Supplementary Material

for

Generation of fluorogen-activating designed ankyrin repeat proteins (FADAs) as versatile sensor tools.

Marco Schütz [a, 1], Alexander Batyuk[a], Christoph Klenk[a], Lutz Kummer[a, 1], Seymour de Picciotto[b], Basri Gülbakan[c, 2], Yufan Wu[a], Gregory A. Newby[a, 3], Franziska Zosel[a], Jendrik Schöppe[a], Erik Sedlák[a, 4], Peer R. E. Mittl[a], Renato Zenobi[c], K. Dane Wittrup[b], and Andreas Plückthun[a].

[a]Department of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland.

[b]Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

[c]Department of Chemistry and Applied Biosciences, ETH Zurich, CH-8093 Zurich, Switzerland.

[1]Present address: G7 Therapeutics AG, Grabenstrasse 11a, CH-8952 Schlieren, Switzerland.

[2]Present address: Institute of Child Health, Hacettepe University, 06100 Ankara, Turkey.

[3]Present address: Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA.

[4]Present address: Centre for Interdisciplinary Biosciences and Department of Biochemistry, Pavol Jozef Šafárik University in Košice, 040 01 Košice, Slovakia.

Supplementary Methods

Fluorescence-activation measurements with 2×FADA-3210 in solution

The 2×FADA-3210 construct was generated by a genetic fusion of two FADA-3210 molecules via a flexible $(G_4S)_6$-linker. Expression of 2×FADA-3210 was carried out with plasmid pDST67 [1] in the *E. coli* strain XL1-Blue (Agilent Technologies, 200249). Pre-cultures were grown in 40 mL 2×YT medium supplemented with 10 g/L glucose and 50 µg/mL ampicillin. For expression 1 L 2×YT medium supplemented with 50 µg/mL ampicillin was inoculated from the pre-culture to OD_{600} = 0.1, growth was continued at 37°C until the culture reached OD_{600} = 0.5 – 0.7, and 0.5 mM IPTG was added with subsequent cultivation at 25°C for 24 h. Then the cultures were harvested by centrifugation, the cell pellets were resuspended once in 25 mL cold TBS_{400} buffer (50 mM Tris-HCl pH 7.4 at 4°C, 400 mM NaCl), centrifuged again, and the pellets were frozen in liquid nitrogen and stored at –80°C until use. Purifications of 2×FADA-3210 and subsequent fluorescence-activation experiments in solution were performed as described in the Materials and Methods of the main text.

Fluorescence-activation titration experiments and binding models

Fluorescence-activation titration experiments were performed identically to the fluorescence-activation assays described in the Materials and Methods of the main text with the exception that concentrations of FADA-3210 or 2×FADA-3210 were varied by serial dilutions from 5 µM to 0.45 nM while the MG-2p concentration was kept constant. Two different experiments with different constant MG-2p concentrations (20 nM and 10 nM) were performed. Since the affinity of the ligand MG-2p is high, the simplification $[L] \approx [L_{tot}]$ cannot be made.

The 1:1 binding model is described by *eq. 1* with $[M]$ being the concentration of free monomeric DARPin, $[L]$ being the concentration of free MG-2p ligand, and $[ML]$ being the concentration of the monomeric DARPin/MG-2p complex (one DARPin binding one MG-2p molecule). Thus, the dissociation constant is defined by *eq. 2*. From mass conservation, *eq. 3* and *eq. 4* follow, with $[M_{tot}]$ and $[L_{tot}]$ being the total concentration of monomeric DARPin and MG-2p in the reaction, respectively. Combining *eq. 2 – 4* gives *eq. 5*, which can be rearranged into the quadratic equation *eq. 6* for $[ML]$. The analytical solution of this quadratic equation is given in *eq. 7*. A scaling factor converting $[ML]$ to fluorescence intensity was introduced into *eq. 7*, which was then used as the fitting equation to determine the dissociation constant by least-squares fitting to experimental data.

