Supporting Information for

De novo Designed Enzymes as Small Molecule-Regulated Fluorescence Imaging Tags and Fluorescent Reporters

Yu Liu[1,2,3,‡], Xin Zhang[1,2,3,‡], Yun Lei Tan[1,2,3], Gira Bhabha[4], Damian C. Ekiert[5], Yakov Kipnis[6], Sinisa Bjelic[6], David Baker[6], and Jeffery W. Kelly[1,2,3,*]

[1]Department of Chemistry, [2]Department of Molecular and Experimental Medicine and [3]The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA. [4]Department of Cellular and Molecular Pharmacology, [5]Department of Microbiology and Immunology, University of California, San Francisco, CA 94158, USA. [6]Department of Biochemistry, University of Washington, WA 98195, USA.

[‡]These authors contribute equally to this work.

*Corresponding author: jkelly@scripps.edu

Table of contents

Supplementary Figures

RAs	Original Design	Size (kD)	Scaffold
RA1	RA114.3	29.6	TIM-Barrel
RA2	RA110.4 (Cys Free)	15.8	KSI-NTF2-like
RA3	RA112	21.1	Rossmann-Fold

Figure S1. X-ray crystal structures of RA1, RA2 and the scaffold of RA3. RA1 is a 29.6 kDa TIM-barrel scaffold protein (PDB code: 4PEK). RA2 is a 15.8 kDa KSI-NTF2-like protein (PDB code: 4PEJ). RA3 is a 21.1 kDa Rossmann-fold protein (Protein design is based on PDB code:1ILW). Table S1 summarizes the data collection and refinement statistics for the X-ray crystal structure determination of Apo-RA1 and Apo-RA2.

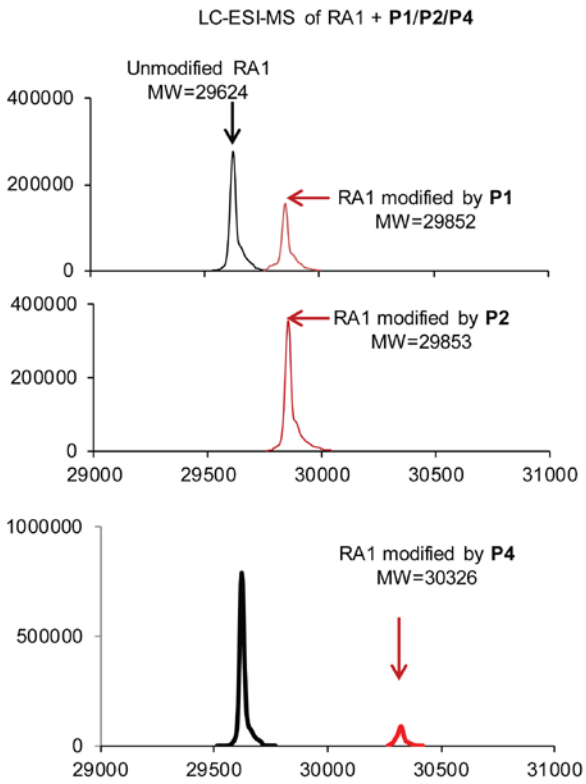


Figure S2. P2 labels RA1 faster than **P1** in buffer. **P2** (25 µM) completely labeled **RA1** (5 µM) within 30 min at 25 °C while **P1** only partially reacted with RA1, as shown by LC-ESI mass spectrometry. **P4** (25 µM) labels RA1 (5 µM) even less efficiently in buffer after 30 min, likely due to the presence of the bulky BODIPY fluorophore π-π stacking with the naphthalene. This intramolecular interaction will slow the reaction with RA1. Interestingly, **P4** reacts much faster with RA1 in a cell (see Figure S5), which may be because endogenous protein (e.g., a macromolecular chaperone) binding to **P4** prevents the π-π stacking, facilitating pre-equilibrium binding.

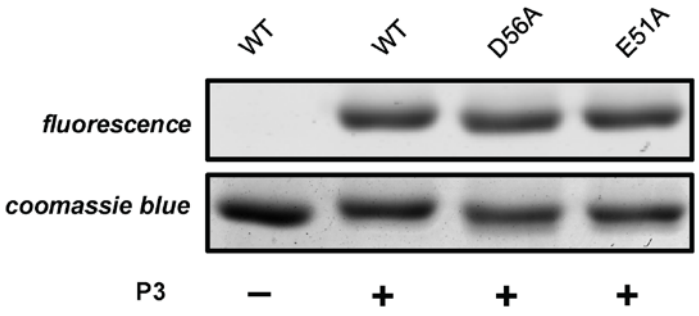


Figure S3. The robust electrostatic interaction network within the catalytic site. Mutating D56 or E51 had negligible effect on the labeling efficiency, indicating that their interactions to mediate **P3** binding and labeling are redundant. RA1 and its mutants (5 μM) were incubated with DMSO or **P3** (200 μM) for 1 h at 25 °C, followed by quenching with SDS loading buffer.

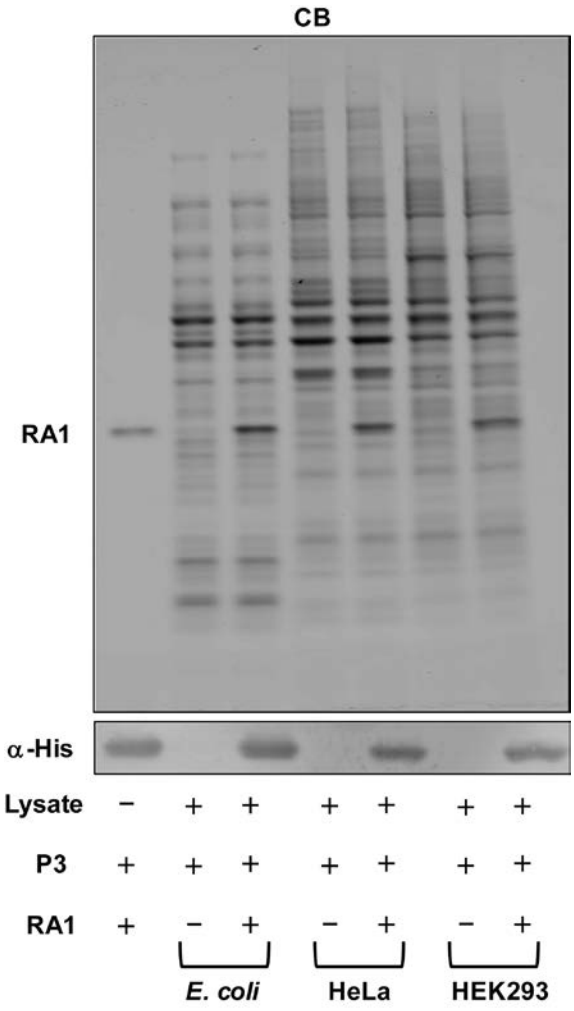
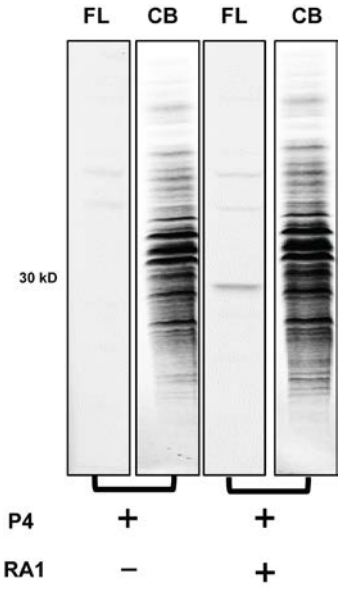


