analysis of fluorogenic dye activation by individual V_H and V_L domains. The specific activity of each parent FAP, identified by category name, was set to 100%. The activity of each of the V_H domains (dark bars) and the V_L domains (light bars) is plotted relative to its parent. Specific activities were measured as fluorogenic dye fluorescence per unit of c-myc epitope expressed by yeast surface display, as described in Materials and methods.

Table 1. Activity of the hybrid scFvs

Hybrid scFv		
V_H domain	V_L domain	Percent activity
H6-MG	–	=100
H6-MG	HL1-TO1	0.7
H6-MG	HL4-MG	79.5
–	HL7-MG	=100
HL4-MG	HL7-MG	4.3
HL9-MG	HL7-MG	45.7
–	HL9-MG	=100
HL4-MG	HL9-MG	5.4
HL7-MG	HL9-MG	95.6
–	L5.1-MG	=100
HL4-MG	L5.1-MG	38.3
HL1-TO1	L5.1-MG	130.1
HL7-MG	L5.1-MG	125.2
HL9-MG	L5.1-MG	87.8
–	HL-M8-DIR	=100
HL1-TO1	HL-M8-DIR	23.4
HL4-MG	HL-M8-DIR	3.1

variable domains of the complementary type, and then measure the activity of these hybrids. For the active single domains in these hybrids, we used the previously described V_H domain H6-MG [4], the V_L domains from HL7-MG, HL9-MG and HL-M8-DIR described above, and L5.1-MG, an affinity-matured variant (S. Capek and C. Woolford, unpublished results) of the V_L domain L5-MG [4].

Hybrid scFv genes were constructed by DNA manipulation of the corresponding gene segments in the yeast surface-display vector pPNL6. Fusion proteins expressed on the surface of yeast were assayed for fluorogenic dye activation by flow cytometry and normalized, as before, by determining the total amount of surface-displayed c-myc-tagged fusion protein. The activity of the parent single-domain FAP (V_H or V_L) was set to 100% and the hybrid protein activity expressed relative to that. As shown in Table 1, the activity of the hybrid scFvs varies considerably depending on the specific partner domain.

Most notably, fluorogenic dye activation by of some of the active single-domain FAPs is significantly inhibited by the covalent attachment of certain partner variable domains. For example, the activation of MG by the V_H domain H6-MG is inhibited >99% by the presence of the V_L domain of HL1-TO1, and the activity of the V_L domains of HL7-MG, HL9-MG and HL-M8-DIR are inhibited ~95% by the V_H domain of HL4-MG. Interestingly, the activity of the V_L domain of L5.1-MG appears to be inhibited by the V_H domain of HL4-MG, stimulated by the V_H domains of HL7-MG and HL1-TO1, and barely affected by the presence of the V_H domain of HL9-MG.

To demonstrate that the activity modulation of active single-domain FAPs seen in Table 1 is not due to an artifact of the yeast surface display of the hybrid protein, we expressed in *E. coli* and purified soluble versions of the active V_H domain of H6-MG and the hybrid fusion protein containing the V_H of H6-MG and the V_L of HL1-TO1. Dye titrations with the soluble proteins were performed to determine the K_d of MG for each protein (Fig. 4).