$$[M] + [L] \rightleftharpoons [ML] \qquad \textit{(eq. 1)}$$

$$K_d = \frac{[M][L]}{[ML]} \qquad \textit{(eq. 2)}$$

$$[M_{tot}] = [M] + [ML] \qquad \textit{(eq. 3)}$$

$$[L_{tot}] = [L] + [ML] \qquad \textit{(eq. 4)}$$

$$K_d = \frac{([M_{tot}] - [ML])([L_{tot}] - [ML])}{[ML]} \quad \textit{(eq. 5)}$$

$$0 = [ML]^2 - ([M_{tot}] + [L_{tot}] + K_d)[ML] + [M_{tot}][L_{tot}] \quad \textit{(eq. 6)}$$

$$[ML] = \frac{[M_{tot}] + [L_{tot}] + K_d}{2} - \sqrt{\left(\frac{[M_{tot}] + [L_{tot}] + K_d}{2}\right)^2 - [M_{tot}][L_{tot}]} \quad \textit{(eq. 7)}$$

The 2:1 binding model is described by *eq. 8* with $[M]$ being the concentration of monomeric DARPin, $[L]$ being the concentration of ligand MG-2p, and $[DL]$ being the concentration of the dimeric DARPin/MG-2p complex (two DARPins binding one MG-2p molecule). Suppl. Fig. S9 shows an overview of the possible whole reaction pathways with the corresponding dissociation constants for each microscopic reaction. However, a model that neglects any intermediate species fits the data already well, leading to the conclusion that these species are not populated to a significant amount under equilibrium conditions (two-state cooperativity). Therefore, the microscopic pathway of the reaction cannot be determined and a dissociation constant can be only given for the overall reaction (*eq. 8*). This dissociation constant is defined in *eq. 9* and has the unit [M^2]. From mass conservation *eq. 10* and *eq. 11* follow with $[M_{tot}]$ and $[L_{tot}]$ being the total concentration of monomeric DARPin and MG-2p in the reaction, respectively. Combining *eq. 9 – 11* gives *eq. 12*, which can be rearranged into a cubic equation for $[DL]$ (*eq. 13*). The equation was analytically solved for $[DL]$ with Mathematica in the domain of real numbers. A scaling factor converting $[DL]$ to fluorescence intensity was introduced into the solution, so the experimental data could be fitted globally to determine the dissociation constant K_d.

$$[M] + [M] + [L] \rightleftharpoons [DL] \quad \textit{(eq. 8)}$$

$$K_d = \frac{[M]^2[L]}{[DL]} \quad \textit{(eq. 9)}$$

$$[M_{tot}] = [M] + 2[DL] \quad \textit{(eq. 10)}$$

$$[L_{tot}] = [L] + [DL] \quad \textit{(eq. 11)}$$

$$K_d = \frac{([M_{tot}] - 2[DL])^2([L_{tot}] - [DL])}{[DL]} \quad \textit{(eq. 12)}$$

$$0 = 4[DL]^3 - 4([M_{tot}] + [L_{tot}])[DL]^2 + (4[M_{tot}][L_{tot}] + [M_{tot}]^2 + K_d)[DL] - [M_{tot}]^2[L_{tot}] \quad \textit{(eq. 13)}$$