Figure S4. Coomassie blue (CB) staining of the lysates of *E. coli*, HeLa, and HEK293 cells. In all lanes, 5 μL of **P3-**treated soluble lysate (3 mg/mL, total protein concentration), with or without His-tagged RA1 spiked in, was resolved by SDS-PAGE.

a)



b)

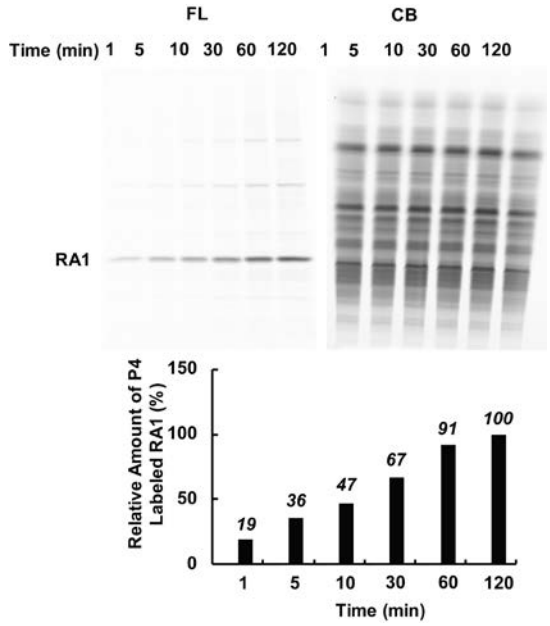


Figure S5. P4 selectively and rapidly labels cytosolic RA1 in live HEK293 cells. a) HEK293 cells transiently transfected with RA1 were incubated with **P4 (**10 μM) for 10 min and then washed twice with DMEM before being harvested with TrypLE Express. The samples were then prepared following Supplementary Method (3) and resolved using electrophoresis. The SDS-PAGE gel was scanned using a fluorescence scanner (FL) to detect the fluorescence of the RA1-**P4** conjugate and Coomassie blue (CB) staining was then used to visualize the cellular proteome. b) HEK293T cells were transiently transfected with RA1. The cells were incubated with **P4** (10 μM) for the indicated time periods (1, 5, 10, 30, 60, 120 min) before harvesting. The cells were lysed and the lysates were normalized to total protein concentration. By quantifying the amount of fluorescent RA1 on the gel, we observed 20% labeling of RA1 in the first minute, indicating rapid labeling in live cells. For most of the experiments in this paper, we used a 10-min labeling period, during which time almost half of the RA1 is labeled. Although the signal is higher for longer incubation times, the background also increases, leading to no significant improvement in the signal-to-noise ratio. Interestingly, **P4** (25 μM) labels RA1 (5 μM) less efficiently in buffer after 30 min, likely due to the presence of the bulky BODIPY fluorophore intramolecularly π-π stacking with the naphthalene, diminishing pre-equilibrium binding with RA1–see Figure S2.

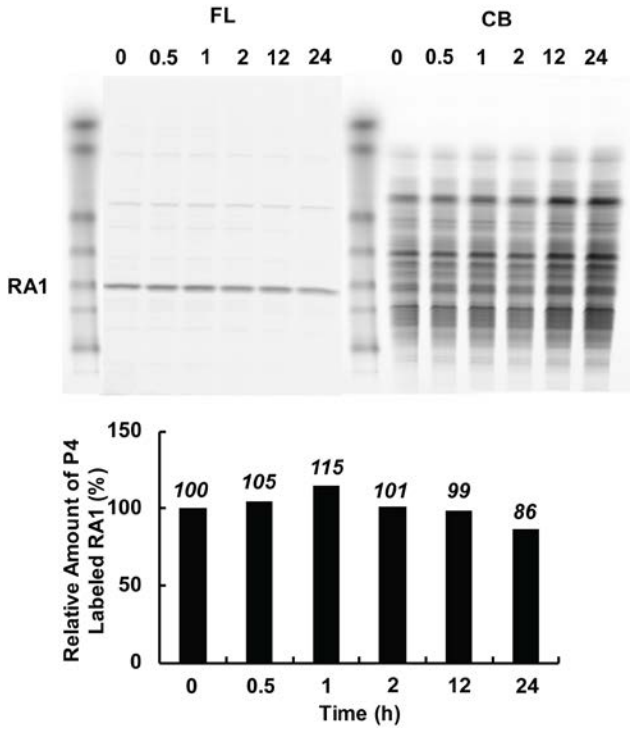


Figure S6. *In vivo* stability of the RA1-**P4** conjugate. HEK293 cells transiently expressing RA1 were labeled with **P4 (**10 μM) for 10 min and then washed twice with DMEM before being harvested with TrypLE Express at the indicated time points (0, 0.5, 1, 2, 12, 24 h). The SDS-PAGE gel was scanned using a fluorescence scanner (FL) to detect the fluorescence of the RA1-**P4** conjugate and Coomassie blue (CB) staining was then used to visualize the cellular proteome. No significant increase in background staining was observed over the longer incubation period. The fluorescence conjugate stability *in vivo* was evaluated by comparing the amount of labeled RA1 remaining as a function of time relative to time 0 (bottom panel).

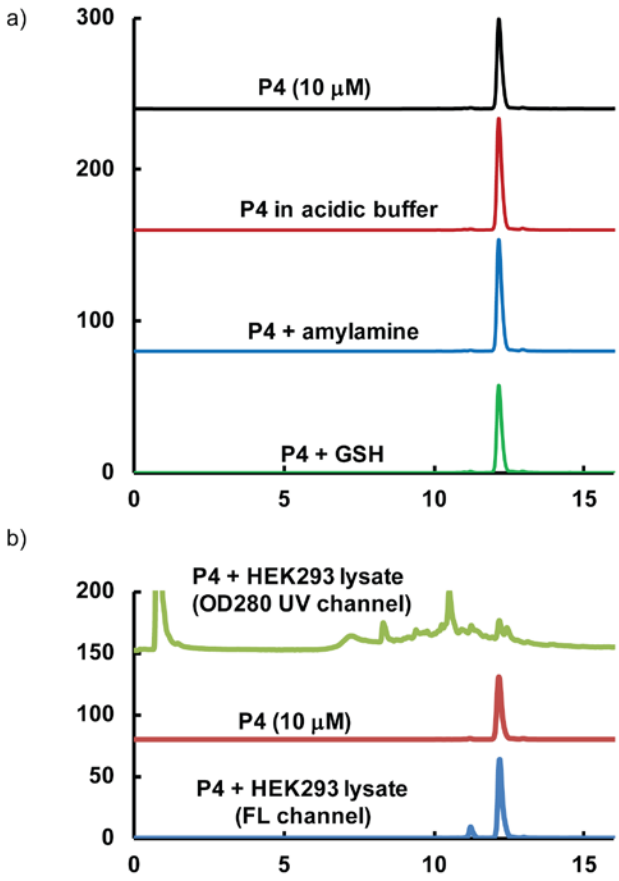


Figure S7. *In vivo* and *in vitro* stability of **P4**. a) **P4** (10 μM) was incubated in acidic buffer (sodium phosphate buffer, pH=4.8, red curve), in amylamine (0.5 mM, blue curve), and in glutathione (GSH, 0.5 mM in argon-purged RA buffer, green curve) for 24 h and the recovery of fluorescent **P4** was analyzed by HPLC-coupled to a fluorescence detector. No significant reaction or hydrolysis of **P4** was observed, as assessed by the lack of a shift in the retention time or a decrease in peak area, indicating lack of chemical reactivity of **P4**. b) **P4** (10 μM, red curve) was incubated in concentrated HEK293 cell lysate (2 mg/mL, blue curve) for 1 h. Minor **P4** loss was observed by fluorescence detection (< 10%). The majority of **P4** remained intact. That the endogenous proteome is present is shown by the green curve, as observed by UV detection.

