BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

A Variable Light Domain Fluorogen Activating Protein Homodimerizes To Activate Dimethylindole Red

Nina Senutovitch,[†] Robyn L. Stanfield,[§] Shantanu Bhattacharyya,[†] Gordon S. Rule,*[,†,‡] Ian A. Wilson,[§,∥] Bruce A. Armitage,[‡,⊥] Alan S. Waggoner,[†,‡] and Peter B. Berget[†,‡,@]

[†]The Department of Biological Sciences and [‡]The Molecular Biosensor and Imaging Center, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States

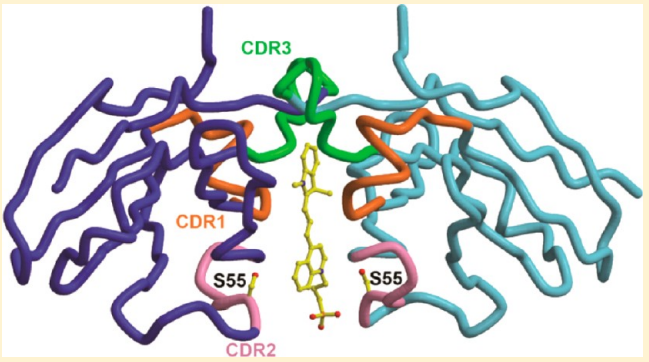
[§]The Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037, United States

[∥]The Skaggs Institute for Chemical Biology, The Scripps Research Institute, La Jolla, California 92037, United States

[⊥]The Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, United States

Supporting Information

ABSTRACT: Novel fluorescent tools such as green fluorescent protein analogues and fluorogen activating proteins (FAPs) are useful in biological imaging for tracking protein dynamics in real time with a low fluorescence background. FAPs are single-chain variable fragments (scFvs) selected from a yeast surface display library that produce fluorescence upon binding a specific dye or fluorogen that is normally not fluorescent when present in solution. FAPs generally consist of human immunoglobulin variable heavy (V_H) and variable light (V_L) domains covalently attached via a glycine- and serine-rich linker. Previously, we determined that the yeast surface clone, V_H-V_L M8, could bind and activate the fluorogen dimethylindole red (DIR) but that the fluorogen activation properties were localized to the M8V_L domain. We report here that both nuclear magnetic resonance and X-ray diffraction methods indicate the M8V_L forms noncovalent, antiparallel homodimers that are the fluorogen activating species. The M8V_L homodimers activate DIR by restriction of internal rotation of the bound dye. These structural results, together with directed evolution experiments with both V_H-V_L M8 and M8V_L, led us to rationally design tandem, covalent homodimers of M8V_L domains joined by a flexible linker that have a high affinity for DIR and good quantum yields.



The development of fluorescent technologies has revolutionized cellular imaging and molecular biology, and the utility of genetically encoded fluorescent proteins, such as green fluorescent protein (GFP), for the detection of particular proteins of interest is well-documented.[1] There is still a need for additional, well-characterized tools that provide real-time, high-signal-to-noise fluorescence and demonstrate a high fluorescence quantum yield (ϕ_f), photostability, and a broad spectral range. Fluorogen activating proteins (FAPs) are part of novel, immunoglobulin-based, fluoromodule platforms that induce fluorescence emission of cognate fluorogenic dyes in solution.[2] FAPs cause a dramatic increase in the ϕ_f or fluorescence enhancement of the fluorogenic dyes that they bind. These fluoromodules have enhanced photostability because of the exchange of bleached dye. Their fluorescence emission spans the visible spectrum from blue to far red, comparable to those of other fluorescence proteins, and often a single FAP can activate multiple dyes.[3−5] The fluorogenic dyes bound by FAPs have low fluorescence background in aqueous solution and undergo increases in fluorescence of as much as 2000-fold upon interaction with a cognate FAP. Previous studies give some insight into the generation of fluorescence from fluorogens, showing that these compounds become fluorescent when the rotation of the aromatic functional groups is restricted by a binding partner, such as upon intercalation into DNA.[6]

Several FAPs that activate the red (640 nm) emitting fluorogenic dye, dimethylindole red (DIR), were isolated.[3] These FAPs are isolated from a naïve human IgG single-chain Fv (scFv) library created in a yeast surface display vector typically consisting of IgG variable heavy (V_H) and light (V_L) chain domains covalently connected by a flexible linker comprised of serine and glycine repeat sequences.[7−9] Two of these FAPs, named V_H-V_L M8 and V_H-V_L K10, were unusual in that they both contained identical V_L domains but different V_H domains. On the basis of the sequence similarity between M8 and K10, it was proposed, and demonstrated experimentally,

Received: September 12, 2011
Revised: March 5, 2012
Published: March 5, 2012

dx.doi.org/10.1021/bi201422g | *Biochemistry* 2012, 51, 2471−2485

that the V_L domain alone was sufficient in the yeast surface display format to bind and activate DIR.[10]

The V_L domain of M8 provides the opportunity to investigate the potential for optimization of DIR fluoromodules. Improvements in the ϕ_f and the affinity of fluoromodules for the dye are desirable, so that they will produce stronger fluorescence intensities for light microscopy at lower dye concentrations. Previously isolated FAPs have ϕ_f values that compare favorably to those of other fluorescent proteins, yet the extent by which the ϕ_f values of DIR activating FAPs can be improved has not yet been determined.[2] To improve the characteristics of FAP−fluorogen pairs, FAP genes can be subjected to directed evolution, utilizing polymerase chain reaction-based mutagenesis, transformation back into yeast, and selection by a fluorescence-activated cell sorter (FACS).[11,12] Because the V_H-V_L M8 FAP is active in both the original isolated V_H-V_L (scFv) format and an isolated V_L domain ($M8V_L$), we were able to compare the directed evolution of this FAP in the two different formats. The aim of this study was to improve our understanding of the mechanism for DIR activation by the $M8V_L$ FAP, in particular determination of the conformational restraints placed on DIR. Following the directed evolution experiments, rational design of the linker region and structural analysis of the active form of the $M8V_L$ FAP were undertaken to improve our understanding of this unusual interaction and the mechanism by which the $M8V_L$ FAP can activate DIR. This study describes the first structural data for FAP-induced fluorescence activation of the environmentally sensitive fluorogen DIR.

EXPERIMENTAL PROCEDURES

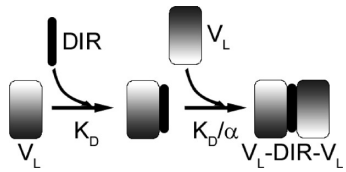
Directed Evolution. Directed evolution of V_H-V_L M8 and $M8V_L$ genes was accomplished through error-prone polymerase chain reaction (PCR), homologous recombination, and FACS enrichment as previously described.[2,11] Three rounds of FACS enrichment of a library of mutants with a diversity of 10^6 were performed at 250 pM DIR for V_H-V_L M8 and 1 nM DIR for the $M8V_L$ domain. During the final round of sorting, cells were autocloned onto induction plates and visually screened for fluorescence enhancement.[3] Further details regarding the directed evolution enrichments for M8 V_HV_L and $M8V_L$ are available in the Figure S1 of the Supporting Information. Plasmid DNA from individual clones was isolated using Zymoprep Yeast Plasmid Miniprep Kit II (Zymogen Inc.) and transformed into bacterial MachT1 cells (Invitrogen). Bacterial mini prep DNA was isolated (Fermentas), and the altered genes were commercially sequenced by Retrogen.

Affinity, Fluorescence Enhancement, and Quantum Yield Determination. The dissociation constants (K_D) of purified soluble FAPs and yeast cell surface-displayed FAPs for DIR were determined by titrations of DIR into samples as previously described.[3] The K_D value for soluble monomeric FAPS (e.g., V_H-V_L) was obtained by fitting the binding data to the standard quadratic binding equation (eq 1) using in-house software that finds both the global and local minima in χ^2.

$$F_{obs} = F_{max}\left[[P_T] + [L_T] + K_d - \sqrt{([P_T] + [L_T] + K_d)^2 - 4[P_T][L_T]}\right]/(2[P_T]) \quad (1)$$

where $[P_T]$ is the total concentration of the protein and $[L_T]$ is the total concentration of the dye. Fitting synthetic data sets indicated that it is possible to distinguish K_D values as low as 0.1 nM with the protein concentrations used in this study (see Figure S11A,B of the Supporting Information). In the case of V_L FAPs that dimerize in solution because of the addition of DIR ($M8V_L$, $M8V_LS^{L55}P$, and Q9), a number of different binding models are possible: (i) formation of V_L-V_L dimers followed by dye binding and (ii) dye binding to one V_L monomer, followed by the addition of the second V_L domain. The second model was selected on the basis of nuclear magnetic resonance (NMR) data that showed that both $M8V_L$ and $M8V_LS^{L55}P$ are monomeric in solution at concentrations of ~500 μM. Consequently, the binding curves were fit to the model illustrated in Scheme 1, using the equations provided by

Scheme 1[a]



[a]DIR binds to the V_L domain with a dissociation constant of K_D, forming the V_L−DIR species. The binding of a second V_L then occurs with a dissociation constant of K_D/α, where α represents a cooperativity parameter.

Mack et al.[13] In this case, the concentration of the V_L−DIR intermediate species was assumed to be small, based on NMR studies with both $M8V_L$ and $M8V_LS^{L55}P$. Consequently, α was set at 10^7 to ensure that $[(V_L)_2DIR] > [V_L-DIR]$. The overall dissociation constant for the reaction, K_D^2/α, was insensitive to the choice of α for values of $\geq 10^5$. Note that the overall dissociation constant for this reaction reports on both the dye−protein interaction (K_D) and the interaction of the free protein with the protein−dye complex (K_D/α). Because we cannot determine α independently, we report the overall affinity, K_D^2/α. For both models, the fitted parameters were the K_D, the fluorescence at saturation, F_{max}, and the total protein concentration, $[P_T]$. In the case of $M8V_L$, a plot of χ^2 versus $[P_T]$ yielded a number of essentially equivalent local minima, in which case the minimum closest (within 10%) to the experimentally measured protein concentration was selected. Errors in K_D were obtained by generating random data sets, assuming the error for each data point followed a normal distribution, and then determining the distribution of the resultant K_D values. Because of the high affinity associated with the synthetic tandem dimers, it was not possible to compare the effect of the serine to proline mutation ($M8V_L$ vs $M8V_LS^{L55}P$) on the affinity of DIR when these FAPs were in the dimer form. However, a quantitative comparison was possible using the isolated V_L domains because the apparent affinity can be controlled by the judicious choice of the protein concentration.

Data for binding of FAPs on the yeast surface were fit using a single-site binding model, assuming that the concentration of free DIR equaled the total amount of DIR

$$F_{obs} = F_{max}\frac{[L_T]}{[L_T] + K_d} \quad (2)$$

A preliminary FACS screen was run prior to all titration assays to determine levels of cell surface FAP expression and

nonviable yeast cells on all FAP constructs, including uninduced control samples. Fluorescence enhancement measurements were performed as described previously.[2]

Quantum yields were determined as previously described[5] utilizing two cross-calibrated reference standards, Cy5.18 and Cresyl Violet, based on the spectral overlap with DIR, and data were analyzed with Origin or GraphPad.[14]

Cloning and Protein Expression and Purification. $M8V_L$ tandem dimer genes were constructed using the wild-type gene and a synthetic, *Escherichia coli* codon-optimized, $M8V_L$ gene (DNA 2.0) that altered the nucleotide sequence but not the amino acid sequence, to prevent homologous recombination of the duplicated genes. The synthetic $M8V_L$ gene was ligated into the 5′ end of wild-type $M8V_L$ using the NheI and BamHI restriction sites in the pPNL6 plasmid containing the wild-type $M8V_L$ gene. Clones were sequenced to confirm the presence of the 5′ synthetic gene and the $(G_4S)_3$ linker region followed by the 3′ wild-type gene.