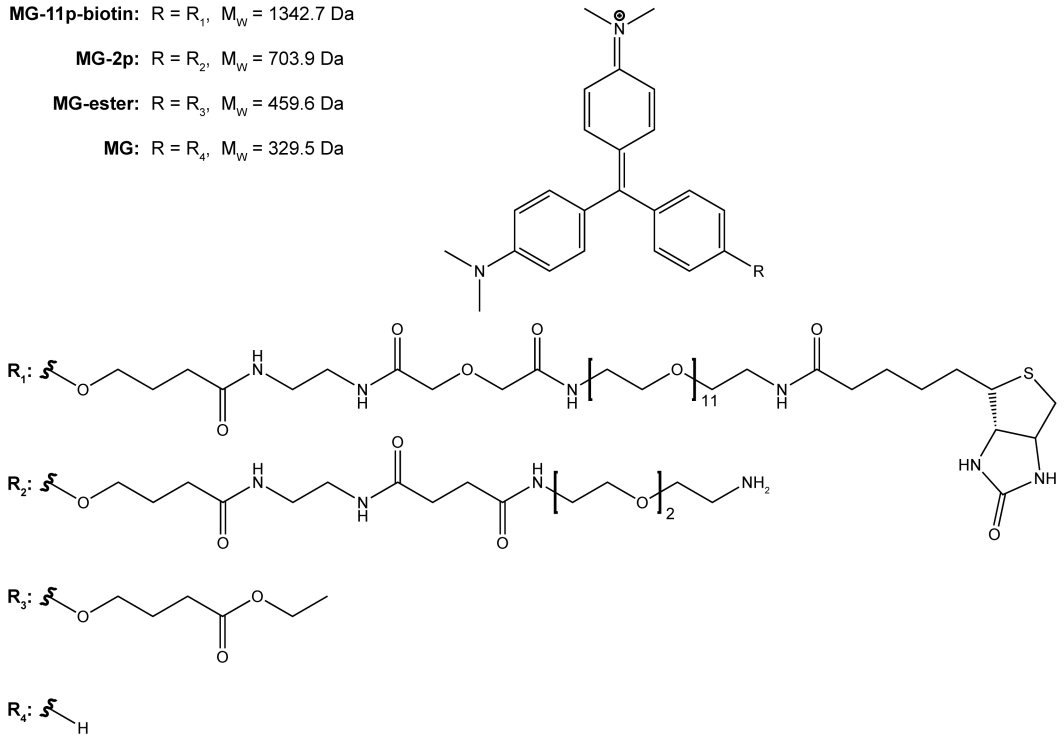
Supplementary Tables

Supplementary Table S1. Data collection and refinement statistics.

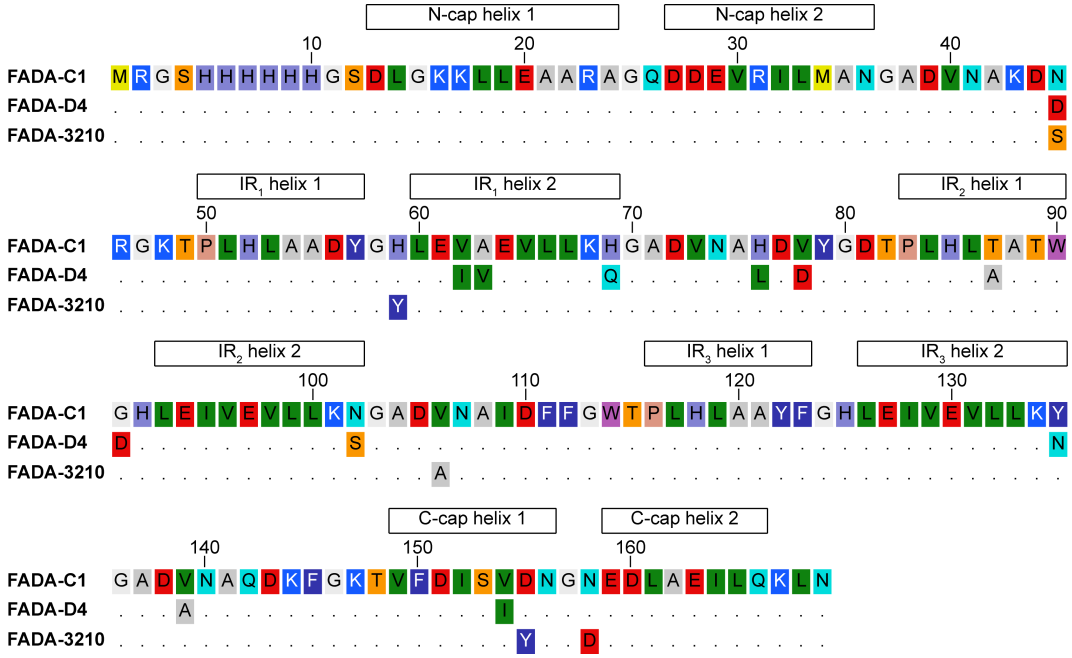
Data collection	
Wavelength [Å]	1.0
Resolution range [Å]	50.0 – 2.6 (2.693 – 2.600)*
Space group	*P* 1 2_1 1
Unit cell	
a, b, c [Å]	93.512, 109.754, 100.301
α, β, γ [°]	90.00, 94.35, 90.00
Total reflections	425989 (40509)*
Unique reflections	62041 (6095)*
Multiplicity	6.9 (6.6)*
Completeness [%]	100 (100)*
Mean I/σ(I)	6.6 (0.86)*
Wilson B-factor	37.84
R_{merge}	0.2733 (2.360)*
R_{meas}	0.2958 (2.564)*
$CC_{1/2}$	0.993 (0.552)*
CC*	0.998 (0.843)*
Refinement	
Resolution range [Å]	19.91 – 2.60 (2.693 – 2.600)*
Reflections used in refinement	61973 (6093)*
Reflections used for R_{free}	3098 (304)*
R_{work}	0.2135 (0.3812)*
R_{free}	0.2622 (0.3995)*
CC_{work}	0.970 (0.458)*
CC_{free}	0.953 (0.384)*
Number of non-hydrogen atoms	
all	14860
macromolecules	14359
ligands	162
water	339
Protein residues	1864
RMS deviations	
bond lengths [Å]	0.001
bond angles [°]	0.46
Ramachandran	
favored [%]	98.2
allowed [%]	1.8
outliers [%]	0
Rotamer outliers [%]	0.69
Clashscore	4.11
Average B-factor	
all	51.86
macromolecules	52.31
ligands	51.10
solvent	33.21
Number of TLS groups	115

