Supplemental Information

‘Turn-On’ Protein Fluorescence: In Situ Formation of Cyanine Dyes
Ipek Yapici, Kin Sing Stephen Lee, Tetyana Berbasova, Meisam Nosrati, Xiaofei Jia, Chrysooula Vasileiou, Wenjing Wang, Elizabeth M. Santos, James H. Geiger,* and Babak Borhan

Department of Chemistry, Michigan State University, East Lansing, MI 48824
babak@chemistry.msu.edu

Table of Contents:
A  Mutagenesis of CRABPII Gene  ................................................................. S2-S3
B  Expression of CRABPII Proteins ................................................................. S3-S4
C  Protein Characterization ............................................................................ S4-S11
D  Cloning and Expression of mRFP .............................................................. S11-S12
E  Live Cell Imaging in Bacteria .................................................................... S12-S14
F  Photobleaching Experiment in Bacteria ...................................................... S14
G  Synthesis of Merocyanine Aldehyde ........................................................ S15-S16
H  Mutagenesis at Position 59 ......................................................................... S17
I  Denaturation Studies of CRABPII Mutants ................................................ S18
J  Crystallization Data ..................................................................................... S19-S22
K  References .................................................................................................. S22
L  NMR Spectra .............................................................................................. S23-S24
A. Mutagenesis of CRABPII Gene

1A. Polymerase Chain Reaction Protocol

Mutagenesis of all CRABPII proteins was performed with PET17b-CRABPII vector, as described below. The polymerase chain reaction (PCR) was done according to specified PCR conditions (see Table below). PCR products were subjected to Dpn I (New England BioLabs®) digestion in order to destroy the original template DNA. 20 units of Dpn I restriction enzyme was added to 50 µL of PCR sample and the reaction mixture was incubated at 37 °C for 1 h. The resulting mixture was transformed to XL-1 competent cells for DNA amplification and purification.

| PCR conditions |                          |
|----------------|---------------------------|
| Total Reaction Volume | 50 µl                   |
| Template (DNA plasmid)  | 70 ng (x µl)         |
| Primer forward          | 20 pmol (y µl)          |
| Primer reverse          | 20 pmol (z µl)          |
| dNTP                    | 1 µl                     |
| DMSO                    | 5 µl                     |
| 10× pfu buffer          | 5 µl                     |
| Pfu Turbo (DNApolymerase) | 1 µl                   |
| DI water                | 38-(x+y+z) µl           |

| PCR program | 
|-------------|
| 1×          | 94 °C          | 3 min          |
|             | 94 °C          | 20 sec         |
| 20×         | temperature 3-5 °C lower than primer melting temperature | 50 sec |
|             | 72 °C          | 4 min 30 sec   |
| 1×          | 72 °C          | 10 min         |
| 1×          | 10 °C          | 5 min          |

2A. List of Primer

Site-directed mutagenesis was performed using KL-CRABPII-PET17b and KLE-CRABPII-PET17b plasmids that were generated in our previous work. The primers used for the mutations described in the manuscript were:

**L28W**
Forward: 5’- GGG GTG AAT GTG ATG **TGG** AGG AAG -3’
Reverse: 5’- CTT CCT **CCA** CAT CAC ATT CAC CCC-3’

**A32W**
Forward: 5’- AGG AAG ATT **TGG** GTG GCT GCA GCG -3’
Reverse: 5’- CGC TGC AGC CAC **CCA** AAT CTT CCT -3’

**S37W**
Forward: 5’- GCA GCG **TGG** AAG CCA GCA GTG-3’
Reverse: 5’- CAC TGC TGG CTT **CCA** CGC TGC-3’

**R59W**
Forward: 5’- CC ACC GTG **TGG** ACC ACA GAG-3’
Reverse: 5’- CTC TGT GGT **CCA** CAC GGT GG-3’
**R59L**
Forward: 5'-CC ACC GTG **TTG** ACC ACA GAG -3’
Reverse: 5'- CTC TGT GGT **CAA** CAC GGT GG-3’

**R59H**
Forward: 5'-CC ACC GTG **CAT** ACC ACA GAG -3’
Reverse: 5'- CTC TGT GGT **ATG** CAC GGT GG -3’

**R59F**
Forward: 5'-CC ACC GTG **TTC** ACC ACA GAG -3’
Reverse: 5'- CTC TGT GGT **GAA** CAC GGT GG -5’

**V76W**
Forward: 5'- **G** GAG CAG ACT **TGG** GAT GGG AGG -3’
Reverse: 5'- CCT CCC ATC **CCA** AGT CTG CTC C-5’

**L121D**
Forward: 5'- **G** GAA CTG ATC **GAT** ACC ATG ACG GC-3’
Reverse: 5'- GC CGT CAT GGT **ATC** GAT CAG TTC C-3’

### 3A. Transformation of PCR product

XL-1 Blue super (Novagen®) competent cells (100 µL) were thawed on ice and 7 µL of PCR product was added to the cell solution. After incubation on ice for 30 minutes, the cells were heat-shocked at 42 °C for 45 seconds and 400 µL of Luria-Bertani broth (LB) was added. The resulting mixture was incubated at 37 °C for 30 minutes. The cells were centrifuged at 8,000 rpm for 1 minute and the supernatant was discarded. The cells were gently spread on LB agar plate containing ampicillin (100 µg/mL) and tetracycline (7.5 µg/mL) and incubated at 37 °C for 16 h.

### 4A. DNA Purification

A single colony was used to inoculate 15 mL of LB solution containing ampicillin (100 µg/mL) and tetracycline (7.5 µg/mL). The cell culture was grown at 37 °C for 12 to 14 h and centrifuged at 13,000 rpm for 1 min. DNA isolation was performed using Promega Wizard® Plus SV Miniprep (A1330) DNA purification kit according to the manufacturers directions, save that 50 µL of Nuclease-Free water was used for DNA elution instead of the recommended 100 µL. DNA concentration was measured by using the Thermo Scientific NanoDrop™ 1000 Spectrophotometer. The average DNA concentration in a 50 µL solution was 120 ng/µL. A sample containing at least 700 ng of DNA was transferred into another eppendorf tube and sequenced by the Research Technology Support Facility at Michigan State University. A primer corresponding to the T7 promoter was used for sequencing.