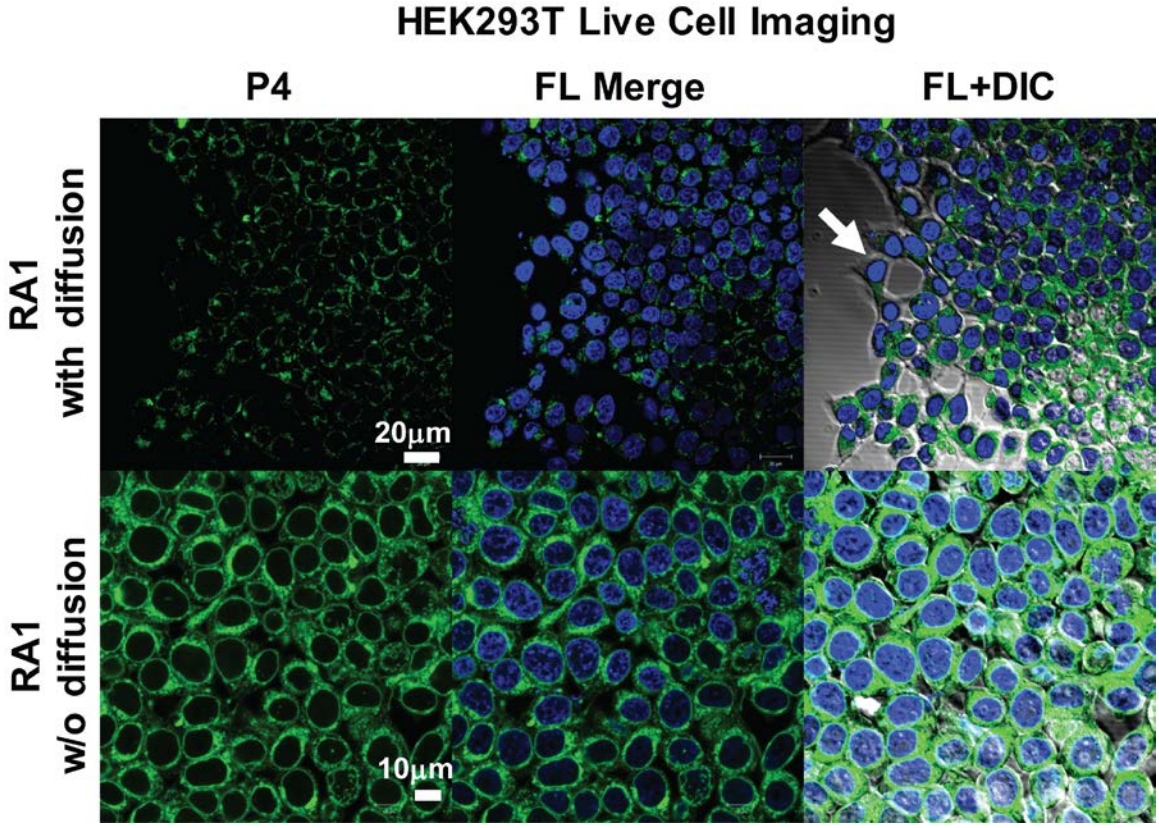


Figure S8. Live cell imaging of the RA1-**P4** conjugate in HEK293T cells. HEK 293T cells were grown to ~50% confluency on chambered coverglasses (Thermo Scientific™ Nunc™ Lab-Tek™) and transfected with RA1 for 24 h. **P4** (10 μM) in DMEM was incubated with the cells for 10 min. The media was then replaced with fresh DMEM at 37 °C. The cells were then washed twice with DPBS after a 30-min diffusion period. Images were taken in DPBS using a confocal microscope. A non-transfected cell is indicated by the white arrow. The lower panel shows the imaging without the additional 30-min diffusion period. The non-transfected cells are not obviously identified. The selectivity of the staining is not comparable to the upper panel with diffusion, showing the necessity of the washing/diffusion step. FLMerge: Green (**P4**) and Blue (DAPI).

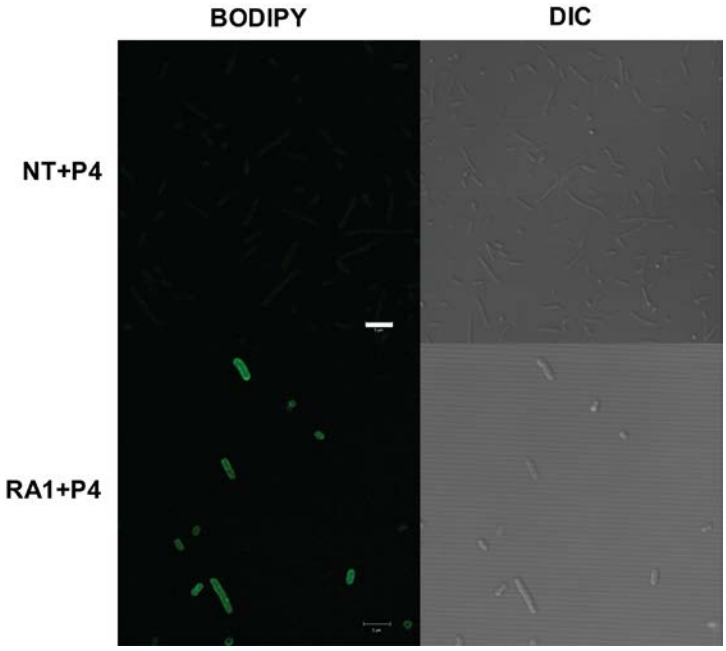


Figure S9. Live cell imaging of the RA1-**P4** conjugate in *E. coli* K12 cells. *E. coli* K12 cells (non-transformed (NT) or transformed with RA1) were grown in LB media to OD_{600} = 0.6 and induced with IPTG (75 μM) for 30 min. The cells were harvested by centrifugation and the pellet was resuspended in DPBS containing **P4** (10 μM). After incubation at 37 °C for 30 min, the cells were washed three times with fresh DPBS before plating on chambered coverglasses (Thermo Scientific™ Nunc™ Lab-Tek™) for imaging. All cells that were transformed with RA1 were stained green by **P4**. See Supplementary Method 7 for further details. Scale bar: 5 μM.

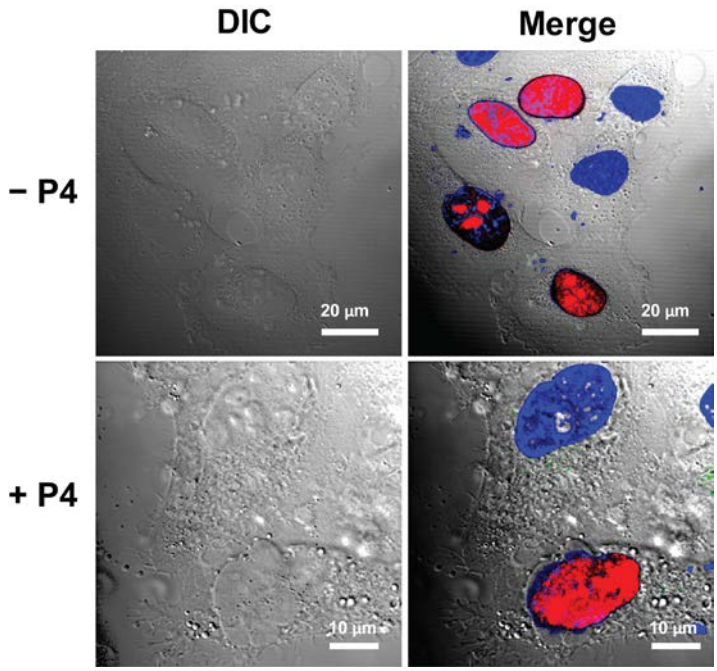


Figure S10. Differential Interference Contrast (DIC) images merge with fluorescence confocal images (Figure 3b). Red = RFP:NLS fluorescence, blue=DAPI staining.