*Statistics for the highest resolution shell are shown in parentheses.

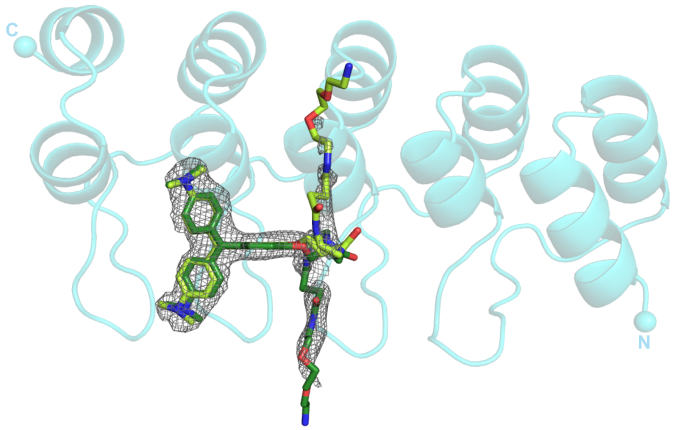
Supplementary Figures



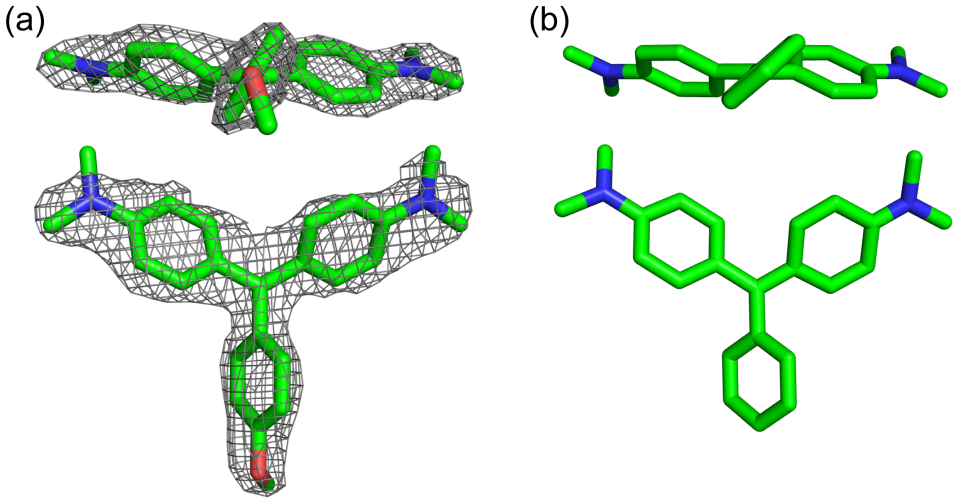
Supp. Fig. S1. Overview of fluorogenic MG-dyes used in this study. The chemical structures of MG and MG-derivatives used in this study are shown. MG-11p-biotin was used for selections with ribosome display, which requires a biotinylated target. MG-2p was used as a competitor in ribosome display off-rate selections, for yeast surface display selections, and for crystallization of the FADA-3210/MG-2p complex. While MG-11p-biotin and MG-2p are membrane impermeant, and thus do not diffuse into cells, MG-ester is membrane permeant and can be used for intracellular staining.