### B. Expression of CRABPII Proteins

The target gene (100 ng of DNA for 100 µL of cell solution) was transformed into BL21(DE3) pLysS (Invitrogen™) E. coli competent cells for protein expression. A protocol similar to that described for transformation of PCR products was used. The only difference was that after heat-shock at 42 °C, the cell solution was spread directly onto an LB plate with the proper antibiotics (ampicillin 100 µg/mL and chloroamphenicol 170 µg/mL).

A single colony was used to inoculate 1 L of modified Luria Broth with ampicillin (100 µg/mL) and chloroamphenicol (170 µg/mL). The modified-LB solution contains: Tryptone (10 g), yeast extract (8 g), NaCl (5 g) and 1 L of dH₂O. The cell culture was grown at 37 °C while shaking for 9 h. The overexpression was induced by the addition of 1 mL of 1M IPTG solution into 1 L cell culture (overall concentration 1.0 mM of IPTG. The culture was shaken at 23 °C for 18 h at 220 rpm.
1B. Protein Isolation and Purification

The cells, from the culture shaken at 23 °C for 18 h, were harvested at 4 °C by centrifugation for 12 min at 5,000 rpm. The supernatant was discarded and the cells were resuspended in 50 mL of Tris buffer (10 mM Tris-HCl, pH=8.0), lysed by ultrasonication (Power 60%, 1 min x 3), treated with DNAse (300 units/ 50 mL suspension) and MgCl₂ (0.12 mmol/ 50 mL suspension) were added and centrifuged at 16 °C (5,000 rpm, 20 min).

The protein was purified by ion exchange chromatography using Q Sepharose™, Fast Flow resin (column diameter: ~ 4 cm; height: ~ 10 cm) at 4 °C. The FastQ column was pre-washed twice with 100 mL of a solution containing 2 M NaCl and 10 mM Tris-HCl, pH=8.0, and then equilibrated with 200 mL of Tris buffer (10 mM Tris-HCl, pH=8.0). The cell supernatant was then loaded onto the column, the column washed with 40 mL of Tris buffer, and the protein eluted with 45 mL of the elution buffer (10 mM Tris-HCl, 200 mM NaCl, pH=8.0). The eluent was desalted with 8.0). The resulting dilute solution was then concentrated to ~25 mL. Further purification was performed with Fast Protein Liquid Chromatography (Biologic Duo Flow, Biorad).

2B. Fast Protein Liquid Chromatography (FPLC) Protocol

Column: Source 15Q (quaternary ammonium strong anion exchange)
Buffer B: 10 mM Tris-HCl, 2 M NaCl, pH=8.1

| Parameters | Description | pH | Volume of protein | Flow rate |
|------------|-------------|----|-------------------|-----------|
| 10.00 mL, 3.00 mL/min | Isocratic Flow | pH=8.1, 0% B | Dynamic Loop | 'Volume of protein' mL, 2.0 mL/min |
| 10.00 mL, 3.00 mL/min | Isocratic Flow | pH=8.1, 0→4% B | Linear Gradient | pH=8.1, 4% B |
| 20.00 mL, 3.00 mL/min | Isocratic Flow | pH=8.1, 4% B | Linear Gradient | pH=8.1, 4→8% B |
| 10.00 mL, 3.00 mL/min | Linear Gradient | pH=8.1, 8% B | Linear Gradient | pH=8.1, 8% B |
| 10.00 mL, 3.00 mL/min | Linear Gradient | pH=8.1, 8→15% B | Linear Gradient | pH=8.1, 8→15% B |
| 15.00 mL, 3.00 mL/min | Linear Gradient | pH=8.1, 15% B | Isocratic Flow | pH=8.1, 15% B |
| 10.00 mL, 3.00 mL/min | Isocratic Flow | pH=8.1, 15→75% B | Isocratic Flow | pH=8.1, 100% B |
| 20.00 mL, 3.00 mL/min | Isocratic Flow | pH=8.1, 100% B | Isocratic Flow | pH=8.1, 0% B |
| 30.00 mL, 3.00 mL/min | Isocratic Flow | pH=8.1, 0% B | End protocol | |

Most of the CRABPII variants were collected between 8% and 15% Buffer B elution.

C. Protein Characterization

Spectroscopic characterizations of purified proteins were carried out using a Fluorolog 3 spectrfluorimeter (Horiba Scientific) and a Cary 300 Bio WinUV UV-Vis spectrophotometer (Varian Instruments).

1C. Extinction Coefficient Determination and Yield Calculation

The extinction coefficients at 280 nm of the mutants were measured according to the method described by Gill and Vonhipple.[2] The theoretical extinction coefficient (ε̂) is calculated based on the following equation:

ε̂ = a*ε_Trp + b*ε_Tyr + c*ε_Cys

where a, b and c are the number of Trp, Tyr and Cys residues, respectively, in the protein sequence. The extinction coefficients of these residues were determined previously (ε_Trp = 5690 M⁻¹ cm⁻¹, ε_Tyr = 1280 M⁻¹
cm$^{-1}$, $\varepsilon_{\text{Cys}} = 120$ M$^{-1}$ cm$^{-1}$). The concentration of the protein was measured in 6 M-guanidine HCl (denaturing solution) solution following Beer's Law:

$$Abs_{\text{280 Den}} = b \cdot c \cdot \varepsilon_{\text{Theor}}$$

where $b$ is the cuvette path length and $c$ is the concentration of protein. The absorbance of the same amount of protein was measured in phosphate buffered saline (0.24 gKH$_2$PO$_4$, 1.45 g Na$_2$HPO$_4$, 0.8 g NaCl, 0.2 g KCl, pH=7.3, total volume 1 L) with the same cuvette. The experimental extinction coefficient of the protein can be derived using:

$$\varepsilon_{\text{Exper}} = (Abs_{\text{280 Nat}} \div Abs_{\text{280 Den}}) \cdot \varepsilon_{\text{Theor}}$$

