Supporting Information

Enzyme-directed Functionalization of Designed, Two-Dimensional Protein Lattices

Rohit H. Subramanian[1], Yuta Suzuki[1,4], Lorillee Tallorin[1], Swagat Sahu[1], Matthew Thompson[1,2], Nathan C. Gianneschi[1,2], Michael D. Burkart[1], F. Akif Tezcan[1,3,*]

[1]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA, USA.

[2]Departments of Chemistry, Materials Science & Engineering, Biomedical Engineering, Chemistry of Life Processes Institute, International Institute for Nanotechnology, Simpson Querrey Institute, Northwestern University, Evanston, Illinois 60208, USA.

[3]Materials Science and Engineering, University of California, San Diego, La Jolla, CA, USA.

[4]Current address: Hakubi Center for Advanced Research, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto, Japan, 606-8501

[*]Correspondence to: tezcan@ucsd.edu

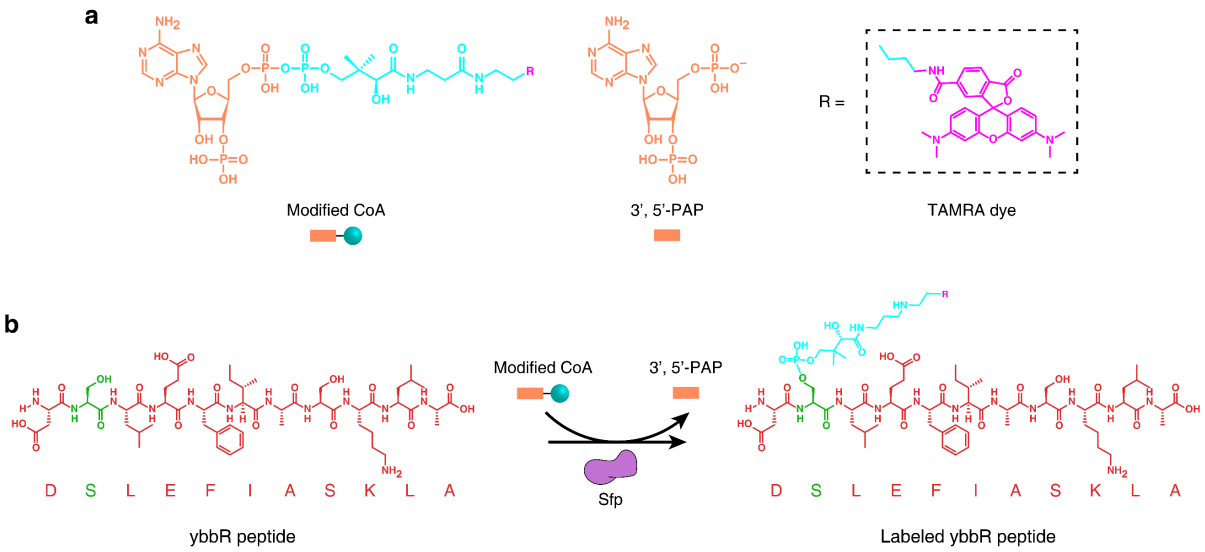


Figure S1. Scheme for enzymatic labeling of ybbR. (a) Chemical structures of coenzyme A (CoA) with enzymatically transferred moiety shaded in cyan. 3', 5'-PAP is 3',5'-phosphoadenosine phosphate (b) Reaction scheme illustrating the site-specific transfer of a modified CoA substrate onto a serine residue on ybbR.