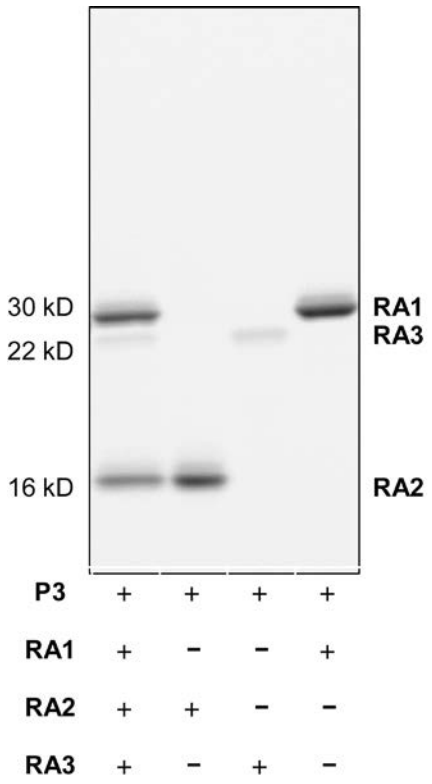
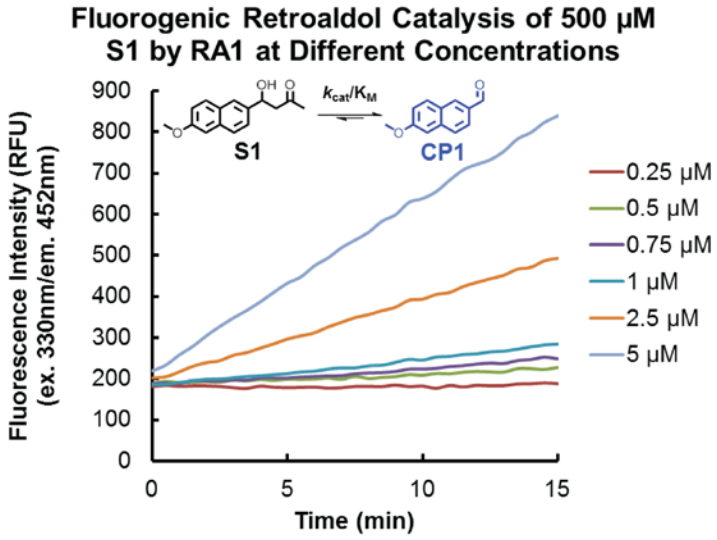


Figure S11. P3 is capable of labeling RAs of different sizes and scaffolds. A 1:1:1 mixture of all RAs (3.3 μM each of RA1, RA2, and RA3; Lane 1), 10 μM RA1 (Lane 4), 10 μM RA2 (Lane 2), and 10 μM RA3 (Lane 3) were labeled with 200 μM **P3** for 1 h at 25 ºC before being quenched with SDS loading buffer. Labeled RAs were resolved on a 4%-20% Tris Glycine gradient SDS PAGE gel. RA1 and RA2 were labeled more efficiently than RA3, as assessed by fluorescence scanning.

a)



b)

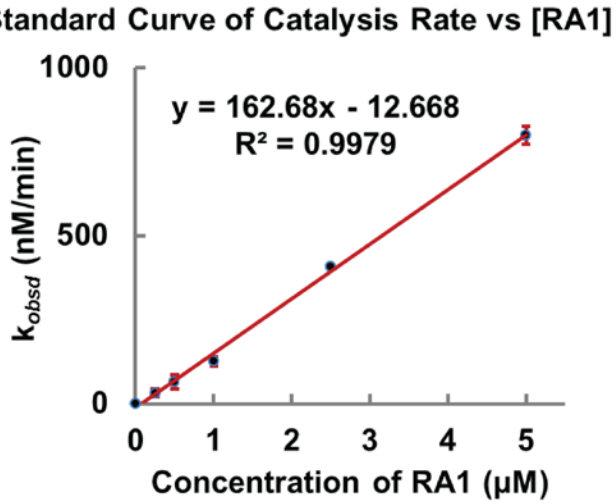


Figure S12. Standard curve of apparent catalysis rates (k_{obsd}) against the concentration of RA1 in buffer under subsaturation conditions. Details of kinetics fitting and calculation are described in Supplementary Method (5).

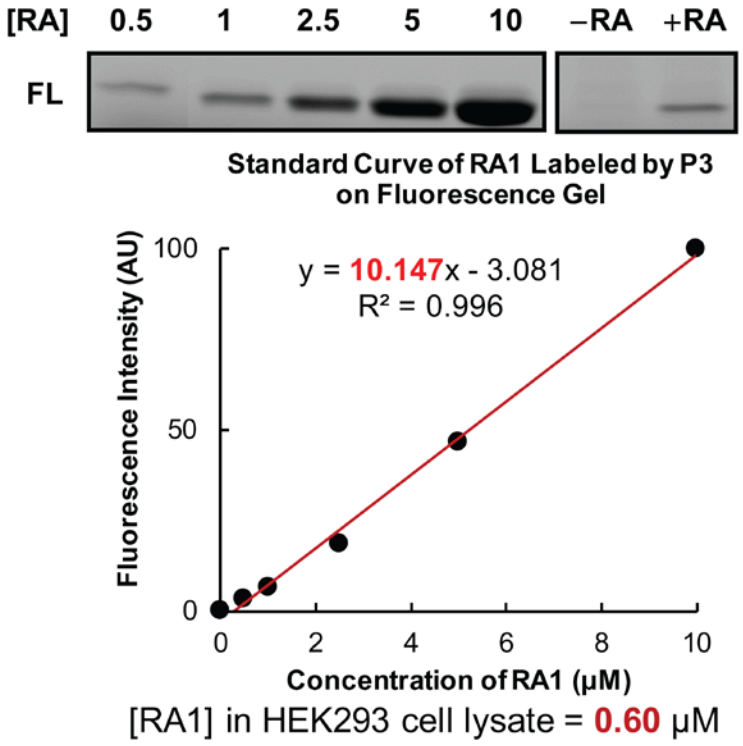


Figure S13. Estimation of the concentration of overexpressed cytosolic RA1in HEK293 cell lysate using **P3**. A standard curve of the fluorescence intensity of **P3**-labeled RA1 at the indicated concentrations was generated. No significant background band was observed in the non-transfected cell lysate (-RA). The concentration of RA1 was estimated to be 0.6 μM by fitting the intensity of the +RA band into the standard curve. This concentration is comparable to the one derived from the fluorogenic assay, indicating that the functional assay can serve as an alternative quantitative readout for the presence of RA1.