Tandem dimers of $M8V_L$ with longer (G_4S) repeats in the linker were constructed by inserting oligomers of differing lengths into the BamHI restriction site at the beginning of the (G_4S) linker of the pPNL6-$M8V_L$ tandem dimer described above. The oligomer pair for adding one (G_4S) repeat included 5′GATCAGGTGGCGGTGGCAGCGGCGGTGGTGGTTCCGGAGGCGGCGGTTCTA3′ and 5′GATCTAGAACCGCCGCCTCCGGAACCACCACCGCCGCTGCCACCGCCACCT3′. Complementary oligomers were annealed by incubating the oligomer for 5 min at 95 °C and then cooling slowly it to 23 °C prior to ligation in buffer [50 mM potassium acetate, 20 mM Tris-acetate (pH 7.9), 10 mM magnesium acetate, and 1 mM dithiothreitol]. DNA sequencing confirmed insertion of the extended linker oligomers.

All genes were amplified via PCR with primers containing nonidentical SfiI restriction sites that are compatible with the SfiI sites in the hexahistidine-containing pAK400 *E. coli* periplasmic expression vector (gift from A. Plückthun). These primers are 5′GGCCCAGCCGGCCATGGCGGGTTCTGCTAGCCAGCCTGTGC3′ and 5′GGCCCCCCAGGCCGCTAGGACGGTGACCTTGGTCC3′. $M8V_L$, a mutant $M8V_LS^{L55}P$, and all tandem dimer SfiI-tailed PCR products were blunt cloned into the pJET plasmid (Fermentas), sequenced, and then inserted into pAK400 via SfiI sites. Selected genes were also cloned into the pPNL9 yeast secretion plasmid by gap repair and transformation into yeast strain YVH10 as previously described.[2]

$M8V_L$ and $M8V_LS^{L55}P$ proteins were produced in milligram quantities for crystallographic studies from the pAK400 vector in *E. coli* MachT1 (Invitrogen) cells. Typically, 3 g of cells from 0.5 L of culture was lysed in 20 mL of buffer [50 mM Tris-HCl (pH 7.5), 750 mM NaCl, 0.1% Triton X-100, and 0.05% Tween 20] for 10 min in a cell homogenizer (Avestin EmulsiFlex-C3). The lysate was centrifuged for 30 min at 28000*g*, and the supernatant fraction batch bound to nickel agarose resin (ThermoFisher). The nickel resin was washed in 10 column volumes of buffer [20 mM imidazole, 10 mM Tris-HCl (pH 8.0), 50 mM KH_2PO_4, and 300 mM NaCl], followed by 10 column volumes of buffer [20 mM imidazole, 10 mM Tris-HCl (pH 8.0), 50 mM KH_2PO_4, and 750 mM NaCl] and 2 column volumes of buffer [50 mM imidazole, 10 mM Tris-HCl (pH 8.0), 50 mM KH_2PO_4, and 300 mM NaCl]. The protein was recovered in elution buffer [250 mM imidazole, 10 mM Tris-HCl (pH 8.0), 50 mM KH_2PO_4, and 300 mM NaCl]. Proteins were further purified by ion exchange chromatography on a HiTrap XP XL column (GE Healthcare) following dialysis into 50 mM MES (pH 6.0) and 50 mM NaCl. A linear gradient from 50 mM to 1 M NaCl was used to elute the protein from the ion exchange column. Protein preparations were analyzed by sodium dodecyl sulfate−polyacrylamide gel electrophoresis to determine purity. $M8V_L$ was further purified by size exclusion chromatography on a Superdex 200, 10/30 column in 0.2 M Tris (pH 7.5) and 0.15 M NaCl.

To allow labeling with ^{13}C for NMR experiments, the $M8V_L$ gene was transferred from the pAK400-$M8V_L$ vector, using flanking NdeI and HindIII sites (enzymes from New England Biolabs), into the pET22b+ plasmid (Novagen). Studier's minimal medium was used to produce labeled $M8V_L$ [50 mM Na_2HPO_4, 50 mM KH_2HPO_4, 25 mM $(^{15}NH_4)_2SO_4$, 0.5% glucose, 1× trace metals, and 2 mM $MgSO_4$].[15] The ^{15}N source was $(^{15}NH_4)_2SO_4$, and the ^{13}C source was glucose, both from Cambridge Isotope Laboratories. The labeled protein was purified essentially as described above.

Crystallization, Data Collection, and Structure Determination. Crystal structures were determined for $M8V_L$, $M8V_L$ with DIR, and $M8V_LS^{L55}P$ with DIR. Data collection and refinement statistics for the three structures are summarized in Table 3. For all three structures, data were processed with HKL-2000,[16] phases were determined by molecular replacement using Phaser,[17] and model building and refinement were conducted with Coot,[18] Refmac5,[19] and Phenix.[20] The coordinates are numbered by the Kabat convention.[21]

$M8V_L$ was cocrystallized with DIR by mixing $M8V_L$ [15 mg/mL in 0.2 M Tris and 0.15 M NaCl (pH 7.5)] with a 5-fold molar excess of DIR [10 mg/mL in 20% DMSO, 0.2 M Tris, and 0.15 M NaCl (pH 7.5)]. Initial crystallization conditions were identified by screening 384 different solutions with the Topaz System (Fluidigm),[22] and the crystal used for data collection was grown at 22 °C in a sitting drop with a well solution of 35% PEG 1500. The crystal was cryoprotected with a mixture of 90% well solution and 10% PEG 200 and cryocooled by being plunged into liquid nitrogen. Data were collected at the Advanced Photon Source, beamline 23-ID-B, to a resolution of 1.50 Å, and the structure was determined by molecular replacement using the λ light chain variable domain from Fab 2219 [Protein Data Bank (PDB) entry 2b0s] as a starting model. The final model (PDB entry 3T0W) contains two $M8V_L$ molecules in the asymmetric unit (residues A1−108 and B1−108), two half-occupancy DIR molecules, two chloride ions, two PEG molecules, and 319 waters and has R_{cryst} and R_{free} values of 14.8 and 18.3%, respectively.

$M8V_L$ [15 mg/mL in 0.2 M Tris and 0.15 M NaCl (pH 7.5)] in the absence of DIR was crystallized at 22 °C in sitting drops with a well solution of 2 M ammonium sulfate, 2% PEG 400, and 0.1 M HEPES [4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid] (pH 7.5) that represented a condition originally identified using the JCSG/IAVI CrystalMation robot (Rigaku) and screening 384 different conditions at two different temperatures. The crystals were cryoprotected with a mixture of 70% well solution and 30% ethylene glycol and cryocooled by being plunged into liquid nitrogen. Data were collected at the Advanced Light Source, beamline 4.2.2, to a resolution of 1.45 Å. The structure was determined by molecular replacement with search model coordinates consisting of one V_L domain from the previously determined $M8V_L$ structure.

The final model (PDB entry 3T0V) contains one unliganded M8V$_L$ molecule in the asymmetric unit (residues A1−110), 127 waters, one Tris, one sulfate ion, four ethylene glycol molecules, and two PEG molecules. The final R_{cryst} and R_{free} values were 17.9 and 20.8%, respectively.

M8V$_L$S^{L55}P cocrystallized with DIR was formed by mixing M8V$_L$S^{L55}P (3 mg/mL) with a 10-fold molar excess of DIR (10 mg/mL in 20% DMSO, 0.2 M Tris, and 0.15 M NaCl) and concentrating the mixture to 15 mg/mL. Initial crystallization screening was conducted with the CrystalMation robot, and the crystal used for data collection was grown in a sitting drop with a well solution of 0.25 M ammonium sulfate and 30% PEG 4000. The crystal was cryoprotected with a mixture of 70% well solution and 30% ethylene glycol and cryocooled by being plunged into liquid nitrogen. Data were collected at the Advanced Photon Source, beamline 23-ID-B, to a resolution of 1.95 Å, and the structure was determined by molecular replacement using one V$_L$ domain from the M8V$_L$ structure as a starting model. The final model (PDB entry 3T0X) contains two M8V$_L$S^{L55}P molecules (residues A1−108 and B1−106), one DIR molecule with two alternate conformations for the indole moiety, one sulfate ion, seven ethylene glycol molecules, and 147 waters. The final R_{cryst} and R_{free} values were 20.9 and 24.0%, respectively.

NMR Analysis. Purified M8V$_L$ protein was dialyzed into buffer [10 mM Na_2HPO_4 (pH 6.3) 250 mM NaCl] with 0.75 mM arginine and 0.75 mM glutamate to reduce the level of aggregation and 0.02% azide.[23]The protein was concentrated to 0.9 mM for NMR relaxation experiments. The M8V$_L$ fluoromodule with DIR was formed using a 4:1 molar ratio of lyophilized DIR that was resuspended in the dialyzed protein sample and incubated overnight. NMR experiments were performed with a Bruker 600 MHz NMR spectrometer fitted with a cryoprobe at 304 K. The spectra were processed with NMRPipe.[24] The T_1, T_2, and heteronuclear NOE experiments[25] with ^{15}N-labeled material were analyzed using NMRViewJ.[26] In-house scripts were used to further analyze the data and prepare the input files for Relax-NMR.[2728−31] The global correlation times, τ_m, were determined using the "model-free.py" script available in Relax-NMR.[32−34]The theoretical values for global correlation time were obtained using HYDRONMR.[35] Standard multinuclear NMR sequences[36,37] were used in conjunction with ^{15}N- and ^{13}C-labeled material to obtain resonance assignments for main chain atoms in the unliganded complex utilizing an in-house Monte Carlo-based assignment program.[38] Interproton NOEs were obtained from ^{15}N-separated NOESY experiments that detected either the ^{1}H frequency or the ^{13}C frequency of the attached carbon of protons that were dipolar coupled to NH protons.[39]

Calculation of Protein−DIR van der Waals Energies. CHARMM version 33b2[40] was used to compare van der Waals energies using the top aa22 parameter file. Bond lengths, angles, and planar torsional angles for DIR were obtained from the X-ray coordinate files described in this paper. Standard nonbonded atom parameters were used for aliphatic and aromatic carbons and hydrogens in DIR. The protein−DIR system was minimized with 100 steps of steepest decent, and then adopted basis set Newton−Raphson minimization was applied until the energy converged. During the minimization, the protein main chain N, C_α, and carbonyl atoms were constrained to positions in the original X-ray structure. The difference between the van der Waals energy for the protein plus DIR minus the protein without DIR was used to determine the contribution of protein−DIR interactions to the van der Waals energy.

RESULTS

Directed Evolution of the V$_H$-V$_L$M8 FAP. To determine if changes in affinity and quantum yield can increase the fluorogenicity of V$_H$-V$_L$ M8, directed evolution was performed using FACS selection for increased fluorescence of yeast surface-displayed FAP at low DIR concentrations.[8] The FACS enrichment selections were conducted at 250 pM DIR after the yeast cell surface affinity of V$_H$-V$_L$ M8 had been determined to be 1.2 nM (Figure S1A−C of the Supporting Information). Each round of enrichment and selection monitored both the amount of yeast surface expression and the intensity of the DIR fluorescence signal of cells in the mutagenized population. Single cells were isolated during the third cycle of enrichment as described previously.[9] Individual clones of interest were identified on the basis of visual inspection for increased fluorescence on induction plates compared to wild-type V$_H$-V$_L$ M8. Amino acid changes were identified by DNA sequencing. DIR ϕ_f and fluorescence enhancement of cell surface-displayed FAPs were then determined by FACS analysis of individual clones. The equilibrium dissociation constants for isolated clones from directed evolution were initially determined to identify mutants with enhanced binding to DIR (Figure S2 of the Supporting Information). These measurements identified three classes, as determined from sequence length and alignment to the original V$_H$-V$_L$ M8, from a total of 12 single clones (Figures S3 and S4 of the Supporting Information).