| CRABPII Variants | Extinction Coefficient $\varepsilon_{\text{Exper}}$ (M$^{-1}$cm$^{-1}$) |
|------------------|---------------------------------|
| 1 R132K:R111L    | 20.072                          |
| 2 R132K:R111L:R59W | 25.983                          |
| 3 R132K:R111L:L121D:R59W | 25.995                          |
| 4 R132K:R111L:L121E:R59W | 23.131                          |
| 5 R132K:R111L:L121E:R59H | 20.216                          |
| 6 R132K:R111L:L121E:R59F | 20.105                          |
| 7 R132K:R111L:L121E:R59L | 19.655                          |
| 8 R132K:R111L:L121E:V76W | 24.845                          |
| 9 R132K:R111L:L121E:S37W | 23.530                          |
| 10 R132K:R111L:L121E:L28W | 23.353                          |
| 11 R132K:R111L:L121E:M123W | 25.153                          |
| 12 R132K:R111L:L121E:A32W | 24.148                          |
| 13 R132K:R111L:L121E:V76W:L28W | 30.839                          |
| 14 R132K:R111L:L121E:R59W:L28W | 27.187                          |
| 15 R132K:R111L:L121E:R59W:M123W | 31.965                          |
| 16 R132K:R111L:L121E:R59W:A32W | 29.252                          |

The yield of protein expression was calculated by measuring the absorbance value in PBS at 280 nm. The protein concentration was diluted by half (500 µL of native buffer mixed with 500 µL of protein solution) and the following calculations were applied:

$$c = (Abs_{\text{280 Nat}} \div \varepsilon_{\text{Exper}}) \cdot 2$$

$$\text{yield (g)} = c \cdot V \cdot M_W$$

where $c$ is the concentration of protein solution (mol/L), $V$ is the volume of protein solution (L) and $M_W$ is the molecular weight of the protein (g/mol). The typical yield for CRABPII mutants is 40 mg/L of modified Luria Broth.

2C. pK$_a$ Measurements

For pK$_a$ measurements, protein (20 µM in PBS) was incubated with merocyanine aldehyde (0.3 equiv) at room temperature for 40 min. The UV spectrum and the pH of the solution were recorded after addition of each portion of 5 M NaOH solution. The $\lambda_{\text{max}}$ versus pH was plotted, a polynomial fit of the data (3$^{rd}$ power) was applied and the pK$_a$ value was calculated when the second derivative of the equation was set to zero (see Figure S1).

$$y = M_3x^3 + M_2x^2 + M_1x + M_0$$
$$y' = 3M_3x^2 + 2M_2x + M_1$$
$$y'' = 6M_3x + 2M_2$$

If $y''=0$ then $pK_a (x) = - [M_2 \div 3M_3]$
3C. Kinetic Measurements

**Determination of Relative Rates of CRABPII Mutants Binding with Merocyanine Aldehyde 1**

Kinetic measurements were performed at 23 °C, with the aid of a temperature controller for the UV-vis spectrophotometer. The protein (20 µM in PBS) was mixed with merocyanine aldehyde 1 (0.3 equiv) and the increase in absorbance of each CRABPII mutant at their respective λ<sub>max</sub> was recorded at 0.5 second intervals for 80 min to 180 min. The data was fit to an exponential rise (pseudo first order) with KaleidaGraph 4.1.3 using the following equation:

\[ A = A_0 \times (1 - e^{-kt}) + c \]

**Figure S1.** Basic titration of the iminium formed between merocyanine aldehyde 1 and R132K=R111L-CRABPII double mutant. The pK<sub>a</sub> was determined through the polynomial fit of the data (Abs at 600 nm vs. pH) as described in the text.

**Figure S2.** Rate of PSB formation with merocyanine aldehyde 1 and CRABPII mutants, exhibiting first order kinetics.
where $A$ is the absorbance value at each recorded time-point, $A_0$ is the final absorbance value after the complex formation is complete, $k$ is the rate constant, $t$ is time after addition, $A$ is the recorded absorbance and $c$ is a free constant, which accounts for the time delay from merocyanine 1 addition to the point when recording was started. This equation was rewritten in KaleidaGraph in the following format:

$$y = m1*(1-exp(-m2*m0))+m3$$

where the calculated value of $m2$ is the rate constant (see Figure S2A for details).

**Determination of Second-Order Rate Constant for KLE:R59W Binding with Merocyanine Aldehyde 1**

Measurement of the second order rate constant for the complexation of KLE:R59W with merocyanine 1 followed previously established protocols.\textsuperscript{[14,15]} Fluorescence measurements were performed by a FluoroLog 3 instrument, equipped with a temperature controller (Horiba Jobin Yvon). KLE:R59W (100 nM) was reacted with an excess amount of merocyanine 1 in PBS buffer (37 °C) at different concentrations (1 to 5 μM), and the increase of fluorescence intensity was monitored for each reaction. The fluorescence intensity of the samples was collected at 0.5 second intervals, with excitation at 565 nm (1 nm slit width) and emission at 616 nm (12 nm slit width). The raw fluorescence intensity data were converted into complexation fraction using the following equation:

$$[\text{complexation fraction}] = \frac{(F_t - F_0)}{(F_{\text{max}} - F_0)}$$

where $F_t$, $F_0$, and $F_{\text{max}}$ are the observed, initial, and maximum fluorescence intensities, respectively. Since the reaction conditions lead to pseudo-first order kinetics $[\text{merocyanine1}] >> [\text{protein}]$, the data at each merocyanine concentration was fit to a first-order kinetic equation as follows:

$$[\text{Complexation fraction}] = 1 - \exp(-k_{\text{obs}}*t)$$

The latter equation was rewritten in KaleidaGraph as shown below:

$$y = m1*(1-exp(-m2*m0))+m3$$

where the calculated value of $m2$ is the pseudo-first-order rate constant ($k_{\text{obs}}$).