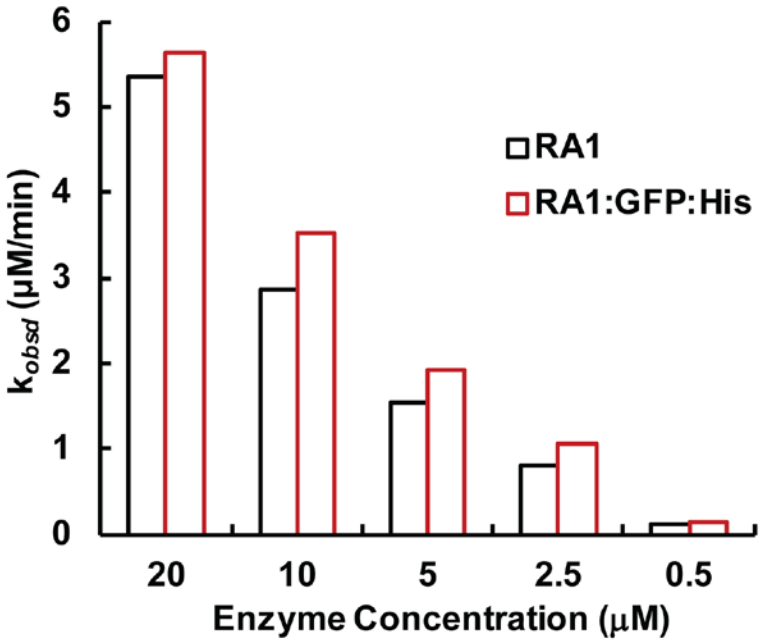


Figure S14. Fusion of RA1 to a POI does not affect the function of RA1. No significant change of the apparent catalysis rates in buffer were observed at the same concentration of RA1 and RA1:GFP:His.

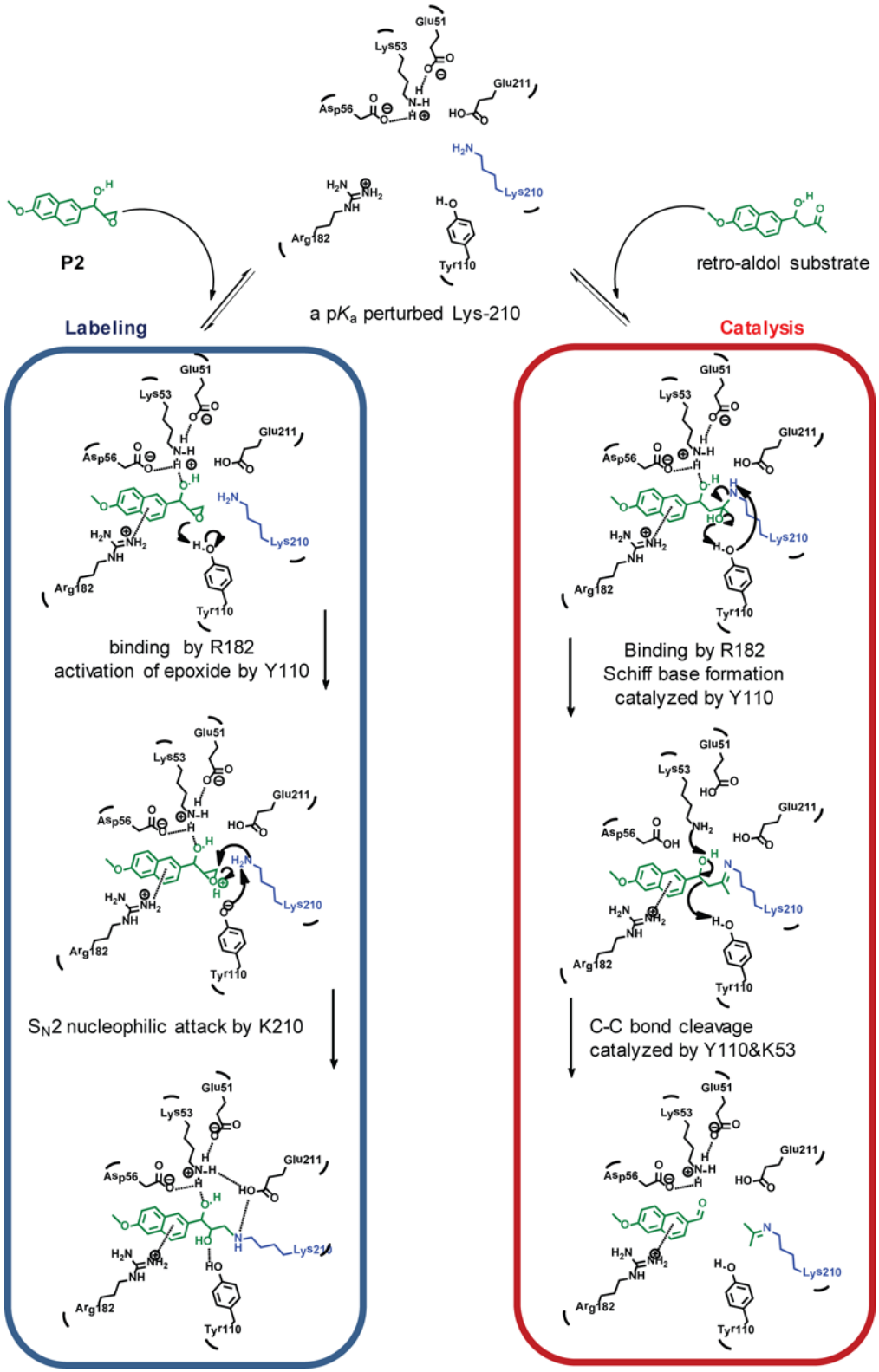


Figure S15. Residues in the RA active site are important for both probe labeling and retro-aldol catalysis reactions, suggesting common mechanistic themes.

Table S1. Data collection and refinement statistics.

Data collection	Apo-RA1	Apo-RA2
	ALS 8.3.1	ALS 8.3.1
Wavelength (Å)	1.12	1.12
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters (Å, °)	a = 53.8, b = 63.08, c = 93.07 Å α = β = γ = 90.0°	A = 35.5, b = 74.2, c = 94.9 Å α = β = γ = 90.0°
Resolution (Å)	50 – 1.60 (1.64-1.60) [a]	50 - 1.85 (1.90-1.85) [a]
Observations	338,565	149,905
Unique Reflections	42,455 (3,071) [a]	21,967 (1,485) [a]
Redundancy	8.0 (7.1) [a]	6.8 (6.5) [a]
Completeness (%)	99.8 (99.9) [a]	99.1 (92.6) [a]
$<I/\sigma_I>$	17.84 (1.43) [a]	15.8 (1.5) [a]
CC1/2	1.00 (0.65)	1.00 (0.77)
R_{sym} [b]	0.05 (1.35) [a]	0.06 (1.54) [a]
Refinement statistics		
Resolution (Å)	50 – 1.60	50 – 1.85
Reflections (work)	40,441	20,855
Refections (test)	1,999	1,099
R_{cryst}(%)[c] / R_{free}(%)[d]	14.7 / 18.6	21.2 / 23.8
Average B (Å^2)	47.1	52.8
Wilson B (Å^2)	26.4	30.8
Protein atoms	2182	1986
Ligand atoms	0	0
Waters	140	37
Other	0	0
RMSD from ideal geometry		
Bond length (Å)	0.014	0.007
Bond angles (°)	1.37	0.019
Ramachandran statistics (%)[e]		
Favored	97.0	96.7
Outliers	0.4	0.4
PDB Code	4PEK	4PEJ

[a] Numbers in parentheses refer to the highest resolution shell.

[b] $R_{sym} = \Sigma_{hkl} \Sigma_i | I_{hkl,i} - <I_{hkl}> | / \Sigma_{hkl} \Sigma_i I_{hkl,I}$, where $I_{hkl,i}$ is the scaled intensity of the i^{th} measurement of relection h, k, l, $< I_{hkl} >$ is the average intensity for that reflection, and n is the redundancy[1]

[c] $R_{cryst} = \Sigma_{hkl} | F_o - F_c | / \Sigma_{hkl} | F_o |$ x 100

[d] R_{free} was calculated as for R_{cryst}, but on a test set comprising 5% of the data excluded from refinement.