The first class of isolated clones, which have both V$_H$ and V$_L$ domains, is represented by L9 in Table 1 (see Figure S2 of the Supporting Information for binding curves of FAPs on the yeast surface and Figure S5 of the Supporting Information for the binding curves of soluble Q9 and J8); the yeast surface display fluorescence enhancement of this clone was increased 5-fold versus the parent M8. Sequence analysis revealed several amino

Table 1. Characteristics of Clones Isolated via Directed Evolution of V$_H$-V$_L$ M8 and M8V$_L$ FAPs

clone	format	fluorescence enhancement (YSD)[a] (*x*-fold)	soluble K_D or K_D^2/α	ϕ_f	mutations in the sequence different from wild type (V$_H$-V$_L$ M8 or M8V$_L$)
M8	V$_H$-V$_L$	1.0	not determined		
L9	V$_H$-V$_L$	5.7	not determined		F^{H29}S, W^{H36}C, A^{H93}V, I^{H113i}T
Q9	V$_L$	3.8	(1.3 ± 0.3) × 10^{-15} M^2	0.19	Q^{L1}R, T^{L14}I, D ^{L30}G, Q^{L37}R, S^{L55}P, F^{L62}L, S^{L80}P
J8	V$_L$-V$_L$	6.4	<0.1 nM[b]		V$_L$1, D^{L85}N, K^{L103}T, L^{L107}S, I^{L108i}T; V$_L$2, S^{L9}P, S^{L32}P
M8V$_L$	V$_L$	3.7	(2.5 ± 0.7) × 10^{-15} M^2	0.71	
M8V$_L$S^{L55}P	V$_L$	8.2	(1.0 ± 0.5) × 10^{-16} M^2	0.58	S^{L55}P

[a]Fluorescence enhancement is calculated as the DIR fluorescence signal normalized for cell surface expression for each individual clone, divided by the equivalent value for the wild-type parent (V$_H$-V$_L$ M8). The concentration of Q9 was 33 nM, that of J8 9 nM, that of M8V$_L$ 10 nM, and that of M8V$_L$S^{L55}P 135 nM. [b]Data were equally well fit to a wide range of K_d values of <0.11 nM.

acid changes in the V_H domain of clone V_H-V_L L9 ($F^{H29}S$, $W^{H36}C$, $A^{H93}V$, and $I^{H113i}T$), according to the Kabat numbering system (H113i is an extension of Kabat numbering based on extra residues on the C-terminus of the M8 heavy chain).[21]

A second class of clones, which have only the V_L domain, is represented by Q9 in Table 1. This clone displayed an increased fluorescence enhancement of 4-fold above the V_H-V_L M8 parent. Sequence analysis revealed that this clone contained seven mutations in the V_L domain ($Q^{L1}R$, $T^{L14}I$, $D^{L30}G$, $Q^{L37}R$, $S^{L55}P$, $F^{L62}L$, and $S^{L80}P$), in addition to suffering deletion of the V_H domain (Figure S4 of the Supporting Information).

A third class represented by one clone, J8, had a 6-fold increase in fluorescence enhancement compared to the parent V_H-V_L M8 (Table 1). This clone is remarkable in that it is a "pseudodimer" gene containing two tandem-linked V_L domains, presumably generated by recombination during the yeast gap repair process used in the creation of the directed evolution library. Each of the two V_L domains is approximately 95% homologous to the V_L domain of the parent V_H-V_L M8 (Figure S6 of the Supporting Information). The V_L-V_L J8 pseudodimer displayed six total amino acid substitutions in both light domains. Four changes were present in the N-terminal light chain, $D^{L85}N$, $K^{L103}T$, $L^{L107}S$, and $I^{L108i}T$, and two additional changes were present in the C-terminal light chain, $S^{L9}P$ and $S^{L32}P$. The pseudodimer contains 12 residues that do not align with the $M8V_L$ sequence and are located prior to the flexible linker. These residues align with the region of V_H M8 that is N-terminal to the linker. They are described as A^{L108a}, S^{L108b}, T^{L108c}, K^{L108d}, G^{L108e}, P^{L108f}, S^{L108g}, G^{L108h}, T^{L108i}, L^{L108j}, and G^{L108k} because they are not found in conventional antibodies and are an addition to the typical Kabat numbering scheme (Figure S7 of the Supporting Information). Homologous recombination is likely to have generated these additional residues as well.

Directed Evolution of the $M8V_L$ FAP. Directed evolution of $M8V_L$ was performed as described above using a DIR concentration of 1 nM during FACS sorting (Figure S1D–F of the Supporting Information). This selection generated a single class of clones, represented by clone A4 in Table 1, with a >8-fold increase in fluorescence enhancement compared to the parent $M8V_L$–DIR species (Table 1). All 16 clones from this class contained only one change of serine to proline at position 55 ($S^{L55}P$). This clone is subsequently termed $M8V_LS^{L55}P$, in accordance with Kabat nomenclature.

Characterization of Tandem Homodimers of $M8V_L$ FAPs on the Yeast Cell Surface. The isolation of a tandem linked V_L-V_L J8 pseudodimer from the directed evolution of V_H-V_L M8 led us to engineer tandem homodimer genes of the $M8V_L$ domain and investigate the effect of different (G_4S) linker lengths on fluorescence enhancement. Yeast cell surface expression and fluorescence enhancement measurements of clones with different linker lengths were performed. There was no significant difference in the DIR affinity of tandem homodimers with extended (G_4S) linkers based on cell surface affinity measurements (Figure S8 of the Supporting Information). None of the linker lengths were less than $(G_4S)_3$ (15 amino acids) and thus not predicted to form diabodies or triabodies.[41,42] Fluorescence enhancement was increased 2-fold for homodimers containing the $(G_4S)_4$ or $(G_4S)_6$ linker. The tandem homodimer gene made from the $M8V_LS^{L55}P$ domain produced a FAP with a 3.5-fold increased fluorescence enhancement compared to the similar tandem homodimer made with the $M8V_L$ domain (Table 2).

Table 2. Results of Altering Glycine- and Serine-Rich Linker Lengths in Tandem Homodimers Created from $M8V_L$

clone	no. of (G_4S) repeats in linker	fluorescence enhancement (YSD)[a]	soluble K_D (nM)	ϕ_f
$dM8V_L(G_4S)_3$	3	1.0	<0.1[b]	0.38
$dM8V_L(G_4S)_4$	4	2.0	<0.1[b]	0.64
$dM8V_L(G_4S)_6$	6	2.0	<0.1[b]	0.55
$dM8V_LS^{L55}P(G_4S)_3$	3	3.4	<0.1[b]	0.14

[a]Fluorescence enhancement is calculated as the DIR fluorescence signal normalized for cell surface expression for each individual clone, divided by the equivalent value for $M8V_L$. The concentration of $dM8V_L(G_4S)_3$ was 3 nM, that of $dM8V_L(G_4S)_4$ 3 nM, that of $dM8V_L(G_4S)_6$ 3 nM, and that of $dM8V_LS^{L55}P$ 12 nM. [b]Data were equally well fit to K_d values of <0.1 nM.

Characterization of V_H -V_L M8, $M8V_L$, and Tandem V_L Homodimer FAPs as Soluble Proteins. To verify that the affinity of the FAPs for DIR is not influenced by the yeast surface display of the protein, we attempted to express soluble versions of these FAPs. The parent V_H-V_L M8 FAP showed a remarkable affinity for DIR on the yeast cell surface (K_D = 1.2 nM); however, a K_D for the soluble protein produced by yeast secretion could not be determined by fluorescence measurements (Figure S9 of the Supporting Information). The purified V_H-V_L M8 protein showed a complete absence of fluorogen activation. To date, no other FAPs that we have characterized have shown strong affinity or activity by cell surface titration measurements but a complete lack of activity when assayed as a soluble protein. The yield of yeast secreted protein from V_H-V_L M8 was unexpectedly low, based on the typically strong correlation of yeast secretion yield and yeast surface display.[43,44] Thus, further biochemical characterization of the soluble V_H-V_L M8 protein was not possible because of the low protein expression levels by either yeast secretion or bacterial expression (data not shown). Similarly, no characterization was possible for purified proteins from any of the clones from directed evolution that contained the V_H domain of V_H-V_L M8 (such as V_H-V_L L9), because of poor protein expression levels (data not shown). However, the purified soluble V_L-V_L J8 pseudodimer protein has a K_D for DIR of <0.1 nM (Table 1). This dissociation constant is similar to that found for the synthetic V_L dimer [$M8V_L(G4S)_3$]; however, the low K_D values for these two proteins preclude an accurate comparison of their K_D values. Determination of ϕ_f for J8 FAP was not possible because of the low protein yield.

The solution binding properties of three different monomer V_L FAPs were studied, Q9, $M8V_L$, and $M8V_LS^{L55}P$. The purified $M8V_L$ protein showed an overall dissociation constant of 2.5×10^{-15} M^2 (Figure 1) and a robust ϕ_f of 71%. The affinity of the $M8V_LS^{L55}P$ mutant increased >10-fold to 1.0×10^{-16} M^2 (Figure 1). Because of the nature of the binding reaction, the amount of DIR required to half-saturate the protein (apparent K_D) depends on the protein concentration; at a protein concentration of 1 nM, the apparent K_D values would be 2.5 μM for $M8V_L$ and 0.1 μM for $M8V_LS^{L55}P$, i.e., an ~25-fold increase in affinity. Although the $M8V_LS^{L55}P$ mutant protein showed an increase in affinity, the ϕ_f decreased to 58% (Table 1). The increase in the affinity of $M8V_LS^{L55}P$ is greater than that observed for the V_L domain isolated during affinity maturation of the original V_L-V_H construct (Q9); the overall dissociation constant of the soluble Q9 protein for DIR is 1.3×10^{-15} M^2. Although the affinity of Q9 for DIR is ~2-fold higher

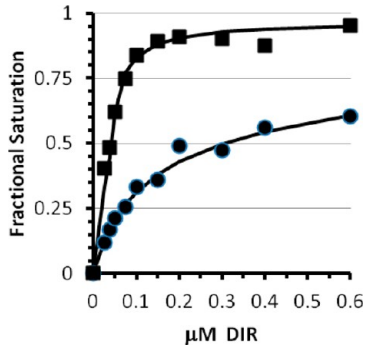


Figure 1. Determination of equilibrium dissociation constants for soluble M8V$_L$ and M8V$_L$S^{L55}P. Fluorescence data for M8V$_L$ (●) and M8V$_L$S^{L55}P (■) were fit to Scheme 1. Solid lines show the best fit to the data. Error bars for the individual data points are within the size of the plotted points. Serial dilutions of DIR were added to 11 nM soluble M8V$_L$ or 135 nM soluble M8V$_L$S^{L55}P protein in buffer [PBS (pH 7.4) with 2 mM EDTA and 0.1% (w/v) Pluronic F-127].

than that of M8V$_L$, the ϕ_f is relatively low, 19% (Table 1). There were no significant differences observed in the DIR absorbance spectra of fluoromodule M8V$_L$ or Q9V$_L$, despite differences in ϕ_f (Figure S10 of the Supporting Information).

We also analyzed the affinities for DIR and ϕ_f values of the soluble proteins made by M8V$_L$ tandem dimers with modified linker lengths. Altering the (G_4S) repeats produced soluble proteins with differing ϕ_f values. The K_D values of the M8V$_L$ dimer with the $(G_4S)_3$, $(G_4S)_4$, or $(G_4S)_6$ linker was <0.1 nM, and thus, it is not possible to compare the K_D values for these different constructs. The quantum yield was somewhat dependent on the length of the linker; a value of 38% was found for the $(G_4S)_3$ linker, 64% for the $(G_4S)_4$ linker, and 55% for the $(G_4S)_6$ linker (Table 2 and Figure S11B of the Supporting Information). A slight decrease in the quantum yield for the longer linker may be due to the formation of intermolecular complexes, similar to diabodies or triabodies, but this was not indicated by size exclusion chromatography (Figure S12 of the Supporting Information). The M8V$_L$S^{L55}P dimer with a $(G_4S)_3$ linker has a K_D of <0.1 nM and the lowest ϕ_f of 14% (Table 2).