The calculated $k_{\text{obs}}$ for each reaction was plotted vs. the merocyanine concentration, leading to a linear fit. The resultant slope yields the second-order rate constant ($k_2$) for the complex formation:

$$k_{\text{obs}} = k_2[\text{merocyanine 1}]$$

**Determination of Half Life of Complexation of KLE:R59W with Merocyanine Aldehyde 1**

Half-life binding measurements ($t_{1/2}$) were determined under stoichiometric conditions (5 μM KLE:R59W and merocyanine) in PBS buffer at 37 °C. The increase in fluorescence intensity, indicative of binding and PSB formation, was recorded at 0.5 second intervals, with excitation at 565 nm (1 nm slit width) and emission at 616 nm (12 nm slit width). The data was fit to a second-order rate equation derived as shown below:
The reaction of $A + B \rightarrow C$ is described by the rate law:

$$\frac{d[C]}{dt} = k_2 [A][B]$$

The initial concentrations of reactants are $[A_0] = [B_0] = 5 \times 10^{-6}$ M. As the reaction continues, both $A$ and $B$ get consumed simultaneously in order to form complex $C$.$[A] = [B] = [5 \times 10^{-6} - C]$

$$\frac{d[C]}{dt} = k_2 (5 \times 10^{-6} - C)^2$$

$$\frac{d[C]}{[5 \times 10^{-6} - C]^2} = k_2 dt$$

$$\int \frac{d[C]}{[5 \times 10^{-6} - C]^2} = \int k_2 dt$$

$$\frac{1}{[5 \times 10^{-6} - C]} = k_2 t + m$$

where $m$ is a constant

$$C = 5 \times 10^{-6} \cdot \frac{1}{[k_2 t + m]}$$

The equation is rewritten in KaleidaGraph:

$$y = m_3 - \frac{1}{m_2 m_0 + m_1}$$

where $m_3 = 5 \times 10^{-6}$, $m_2 = k_2$, $m_0 = t$ and $m_1 = 1$

$C$ is directly proportional to fluorescence intensity (cps), plotted vs. time (s). The maximum fluorescence intensity is adjusted to 5 μM, which is the final concentration of the complex, in order to fit the data shown below (Figure S3). From the plot ($t_{1/2}$) was calculated for a second-order reaction where $t_{1/2} = 1/k_2[\text{concentration}]$. 

S8
Rate Comparison for Iminium Formation between KLE:R59W vs. BSA

Combination of the high affinity binding of hydrophobic, retinal-like molecules, and the active site Lys residue engineered within the CRABPII mutants, leads to a higher rate of iminium formation with the target proteins as compared to non-specific binding with other proteins. This is demonstrated below (Figure S4a), where upon incubation of merocyanine 1 (5 \( \mu \text{M} \)) with BSA (5 \( \mu \text{M} \)) in comparison to KLE:R59W (5 \( \mu \text{M} \)) in PBS (37 °C), the latter achieves 7% of the fluorescence output as compared to the former within the first minute of the reaction (excitation, 565 nm, 1 nm slit width; emission 616 nm, 12 nm slit width). During the same time period (1 min), KLE:R59W has reached over 66% of its total fluorescence output, clearly indicating the higher level of specificity for binding merocyanine 1. The higher specificity for binding KLE:R59W is maintained even at much higher BSA concentration (10x) as illustrated in Figure S4b.

![Figure S3](image)

**Figure S3.** The fit of fluorescence data correlated to the concentration of complex formation with merocyanine aldehyde 1 and KLE:R59W.

![Figure S4](image)

**Figure S4.** Comparison for rate of iminium formation between KLE:R59W and BSA with merocyanine 1. a. Time course binding of 1 with KLE:R59W and BSA (5 \( \mu \text{M} \) each) is illustrated. b. The previous experiment is repeated, however with 10 fold excess of BSA (50 \( \mu \text{M} \)).
4C. Quantum Yield Measurements

All quantum yields were determined by comparing the integrated area of the corrected emission spectra for each CRABPII-merocyanine complex with the corresponding integrated area obtained from a solution of fluorescent standards Oxazine-170 (purchased from Across Organics, lot# A0098689) or Oxazine-1 (purchased from Exciton, lot# 26424). Integrals at various concentrations were plotted against the absorbance obtained at the wavelength corresponding to the excitation wavelength and the slope of the curve obtained for protein-merocyanine complexes was compared to the slope of the curves found for reference fluorophores. We kept all the absorbance values at excitation wavelength under 0.1 absorbance unit. All samples were excited at 565 nm and the emission was collected from 575 nm to 800 nm. Temperature was set to 22 °C and kept constant with a temperature controller during measurements. The absolute quantum yields for reference fluorophores were taken to be 0.579 for Oxazine-170 and 0.141 for Oxazine-1 in ethanol. Quantum yield calculations were corrected for the refractive index.

**Figure S5.** Quantum yield calculations based on total fluorescence emission (integrated area from 575 nm to 800 nm, excitation at 565 nm) as a function of absorption at 565 nm. Slope of the line is correlated to the quantum efficiency as described in the text.
between 575 nm to 800 nm. As described above, the excitation wavelength was set to 565 nm and the emission spectrum was collected from 2D solution was concentrated back to cellulose, filter code: YM10, diameter 63.5 mm, NMWL: 10,000). The protein was first concentrated to buffer containing 250 mM 

μM in PBS were incubated with merocyanine aldehyde 1 (0.8 mM stock solution in EtOH) to a final concentration of 2 μM. The solution was kept at room temperature for 1.5 h by which time maximum PSB absorbance is observed. The absorbance value of each CRABPII/1 complex was recorded at the corresponding λmax and the extinction coefficient was determined via application of the Beer’s law:

\[
\frac{\text{Abs}_{\text{CRABPII/1}}}{\epsilon_{\text{CRABPII/1}}} = b \cdot c \cdot \frac{\epsilon_{\text{CRABPII/1}}}{\text{Abs}_{\text{CRABPII/1}}} = (\text{Abs}_{\text{CRABPII/1}} + 2 \cdot 10^{-6}) \cdot \text{M}^{-1} \cdot \text{cm}^{-1}
\]

where b is the cuvette path length (1 cm), c is the concentration of CRABPII/1 complex (2 μM), and \(\text{Abs}_{\text{CRABPII/1}}\) is the absorbance value of PSB at λmax of CRABPII/1 complex.