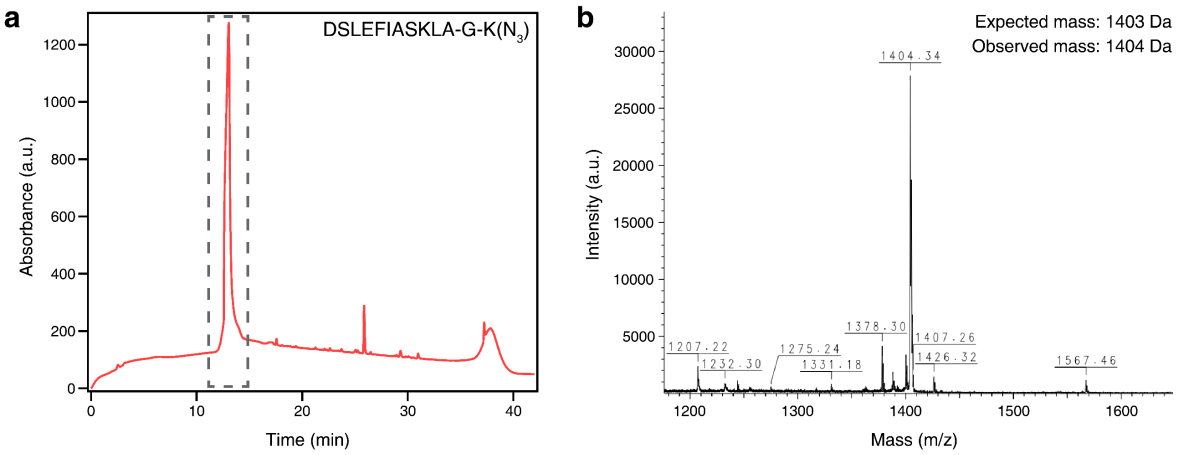
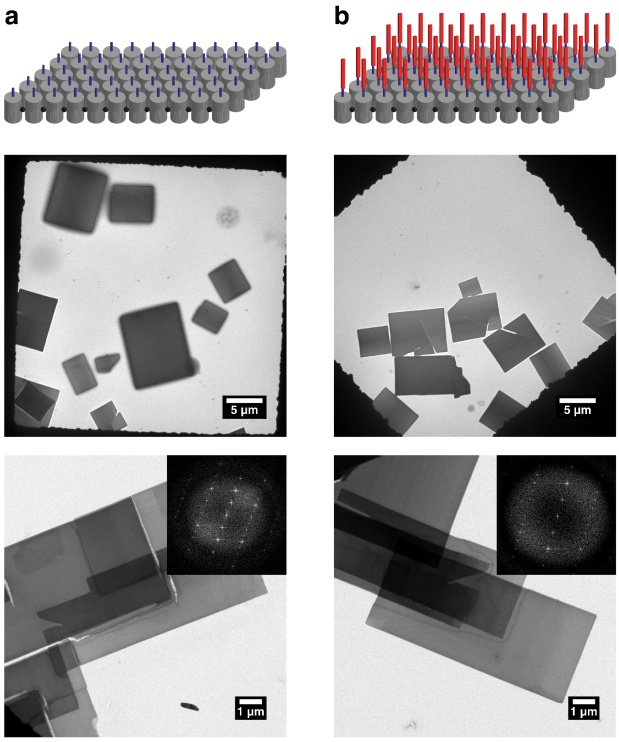


Figure S2. Characterization of synthesized ybbR-N_3 peptide. (a) HPLC chromatogram showing elution profile of ybbR-N_3 purification. (b) MALDI spectra of boxed region in (a) shows the expected mass for pure ybbR-N_3.



c

Incorporation of DBCO

$$\text{Number of DBCO per RIDC3} = \frac{[\text{DBCO}]}{[\text{RIDC3}]}$$

$$[\text{DBCO}] = \frac{A_{309}\ \text{DBCO}}{\varepsilon_{309}\ \text{DBCO}}$$

(Corrected DBCO absorbance)

$$A_{309}\ \text{DBCO} = A_{309}\ \text{RIDC3+DBCO} - A_{309}\ \text{RIDC3}$$

$$\varepsilon_{309}\ \text{DBCO} = 12{,}000\ M^{-1}\cdot L\cdot cm^{-1}$$

$$[\text{RIDC3}] = \frac{A_{415}\ \text{RIDC3}}{\varepsilon_{415}\ \text{RIDC3}}$$

$$\varepsilon_{415}\ \text{RIDC3} = 148{,}000\ M^{-1}\cdot L\cdot cm^{-1}$$

Incorporation of ybbR

$$\text{Number of ybbR per RIDC3} = \frac{[\text{ybbR}]}{[\text{RIDC3}]}$$

[ybbR] = DBCO consumed
= [DBCO before click] - [DBCO after click]

$$[\text{DBCO after click}] = \frac{A_{309}\ \text{RIDC3+DBCO} - A_{309}\ \text{RIDC3+ybbR}}{\varepsilon_{309}\ \text{DBCO}}$$

[DBCO before click] = [DBCO] (see above section)

Figure S3. TEM characterization and quantification of modified RIDC3 crystals. Cartoon representation and TEM micrographs for (a) RIDC3 crystals modified with DBCO and (b) RIDC3-ybbR chemically modified crystals. FFT inserts show that lattice crystallinity is retained throughout the process. (c) Calculations for quantifying DBCO and ybbR incorporation onto RIDC3 arrays.