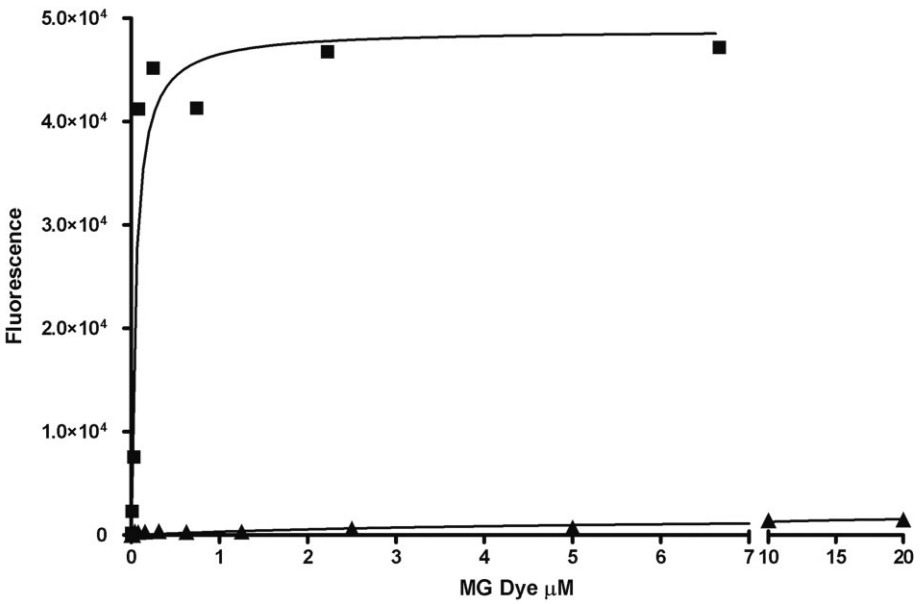

Figure 4. Dye titration analysis of soluble proteins. Increasing amounts of malachite green dye were added to 100-μL reactions containing 500 ng H6-MG protein or 1 μg H6-MG/HL1-TO1 hybrid protein. Fluorescence at 660 nm in arbitrary units was measured on a Tecan fluorimeter. Square symbols are measurements for H6-MG protein; triangle symbols are measurements for the H6-MG/HL1-TO1 inhibited hybrid protein.

In these experiments, the K_d of H6-MG was measured to be 50 nM, which is in reasonable agreement with the published value of 38 nM [4]. The apparent K_d of the inhibited hybrid construct was at least 4.9 μM; this number is difficult to measure, however, because the high MG dye concentration required to observe any fluorescence signal causes significant inner filter effects – self-absorption of the incident and fluorescently emitted light. Nevertheless, the solution Kd of the inhibited hybrid protein is at least two orders of magnitude higher than the active V_H domain that is contained in it. These results confirm that it is the partnering of the V_H and V_L domains in this hybrid that inhibits the activity of the V_H H6-MG domain and not some artifact of the protein's location on the yeast cell surface.

4 Discussion

Our scFv-based FAPs have two distinct activities. Like other antibody molecules, they bind their target antigens with varying affinities depending on the specific antibody-antigen complex. However, because the target antigens are members of a set of environmentally sensitive dye molecules, binding leads to dye conversion or activation from a non-fluorescent to a fluorescent state. It is this second activity that allowed us to use simple cytometric and fluorometric assays to explore V_H and V_L domain requirements and domain organization of the scFv antibodies that we have reported here.

The fluorogenic behavior of the dyes TO, MG and DIR is due to an environmentally sensitive conformational relaxation pathway [8–11]. In solution, excitation of these dyes with visible light allows them to undergo rotation or twisting around one or more of the bonds separating the different ring substituents. This rotation leads to non-radiative decay to the ground state, *i.e.*, the dye exhibits very low fluorescence. However, when the dye is conformationally constrained, such as in viscous solutions or a binding pocket within a protein, this twisting motion is slowed or inhibited and the dye is forced to relax to the ground state radiatively. This constraint can lead to as much as a 1000-fold enhancement of fluorescence and has resulted in the use of TO, for example, in a variety of fluorescent sensors [12–15].

4.1 Structural basis of dye activation and domain inhibition

mAbs that have properties similar to our scFv-based FAPs have been isolated against trans-stilbene [16]. These mAbs prevent the cis-trans isomerization of the bound stilbene causing it to fluoresce, and in one case luminesce, at different wavelengths depending on which particular antibody stilbene is bound to. X-ray crystallographic studies have determined what structural elements constrain the stilbene molecule and suggest molecular and electronic mechanisms for the emission of different colors of light [17]. In these different antibodies and their scFv derivatives, the stilbene molecule is held between the V_H and V_L domains by two different mechanisms. One involves stilbene interactions with complementarity determining region (CDR) loops from both domains, and the other involves a deeper binding cleft where, in addition to the CDR loops, the hydrophobic interface between the V_H and V_L domains is involved in the binding. In both modes, however, binding of stilbene involves structural elements of both domains. It is particularly remarkable, therefore, that more than half of the FAPs we have isolated that bind to and fluorogenically activate our environmentally sensitive dyes are composed of single domain antibody fragments, either V_H or V_L. The molecular mechanisms of their dye binding, leading to fluorescence, may involve unusual interactions of the CDR loops with framework residues, or the V_H or V_L domains acting as dimers, to trap and constrain the different dye molecules. X-ray crystallographic analyses, now underway, should reveal the interesting details of these interactions.