Structure of Unliganded M8V$_L$ and Cocrystals of M8V$_L$ and M8V$_L$S^{L55}P with DIR. To investigate the mechanism of DIR fluorescence activation by the M8V$_L$ FAPs and to

Table 3. Data Collection and Refinement Statistics for M8V$_L$ and M8V$_L$S^{L55}P

	M8V$_L$ (no DIR)	M8V$_L$–DIR	M8V$_L$S^{L55}P–DIR
		Data Collection	
beamline	ALS 4.2.2	APS 23-ID-B	APS 23-ID-B
wavelength (Å)	1.000	1.033	0.722
resolution (Å)[a]	1.45 (1.48–1.45)	1.50 (1.53–1.50)	1.96 (1.98–1.96)
space group	$C222_1$	$P2_12_12$	$P4_32_12$
a, *b*, *c* (Å)	83.66, 93.64, 30.57	73.90, 83.53, 35.14	83.49, 83.49, 76.35
no. of observations[a]	80786 (2233)	250067 (11950)	240775 (12015)
no. of unique reflections[a]	21290 (830)	35526 (1758)	20018 (977)
completeness (%)[a]	97.7 (79.2)	99.9 (100.0)	100.0 (100.0)
R_{sym} (%)[a,b]	4.3 (31.7)	8.7 (53.7)	8.7 (54.9)
average I/σ[a]	30.1 (2.4)	28.2 (4.2)	35.5 (4.9)
		Refinement (all reflections of >0.0σF)	
resolution (Å)	29.2–1.45	41.8–1.50	37.3–1.96
no. of reflections (working set)	20190	33690	18888
no. of reflections (test set)	1090	1779	1048
R_{cryst} (%)[c]	17.9	14.8	20.9
R_{free} (%)[d]	20.8	18.3	24.2
no. of protein atoms	845 (29 in alternate conformations)	1670 (70 in alternate conformations)	1662 (14 in alternate conformations)
no. of DIR atoms	0	32 (32 in alternate conformations)	32 (18 in alternate conformations)
no. of waters	127	319 (8 in alternate conformations)	138
no. of solvent atoms	51	12	33
average *B* value (Å^2)			
chain A	18.9	11.7	32.5
chain B	not available	15.5	46.6
DIR	not available	16.0	27.6
Wilson *B* value (Å^2)	15.1	11.7	26.6
Ramachandran plot			
most favored	90.3	90.7	85.4
additionally allowed	8.6	8.2	12.9
generously allowed	0.0	0.0	0.6
disallowed[e]	1.1	1.1	1.1
rmsd			
bond lengths (Å)	0.006	0.005	0.011
bond angles (deg)	1.12	1.30	1.45

[a]Numbers in parentheses are for the highest-resolution shell of data. [b]$R_{sym} = \sum_{hkl}|I - \langle I\rangle|/\sum_{hkl}I$. [c]$R_{cryst} = \sum_{hkl}|F_o - F_c|/\sum_{hkl}|F_o|$. [d]$R_{free}$ is the same as R_{cryst}, but for 5% of the data excluded from the refinement. [e]Residue Asn^{51}, which is in a conserved γ-turn and is almost always found in this region in antibody structures.

dx.doi.org/10.1021/bi201422g | *Biochemistry* 2012, 51, 2471–2485

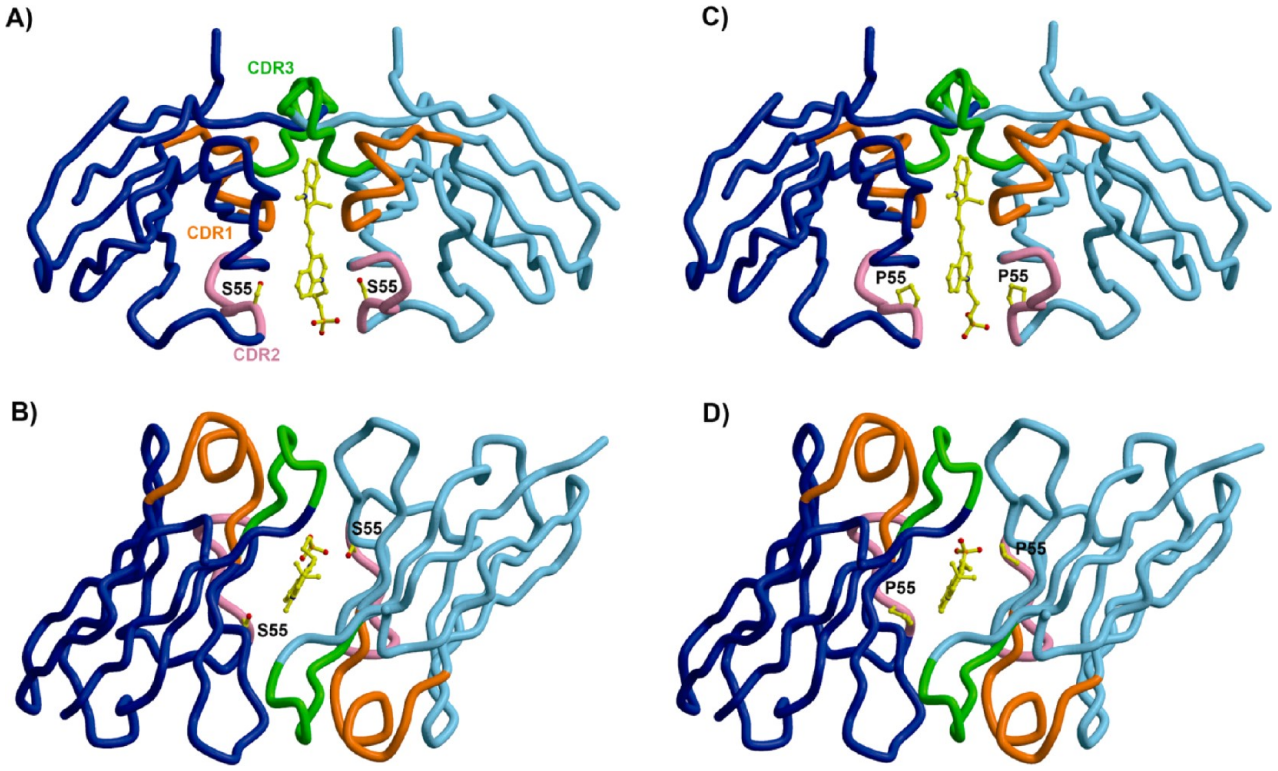


Figure 2. Overall topology of M8V$_L$ and M8V$_L$S^{L55}P bound to DIR. (A and B) M8V$_L$ binds DIR (only one DIR conformation is shown for the sake of clarity) sandwiched between two identical V$_L$ domains (blue and light blue). The CDR1, CDR2, and CDR3 loops are colored orange, pink, and green, respectively, and are labeled in panel A. The view in panel B is rotated 90° about a horizontal axis. (C and D) M8V$_L$S^{L55}P binds the DIR in a manner almost identical to that of M8V$_L$. The S^{L55}P mutation is the only difference in sequence between the two structures. The structures were generated with Molscript[59] and Bobscript.[60]

understand the increased affinity and altered fluorescence enhancement of the M8V$_L$S^{L55}P mutant, we determined the crystal structure of two of these proteins (Table 3). Crystal structures were obtained for both M8V$_L$ and M8V$_L$S^{L55}P in complex with DIR (1.5 and 1.96 Å resolution, respectively), as well as M8V$_L$ in the absence of DIR (1.45 Å resolution). In all three crystal structures, the M8V$_L$ or M8V$_L$S^{L55}P domains adopt a typical Ig fold. CDR loops from V$_L$ domains from all three structures adopt the expected canonical structures and are classified as L1 = 5λ/13A, L2 = 1/7A, and L3 = 5λ/11A.[45,46] Both M8V$_L$ and M8V$_L$S^{L55}P crystallize as a noncovalent homodimer of V$_L$ domains, with the DIR molecule packed tightly at the homodimer interface (Figure 2). In both structures, the V$_L$ domains are related by an approximate noncrystallographic 2-fold axis, with an angle of ~178° relating the two V$_L$ domains in each structure, and the DIR molecules bound on the noncrystallographic symmetry (NCS) axis (Figure 3 and Table 4). In the M8V$_L$ structure, the DIR occupies two conformations, related by the approximate NCS. In the M8V$_L$S^{L55}P structure, only the indole ring adopts alternate conformations, also related by the NCS 2-fold axis (Figure 4). The alternate conformations of the DIR are a crystallization artifact, where the V$_L$ dimer binds DIR in one single conformation but the complex then crystallizes in two orientations around the NCS 2-fold axis. In the absence of the DIR, M8V$_L$ crystallizes with one monomer in the asymmetric unit. This monomer does not closely associate with any symmetry-related neighbors to form a dimer in the crystal (Figure 7A). The homodimeric arrangement seen for M8V$_L$ and M8V$_L$S^{L55}P differs strikingly from that seen for a V$_L$-V$_H$ dimer in an Fab molecule and also differs from any V$_L$-V$_L$ dimer structures found in the PDB in that the two V$_L$ domains are antiparallel (Figure S13 of the Supporting Information), while still burying the hydrophobic interface that would normally be buried in a V$_L$-V$_H$ interaction in a typical antibody. The DIR is buried in a deep pocket and forms contacts primarily with the CDR loops for both M8V$_L$ and M8V$_L$S^{L55}P (Figure 2). DIR contact residues are contributed by all CDR loops: Tyr^{L34}, Asn^{L50}, Arg^{L53}, Ser/Pro^{L55}, Ser^{L56}, Leu^{L89}, and Trp^{L96} (Table 4). Tyr^{L34} participates in π-stacking with the conjugated DIR polymethine bridge for both M8V$_L$ and M8V$_L$S^{L55}P (Figures 3 and 4), with the centroid of the Tyr^{L34} rings located 5.6−6.1 Å from the plane of the DIR, and angles between Tyr^{L34} and DIR between 31.7° and 36.7° (Figure S14 of the Supporting Information). Tyr^{L49} also participates in parallel π-stacking with the DIR quinoline ring for both M8V$_L$ and M8V$_L$S^{L55}P (Figures 3 and 4), with the centroid of the Tyr^{L49} rings located 3.6−4.0 Å from the plane of the DIR, and angles between Tyr and DIR ranging from 5.0° to 10.7° (Figure S14 of the Supporting Information). Trp^{L96} packs perpendicular to the indole ring for both M8V$_L$ and M8V$_L$S^{L55}P (Figures 3 and 4), with the centroid of the Trp^{L96} ring located 5.5−5.7 Å from the plane of the DIR with angles between the Trp^{L96} and indole ring ranging from 70.3° to 78.6° (Figure S14 of the Supporting Information). The DIR molecules are bound with nearly coplanar quinoline and indole rings, with angles between the two ring systems (for the two DIR conformers) in the M8V$_L$−DIR protein of 8.5° and 4.4°, respectively and in the M8V$_L$S^{L55}P−DIR protein of 4.2° and 3.0°, respectively (Figures S14 and S15 of the Supporting Information).

dx.doi.org/10.1021/bi201422g | *Biochemistry* 2012, 51, 2471−2485

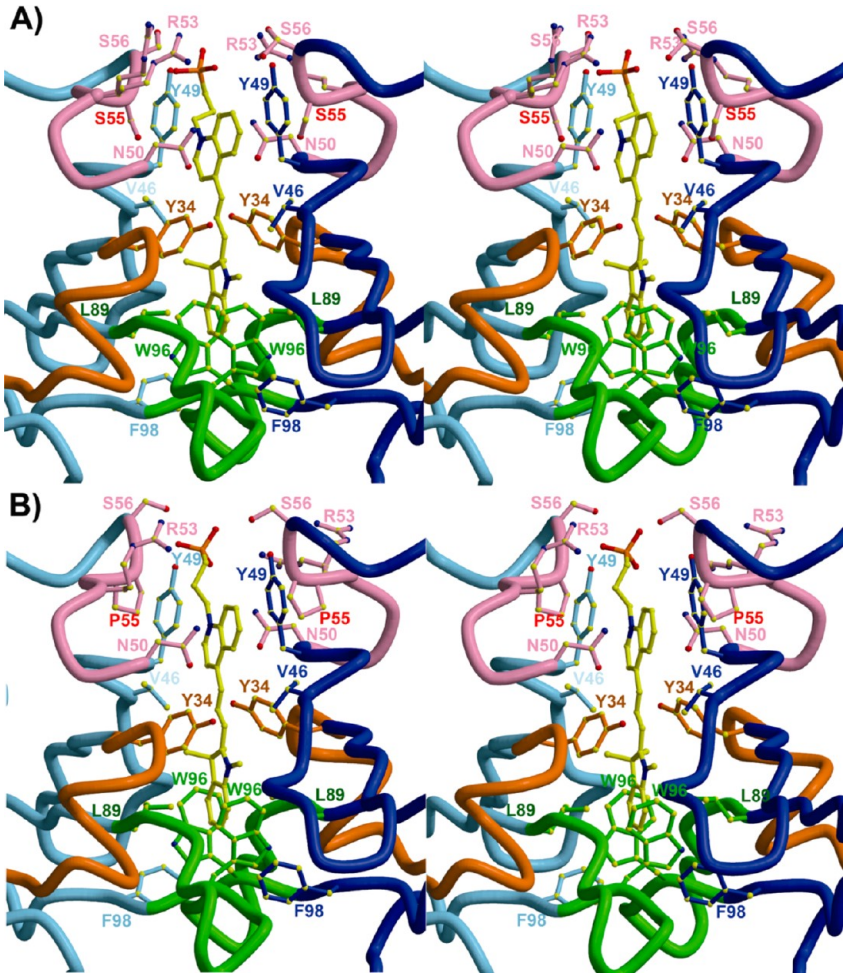


Figure 3. Environment around the DIR ligand in M8V$_L$ and M8V$_L$S^{L55}P. The two identical V$_L$ domains are colored blue and light blue. The CDR1, CDR2, and CDR3 loops are colored orange, pink, and green, respectively. (A) Stereoview of the M8V$_L$ binding site with contacting residues labeled. Serine 55 is labeled in red. (B) Stereoview of the M8V$_L$S^{L55}P binding site. Proline 55 is labeled in red. Only one DIR conformation is shown for the sake of clarity.