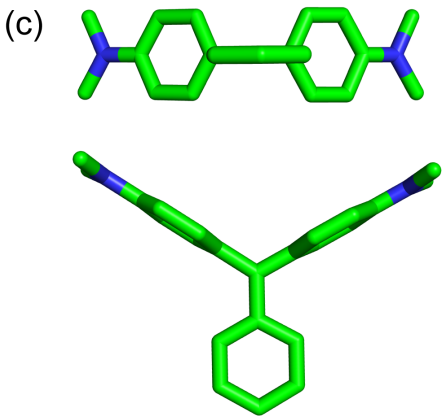


Supp. Fig. S2. Sequence alignment of FADAs. Sequences of fluorogen-activating DARPins obtained from ribosome display selections (FADA-C1 and FADA-D4) are aligned with the sequence of FADA-3210 which was obtained from maturation with yeast surface display. FADA-3210 is a descendant of FADA-C1 and has acquired five additional mutations during maturation (S^{45}, Y^{59}, A^{106}, Y^{155}, and D^{158}). Each of the two α-helices of the N-capping repeat (N-cap), of the three internal repeats (IR_1, IR_2, and IR_3), and of the C-capping repeat (C-cap) are indicated in the alignment.

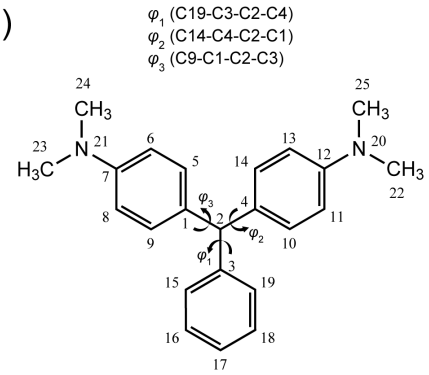


Supp. Fig. S3. For clarity, only one FADA-3210 molecule (cyan) of the dimer is shown. Crystal structure of FADA-3210 in complex with MG-2p at 2.6 Å resolution. The structure was refined with the complete MG-2p ligand in two alternative conformations (colored in light and dark green, respectively). The MG moiety of MG-2p shows well-defined electron density (contoured at 1.0 σ), whereas electron density for the R_2 moiety (defined as shown in Supp. Fig. S1) is much weaker. This observation can be explained by a symmetry mismatch. While the *C*2 symmetry of the MG moiety matches the symmetry of the FADA-3210 dimer, the R_2 moiety does not obey *C*2 symmetry. This symmetry mismatch imposes an occupancy decrease of the R_2 moiety. Since both orientations of the R_2 moieties show similar electron density, it is assumed that both alternative conformations possess approximately equal occupancies.