**5C. Calculation of Extinction Coefficient of CRABPII/1 Complexes**

The extinction coefficient of merocyanine aldehyde 1 was measured by UV-vis spectroscopy and found to be 77,000 M⁻¹cm⁻¹. Protein solutions (10 μM in PBS) were incubated with merocyanine aldehyde 1 (0.8 mM stock solution in EtOH) to a final concentration of 2 μM. The solution was kept at room temperature for 1.5 h by which time maximum PSB absorbance is observed. The absorbance value of each CRABPII/1 complex was recorded at the corresponding λmax and the extinction coefficient was determined via application of the Beer’s law:

\[
\frac{\text{Abs}_{\text{CRABPII/1}}}{\epsilon_{\text{CRABPII/1}}} = b \cdot c \cdot \frac{\epsilon_{\text{CRABPII/1}}}{\text{Abs}_{\text{CRABPII/1}}} = (\text{Abs}_{\text{CRABPII/1}} + 2 \cdot 10^{-6}) \cdot \text{M}^{-1} \cdot \text{cm}^{-1}
\]

where b is the cuvette path length (1 cm), c is the concentration of CRABPII/1 complex (2 μM), and \(\text{Abs}_{\text{CRABPII/1}}\) is the absorbance value of PSB at λmax of CRABPII/1 complex.

**D. Cloning and Expression of mRFP**

The plasmid mRFP-Rab5 containing the mRFP1 sequence was obtained from Addgene (plasmid #14437). This plasmid was used as a template for PCR amplification with the following primers: forward- NdeI-(5'- ccg gtc gcc cat atg tcc tcc-3') and reverse- Xhol-(5'-gag atc tga gta ctt ctc gag ggc gcc ggt g-3'). The amplified DNA fragment was digested with NdeI and Xhol (New England Biolabs), and cloned into PET22b expression vector (Novagen), resulting in the plasmid PET22b-mRFP1-His6. The mRFP1 gene was sequenced by the first tier/DMSO sequencing method and the results were identical to those reported by Campbell et al.[5] mRFP1 expression followed the same procedures as described above for CRABPII mutants.

**1D. Protein Isolation and Purification**

Protein isolation and purification were performed as described previously, with modifications.[6] Cells, harvested from culture, were centrifuged at 4 °C for 12 min at 5,000 rpm. The supernatant was discarded and the bacterial pellet was resuspended in 50 mM KH₂PO₄, pH 8.0, and 0.02% sodium azide (buffer A). The resulting suspension was ultrasonicated (Power 60%, 1 min x 3) and centrifuged at 5,000 rpm for 20 min. The following steps were carried under red light to prevent photobleaching of the protein.

The supernatant was applied to a Ni²⁺-chelating Sepharose Fast Flow column and incubated for 15 min at 4 °C. The beads were washed with 40 mL of buffer A. The protein was eluted with 50 mL of elution buffer containing 250 mM of imidazole, 20 mM of Tris, pH=8.0. Buffer exchange was performed by an ultrafiltration cell under nitrogen pressure (~20 psi) equipped with a 10,000 MW cutoff filter (regenerated cellulose, filter code: YM10, diameter 63.5 mm, NMWL: 10,000). The protein was first concentrated to ~25 mL and then diluted to 160 mL with Tris buffer (10 mM Tris-HCl, pH=8.0). The resulting dilute solution was concentrated back to ~25 mL. Protein was stored at 4 °C in a sterilized tube covered with aluminum foil.

**2D. Quantum Yield Measurement**

Quantum yield of mRFP1 was measured under identical conditions with CRABPII variants. As described above, the excitation wavelength was set to 565 nm and the emission spectrum was collected between 575 nm to 800 nm. The quantum yield was calculated as 27.2% based on comparison with standards, which is in agreement with published data.[6]
BL21(DE3) pLysS (Invitrogen™) were transformed with 100 ng of plasmid DNA expressing CRABPII variants under the T7 promoter. Cells were plated, grown overnight and single colonies were used for inoculation of 2 mL of Luria Broth for 6 h containing ampicillin and chloroamphenicol (ampicillin 100 µg/mL and chloroamphenicol 170 µg/mL). Expression was induced with IPTG (final concentration 1 mM) and the culture was shaken at room temperature overnight. For control experiment, 5 µL of competent BL21(DE3) pLysS cells were inoculated in 2 mL Luria Broth containing chloroamphenicol (170 µg/mL) and the culture was shaken at 37 °C for 6 h.

The cultures were heated to 37 °C for 10 min and merocyanine aldehyde, at a final concentration of 1 µM, was added to the pre-warmed cultures. The cells were immediately harvested by centrifugation at 5,000 rpm for 1 min and the supernatant was discarded. The pellets were washed three times with PBS (preheated to 37 °C) followed by 1 min centrifugation. After the last wash the cells were resuspended in PBS and plated on a glass coverslip. In order to stop bacteria from floating, the coverslip was heated to 40 °C for 1 min and covered with no.1 cover glass. Confocal imaging was performed using a Zeiss 510 Meta FCS inverted microscope with a 63x oil immersed objective. The sample was placed with the cover glass facing the objective and imaged with a 594 nm laser and a 615 nm long pass filter corresponding to the CRABPII-merocyanine complexes absorbance and emission. Kalman averaging 8 was applied during imaging. All the images have pseudo color. Figure S8 depicts images from E. coli cells incubated with merocyanine 1. The bright red fluorescence obtained within minutes of incubation illustrate the speed of complexation. The lack of fluorescence in the control panel (Figure S8e) indicates the specificity of merocyanine binding with the target protein.
Live Cell Imaging of Basal Expression of CRABPII Mutant in Bacteria

In the previous example expression of the CRABPII mutants in *E. coli* was induced with the strong T7 promoter, leading to high expression yields of the target protein. In order to illustrate the feasibility of visualizing low expressed CRABPII protein, we resorted to investigate basal, uninduced, E. coli systems that have low yields of the CRABPII mutants as a result of leaky expression.

BL21(DE3) pLysS (Invitrogen™) were transformed with 100 ng of plasmid DNA expressing CRABPII KLE:R59W variants under the T7 promoter. Cells were plated, grown overnight and single colonies were used for inoculation of 10 mL of Luria Broth for 6 h containing ampicillin and chloroamphenicol (ampicillin 100 µg/mL and chloroamphenicol 170 µg/mL). The culture was cooled down to room temperature and shaken for additional 8 h. Same protocol was followed for control experiment in which 5 µL of competent BL21(DE3) pLysS cells were inoculated in 10 mL Luria Broth containing chloroamphenicol (170 µg/mL).