The details of the molecular mechanism of inhibition or enhancement of fluorogen-activating single V domains by other V domains will be equally interesting. Pairing of unrelated V_H and V_L domains through molecular cloning procedures produced the hybrid scFvs we have examined in this study. We interpret the remarkable change of fluorogenic dye activation in some of these hybrids (Table 1) to be the result of the natural association of some V_H and V_L domains and the interaction of their CDR loops in the classical scFv architecture. These interactions can either interfere with dye binding and fluorogenic activation by the active single domain in the hybrid, or they can enhance the dye binding and fluorogenic activation. The wide range of measured fluorogenic activity, even among hybrids containing a common active domain, suggests several possible molecular mechanisms could be involved. We can imagine one scenario in which the formation of a natural interface between V_H and V_L domains actually occludes a part of the dye-binding site that is located either, wholly or partially, in the V_H/V_L interface region in the active V domain. Natural V_H/V_L association would inhibit the dye molecule from binding to the active domain.

In another scenario, we can imagine that the formation of the V_H/V_L hybrid molecule forces a rearrangement of the three-dimensional architecture of the CDR loops in the active domain. If the loops in the active domain were involved in the binding and constraining of the dye, the fluorogen binding or activation could be inhibited or enhanced. For example, the activity of V_L domain FAP, L5.1-MG, is inhibited 15–60% by the V_H partners from HL4-MG and HL9-MG. Its activity is also enhanced 25–30% by the V_H partners from HL1-TO1 and HL7-MG (Table 1). In these contrasting cases, the CDR loops of the V_H partners may either interfere or assist the active domain in binding the dye, thus modulating the activity down or up. Such CDR loop rearrangements in other hybrid scFvs have been documented by X-ray crystallography [18].

The L5.1-MG FAP is a derivative of the parent L5-MG FAP, the isolation of which we have reported previously [4]. L5.1-MG was created by mutagenesis and selection for increased brightness upon dye binding [19]. The L5.1-MG FAP has an increased quantum yield of fluorescence approximately five times higher than its parent, L5-MG (Chris Szent-Gyorgyi, personal communication). This increase in apparent fluorescence is at least an order of magnitude larger that the increase seen when L5.1-MG is partnered with the V_H domains from HL1-TO1 and HL7-MG (Table 1). Nevertheless, because the parent L5-MG FAP was isolated from the yeast surface-display library without a complete partner V_H domain, it seems plausible that its activity could be modulated in either direction by an artificial partner V_H domain.

Lastly, homodimerization of an active single V domain might be required for dye binding and fluorogenic activation. Forced partnering with a complementary V domain may prevent this dimerization and thus cause the inhibition seen for some of our hybrids. Homodimerization of isolated V domains is well known [20, 21]. Dimerization, in some cases, leads to homodimers with V domain structures that differ from the same domain found in the classic arrangement of V_H and V_L domains in either Fab or scFv structures [22, 23]. This might help explain the variation in activities seen in our hybrids. Dimerization of V domains in the yeast surface-display system should be possible. The Aga2-scFv fusion proteins are in extremely close proximity as they pass through the yeast secretory apparatus. The natural affinity of V domains for each other might drive their interaction at any time during the transit through the endoplasmic reticulum, Golgi apparatus and finally secretory vesicles that fuse with the yeast membrane, depositing the fusion proteins on the surface of yeast. In experiments to be reported elsewhere, we have demonstrated FAP activity by complementation of physically separated HL1-TO1 V_H and V_L domains co-expressed on the yeast cell surface using this yeast surface-display system. Thus, it seems likely that V domain dimerization could occur and might be disrupted in the hybrids in Table 1. Preliminary X-ray crystallography structural analyses of the H6-MG and the L5.1-MG FAPs indicate two protein molecules in the unit cell, suggesting that dimerization may be involved in the active form of these FAPs (Robyn Stanfield and Ian Wilson, personal communication).