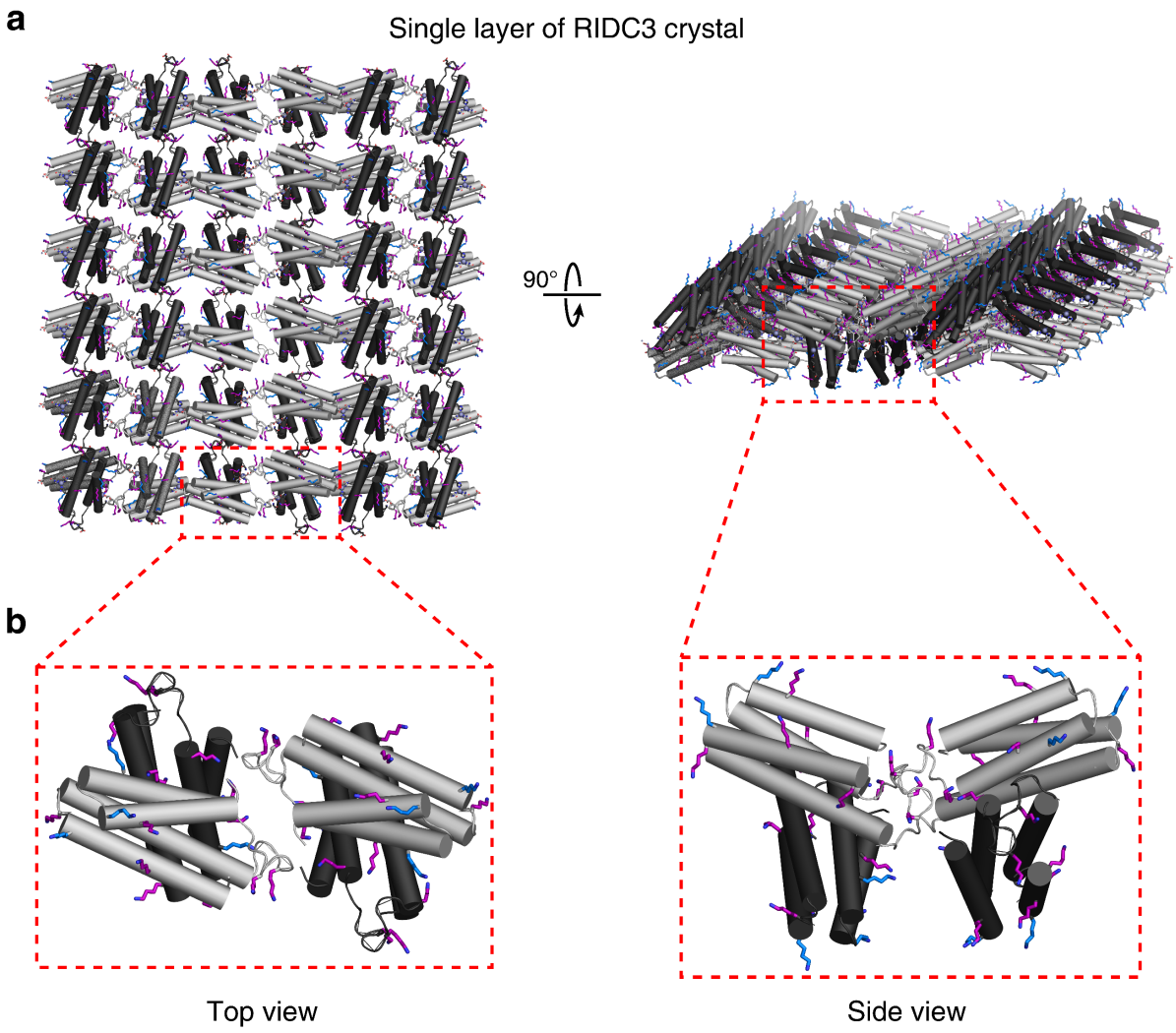


Figure S4. Identification of lysines on the RIDC3 crystal surface. (a) A single layer of the RIDC3 2D crystal packing is shown with (b) zoomed-in insets of the asymmetric unit. Lysine resides are shown as sticks colored in purple. Lysine residues identified as ybbR-modified via MS/MS analysis are colored in blue.