Merocyanine aldehyde, at a final concentration of 1 µM, was added to 1 mL of grown cultures and the cells were immediately harvested by centrifugation at 5,000 rpm for 1 min and the supernatant was discarded. The pellets were washed three times with PBS (preheated to 37 °C) followed by 1 min centrifugation. The cells were resuspended in PBS and a small sample was plated on a glass coverslip. The sample was covered with no.1 cover glass and heated to 40 °C for a minute to fix the bacterial cells. Confocal imaging was performed using a Zeiss 510 Meta FCS inverted microscope with a 63x oil immersed objective. The sample was placed with the cover glass facing objective and imaged with a 594 nm laser and a 615 nm long pass filter. Kalman averaging 8 was applied during imaging. All the images have pseudo color. The lack of observable fluorescence in Figure S9 is again illustrative of insignificant background signal arising from nonspecific binding. On the other hand, bright red fluorescence is observed for basal expression of KLE:R59W.

**Figure S8.** Imaging CRABPII-merocyanine fluorescence in *E. coli* cells. Red fluorescence of KL-CRABPII mutant complexed with merocyanine 1 at 594 nm excitation (scale bar is 10 µm). b. Overlay of fluorescence and bright field images of KL-CRABPII mutant in *E. coli*. c. Red fluorescence of KL-CRABPII/1 at 594 nm excitation was imaged by 63x objective with five fold digitally enlarged and d. its overlay with bright field image (scale bar is 5 µm). e. Control panel shows no fluorescence upon merocyanine incubation with non-transformed BL21 *E. coli* cells (scale bar is 10 µm). f. Overlay of control fluorescence and bright field image.
Figure S9. Imaging KLE:R59W-merocyanine fluorescence in *E. coli* cells. a. Control panel shows no fluorescence upon merocyanine incubation with non-transformed BL21 *E. coli* cells. b. Overlay of control fluorescence and bright field image. c. Red fluorescence of KLE:R59W-CRABPII mutant complexed with merocyanine 1 at 594 nm excitation. d. Overlay of fluorescence and bright field images of KLE:R59W/1 complex in *E. coli*. 
F. Photobleaching Experiment in Bacteria

Laser scanning confocal microscopy (LSCM) photobleaching experiments were conducted with BL21(DE3) pLysS cells after expressing CRABPII mutants (R132K:R111L, R132K:R111L:L121E:R59W and R132K:R111L:L121E:R59W:L28W) and labeling with merocyanine. Cells expressing mRFP with no additional labeling were used for control experiment. The glass coverslips were prepared as described above. Photobleaching was performed using a 63x oil immersion objective (Zeiss 510 Meta FCS). Laser line HeNe594 was adjusted to 100% of output power (2.0 mW). Detector gain and amplifier offset were brought to maximum (1250 and 0.5, respectively). Selected frame size had 1024x1024 pixels and 5x digital zoom. The number of collected frames for each experiment was 200. A single cell having approximately the same dimensions and intensity under the fixed instrument settings was chosen for photobleaching assays. Fluorescence using the 594 nm laser was recorded with a 488/594 main dichroic mirror, 545 nm dichroic beam splitter and 615 nm longpass barrier filter. To produce comparable bleaching curves, we simply scale the raw time coordinates to normalize the intensity to 100 photons/sec of initial emission.

G. Synthesis of Merocyanine Aldehyde

The synthesis of ethyl phosphonate, compound 3 and Fisher’s base aldehyde, compound 4 were described previously. Both compound 3 and 4 can also be purchased from Sigma Aldrich (cas number: 87549-50-6 and 84-83-3, respectively). All the syntheses shown below were performed under red light and the products were stored at -20 °C, under an atmosphere of nitrogen gas.

\[ \text{O} \quad \begin{array}{c} \text{O} \quad \text{P(=O)(=O)} \quad \text{CN} \\ 3 \end{array} \]

\[ \begin{array}{c} \text{1. NaH, THF, 0 °C} \\ \text{2. room temp., 1h} \\ \text{CN} \\ 5 \end{array} \]

NaH (60% by weight, 0.65 g suspension containing 0.39 g NaH, 1.62 mmol) was placed in a dry 100 mL round bottom flask. Hexane (∼15 mL) was added and the mixture was stirred briefly, the solid particles were allowed to settle. The solvent-oil solution was removed with a glass pipette. Washing was performed twice. The round bottom flask was then capped and purged with nitrogen. The washed NaH was resuspended in anhydrous THF (35 mL). The stabilized nitrile ylide 3 (0.3 g, 1.39 mmol) was added to the NaH/THF suspension at 0 °C with vigorous stirring during 5 min. The reaction was left stirring in the ice-bath for 10 min and then slowly warmed to room temperature for an hour. Fisher’s base aldehyde, 4, (0.23 g, 1.16 mmol) was added dropwise during 5 min. Once the addition of the aldehyde was completed, the reaction was kept under nitrogen atmosphere and was left stirring for 8 h.

**Work up:** Brine solution was added to reaction mixture and the resulting aqueous layer was extracted three times with diethyl ether. Organic layers were combined and washed with deionized water. The combined organic phase was dried over anhydrous Na₂SO₄. The solvent was removed under reduced pressure to yield an oily mixture. Purification was performed with a 0.1% triethylamine basified silica gel column with 20% ethyl acetate/hexane mixture to yield the conjugated nitrile 5 as a mixture of two isomers (0.2 g overall yield, 65%).
Mutagenesis at Position 59

A number of mutants of Arg59 were prepared in order to examine the effect of changing amino acid size at the mouth of the protein. The following table lists a number of mutants that were prepared and examined. An exchange between Trp and His resulted in decreased quantum yield, while shifting the absorption maximum to 589 nm. The much smaller and more polar imidazole side chain most likely cannot provide sufficient favorable van der Waals contact with the bound ligand. It might also effect better solvation of the loop environment, thus reducing the hyrophobicity in the ligand’s neighborhood. In contrast, the replacement of Trp by Leu gave rise to a reduction in quantum yield while showing almost no change on absorption maximum. Leu could still participate in hydrophobic interactions with the bound merocyanine, similarly to Trp, but it might adopt more rotameric structures and result in poor rigidity around the ligand. The replacement of Trp by Phe, which can be considered as an equivalent exchange, was finally investigated in the context of the CRABPII variant KLE:R59F. A substitution of Trp by Phe resulted in a moderate change of absorption maxima and a slight decrease in quantum efficiency. In
general, larger hydrophobic amino acids led to protein complexes with merocyanine that exhibited larger extinction coefficients and quantum efficiencies, consistent with the idea that larger hydrophobic residues reduce conformational flexibility of the fluorophore, resulting in both higher extinction coefficients and quantum yields.