Table 4. Total Contacts (van der Waals and hydrogen bonds) from M8V$_L$ and M8V$_L$S^{L55}P to the DIR (alternate conformation 1/2)[a]

	M8V$_L$		M8V$_L$S^{L55}P	
residue	protein chain A	protein chain B	protein chain A	protein chain B
*Tyr^{L34}	5/10	8/5	6/6	7/7
Val^{L46}	3/1	1/2	0/0	0/4
Tyr^{L49}	8/21	16/3	10/20	6/2
*Asn^{L50}	0/2	1/0	0/0	0/0
*Arg^{L53}	0/8	2/0	0/3	0/0
*Ser/Pro^{L55}	3/0	0/3	3/0	0/7
*Ser^{L56}	5/0	0/12	4/0	0/9
*Leu^{L89}	2/2	2/4	4/4	2/2
*Trp^{L96}	4/5	5/3	5/5	5/5
Phe^{L98}	2/2	0/0	0/0	1/7
total	32/51	35/32	32/38	21/37

[a]Residues in the CDR loops are demoted with asterisks. Contacts calculated with Contacsym.[58]

M8V$_L$ and M8V$_L$S^{L55}P complexes with the DIR bury a similar amount of molecular surface upon binding the DIR, with 573 Å^2 of the protein surface and 482 Å^2 of the DIR surface buried in the M8V$_L$ complex and 581 Å^2 of the protein surface and 456 Å^2 of the DIR surface buried in the M8V$_L$S^{L55}P binding pocket (Figure 3). Contacts between the DIR and protein are all van der Waals except for hydrogen bonds between the protein and the sulfonate groups of the DIR (Figure 4). The M8V$_L$S^{L55}P V$_L$ domains do not associate as closely with DIR, with only 53 of 75 total contacts (van der Waals and hydrogen bonds, to each alternate conformation of DIR) to the DIR as opposed to 67 of 83 contacts seen for the M8V$_L$–DIR complex (Table 4). In particular, the Tyr^{L49} of the M8V$_L$–DIR complex has slightly more van der Waals (10) contacts to the DIR quinoline ring than Tyr^{L49} of the M8V$_L$S^{L55}P–DIR complex. The effect of the reduced number of van der Waals contacts in the M8V$_L$S^{L55}P–DIR structure compared to the number in the M8V$_L$–DIR structure on the binding energy was investigated by calculation of the differences in DIR binding energies between the two structures using CHARMM. In spite of the reduced number of van der Waals contacts in the M8V$_L$S^{L55}P–DIR structure, the overall contribution of van der Waals interactions to the DIR binding energy is more negative in the M8V$_L$S^{L55}P–DIR complex at −48.8 kcal/mol, while the van der Waals energy is less negative and thus less favorable at −43.5 kcal/mol for the M8V$_L$–DIR complex. Consequently, the replacement of serine with proline leads to an additional stabilization of the complex by 5.3 kcal/mol. Superimposition of the M8V$_L$ and M8V$_L$S^{L55}P

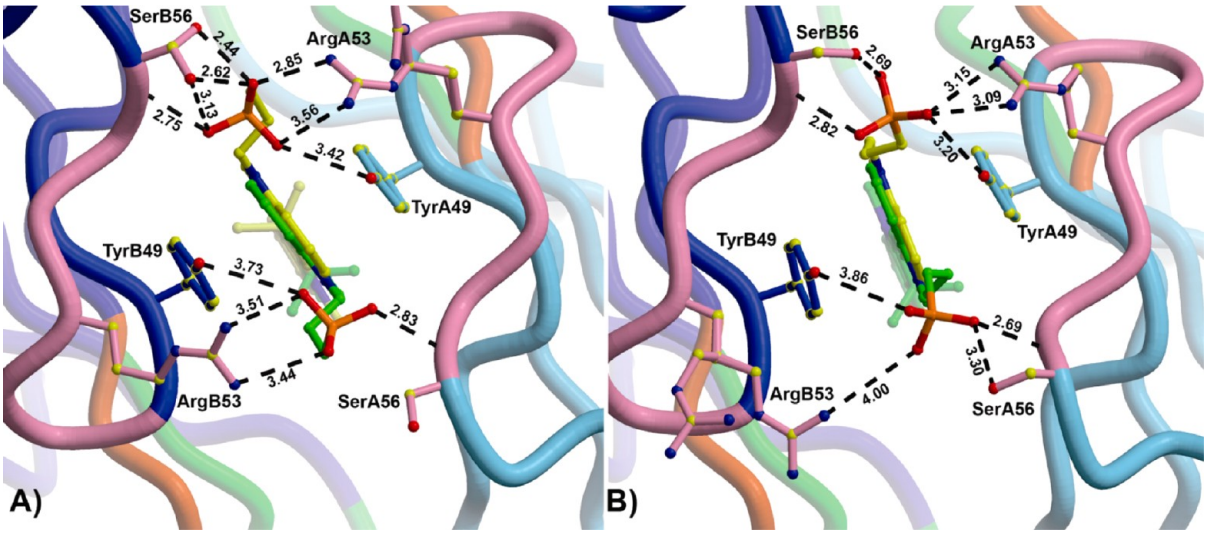


Figure 4. Sulfonate−protein interactions in M8V$_L$ and M8V$_L$S^{L55}P. (A) The interactions between the DIR sulfonate and M8V$_L$ show the two orientations of the bound DIR in the crystal structure (green and yellow), and the A (light blue) and B (blue) V$_L$ chains. CDR2 is colored pink. Ser^{B56} and Arg^{A53} also sample two alternate conformations. (B) Interactions between DIR sulfonates and M8V$_L$S^{L55}P. The indole ring of the DIR is bound in two alternate conformations, and Arg^{B53} has two alternate conformations. In both M8V$_L$ and M8V$_L$S^{L55}P, one sulfonate (top in both of these views) has more interactions with the protein than the alternate (bottom) sulfonate. Note that S^{L55} or P^{L55} is not shown in either panel.

structures reveals no significant conformational changes between the structures, despite a >10-fold decrease in the K_D of the M8V$_L$S^{L55}P FAP for the DIR compared to that of M8V$_L$. The rmsds for the C$_\alpha$ atoms of residues 1−108 are 0.43 Å for A chains, 0.46 Å for B chains, and 0.52 Å for both A and B chains (A and B chains designate the two light chains). When the A chains from each structure are superimposed, it takes an only 1.6° rotation to overlap the B chains. Thus, the M8V$_L$ and M8V$_L$S^{L55}P structures have similar tertiary and quaternary structures. Three of the four aromatic residues that surround the DIR are slightly closer in the M8V$_L$ complex than in the M8V$_L$S^{L55}P complex, with distances between the ring centers of Tyr$^{L34A-34B}$, Tyr$^{L49A-49B}$, Trp$^{L96A-96B}$, and Phe$^{L98A-98B}$ of 8.6, 7.2, 8.8, and 7.3 Å, respectively, for M8VL and 8.8, 7.4, 8.6, and 7.5 Å, respectively, for M8V$_L$S^{L55}P. This observation is in good agreement with the slightly smaller number of van der Waals contacts to the DIR for M8V$_L$S^{L55}P. Pro^{L55} is close to residue Tyr^{L49} that packs next to the DIR quinoline ring system. The Pro^{L55} residue in M8V$_L$S^{L55}P packs against Tyr^{L49} in an edge-to-face manner, with the plane of the proline approximately 4 Å from the edge of Tyr^{L49} (Figure 5), and engages in stronger van der Waals interactions with the DIR than the corresponding Ser^{L55} of M8V$_L$.

The structure of unliganded M8V$_L$ is very similar to that of the M8V$_L$−DIR complex. The rmsd for backbone atoms is 0.95 Å, with the largest difference in backbone configuration being in the region of P^{L8}, which shows a 4.1 Å displacement between the C$_\alpha$ atoms. Two of the aromatic side chains that π-stack on the bound DIR, Tyr^{L34} and Tyr^{L49}, can be superimposed in both structures (Figure 6). In contrast, Tyr^{L36} and Phe^{L98}, which also interact with the bound DIR, show large changes in side chain orientation between the unliganded form and the M8V$_L$−DIR complex (Figure 6). The X-ray-derived structure of unliganded M8V$_L$ appears to reflect the solution conformation of monomeric M8V$_L$. A comparison of 108 H$_N$−H$_N$ and H$_N$−aliphatic NOE-derived distances shows only three violations of >0.1 Å and no violations of >0.5 Å.

NMR Data Show That Dimerization of Two Monomeric M8V$_L$ Molecules Is Dependent on the DIR. Attempts to measure the oligomeric state of M8V$_L$ by size exclusion were unsuccessful because of the interaction of the protein with the column (Figure S16 of the Supporting Information). Consequently, solution NMR experiments were performed to determine the ^{15}N T_2 relaxation times in the presence and absence of the DIR to confirm that M8V$_L$ dimerization is not an artifact of crystallization (Figure 7B). The mean transverse relaxation time (T_2) for M8V$_L$ in the absence of the DIR was found to be ~95.8 ms. Similar values were obtained for M8V$_L$S^{L55}P (data not shown). The mean T_2 value for M8V$_L$ in the presence of DIR was 57.15 ms (Figure S17 of the Supporting Information), and a similar decrease was observed for M8V$_L$S^{L55}P (data not shown). To verify the oligomeric state as suggested by the T_2 relaxation time, we determined the theoretical and experimental global correlation time (τ_m) using T_1, T_2, and heteronuclear NOE data. The τ_m values for the monomer and homodimer were theoretically determined by HYDRONMR using the appropriate PDB file. The τ_m for M8V$_L$ in the absence of the DIR was 9.36 ns compared to the theoretical τ_m for M8V$_L$ in the presence of the DIR (16.38 ns). The τ_m values determined from the experimental relaxation data were 9.90 ns for M8V$_L$ in the absence of the DIR and 15.87 ns for M8V$_L$ in a complex with DIR. These experimental data are consistent with the formation of a dimer in solution when DIR is present.