(d)



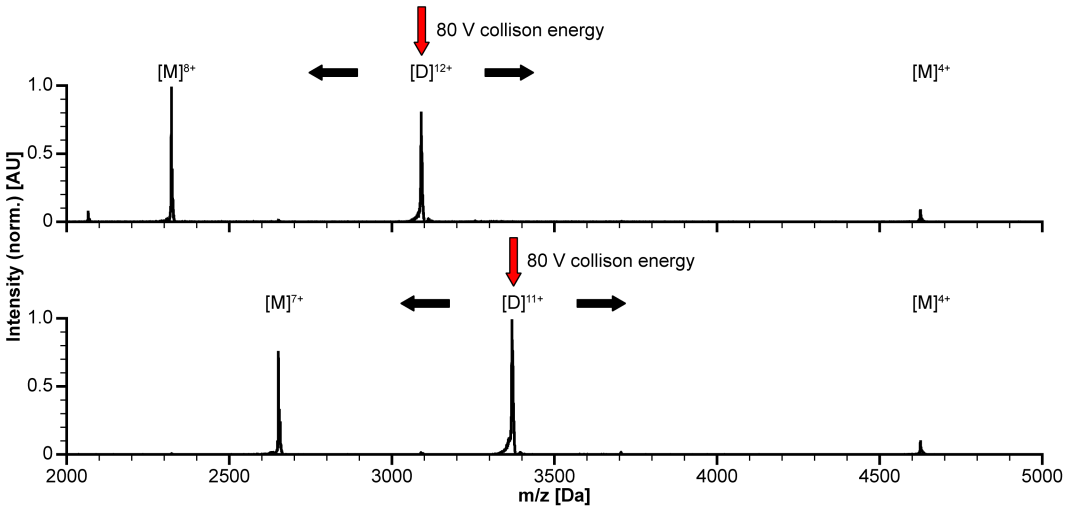
(e)

Structure	φ_1	φ_2	φ_3
3btl	-57.1°	-21.6°	-59.1°
3hti	-80.9°	-0.4°	-55.7°
3br0	-65.3°	-38.9°	-46.2°
3btc	-59.6°	-15.7°	-58.2°
3bqz	-66.3°	-30.8°	-25.7°
1jup	-42.4°	-17.3°	-45.4°
4k3h	-41.3°	-29.3°	-38.2°
1q8n	74.4°	26.4°	26.1°
5aao	55.4°	32.2°	30.7°
MG2 MP2*	41.5°	28.8°	28.8°

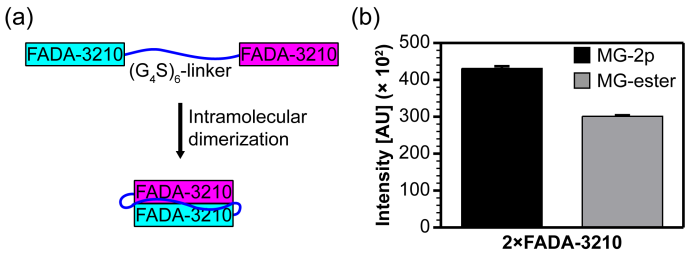
* Equilibrium structure of MG in the electronic ground state optimized by the MP2 method

Supp. Fig. S4. Structural conformation of MG-2p bound by the FADA-3210 dimer compared to the S_0 electronic ground state and the optimized structure of the S_1 state and comparison to other MG complexes. (a) Two perpendicular views of the conformation of MG-2p bound by the FADA-3210 dimer including the electron density map (contoured at 1.6 σ) are shown. MG-2p is fixed by the FADA-3210 dimer in a right-handed propeller conformation. (b) Shown are two perpendicular views of MG in the electronic S_0 ground state derived from calculation with the MP2 method and the DZP basis set [2]. (c) Shown are two perpendicular views of MG in the electronically excited S_1 state after relaxation, obtained by SA-CASSCF