| Fluorescent Protein | Absorption $\lambda_{\text{max}}$ (nm) | Emission $\lambda_{\text{max}}$ (nm) | Quantum Yield (%) | Extinction Coefficient (M$^{-1}$cm$^{-1}$) |
|---------------------|----------------------------------------|-------------------------------------|-------------------|--------------------------------------------|
| 1 KLE:R59W          | 595                                    | 616                                 | 38.5              | 169,800                                    |
| 2 KLE:R59H          | 589                                    | 615                                 | 29.3              | 124,600                                    |
| 3 KLE:R59F          | 593                                    | 613                                 | 37.1              | 160,800                                    |
| 4 KLE:R59L          | 594                                    | 613                                 | 32.6              | 142,000                                    |

Table S1. Mutations of Arg59 and their effect on quantum efficiency of the complex with 1
I. Denaturation Studies of CRABPII Mutants

CRABPII mutants (4 µM) were incubated with 0.2 equiv of merocyanine aldehyde 1 in phosphate buffered saline (PBS) at room temperature. Once the CRABPII-merocyanine complex formation was complete, 5 µL of 0.2 M SDS solution was added to the protein solution. UV-vis spectra were taken before and after SDS treatment. Since denatured proteins no longer encapsulate the covalently bound ligand (lose tertiary structure), the maximum absorption of the resulting water exposed ligand shifts from the wavelength native to the complex to 584 nm. The latter observed $\lambda_{\text{max}}$ corresponds exactly to that observed for iminium 2 (Figure 1 in the main text) in PBS buffer with BSA (20 µM).

**Figure S11.** UV-vis spectra of a) KLE:R59W:L28W and b) KLE:V76W:L28W, before and after denaturation with SDS, illustrating the effect of protein environment on the absorption of the bound merocyanine dye. c) UV-vis spectra of 1 (492 nm), the SB of 1 with n-butyl amine (425 nm), and its corresponding PSB (574 nm). Addition of 20 µM BSA redshifts the $\lambda_{\text{max}}$ of the latter PSB to 583 nm, similar to the spectra observed in a and b after denaturation of CRABPII mutant complexes with SDS.
J. Crystallization Data

Table S2. X-ray Crystallographic Data and Refinement Statistics*

|                | KL/1 | KLD/1 | KLE/1 | KLE-R59W/1 |
|----------------|------|-------|-------|------------|
| Spacegroup     | P1   | P1    | P1    | P3 21      |
| Z (mol/asymm. unit) | 2    | 2     | 2     | 1          |
| Unit cell dimensions |      |       |       |            |
| a (Å)          | 34.8 | 34.7  | 34.6  | 55.8       |
| b (Å)          | 37.4 | 37.3  | 36.9  | 55.8       |
| c (Å)          | 60.3 | 60.3  | 58.5  | 108.2      |
| α (°)          | 104.3| 105.2 | 72.61 | 90.0       |
| β (°)          | 106.1| 105.3 | 76.37 | 90.0       |
| γ (°)          | 90.9 | 90.4  | 88.14 | 120.0      |
| Wavelength (Å) | 1.1272| 0.97850| 0.97850| 0.97850    |
| Resolution range (Å) | 50-1.73(1.77-1.73) | 50-1.77(1.80-1.77) | 50-1.47(1.50-1.47) | 48.34-2.60 (2.69-2.60) |
| Redundancy     | 3.9 (3.8) | 4.0 (3.8) | 4.0 (3.8) | 9.2 (4.5) |
| Average l/σ    | 36.9(3.4) | 33.6(3.5) | 31.5(2.8) | 43.3 (1.9) |
| Total reflections | 134315 | 105465 | 174571 | 52112 |
| Unique reflections | 34694 | 27236 | 45647 | 5671 |
| Completeness   | 95.2(94.6) | 97.6(95.5) | 96.5(93.6) | 88.4 (41.3) |
| Rmerge (%)     | 0.045(0.423) | 0.045(0.334) | 0.045(0.402) | 0.058 (0.544) |
| Average B-Factor (Å) | 32.54 | 27.93 | 19.62 | 38.856 |
| Rwork (%)      | 19.66 | 19.26 | 18.87 | 23.0 |
| Rfree (%)      | 24.01 | 24.31 | 22.04 | 29.9 |
| Num. of water  | 212  | 209   | 333   | 78     |
| Total reflections used | 28219 | 26569 | 44036 | 5394 |

RMSD from Ideality

| Bond length (Å) | 0.007 | 0.007 | 0.006 | 0.006 |
| Bond angle (°)  | 1.224 | 1.254 | 1.252 | 0.887 |

Ramachandran Plot

| Most favored (%) | 96.93 | 96.17 | 98.06 | 83.9 |
| Allowed (%)      | 0     | 0     | 0     | 15.3 |
| Generously allowed (%) | 1.92 | 3.07 | 0.78 | 0 |
| Disallowed (%)   | 1.15  | 0.77  | 1.16  | 0.8 |
| PDB code         | 4QGV  | 4QGW  | 4QGX  | 3FEP |

*values in the parentheses refer to the last resolution shell.

Crystallization Protocol:

The light sensitive merocyanine was dissolved in ethanol, stored at -20°C, and protected from light in a covered vial. The protein/merocyanine complex formation was achieved via addition of the merocyanine solution (up to 4 equivalents) to the concentrated protein solution (~20 mg/mL). It is important to keep the ethanolic merocyanine solution less than 10% in the concentrated protein solution in order to prevent protein precipitation. The protein/merocyanine sample was incubated at room temperature in the dark for 1h to reach complete PSB formation. All crystals were grown by hanging drop vapor diffusion at 4 °C, with the drop an equal mixture of crystallization and complex solution. Crystallizations were set up under red light, and the boxes were kept in the dark. Crystals appear within 3-4 days and grow to their maximum size in two weeks.