The theoretically predicted τ_m for monomeric M8V$_L$ of 9.36 ns closely matches the experimental τ_m value of 9.90 ns. The experimental τ_m value for dimeric M8V$_L$ using the model-free approach is 12.37 ns. The model-free spectral density function that was used assumes a spherical shape for the molecule, while the actual M8V$_L$ structure is more consistent with a prolate ellipsoid with a 2:1 axial ratio. The experimental τ_m value for the M8V$_L$−DIR complex, when adjusted for this shape, is 15.87 ns, which is consistent with the predicted value of 16.38 ns for the homodimer.[47,48]

dx.doi.org/10.1021/bi201422g | *Biochemistry* 2012, 51, 2471−2485

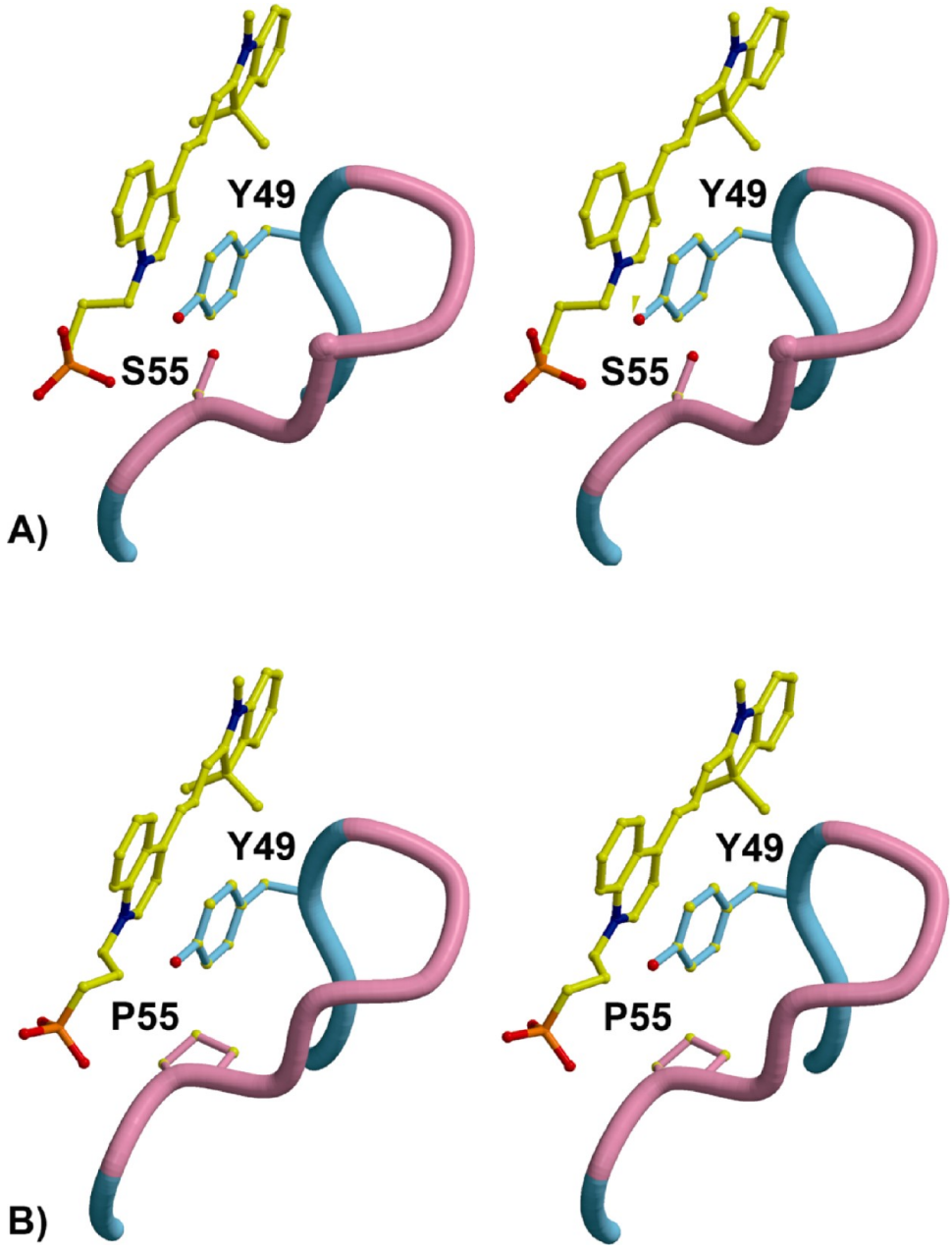


Figure 5. Close-up of the M8V$_L$S^{L55}P mutant. (A) Stereoview of M8V$_L$ and the interaction among the DIR, Tyr^{L49}, and Ser^{L55}. Ser^{L55} in M8V$_L$ has six van der Waals contacts with the DIR. (B) Same view of the M8V$_L$S^{L55}P mutant. The proline ring is approximately 4 Å from the edge of the Tyr^{L49} ring and makes 10 van der Waals contacts with the DIR.

DISCUSSION

We have described the structure–function relationship of a novel fluoromodule based on a scFv that selectively binds the environmentally sensitive fluorogen DIR. We were interested in the dramatic fluorescence activation of the DIR by the V$_L$ domain when separated from the V$_H$ domain and how it related to the fluorescence-generating function of these domains. Affinities and ϕ_f values were determined for several tandem homodimers designed to covalently link two V$_L$ domains containing differing serine-glycine linker repeats.

Comparison of Directed Evolution of V$_H$-V$_L$ M8 and M8V$_L$ FAPs. The directed evolution of V$_H$-V$_L$ M8 produced three distinct families of clones as opposed to the directed evolution of M8V$_L$ that yielded a single clone. V$_H$-V$_L$ M8 provided more genetic material for directed evolution, and more variants were isolated, underscoring the utility of the directed evolution approach in selecting desired protein properties without previous structural information.[8] These results also suggest that there are limitations to the usefulness of rational design approaches, which may fail to account for unusual protein conformations, protein–protein interactions, or ligand interactions.

The finding that V$_H$-V$_L$ M8 did not activate DIR when purified as a soluble protein led us to speculate that interactions may be occurring between neighboring FAPs on the surface of the yeast cell. We suggest that the tight affinity of yeast surface-displayed V$_H$-V$_L$ M8 for the DIR is due primarily to the proximity of two nearby V$_L$ domains as a result of the pPNL6 yeast surface display system. This proximity would allow two nearby V$_L$ domains from two V$_H$-V$_L$ chains to readily dimerize

dx.doi.org/10.1021/bi201422g | *Biochemistry* 2012, 51, 2471–2485

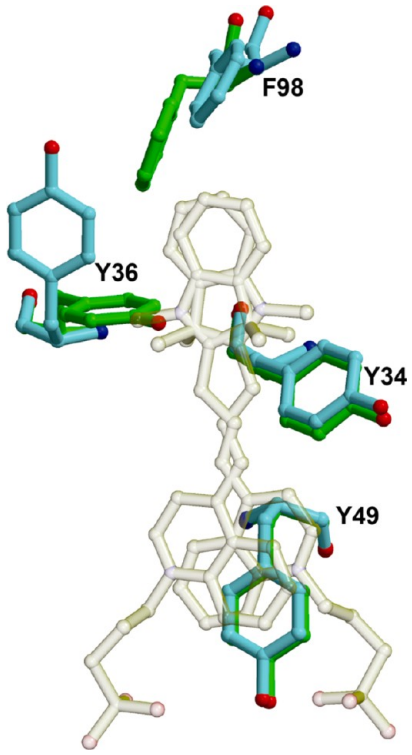


Figure 6. Comparison of the DIR binding site in unliganded M8V_L and the M8V_L–DIR complex. The bound DIR is transparent to allow for a better view of the protein side chains. Side chains from Tyr^{L34}, Tyr^{L36}, Tyr^{L49}, and Phe^{L98} are shown for the unliganded (green) and liganded (light blue) complexes.

after addition of the DIR, as the yeast cell surface is highly studded with expressed scFv.[49,50] Homodimerization of V_L domains occurs in additional V_L domains that bind the fluorogen Malachite Green (personal communication with C. Szent-Gyorgyi).

The two directed evolution approaches indicate that yeast surface display favors the dimerization of M8V_L, regardless of the starting FAP, as both monomeric V_L domain and covalent V_L-V_L FAP clones were isolated from the directed evolution of V_H-V_L M8. The strong selective pressure of 250 pM DIR is likely responsible for the isolation of covalently linked M8V_L domains from the V_H-V_L M8 parent. Additionally, three V_L clones isolated without associated V_H domains from V_H-V_L M8-directed evolution contained cysteine substitutions in the flexible linker normally between the V_H and V_L domains. These cysteine residues might enhance dimerization on the yeast cell surface through the formation of disulfide bonds between the linker regions of nearby V_L domains. This hypothesis is difficult to test without disruption of disulfide bonds that anchor the FAP to the yeast cell surface via the linkage to the Aga1p and Aga2p proteins on the yeast cell.

The fact that directed evolution of M8V_L would produce a single clonal population was unexpected. This result is an indication that either M8V_L required less optimization to increase affinity and ϕ_f than V_H-V_L M8 or, when displayed as a single light domain, M8V_L has more stringent structural requirements. In the latter case, single-amino acid changes in the V_L domain have to be accommodated across the V_L-V_L homodimer to allow the tight interaction of both the protein and dye component of the fluoromodule.

Yeast Surface Display Analysis of Two-Domain FAPs. Two clones isolated from the directed evolution of V_H-V_L M8 underscore the novel sequence variations that these FAPs can undergo to cause enhancement of the fluorescence of the DIR. Clone V_H-V_L L9 that resulted from V_H-V_L M8 enrichment retained both V_H-V_L domains but contained several mutations only in the V_H domain ($F^{H29}S$, $W^{H36}C$, $A^{H93}V$, and $I^{H113i}T$). The increased fluorescence enhancement of this clone could be a result of V_L dimer formation that is enhanced by the $W^{H36}C$ mutation. This mutation might allow the formation of a disulfide bridge between neighboring V_H domains on the yeast cell surface, facilitating greater proximity of their connected V_L domains. The V_L-V_L J8 tandem V_L dimer isolated from directed evolution of V_H-V_L M8 covalently links two M8V_L domains, resulting in a substantial decrease in K_D that is comparable to that of the synthetically constructed homodimeric M8$V_L$$(G_4S)_3$. It is noteworthy that V_L-V_L J8 also contains residue insertions prior to the (G_4S) linker region that may impart increased affinity or ϕ_f. We hypothesize that residue insertions that originated from the V_H domain act to extend the linker length and provide a more suitable dimer structure for the activation of the DIR. However, the additional amino acid mutations present in both light domains of V_L-V_L J8 might have an additional effect on the affinity of V_L-V_L J8 for the DIR.

The isolation of the V_L-V_L J8 tandem V_L homodimer was novel and unexpected, which led to the investigation of engineered tandem M8V_L dimers with different linker lengths. The additional amino acids prior to the linker region in V_L-V_L J8, as well as measurements from the crystal structure, suggested that a six-repeat (G_4S) linker would have the most

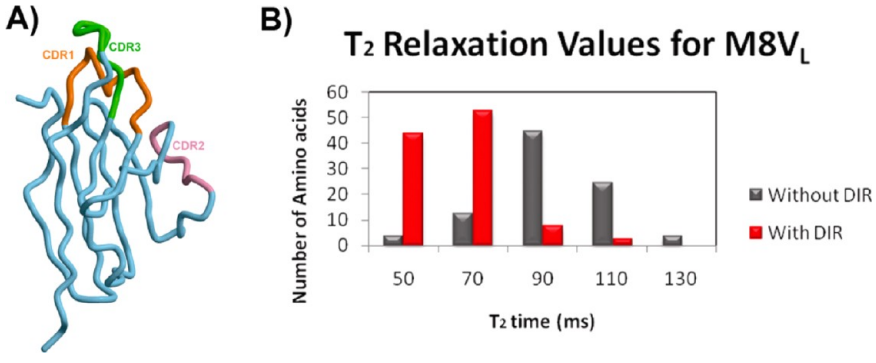


Figure 7. Crystal structure and solution phase ^{15}N relaxation data for monomeric M8V_L. (A) Overall crystal structure of M8V_L in the absence of the DIR. This V_L domain is monomeric in the crystal structure. (B) Mean NMR T_2 relaxation times in the absence (gray bars) or presence (red bars) of DIR.

dx.doi.org/10.1021/bi201422g | *Biochemistry* 2012, 51, 2471–2485

relaxed structural constraints for the DIR. The yeast surface-displayed tandem dimers showed no significant difference in DIR affinity regardless of linker length or inclusion of the $S^{L55}P$ mutation of $M8V_L$. Thus, the linker length between two tandem dimers on yeast surface display does not alter the affinity of the tandem dimers for the DIR; however, linking the two monomers together does increase the affinity of the soluble proteins. These data indicate that the yeast surface display platform may allow V_L-V_L interactions of independently displayed V_L domains.