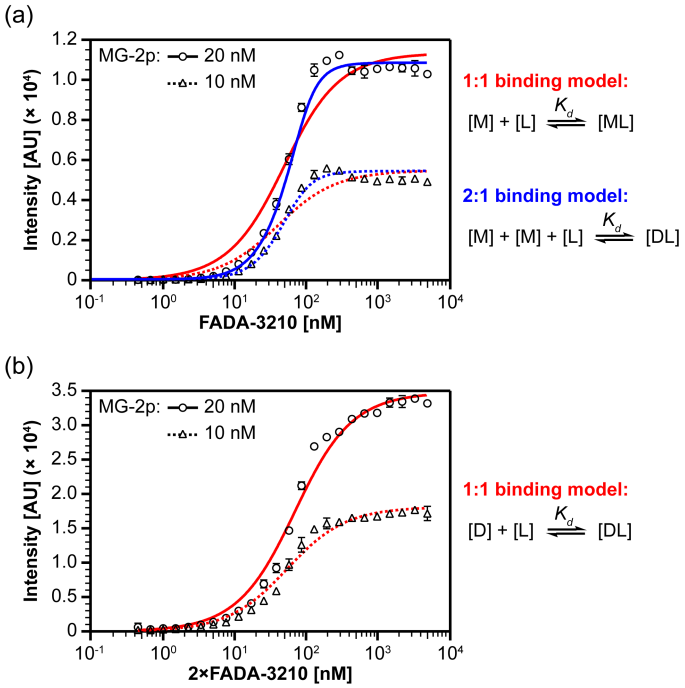
calculations, according to Nakayma et al. [2]. (d) Atom numbering of MG and definition of torsion angles. (e) Torsion angles within MG in different complexes. 5aao is the FADA-3210 structure, corresponding to the molecules shown in (a). MG2 MP2 corresponds to the calculated S_0 electronic ground state as shown in (b).



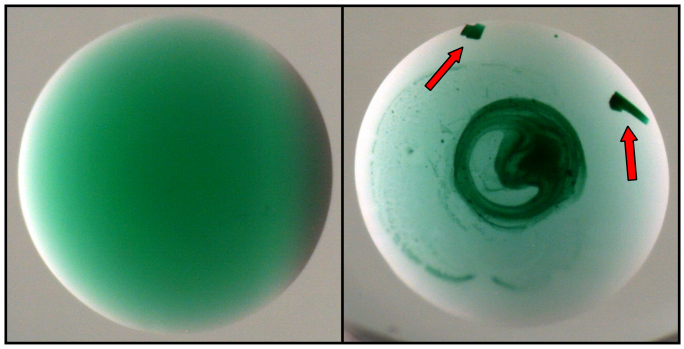
Supp. Fig. S5. Collision-induced dissociation (CID) of FADA-3210 dimers analyzed by tandem mass spectrometry (MS/MS). The CID spectra of the $[D]^{+12}$ (top) and the $[D]^{+11}$ (bottom) dimeric charge states at a collision energy offset of 80 V are shown. Dissociation of the dimeric ion results in monomeric ions with typical asymmetric charge partitioning [3,4]. $[D]^{+12}$ dissociates into $[M]^{+8}$ and $[M]^{+4}$, while $[D]^{+11}$ dissociates into $[M]^{+7}$ and $[M]^{+4}$.



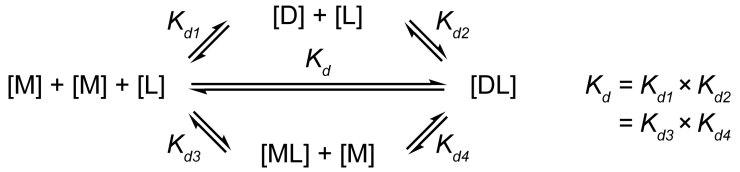
Supp. Fig. S6. Fluorescence-activation of purified 2×FADA-3210 in solution. (a) Design of the 2×FADA-3210 construct. The long and flexible $(G_4S)_6$-linker enables the intramolecular dimerization of the two linked FADA-3210 molecules. In contrast to FADA-3210, which forms intermolecular dimers, it can be assumed that intramolecular dimers of 2×FADA-3210 are present even at very low DARPin concentrations. (b) Fluorogen-activation of MG-2p (black bar) and MG-ester (grey bar) by 2×FADA-3210. 2×FADA-3210 (2 µM) was incubated with 200 nM of MG-2p or MG-ester. Fluorescence-activation of MG-2p by 2×FADA-3210 is identical to the activation obtained with FADA-3210. Error bars indicate standard deviations from triplicates.