At this time we have no specific experimental data that leads us to favor any of these suggested molecular explanations for the modulation of FAP activity by partner V domains in our hybrids. Indeed, the remarkable results showing that the active L5.1-MG FAP, can be inhibited, enhanced, or unaffected by partnering with different V_H domains (Table 1) suggests that X-ray structural data will be needed to understand the different mechanisms of this activity modulation. However, these hybrids with modulated fluorogen activation activity will be useful platforms for sensor development (see below).

In an attempt to categorize which V_H and V_L domains inhibit the activities of domains with which they are partnered, we tabulated the lengths of the CDR loops and the V, J and D gene families used to make the various V domains. We have observed no patterns that would lead us to predict which V_H and V_L domains will inhibit other domains, nor have we been able to observe any pattern that would enable us to predict which active V_H and V_L domains would be susceptible to such inhibition. However, the number of hybrid pairs we have been able to create, test and analyze is rather small and, thus, any meaningful correlations may remain hidden. Nevertheless, by simply combining all of the active V_H or V_L FAPs we have isolated with inactive partner domains of the complementary type, we have been able to find at least one pairing in an scFv structure where significant inhibition of the active domain is observed.

4.2 Potential uses of inhibited hybrids

Our scFv hybrids that have low fluorogen-activating activity still contain an active V domain that is being masked by the attachment of another V domain through the flexible $(Gly_4Ser)_3$ linker. These hybrids are excellent platforms for the development of genetically encoded "light-up" biosensors. If the three-dimensional structure of this two-domain protein could be altered by some post-trans-

lational modification to expose or reveal the dye-binding site in the active domain, the fluorogenic dye activation activity should reappear. A wide range of possible uses for this platform can be imagined. For example, if a site-specific protease recognition sequence were inserted into the flexible linker of an inhibited hybrid, cleavage by the cognate protease would allow the inhibiting and active domains to dissociate, a separation driven by the increase in translational entropy. This dissociation should remove the inhibition, allow the active domain to bind a fluorogenic dye and cause an increase in a fluorescence signal. In effect, the engineered hybrid protein could be a genetically encoded, protein-based, fluorescence-generating protease biosensor. In a similar fashion, amino acid sequences that are targets for other post-translational modifications, such as phosphorylation, acetylation, *etc*., can be incorporated into the inhibited hybrid. Enzymatic action at those sites might change the tertiary structure of the hybrid by a sufficient amount to allow fluorogenic dye binding and activation. Some of these types of applications to create biosensors will be discussed in a later report.

The authors wish to thank Liz Chen, John Holleran and Greg Newby for plasmid constructions; and Yehuda Creeger for assistance with flow cytometry. Bruce Armitage, Gordon Rule and V. Emily Stark provided critical review of the manuscript prior to submission. This work was supported by the NIH grant (U54 RR022241) and by the Howard Hughes Medical Institute Undergraduate Education Grant 52005865 to Carnegie Mellon University.

The authors have declared no conflict of interest.

5 References

[1] Deng, X. K., Nesbit, L. A., Morrow, K. J., Jr., Recombinant single-chain variable fragment antibodies directed against Clostridium difficile toxin B produced by use of an optimized phage display system. *Clin. Diagn. Lab. Immunol.* 2003, *10*, 587–595.

[2] Boder, E. T., Wittrup, K. D., Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol.* 1997, *15*, 553–557.

[3] Feldhaus, M. J., Siegel, R. W., Opresko, L. K., Coleman, J. R. *et al.*, Flow-cytometric isolation of human antibodies from a nonimmune *Saccharomyces cerevisiae* surface display library. *Nat. Biotechnol.* 2003, *21*, 163–170.

[4] Szent-Gyorgyi, C., Schmidt, B. F., Creeger, Y., Fisher, G. W. *et al.*, Fluorogen-activating single-chain antibodies for imaging cell surface proteins. *Nat. Biotechnol.* 2008, *26*, 235–240.

[5] Ozhalici-Unal, H., Pow, C. L., Marks, S. A., Jesper, L. D. *et al.*, A rainbow of fluoromodules: A promiscuous scFv protein binds to and activates a diverse set of fluorogenic cyanine dyes. *J. Am. Chem. Soc.* 2008, *130*, 12620–12621.