Differences in Fluorescence Activity for Yeast Surface Display versus Solution. A consistent observation is that affinities and fluorescence enhancement of fluorgen activating proteins on the surface of yeast can be markedly different than those obtained for the same protein free in solution. In particular, K_D values for the homodimeric V_L-V_L proteins in solution were consistently lower than the affinity found on the surface of the yeast cell. For example, the K_D value for $M8V_L(G_4S)_3$ was 1.7 nM on the surface but substantially less than 1 nM in solution. Similarly, J8 exhibited a surface K_D of ~10 nM, yet the K_D in solution was also less than 1 nM. One possible explanation for this effect is that the high density of FAPs on the surface of the protein can result in novel protein−protein interactions. An extreme example of this is the discovery that the soluble V_H-V_L M8 protein failed to activate DIR while showing a strong fluorescence enhancement on the yeast surface. This result suggests that in solution the V_H domain may have a high affinity for the attached V_L domain, thus preventing V_L-V_L homodimerization that likely occurs on the surface of the yeast cell. Alternatively, soluble V_H-V_L M8 may have a favorable affinity for the DIR but fails to restrain the DIR in an appropriate conformation for fluorescence emission, similar to what is predicted for other fluorogenic cyanine dyes.[51]

Relationship between Mutations and Dimerization with Respect to Affinity and Quantum Yield. The soluble Q9 protein has a strong affinity for the DIR, but a low ϕ_f of 19%. The numerous mutations isolated in Q9 ($Q^{L1}R$, $T^{L14}I$, $D^{L30}G$, $Q^{L37}R$, $S^{L55}P$, $F^{L62}L$, and $S^{L80}P$) may be responsible for the decrease in ϕ_f, although this effect may be combinatorial. Both $M8V_LS^{L55}P$ and Q9 contain the $S^{L55}P$ mutation that may cause a decrease in ϕ_f and potentially improve the respective affinities. However, the ϕ_f of Q9 is significantly lower than that of either $M8V_L$ or $M8V_LS^{L55}P$. Thus, Q9 and $M8V_LS^{L55}P$ suggest that directed evolution of FAP complexes with the DIR does not necessarily yield simultaneous increases in both affinity and ϕ_f. Further characterization of FAPs selected for other fluorogens will determine if decreased ϕ_f values coupled with enhanced affinity can occur among FAPs subjected to directed evolution.

The linking of $M8V_L$ and $M8V_LS^{L55}P$ by tandem $(G_4S)_3$ linkers generated FAPs with high affinity for DIR, as expected by reducing the entropic penalty for bringing two protein domains together with a covalent linkage. Although the affinity increased, the quantum yield decreased for both $M8V_L$ and $M8V_LS^{L55}P$. Alternatively, the quantum yields of the extended dimers are larger, and thus, they are more efficient at emitting fluorescence than the tandem $(G_4S)_3M8V_L$. Thus, the affinity and ϕ_f do not simultaneously increase for the different linker constructions of the $M8V_L$ tandem homodimers. These data taken together indicate that the conformational requirements for the increased affinity of the fluoromodule do not produce a robust ϕ_f.

X-ray Crystallography and NMR Data Support the Homodimerization of Two $M8V_L$ Domains in the Presence of the DIR and Elucidate the Mechanism for Fluoromodule Fluorescence Generation. The genetic and structural data support the conclusion that $M8V_L$ forms a homodimer in the presence of the DIR. The crystal structures of both $M8V_L$ and $M8V_LS^{L55}P$ reveal a dimer of V_L domains sandwiching the DIR and constraining the rotation of the two DIR heterocycle rings, so that the quinoline and indole rings are essentially planar. This nearly planar orientation may be optimal for fluorescence decay as it allows the DIR to emit fluorescence as it relaxes from the excited state to the ground state while the heterocycles are fixed around the conjugated polymethine bridge.[52]

NMR experiments are consistent with both $M8V_L$ and $M8V_LS^{L55}P$ forming a homodimer only in the presence of DIR. On the basis of the canonical scFv interaction between V_H and V_L domains, it would be expected that the domains would have a low micromolar K_D for each other;[53] however, even at ~0.5 mM $M8V_L$ and $M8V_LS^{L55}P$ in the NMR experiments, there is no evidence of a strong interaction between light domains. The average transverse relaxation time, T_2, for monomeric $M8V_L$ was approximately double that of the $M8V_L$ homodimer in the presence of DIR. This result is consistent with the homodimerization of $M8V_L$ only in the presence of the DIR.

The isolation of functional V_L domains that possess strong affinity for small molecule fluorogens is intriguing and shows similarities to previous reports describing Bence-Jones proteins isolated from multiple myeloma patients.[54] Bence-Jones proteins are immunoglobulin variable light chains that form homodimers. It has been demonstrated that small molecule haptens in solution penetrate into preformed crystals of Bence-Jones proteins and the haptens interact with the hydrophobic regions formed between two variable light domains.[55,56] Although the FAPs described here are distinct from Bence-Jones proteins as they lack a constant light chain domain, they display low nanomolar affinity for fluorogenic compounds.

$M8V_LS^{L55}P$ displays a >10-fold enhancement in affinity for DIR compared to that of wild-type $M8V_L$, despite having fewer van der Waals contacts with the DIR (Tables 1 and 4). Calculations indicate that the complex between $M8V_LS^{L55}P$ and the DIR is stabilized by stronger van der Waals contacts. In the absence of detailed solution phase structural and dynamic data, we cannot conclude precisely how the $S^{L55}P$ mutation increases the affinity for DIR. One hypothesis is that the serine to proline mutation imparts more rigidity to the $M8V_LS^{L55}P$ protein and ultimately creates a less flexible dimer binding pocket for retaining the DIR. This increased protein rigidity may reduce entropic changes during binding, thus modulating the affinity. These interactions will be the focus of future studies to determine the energetics and solution dynamics that underlie the changes in the affinity of these FAPs with the DIR.

Both $M8V_L$ and $M8V_LS^{L55}P$ significantly enhance the fluorescence of the DIR, likely by constraining the angle between the planes of the indole and quinoline groups to <10°. Previous theoretical predictions of fluorescence generation for torsionally responsive fluorogens such as thiazole orange indicate that optimal radiative decay occurs when the heterocyclic groups are restricted to an interplanar angle of 0−60°.[51] There is no corresponding model for the DIR. However, our data indicate that the ϕ_f is robust when DIR is held in a nearly planar conformation. The π-stacking interactions between tyrosines in $M8V_L$ and $M8V_LS^{L55}P$ with the DIR may also

stabilize the fluorescence signal from the DIR based on studies of unsymmetrical cyanine dyes that π-stack during DNA intercalation.[6,57] The structural data suggest that conserved π-stacking interactions with Tyr^{L34} and Tyr^{L49} are important for fluorescence activation of both the $M8V_L$ and $M8V_LS^{L55}P$ homodimers and will be the focus of future studies.

The ϕ_f of $M8V_LS^{L55}P$ is lower than that of $M8V_L$. The $S^{L55}P$ mutation in $M8V_L$ may contribute additional effects to the fluorescence of DIR that are unrelated to a global increase in protein rigidity. For example, direct interactions between the proline and the DIR in solution may only weakly constrain the conformation of DIR, thereby reducing the ϕ_f. The reduced quantum yield associated with other FAPs, Q9, for example, may be due to restriction of the planar groups in the DIR at an angle greater than what is optimal for fluorescence emission. Finally, FAPs may show reduced quantum yields because of quenching of the DIR in the excited state caused by electron transfer involving nearby redox-active side chains. Further structural studies of other DIR binding FAPs will help to clarify the interactions required for ϕ_f.

CONCLUSIONS

Fluoromodules consisting of a specific fluorescence activating protein and an environmentally sensitive fluorogen demonstrate unique properties such as homodimerization induced by the fluorogen DIR. Experimental data show that the DIR fluoresces when it is rigidly held between two immunoglobulin variable light domains that dimerize in the presence of the DIR. The structural data suggest that conserved π-stacking interactions with tyrosine residues are important for fluorescence activation of both $M8V_L$ and $M8V_LS^{L55}P$ homodimers. $M8V_LS^{L55}P$ holds the DIR in a potentially less flexible binding pocket compared to that of $M8V_L$. This change is caused by a single serine to proline mutation that also alters ϕ_f and affinity. Our results demonstrate the challenge of predicting enhanced FAP fluorescence via mutagenic approaches in the absence of directed evolution. Linker scanning, for example, may provide information about tolerated structural changes to a particular protein but fail to successfully constrain a fluorogen to produce fluorescence. Furthermore, it is reasonable that even with high-resolution crystal structures and subsequent rational design we may not correctly predict the mutations that increase both DIR affinity and ϕ_f, given that a single mutation in $M8V_LS^{L55}P$ enhanced binding but not ϕ_f. However, combining detailed structural information with computational or rational design approaches may yield significant improvements in the fluorescence quantum yield and affinity. For example, performing targeted mutagenesis on the residues that surround the fluorogen in the binding pocket prior to enrichment might increase the probability of obtaining mutants with improved fluorescence. Such an approach may also include computational modeling of the altered protein in addition to random mutagenesis to enhance the FAP complex.

ASSOCIATED CONTENT

Supporting Information

Sequence information, absorbance spectra, and details of crystallography data. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

X-ray structure coordinates and structure factors for $M8V_L$, the $M8V_L$–DIR complex, and the $M8V_LS^{L55}P$–DIR complex have been deposited as Protein Data Bank entries 3T0V, 3T0W, and 3T0X, respectively.

AUTHOR INFORMATION

Corresponding Author

*Address: 4400 5th Ave., Carnegie Mellon University, Mellon Institute, Pittsburgh, PA 15213. E-mail: rule@andrew.cmu.edu. Fax: (412) 268-7129. Phone: (412) 268-1839.

Present Address

[@]600 S. 43rd St., McNeil Science & Technology Center, STC 342, University of the Sciences, Philadelphia, PA 19104.

Funding

This work was supported by National Institutes of Health Grant U54 RR022241 and Howard Hughes Medical Institute Grant 52005865.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank Yehuda Creeger for assistance with FACS, Andreas Plückthun for the gift of the pAK-400 plasmid, Gloria Silva and Nathaniel Shank for the DIR, Josef Franke and Cheryl Telmer for critical review of the manuscript, students in the Howard Hughes Medical Institute-supported summer research institute for preliminary NMR data, Paul Barton for help in designing the synthetic $M8V_L$ gene, V. Simplaceanu for help in acquiring the NMR relaxation data, and Dr. P. R. Gooley for help with the analysis of the relaxation data. Portions of this work were conducted at GM/CA CAT at the Advanced Photon Source and the Molecular Biology Consortium at the Advanced Light Source. GM/CA CAT has been funded, in whole or in part, by the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Sciences (Y1-GM-1104). Use of the Advanced Photon Source was supported by the U.S. Department of Energy (DE-AC02-06CH1135). The Advanced Light Source is supported by U.S. Department of Energy Contract DE-AC02-05CH11231.

ABBREVIATIONS

FAP, fluorogen activating protein; DIR, dimethylindole red; K_D, dissociation constant; ϕ_f, fluorescence quantum yield; V_H, immunoglobulin variable heavy domain; V_L, immunoglobulin variable light domain; GFP, green fluorescence protein; FACS, fluorescence-activated cell sorting; CDR, complementarity-determining region; NOE, nuclear Overhauser effect; rmsd, root-mean-square deviation; τ_m, global rotational correlation time; T_2, transverse relaxation time.

REFERENCES

(1) Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994) Green fluorescent protein as a marker for gene expression. *Science 263*, 802–805.

(2) Szent-Gyorgyi, C., Schmidt, B. F., Creeger, Y., Fisher, G. W., Zakel, K. L., Adler, S., Fitzpatrick, J. A., Woolford, C. A., Yan, Q., Vasilev, K. V., Berget, P. B., Bruchez, M. P., Jarvik, J. W., and Waggoner, A. (2008) Fluorogen-activating single-chain antibodies for imaging cell surface proteins. *Nat. Biotechnol. 26*, 235–240.

(3) Ozhalici-Unal, H., Pow, C. L., Marks, S. A., Jesper, L. D., Silva, G. L., Shank, N. I., Jones, E. W., Burnette, J. M. III, Berget, P. B., and Armitage, B. A. (2008) A rainbow of fluoromodules: A promiscuous scFv protein binds to and activates a diverse set of fluorogenic cyanine dyes. *J. Am. Chem. Soc. 130*, 12620–12621.