[e] Calculated using Molprobity[2]

Table S2. Comparison of Exisiting Protein Tags with RA-tag[3]

Tags	Size (kDa)	Fluorogenic Labeling	Fluorogenic Catalysis	Kinetics ($M^{-1}s^{-1}$)
SNAP	20	Yes	No	2.8×10^4
CLIP	20	No	No	1.1×10^3
Halo	34	No	No	2.7×10^6
TMP	18	Yes	No	NA
BL	29	Yes	Yes	7.8×10^4
RA	16-31	No	Yes	$\sim 10^2$

Supplementary Methods

(1) Crystallization of the apo-RA1, apo-RA2.

Apo-RA1 and apo-RA2 were purified using anion exchange followed by size exclusion chromatography as described previously.[4] Pure fractions were pooled and concentrated to 9.1 mg/mL for apo-RA1 and 17.2 mg/mL for apo-RA2. Initial crystallization trials were performed using the JCSG core I-IV screens at 22 °C. Drops were set up with the Mosquito HTS using 100 nL protein and 100 nL of the well solution. Apo-RA1 was crystallized in 0.2 M ammonium acetate, 0.1 M sodium citrate, pH 5.6 and 30 % (w/v) PEG 4000. Apo-RA2 was crystallized in 0.1 M sodium acetate, pH 4.5 and 40% (v/v) PEG 400. Crystals were cryoprotected in the reservoir solution supplemented with ethylene glycol if necessary, then flash cooled and stored in liquid nitrogen until data collection. For apo-RA1, a 1.65 Å dataset was indexed and integrated in space group $P2_12_12_1$. For apo-RA2, a 1.85 Å dataset was indexed and integrated in space group $P2_12_12_1$. All diffraction data were collected at the Advanced Light Source (ALS) beamline 8.3.1. Data collection and structure refinement statistics are shown in Table S1.

Data reduction was carried out using XDS.[5] The structures were solved by molecular replacement using the Phaser module in Phenix with only the protein coordinates of PDB ID 3TC7 as a search model for apo-RA1, and 1W02 as a search model for apo-RA2. Rigid body, restrained refinement with TLS and simulated annealing were carried out in Phenix.[6] Manual adjustment of the model was carried out in Coot.[7] The structures were validated using the Quality Control Check v2.8 developed by the JCSG, which included Molprobity[2] (publicly available at http://smb.slac.stanford.edu/jcsg/QC/). Final refinement statistics are shown in Table S1.

(2) ***E. coli* strain and lysate preparation.**

E. coli strain HMS174 was used to show the selectivity of **P3** in *E. coli* lysate. The lysate was obtained by sonication of resuspended cells in RA buffer (25 mM HEPES, 100 mM NaCl, pH 7.5) supplemented with a protease inhibitor cocktail (Roche). The protease inhibitor cocktail was used to inhibit proteolytic reaction in lysate. Importantly, components in the cocktail were shown not to react with RA or affect RA activity.[4] The lysates were centrifuged at 16,000 g for 20 min at 4 °C and the supernatant collected as soluble lysates. Total protein concentrations of all soluble lysates were measured using the BCA assay (Thermo Scientific).

(3) Fluorescence electrophoresis gel and western blot analysis.

HEK293 or HeLa cells prepared under different transfection and labeling conditions were harvested using TrypLE Express (Life Technologies), washed twice with DPBS, and stored as pellets at −80 °C. When needed, cells were thawed on ice and lysed using cOmplete Lysis M Buffer supplemented with a protease inhibitor cocktail (Roche) for 20 min on ice. The lysates were centrifuged at 16,000 g for 20 min at 4 °C and the supernatant collected as soluble lysates. Total protein concentrations of all soluble lysates were measured using the BCA assay (Thermo Scientific). To analyze H2B:RA1:His and H2B:RA2:His in the nucleus, the pellet was collected after centrifugation and redissolved in the lysis buffer.

SDS loading buffer was added to each sample, and the samples were boiled for 5 min. Proteins were resolved by SDS-PAGE gel (12%). For fluorescence visualization, wet slab gels were scanned using a Typhoon Variable Mode Imager (Amersham

Biosciences) with λ_{ex} of 488 nm and λ_{em} of 520 nm (band pass of 40 nm) to visualize the RA-**P3** and RA-**P4** conjugates.

For western blot analysis, the wet slab gels were transferred to a PVDF membrane, which was subsequently blocked with 5% non-fat milk in PBST (PBS containing 0.05% Tween-20) for 1 h at 25 °C. Western blot analysis was performed with a mouse monoclonal α-His antibody (372900, Invitrogen) against His-tagged RA1 and RA2 or a goat polyclonal α-Histone H2B antibody (N-20, sc-8650, Santa Cruz Biotech.) against the endogenous histone H2B, followed by secondary α-mouse or α-goat antibodies (IRDye 800, LI-COR Corp.). The membrane was scanned using the LI-COR Odyssey Imager.

(4) Gel shift to visualize both labeled and unlabeled RA1 for labeling efficiency evaluations.

To directly visualize the fluorescence labeling in Figure 2a and 2b, RA1 (5 μM) was incubated with **P3** (200 μM) at 25 °C for 1 h. The labeling reaction was quenched by the addition of SDS loading buffer and boiling for 5 min. Proteins were resolved by SDS PAGE gel (11%) and visualized by coommassie blue staining. Due to the conjugation of the ~700 Da **P3**, the band of RA1-**P3** is slightly higher than the unlabeled RA1.

(5) Retro-aldol catalysis and P3-labeling kinetics.

The retro-aldol reaction was performed using 4-hydroxy-4-(6-methoxy-2-naphthyl)-2-butanone (**S1** in Figure 4b) as substrate in RA buffer at 25 ºC. Accumulation of product (**CP1** in Figure 4b) was monitored using a Gemini™ EM Fluorescence

Microplate Reader with an excitation wavelength of 330 nm and an emission wavelength of 452 nm. The fluorescence signal allowed us to measure the initial velocity (v_0; $v_0 = \Delta[P] / \Delta t$, wherein [P] is the product concentration and t is the reaction time) from the initial linear portion of a pseudo-first order reaction, which was achieved when the substrate concentration ([S]) was > 20-fold greater than the enzyme concentration ([E]) and when the product accumulation is negligible. In this situation, the initial velocity v_0 is linearly related to the enzyme concentration at a fixed, saturating substrate concentration. Because the initial velocity is affected by the enzyme concentration but not by the substrate concentration, these velocities can be used to derive the enzyme concentration given that a linear standard curve is available.

To quantify RA1 concentration in lysates, we incubated HEK293 cell lysates expressing RA1 with substrate S1 (500 μM) to measure the initial velocity (v_{HEK}; Figure 4c). A linear standard curve was generated from the initial velocities of reactions containing varying, subsaturating concentrations of RA1 (0.25, 0.5, 0.75, 1, 2.5, and 5 μM) and 500 μM substrate S1 (Figure S12a-b). Finally, v_{HEK} was compared to the standard curve to derive the concentration of RA1 in HEK293 cell lysates.

Labeling kinetics (Figure 2c and 2d) were measured by incubating RA1 (5 μM) with varying concentrations of **P3** (25 ºC) in RA buffer. At different time points (t), the reaction was quenched by the addition of SDS loading buffer and boiling for 5 min. Proteins were resolved by SDS polyacrylamide gels (15% or 11%). Wet slab gels were scanned to visualize the fluorescence signal from the RA1-**P3** conjugate using a Typhoon Variable Mode Imager (Amersham Biosciences) using blue laser excitation (λ_{ex} 488 nm) and fluorescein emission channel (λ_{em} 520 nm, bandpass 40) with 100 micron resolution.