[6] Krebber, A., Bornhauser, S., Burmester, J., Honegger, A. *et al.*, Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *J. Immunol. Methods* 1997, *201*, 35–55.

[7] Maynard, J., Adams, E. J., Krogsgaard, M., Petersson, K. *et al.*, High-level bacterial secretion of single-chain alphabeta T-cell receptors. *J. Immunol. Methods* 2005, *306*, 51–67.

[8] Furstenberg, A., Julliard, M. D., Deligeorgiev, T. G., Gadjev, N. I. *et al.*, Ultrafast excited-state dynamics of DNA fluorescent intercalators: New insight into the fluorescence enhancement mechanism. *J. Am. Chem. Soc.* 2006, *128*, 7661–7669.

[9] Silva, G. L., Ediz, V., Yaron, D., Armitage, B. A., Experimental and computational investigation of unsymmetrical cyanine dyes: Understanding torsionally responsive fluorogenic dyes. *J. Am. Chem. Soc.* 2007, *129*, 5710–5718.

[10] Duxbury, D. F., The photochemistry and photophysics of triphenylmethane dyes in solid and liquid media. *Chem. Rev.* 1993, *93*, 381–433.

[11] Magde, D., Windsor, M. W., Picosecond internal conversion in crystal violet. *Chem. Phys. Lett.* 1974, *24*, 144–148.

[12] Babendure, J., Liddell, P. A., Bash, R., LoVullo, D. *et al.*, Development of a fluorescent probe for the study of nucleosome assembly and dynamics. *Anal. Biochem.* 2003, *317*, 1–11.

[13] Carreon, J. R., Mahon, K. P., Jr., Kelley, S. O., Thiazole orange-peptide conjugates: Sensitivity of DNA binding to chemical structure. *Org. Lett.* 2004, *6*, 517–519.

[14] Kohler, O., Jarikote, D. V., Seitz, O., Forced intercalation probes (FIT Probes): thiazole orange as a fluorescent base in peptide nucleic acids for homogeneous single-nucleotide-polymorphism detection. *Chembiochem* 2005, *6*, 69–77.

[15] Svanvik, N., Westman, G., Wang, D., Kubista, M., Light-up probes: Thiazole orange-conjugated peptide nucleic acid for detection of target nucleic acid in homogeneous solution. *Anal. Biochem.* 2000, *281*, 26–35.

[16] Simeonov, A., Matsushita, M., Juban, E. A., Thompson, E. H. *et al.*, Blue-fluorescent antibodies. *Science* 2000, *290*, 307–313.

[17] Debler, E. W., Kaufmann, G. F., Meijler, M. M., Heine, A. *et al.*, Deeply inverted electron-hole recombination in a luminescent antibody-stilbene complex. *Science* 2008, *319*, 1232–1235.

[18] Pei, X. Y., Holliger, P., Murzin, A. G., Williams, R. L., The 2.0-Å resolution crystal structure of a trimeric antibody fragment with noncognate VH-VL domain pairs shows a rearrangement of VH CDR3. *Proc. Natl. Acad. Sci. USA* 1997, *94*, 9637–9642.

[19] Chao, G., Lau, W. L., Hackel, B. J., Sazinsky, S. L. *et al.*, Isolating and engineering human antibodies using yeast surface display. *Nat. Protoc.* 2006, *1*, 755–768.

[20] Novotny, J., Haber, E., Structural invariants of antigen binding: comparison of immunoglobulin VL-VH and VL-VL domain dimers. *Proc. Natl. Acad. Sci. USA* 1985, *82*, 4592–4596.

[21] Dottorini, T., Vaughan, C. K., Walsh, M. A., LoSurdo, P., Sollazzo, M., Crystal structure of a human VH: Requirements for maintaining a monomeric fragment. *Biochemistry* 2004, *43*, 622–628.

[22] Steipe, B., Pluckthun, A., Huber, R., Refined crystal structure of a recombinant immunoglobulin domain and a complementarity-determining region 1-grafted mutant. *J. Mol. Biol.* 1992, *225*, 739–753.

[23] Nymalm, Y., Kravchuk, Z., Salminen, T., Chumanevich, A. A. *et al.*, Antiferritin VL homodimer binds human spleen ferritin with high specificity. *J. Struct. Biol.* 2002, *138*, 171–186.