Supp. Fig. S7. Fluorescence-activation titration experiments with FADA-3210 and 2×FADA-3210 activating MG-2p. (a) Fluorescence-activation assays with constant MG-2p concentrations (20 nM MG-2p, circles, solid lines; 10 nM MG-2p, triangles, dashed lines) and increasing FADA-3120 concentrations (0.45 nM – 5 µM) are shown. Fitting the experimental data to a 1:1 binding model (red lines) does not account for the steepness of the fluorescence activation (χ^2 = 0.472), whereas the data can be described well by a 2:1 binding model (χ^2 = 0.057, blue lines) which assumes that two FADA-3210 molecules activate one MG-2p molecule, resulting a dissociation constant of 840 ± 70 nM^2. Error bars indicate standard deviations from quadruplicates. (b) Fluorescence-activation assays with constant MG-2p concentrations (20 nM MG-2p, circles, solid lines; 10 nM MG-2p, triangles, dashed lines) and increasing 2×FADA-3120 concentrations (0.45 nM – 5 µM) are shown. The experimental data are described well by a 1:1 binding model, which assumes that one 2×FADA-3120 molecule (intramolecular dimer) activates one MG-2p molecule. At the higher 2×FADA-3120 concentration a small amount of intermolecular dimer, i.e. two molecules of 2×FADA-3120 binding one molecule of MG-2p, may contribute to the deviation from the fit. Error bars indicated standard deviations from quadruplicates. All binding models used are explained in more detail in the Supplementary Methods.



Supp. Fig. S8. FADA-3210/MG-2p crystals obtained by sitting drop vapor diffusion. Shown is the well in which FADA-3210/MG-2p crystals were obtained (20% (w/v) polyethylene glycol 3,350, 0.2 M sodium malonate, pH 5.0) after setting up the crystallization experiment (left) and with grown protein crystals (right). The obtained crystals (indicated by red arrows) appeared in two days as thin green-colored plates and grew to the full size within a week.



Supp. Fig. S9. Reaction pathways for the formation of a dimeric DARPin complex binding one dye molecule. The whole reaction pathway for the formation of a dimeric DARPin complex binding one dye molecule with the corresponding dissociation constants for each microscopic reaction is shown. [M] and [D] are the concentrations of monomeric and dimeric DARPin, respectively, [L] is the concentration of free MG-2p ligand, and [ML] and [DL] are the concentrations of monomeric DARPin/MG-2p and dimeric DARPin/MG-2p complexes, respectively. Two microscopic reaction routes are possible: (i) first the monomeric DARPins dimerize and then bind MG-2p (described by K_{d1} and K_{d2}) and (ii) a monomeric DARPin first binds MG-2p and then dimerizes with another DARPin molecule (described by K_{d3} and K_{d4}). In the case where the intermediate products are not populated at equilibrium ($K_{d1} >> K_{d2,}$ $K_{d3} >> K_{d4}$), a reliable fit is only possible for the dissociation constant of the overall reaction (K_d), which is the product of the dissociation constants of the microscopic reactions.

References

[1] D. Steiner, P. Forrer, A. Plückthun, Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display, J Mol Biol. 382 (2008) 1211–1227.

[2] A. Nakayama, T. Taketsugu, Ultrafast nonradiative decay of electronically excited states of malachite green: ab initio calculations, J Phys Chem A. 115 (2011) 8808–8815.

[3] J.C. Jurchen, E.R. Williams, Origin of asymmetric charge partitioning in the dissociation of gas-phase protein homodimers, J Am Chem Soc. 125 (2003) 2817–2826.

[4] J.C. Jurchen, D.E. Garcia, E.R. Williams, Further studies on the origins of asymmetric charge partitioning in protein homodimers, J Am Soc Mass Spectrom. 15 (2004) 1408–1415.