The intensity of the RA1-**P3** conjugate fluorescence (F_t) was quantified with ImageJ and plotted as a function of time to derive the observed rate constant (k_{obsd}) using the equation,

$$F_t = (F_\infty) \bullet \exp(-k_{obsd} \bullet t) \quad \text{(Eq. 1)}$$

where F_∞ is the fluorescence intensity of the fully labeled RA1-**P3** conjugate (5 μM). Steady-state kinetic parameters were derived by fitting k_{obsd} to the equation,

$$k_{obsd} = k_{inact} \bullet [\mathbf{P3}]/(K_i + [\mathbf{P3}]) \quad \text{(Eq. 2)}$$

where k_{inact} is the rate constant of the chemical step, and K_i is the **P3** concentration that provides half of the maximal rate.

(6) Protein expression and purification.

The pET29b (+) vector encoding RA1, RA2 or RA3 was transformed into BL21 DE3 Star competent cells (Invitrogen, Carlsbad, CA). When the bacteria reached an OD_{600} of 0.7, protein expression was induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG, BioPioneer Inc., San Diego, CA) for 4 h at 37 ºC. His-tagged RAs were affinity purified using 10 mL of Ni-NTA His•Bind resin (Qiagen, Valencia, CA) in 25 mM HEPES, 100 mM NaCl, 10 mM imidazole, pH 7.5, which was eluted with an imidazole gradient (10 mM to 1 M imidazole over 100 mL volume). RA-containing fractions were pooled and further purified by size exclusion chromatography (Superdex 75, GE Healthcare) equilibrated in RA buffer (25 mM HEPES, 100 mM NaCl, pH 7.5). The purity of the enzymes was determined to be > 98% by LC-ESI-MS and electrophoresis analyses.

(7) Cell labeling and confocal microscopy imaging.

HEK293 or HeLa cells were grown on cover slips in DMEM and transiently transfected with RA1, RA1:RFP:NLS, H2B:RA1:His or H2B:RA2:His using X-treme Gene 9 DNA transfection reagent (Roche) at 65-70% confluency. The proteins were overexpressed for 48 h, except for RA1 which was overexpressed for 24 h. The medium was replaced with DMEM media containing 10 μM **P4** and incubated for 10 min at 37°C. The cells were then washed twice with fresh DMEM media, incubated for additional 15 min in DMEM and fixed using 4% paraformaldehyde for 15 min. The sample slips were then washed with DPBS and stained with DAPI nuclear staining reagent (NucBlue Fixed Cell ReadyProbes Reagent, Molecular Probes) for 5 min. *SlowFade* Gold Antifade Reagent (Molecular Probes) was used to prepare the imaging samples. Confocal images were taken using a Zeiss LSM 710 laser scanning confocal microscope (LSCM).

For live cell imaging of mammalian cells without fixation, the cells were grown to ~50% confluency on chambered coverglasses (Thermo Scientific™ Nunc™ Lab-Tek™) and transfected with RA1 for 24 h. **P4** (10 μM) in DMEM was incubated with the cells for 10 min. The media was then replaced with fresh DMEM at 37 °C. Two drops of nuclear staining (NucBlue Live Cell ReadyProbes, Life Technologies) were added in each chamber. The cells were then washed twice with DPBS after a 30-min diffusion period. Images were taken in DPBS using a confocal microscope.

For live cell imaging with *E. coli* K12 cells, *E. coli* K12 cells (non-transformed (NT) or transformed with RA1) were grown in LB media to OD_{600} = 0.6 and induced with IPTG (75 μM) for 30 min. The cells were harvested by centrifugation and the pellet was resuspended in DPBS containing **P4** (10 μM). After incubation at 37 °C for 30 min,

the cells were washed three times with fresh DPBS before plating on chambered coverglasses (Thermo Scientific™ Nunc™ Lab-Tek™) for imaging.

(8) Plasmids and vectors constructions.

a) *Construction of **RA1** vector for mammalian cell expression.* The RA1 gene was subcloned from the previously reported pET29b (+) vector[4] to the pENTR1A vector. Subsequently, the RA1 gene was shuttled from pENTR1A into the pT-Rex-DEST30 vector using LR clonase II (Invitrogen, Inc.) recombination to generate the pT-Rex-RA1 vector for expression in mammalian cells.

b) *Construction of **RA1:RFP:NLS** vector.* The previously reported vector peDHFR:L28C:RFP:NLS was a generous gift from Prof. Virginia Cornish at Columbia University.[8] The RA1 gene was subcloned in the vector peDHFR:L28C:RFP:NLS to generate pRA1:RFP:NLS using the Polymerase Incomplete Primer Extension (PIPE) cloning method.[9]

c) *Construction of **H2B:RA1:His** and **H2B:RA2:His** vectors*. The previously reported vector pH2B:eDHFR:L28C was provided by Prof. Virginia Cornish at Columbia University.[10] The PIPE cloning method was used to subclone the RA1:His and RA2:His genes into this vector to generate the pH2B:RA1:His and pH2B:RA2:His vectors, respectively.

All tagged POIs are separated from RA1 or RA2 by a flexible linker to prevent possible perturbation of function, as previously reported.[8,10]

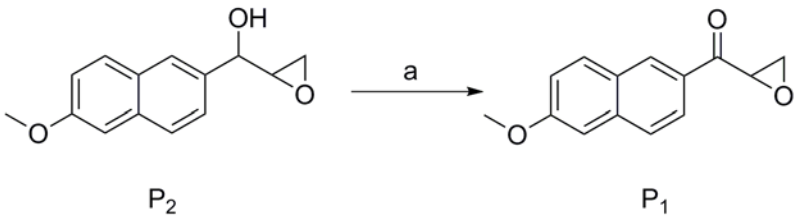
(9) General Synthetic and Chromatographic Methods.

All reagents and anhydrous solvents of commercial grade were used as received unless stated otherwise. All reactions were carried out under Argon gas unless stated otherwise. Reactions were monitored by TLC analysis with detection by UV-absorption (254 nm) using Merck glass sheets precoated with silica gel 60. Flash column chromatography was performed using Silica Flash P60 silica gel in the indicated solvent systems. ^{1}H NMR and ^{13}C NMR spectra were obtained on a Bruker DRX-600 with CryoProbe (600/150 MHz) or an Inova-400 (400/100 MHz) spectrometer in the given solvent. Chemical shifts are reported as δ- values in ppm relative to the DMSO residual solvent peak or tetramethylsilane (TMS) as an internal standard. Coupling constants are provided in Hz. All given ^{13}C spectra are proton-decoupled. High resolution mass spectra were recorded with an Agilent 6210 ESI-TOF system coupled to an Agilent 1100 LC stack. LC/MS analysis was performed on Agilent G1956 single quadrapole mass spectrometer coupled with an Agilent 1100 LC stack. Solvents used are H_2O/0.1% formic acid and acetonitrile/0.1% formic acid. Preparative reverse phase HPLC was used for small scale purification on a C18 column. Solvents used are H_2O/0.2% trifluoroacetic acid and acetonitrile/0.2% trifluoroacetic acid.

(10) Synthesis of RA probes.