Crystals of R132K:R111L (KL), R132K:R111L:L121D (KLD) and R132K:R111L:L121E (KLE) were grown in 0.1 M bis tris propane pH = 6.5, 0.2 M sodium fluoride, 16 % PEG 3350.

Crystals of R132K:R111L:L121E:R59W were grown in 0.1 M MES pH = 6.3, 0.2 M ammonium sulfate and 30% PEG mono-methyl-ether (PEGMME) 5000.

Crystals were briefly soaked in a cryoprotectant solution containing the mother liquor and 20% glycerol and flash frozen in liquid nitrogen under red light.
Diffraction data were collected at the Advanced Photon Source (APS) (Argonne IL) at the LSCAT (sector 21) using 1.00 Å wavelength radiation at 100 K. Data reduction and scaling were performed using the HKL2000 program package.\cite{11} The structures were determined using the Molecular Replacement program in the CCP4 program suite, and refined using the CCP4 and PHENIX program packages.\cite{12} The search model was the R132K:R111L:L121E mutant of Cellular Retinoic Acid Binding Protein II (protein data bank code 2G7B). Several cycles of refinement, model adjustment and placement of ordered water molecules were performed for each structure. All model building and placement of water molecules were performed using COOT (0.6.1).\cite{13} The chromophore was manually fitted in the electron density near the end of the refinement.

**Structural Comparisons of KL/1 complex with CRABPII/retinylidene Complexes**

![Figure S12](image)

**Figure S12.** a) Overlay of crystal structures of KLE:59W complexed with C15-retinal analog in yellow (3F8A) and KL/1 complex in magenta (4QGV). b) Overlay of crystal structures of KLE complexed with all-trans-retinal in cyan (2G7B) and KL/1 complex in magenta (4QGV).

Figure S12 shows that the bound merocyanine in the KL/1 complex is not found in the same location within the binding pocket as compared to our previously obtained CRABPII crystal structures bound to retinal or retinal analogs.\cite{1,10} The indoline ring of the merocyanine which is tucked into a hydrophobic cavity stays far from the R59 residue located at the portal of the binding pocket. The resulting change in position of the indoline ring gives more room for the R59 residue to slide deeper into the binding pocket.
The overlaid figure shows that the merocyanine fluorophore follows a trajectory through the protein binding cavity similar to that seen in the C15-bound KLE:R59W and retinal bound KLE complexes as we published earlier.[1,10] The trajectory is almost identical to that of C15, with both ligands forming the more stable trans-iminium to Lys132. The formation of the cis-iminium in the retinal-bound KLE complex leads to a significant difference in binding and changes the loop conformation where R59 residue resides.

Figure S13. a) Overlay of crystal structures of KLE:R59W complexed with C15-retinal analog in yellow (3F8A) and merocyanine in magenta (3FEP). b) Overlay of crystal structures of KLE complexed with all-trans-retinal in cyan (2G7B) and merocyanine in magenta (3FEP).

Figure S14. a) Overlay of crystal structures of KLE/1 (in cyan, PDB ID: 4QGX) and KLD/1 (in purple, PDB ID: 4QGW) complexes. b) Exploded view of bound chromophore in KLD with the nearby residues highlighted.
The crystal structures of both KLE/1 and KLD/1 complexes display identical binding mode of action for the merocyanine aldehyde. The side access to the nucleophilic Lys132 generates a trans-iminium which is stabilized by the nearby Ser12 residue. The s-cis conformation around the C3-C4 bond shows a slight twist on polyene chain with a dihedral angle ($\psi_3$) of 7.2º, similar to KLE/1 complex (Figure 3b in main text).

K. References

1. Vasileiou, C.; Vaezeslami, S.; Crist, R. M.; Rabago-Smith, M.; Geiger, J. H.; Borhan, B. J. Am. Chem. Soc. 2007, 129, 6140.
2. Gill, S.; Vonhipple, P. H. Anal. Biochem. 1989, 182, 319.
3. Rurack, K.; Spieles, M. Anal. Chem. 2011, 83, 1232.
4. Vonderheyt, A.; Helenius, A. PLoS Biol. 2005, 3, e233.
5. Campbell, R. E.; Tour, O.; Palmer, A. E.; Steinbach, P. A.; Baird, G. S.; Zacharias, D. A.; Tsien, R. Y. Proc. Natl. Acad. Sci. USA 2002, 99, 7877.
6. Fischer, M.; Haase, I.; Simm, E.; Gerisch, G.; Muller-Taubenberger, A. FEBS Lett. 2004, 577, 227.
7. Sovdat, T.; Bassolino, G.; Liebel, M.; Schnedermann, C.; Fletcher, S. P.; Kukura, P. J. Am. Chem. Soc. 2012, 134, 8318.
8. Wang, J.; Cao, W.F.; Su, J.H.; Tian, H.; Huang, Y.H.; Sun, Z. R. Dyes and Pigments 2003, 57, 171.
9. Meguellati, K.; Spichty, M.; Ladame, S. Org. Lett. 2009, 11, 1123.
10. Lee, K. S. S.; Berbasova, T.; Vasileiou, C.; Jia, X.; Wang, W.; Choi, Y.; Nossoni, F.; Geiger, J. H.; Borhan, B. ChemPlusChem 2012, 77, 273.
11. Otwinowski, Z.; Minor, W. Macromolecular Crystallography 1997, part A, 276, 307.
12. a) 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. M.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. Acta Crystallogr. D Biol. Crystallogr. 2010, 66, 213. b) Winn, M. D. et al. Acta. Cryst. 2011, D67, 235. c) Murshudov, G. N.; Skubak, P.; Lebedeva, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. Acta Crystallographica, 2011, D67, 355.
13. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Acta Crystallogr. D Biol. Crystallogr. 2010, D66, 486.
14. Hori, Y.; Norinobu, T.; Sato, M.; Arita, K.; Shirakawa, M.; Kikuchi, K. J. Am. Chem. Soc. 2013, 135, 12360.
15. Liu, T. K.; Hsieh, P. Y.; Zhuang, Y. D.; Hsia, C. Y.; Huang, C. L.; Lai, H. P.; Lin, H. S.; Chen, I. C.; Hsu, H. Y.; Tan, K. T. ACS Chem Biol 2014, In Press.
L. NMR Spectra

[Image of NMR spectrum with chemical structure and peaks labeled]