(4) Zanotti, K. J., Silva, G. L., Creeger, Y., Robertson, K. L., Waggoner, A. S., Berget, P. B., and Armitage, B. A. (2011) Blue fluorescent dye-protein complexes based on fluorogenic cyanine dyes and single chain antibody fragments. *Org. Biomol. Chem. 9*, 1012−1020.
(5) Shank, N. I., Zanotti, K. J., Lanni, F., Berget, P. B., and Armitage, B. A. (2009) Enhanced photostability of genetically encodable fluoromodules based on fluorogenic cyanine dyes and a promiscuous protein partner. *J. Am. Chem. Soc. 131*, 12960−12969.
(6) Lee, L. G., Chen, C. H., and Chiu, L. A. (1986) Thiazole orange: A new dye for reticulocyte analysis. *Cytometry 7*, 508−517.
(7) Alfthan, K., Takkinen, K., Sizmann, D., Soderlund, H., and Teeri, T. T. (1995) Properties of a single-chain antibody containing different linker peptides. *Protein Eng. 8*, 725−731.
(8) Colby, D. W., Kellogg, B. A., Graff, C. P., Yeung, Y. A., Swers, J. S., and Wittrup, K. D. (2004) Engineering antibody affinity by yeast surface display. *Methods Enzymol. 388*, 348−358.
(9) Feldhaus, M. J., Siegel, R. W., Opresko, L. K., Coleman, J. R., Feldhaus, J. M., Yeung, Y. A., Cochran, J. R., Heinzelman, P., Colby, D., Swers, J., Graff, C., Wiley, H. S., and Wittrup, K. D. (2003) Flow-cytometric isolation of human antibodies from a nonimmune *Saccharomyces cerevisiae* surface display library. *Nat. Biotechnol. 21*, 163−170.
(10) Falco, C. N., Dykstra, K. M., Yates, B. P., and Berget, P. B. (2009) scFv-based fluorogen activating proteins and variable domain inhibitors as fluorescent biosensor platforms. *Biotechnol. J. 4*, 1328−1336.
(11) Chao, G., Lau, W. L., Hackel, B. J., Sazinsky, S. L., Lippow, S. M., and Wittrup, K. D. (2006) Isolating and engineering human antibodies using yeast surface display. *Nat. Protoc. 1*, 755−768.
(12) Boder, E. T., Midelfort, K. S., and Wittrup, K. D. (2000) Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. *Proc. Natl. Acad. Sci. U.S.A. 97*, 10701−10705.
(13) Mack, E. T., Perez-Castillejos, R., Suo, Z., and Whitesides, G. M. (2008) Exact analysis of ligand-induced dimerization of monomeric receptors. *Anal. Chem. 80*, 5550−5555.
(14) Southwick, P. L., Ernst, L. A., Tauriello, E. W., Parker, S. R., Mujumdar, R. B., Mujumdar, S. R., Clever, H. A., and Waggoner, A. S. (1990) Cyanine dye labeling reagents: Carboxymethylindocyanine succinimidyl esters. *Cytometry 11*, 418−430.
(15) Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expression Purif. 41*, 207−234.
(16) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276A*, 307−326.
(17) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658−674.
(18) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. 66*, 486−501.
(19) Winn, M. D., Murshudov, G. N., and Papiz, M. Z. (2003) Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol. 374*, 300−321.
(20) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. 66*, 213−221.
(21) Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S., and Foeller, C. (1991) *Sequences of proteins of immunological interest*, Vol. *1*, 5th ed., U.S. Department of Health and Human Services, Washington, DC.
(22) Hansen, C. L., Skordalakes, E., Berger, J. M., and Quake, S. R. (2002) A robust and scalable microfluidic metering method that allows protein crystal growth by free interface diffusion. *Proc. Natl. Acad. Sci. U.S.A. 99*, 16531−16536.
(23) Golovanov, A. P., Hautbergue, G. M., Wilson, S. A., and Lian, L. Y. (2004) A simple method for improving protein solubility and long-term stability. *J. Am. Chem. Soc. 126*, 8933−8939.
(24) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277−293.
(25) Farrow, N. A., Zhang, O., Forman-Kay, J. D., and Kay, L. E. (1997) Characterization of the backbone dynamics of folded and denatured states of an SH3 domain. *Biochemistry 36*, 2390−2402.
(26) Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol. 278*, 313−352.
(27) d'Auvergne, E. J., and Gooley, P. R. (2003) The use of model selection in the model-free analysis of protein dynamics. *J. Biomol. NMR 25*, 25−39.
(28) d'Auvergne, E. J., and Gooley, P. R. (2006) Model-free model elimination: A new step in the model-free dynamic analysis of NMR relaxation data. *J. Biomol. NMR 35*, 117−135.
(29) d'Auvergne, E. J., and Gooley, P. R. (2007) Set theory formulation of the model-free problem and the diffusion seeded model-free paradigm. *Mol. BioSyst. 3*, 483−494.
(30) d'Auvergne, E. J., and Gooley, P. R. (2008) Optimisation of NMR dynamic models II. A new methodology for the dual optimization of the model-free parameters and the Brownian rotational diffusion tensor. *J. Biomol. NMR 40*, 121−133.
(31) d'Auvergne, E. J., and Gooley, P. R. (2008) Optimisation of NMR dynamic models. I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. *J. Biomol. NMR 40*, 107−119.
(32) Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. II. Analysis of experimental results. *J. Am. Chem. Soc. 104*, 4559−4570.
(33) Lipari, G., and Szabo, A. (1982) Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. I. Theory and range of validity. *J. Am. Chem. Soc. 104*, 4546−4559.
(34) Clore, G. M., et al. (1990) Deviations from the simple 2-parameter model-free approach to the interpretation of N-15 nuclear magnetic-relaxation of proteins. *J. Am. Chem. Soc. 112*, 4989−4991.
(35) Garcia de la Torre, J., Navarro, S., Lopez Martinez, M. C., Diaz, F. G., and Lopez Cascales, J. J. (1994) HYDRO: A computer program for the prediction of hydrodynamic properties of macromolecules. *Biophys. J. 67*, 530−531.
(36) Ikura, M., Kay, L. E., and Bax, A. (1990) A novel approach for sequential assignment of ^{1}H, ^{13}C, and ^{15}N spectra of proteins: Heteronuclear triple-resonance three-dimensional NMR spectroscopy. Application to calmodulin. *Biochemistry 29*, 4659−4667.
(37) Clubb, R. T., Thanabal, V., and Wagner, G. (1992) A new 3D HN(CA)HA experiment for obtaining fingerprint HN-Hα peaks in ^{15}N- and ^{13}C-labeled proteins. *J. Biomol. NMR 2*, 203−210.
(38) Hitchens, T. K., Lukin, J. A., Zhan, Y., McCallum, S. A., and Rule, G. S. (2003) MONTE: An automated Monte Carlo based approach to nuclear magnetic resonance assignment of proteins. *J. Biomol. NMR 25*, 1−9.
(39) Briercheck, D. M., Wood, T. C., Allison, T. J., Richardson, J. P., and Rule, G. S. (1998) The NMR structure of the RNA binding domain of *E. coli* rho factor suggests possible RNA-protein interactions. *Nat. Struct. Biol. 5*, 393−399.
(40) Brooks, B. R., Brooks, C. L. III, Mackerell, A. D. Jr., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., and Karplus, M. (2009) CHARMM: The biomolecular simulation program. *J. Comput. Chem. 30*, 1545−1614.
(41) Atwell, J. L., Breheney, K. A., Lawrence, L. J., McCoy, A. J., Kortt, A. A., and Hudson, P. J. (1999) scFv multimers of the

anti-neuraminidase antibody NC10: Length of the linker between VH and VL domains dictates precisely the transition between diabodies and triabodies. *Protein Eng. 12*, 597−604.
(42) Albrecht, H., Denardo, G. L., and Denardo, S. J. (2006) Monospecific bivalent scFv-SH: Effects of linker length and location of an engineered cysteine on production, antigen binding activity and free SH accessibility. *J. Immunol. Methods 310*, 100−116.
(43) Rakestraw, J. A., Sazinsky, S. L., Piatesi, A., Antipov, E., and Wittrup, K. D. (2009) Directed evolution of a secretory leader for the improved expression of heterologous proteins and full-length antibodies in *Saccharomyces cerevisiae*. *Biotechnol. Bioeng. 103*, 1192−1201.
(44) Shusta, E. V., Kieke, M. C., Parke, E., Kranz, D. M., and Wittrup, K. D. (1999) Yeast polypeptide fusion surface display levels predict thermal stability and soluble secretion efficiency. *J. Mol. Biol. 292*, 949−956.
(45) Al-Lazikani, B., Lesk, A. M., and Chothia, C. (1997) Standard conformations for the canonical structures of immunoglobulins. *J. Mol. Biol. 273*, 927−948.
(46) Martin, A. C., and Thornton, J. M. (1996) Structural families in loops of homologous proteins: Automatic classification, modelling and application to antibodies. *J. Mol. Biol. 263*, 800−815.
(47) Weber, G. (1952) Polarization of the fluorescence of macromolecules. II. Fluorescent conjugates of ovalbumin and bovine serum albumin. *Biochem. J. 51*, 155−167.
(48) Lakowicz, J. (1999) *Principles of Fluorescence Spectroscopy*, 2nd ed., Springer, New York.
(49) Boder, E. T., and Wittrup, K. D. (1997) Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotechnol. 15*, 553−557.
(50) Boder, E. T., and Wittrup, K. D. (2000) Yeast surface display for directed evolution of protein expression, affinity, and stability. *Methods Enzymol. 328*, 430−444.
(51) Silva, G. L., Ediz, V., Yaron, D., and Armitage, B. A. (2007) Experimental and computational investigation of unsymmetrical cyanine dyes: Understanding torsionally responsive fluorogenic dyes. *J. Am. Chem. Soc. 129*, 5710−5718.
(52) Constantin, T. P., Silva, G. L., Robertson, K. L., Hamilton, T. P., Fague, K., Waggoner, A. S., and Armitage, B. A. (2008) Synthesis of new fluorogenic cyanine dyes and incorporation into RNA fluoromodules. *Org. Lett. 10*, 1561−1564.
(53) Mallender, W. D., Carrero, J., and Voss, E. W. Jr. (1996) Comparative properties of the single chain antibody and Fv derivatives of mAb 4-4-20. Relationship between interdomain interactions and the high affinity for fluorescein ligand. *J. Biol. Chem. 271*, 5338−5346.
(54) Epp, O., Lattman, E. E., Schiffer, M., Huber, R., and Palm, W. (1975) The molecular structure of a dimer composed of the variable portions of the Bence-Jones protein REI refined at 2.0-Å resolution. *Biochemistry 14*, 4943−4952.
(55) Edmundson, A. B., Ely, K. R., Girling, R. L., Abola, E. E., Schiffer, M., Westholm, F. A., Fausch, M. D., and Deutsch, H. F. (1974) Binding of 2,4-dinitrophenyl compounds and other small molecules to a crystalline λ-type Bence-Jones dimer. *Biochemistry 13*, 3816−3827.
(56) Schiffer, M., Girling, R. L., Ely, K. R., and Edmundson, A. B. (1973) Structure of a λ-type Bence-Jones protein at 3.5-Å resolution. *Biochemistry 12*, 4620−4631.
(57) Rye, H. S., Yue, S., Wemmer, D. E., Quesada, M. A., Haugland, R. P., Mathies, R. A., and Glazer, A. N. (1992) Stable fluorescent complexes of double-stranded DNA with bis-intercalating asymmetric cyanine dyes: Properties and applications. *Nucleic Acids Res. 20*, 2803−2812.
(58) Sheriff, S., Hendrickson, W. A., and Smith, J. L. (1987) Structure of myohemerythrin in the azidomet state at 1.7/1.3 Å resolution. *J. Mol. Biol. 197*, 273−296.
(59) Kraulis, P. (1991) MOLSCRIPT: A program to produce both detailed and schematic plots of protein strucutres. *J. Appl. Crystallogr. 24*, 846−950.
(60) Esnouf, R. M. (1999) Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. *Acta Crystallogr. 55*, 938−940.