The syntheses and characterization of Probes **P2**, **P3** were previously reported.[4] **P1** and **P4** were obtained as follows. To generate probe **P1**, **P2** was first synthesized and further oxidized into ketone (**Scheme S1**). **P4** was generated by conjugating a cell

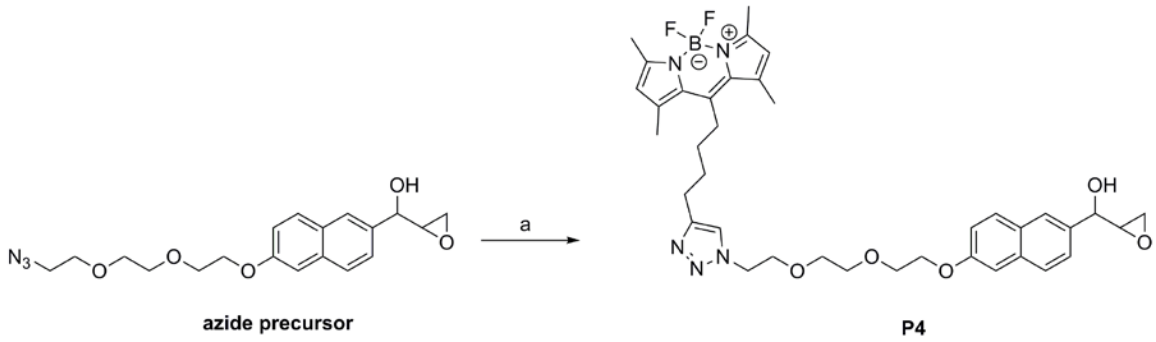
permeable fluorophore BODIPY-alkyne[11] with the azide precursor through "click" chemistry. (**Scheme S2**)



Scheme S1. Synthetic approach for **P1**. Reagents and conditions: (a) TPAP, NMO, 4 Å MS in CH_2Cl_2.

Probe **P1**

To an anhydrous CH_2Cl_2 solution (0.5 mL), the epoxide **P2** (46.1 mg, 0.2 mmol), *N*-methyl morpholine oxide (35.1 mg, 0.3 mmol), TPAP (3.5 mg, 0.01 mmol), and powdered 4 Å molecular sieves (0.1 g) were added and stirred at 25 °C for 1 h. The mixture was then filtered, concentrated *in vacuo*, and purified by flash column (Hexane: Ethyl acetate=1:1), affording white crystalline solid (34 mg, 75%) as **P1**. ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.54 (s, 1H), 8.03 (d, *J* = 8.6Hz, 1H), 7.87 (d, *J* = 9.0 Hz, 1H), 7.79 (d, *J* = 8.6 Hz, 1H), 7.22 (dd, *J* = 9.0, 2.5 Hz, 1H), 7.16 (d, *J* = 2.5 Hz, 1H), 4.37 (dd, *J* = 4.5, 2.5 Hz, 1H), 3.95 (s, 3H), 3.17 (dd, *J* = 6.5, 4.5 Hz, 1H), 3.04 (dd, *J* = 6.5, 2.5 Hz, 1H). ^{13}C NMR (101 MHz, $CDCl_3$) δ 194.24, 160.38, 137.93, 131.50, 131.04, 130.51, 127.89, 127.66, 124.62, 120.24, 106.03, 55.69, 51.32, 47.88. HRMS: ($M+H^+$) calcd for $C_{14}H_{13}O_3$ 229.0859 found 229.0861.

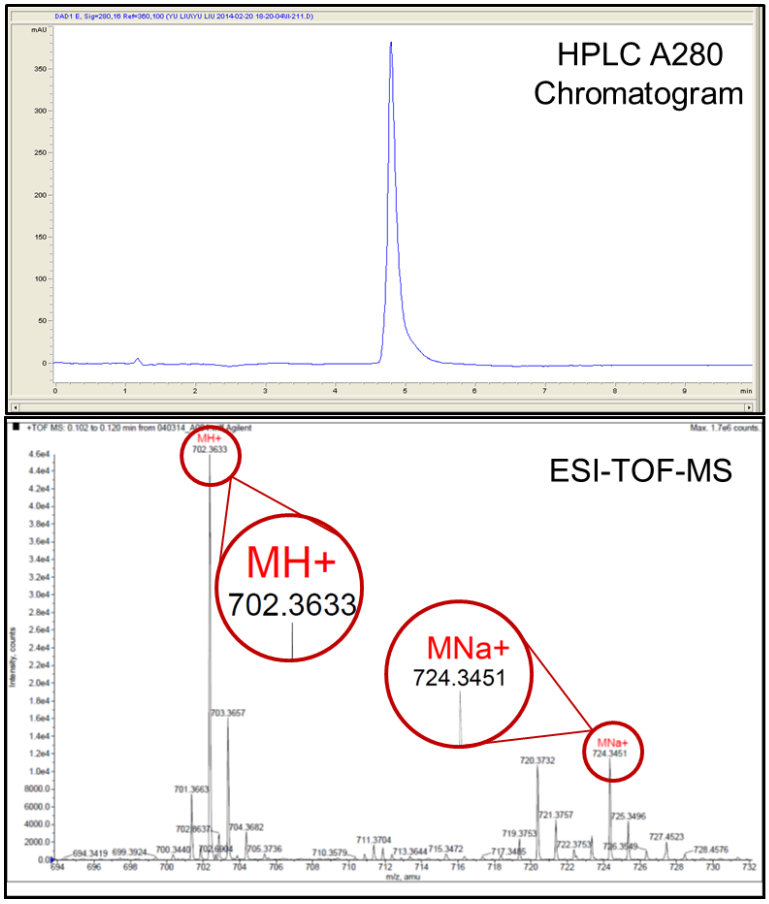


Scheme S2. Synthetic approach for **P4**. Reagents and conditions: (a) $CuSO_4$, Sodium ascorbate, BODIPY-alkyne dye, Tol/tert-BuOH/H_2O.

Probe **P4**

Compound BODIPY-alkyne (99 mg, 0.3 mmol) and the azide precursor (37.5 mg, 0.1 mmol) were dissolved in Tol/tert-BuOH/water (1:1:1) solvent system. $CuSO_4$ aq. solution (0.02 mmol) and sodium ascorbate aq. solution (0.03 mmol) were added into the reaction. The reaction mixture was stirred vigorously at 80 °C for 1 hour. The reaction was monitored by TLC until the azide precursor was completely consumed. The reaction mixture was then loaded on prep-HPLC. ~ 7 mg dark orange (~10%) compound was obtained. The purity of **P4** was determined by LCMS and HRMS. HRMS: ($M+H^+$) calcd for $C_{38}H_{47}BF_2N_5O_5$ 702.3633 found 702.3633. (No hydrolytic product of epoxide mass was observed).

Purity Determination of **P4**



References:

(1) Weiss, M. S.; Hilgenfeld, R. *J Appl Crystallogr* **1997**, *30*, 203.
(2) Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta crystallographica. Section D, Biological crystallography* **2010**, *66*, 12.
(3) Mizukami, S.; Watanabe, S.; Akimoto, Y.; Kikuchi, K. *J Am Chem Soc* **2012**, *134*, 1623.
(4) Liu, Y.; Tan, Y. L.; Zhang, X.; Bhabha, G.; Ekiert, D. C.; Genereux, J. C.; Cho, Y.; Kipnis, Y.; Bjelic, S.; Baker, D.; Kelly, J. W. *Proc. Natl. Acad. Sci. U S A* **2014**.
(5) Kabsch, W. *Acta crystallographica. Section D, Biological crystallography* **2010**, *66*, 125.
(6) Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. *Acta crystallographica. Section D, Biological crystallography* **2002**, *58*, 1948.
(7) Emsley, P.; Cowtan, K. *Acta crystallographica. Section D, Biological crystallography* **2004**, *60*, 2126.
(8) Gallagher, S. S.; Sable, J. E.; Sheetz, M. P.; Cornish, V. W. *Acs Chem Biol* **2009**, *4*, 547.
(9) Klock, H. E.; Lesley, S. A. *Methods Mol Biol* **2009**, *498*, 91.
(10) Chen, Z.; Jing, C.; Gallagher, S. S.; Sheetz, M. P.; Cornish, V. W. *J. Am. Chem. Soc.* **2012**, *134*, 13692.
(11) Verdoes, M.; Hillaert, U.; Florea, B. I.; Sae-Heng, M.; Risseeuw, M. D. P.; Filippov, D. V.; van der Marel, G. A.; Overkleeft, H. S. *Bioorg Med Chem Lett* **2007**, *17*, 6169.