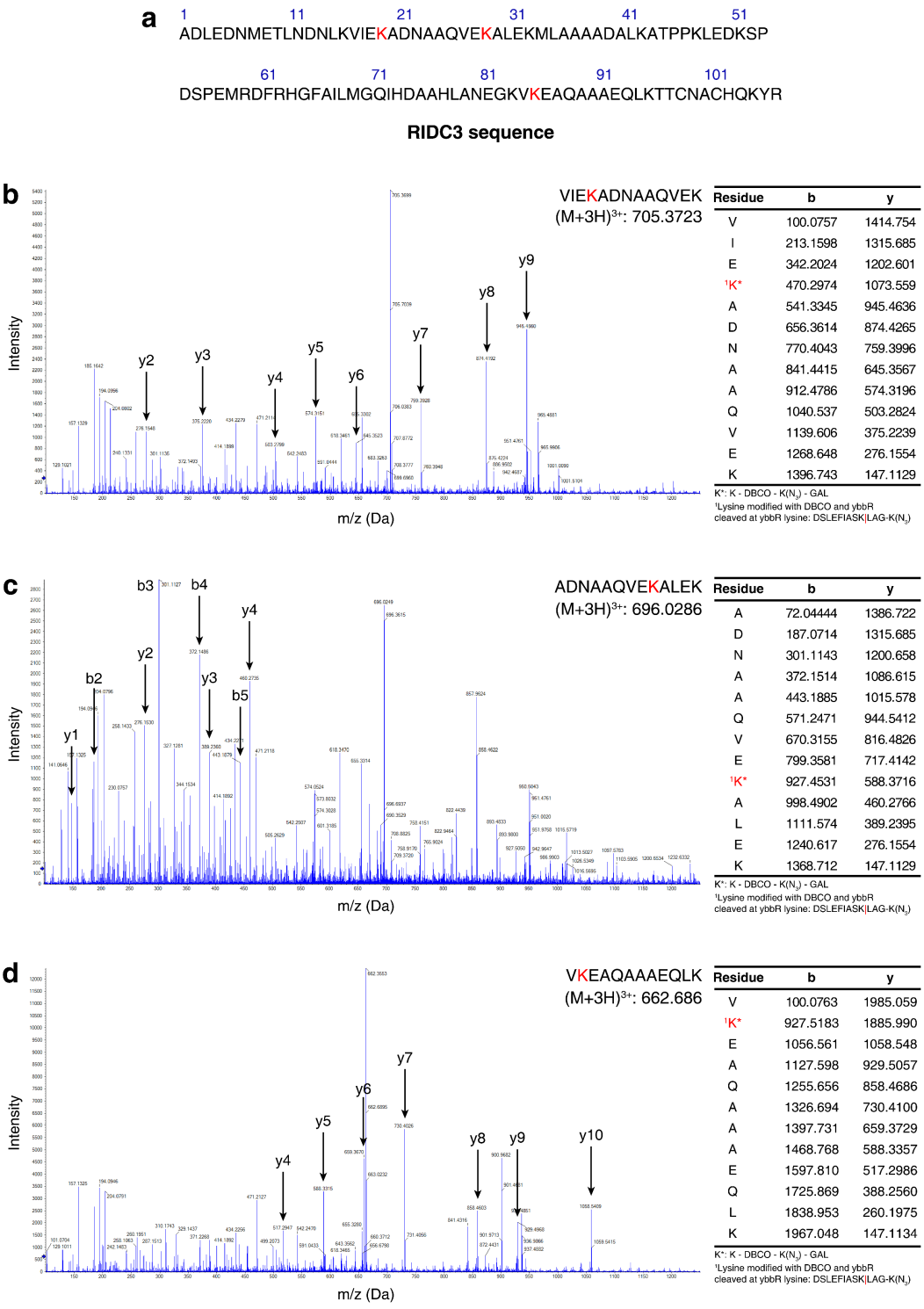


Figure S5. MS/MS analysis of ybbR conjugation onto RIDC3 arrays. **(a)** RIDC3 sequence with lysines identified as modified with ybbR colored red. **(b), (c), (d)** MS/MS spectra with y ions identified for each peptide fragment. Expected masses for b and y ions are shown on the right. The major peak in all three spectra correspond to the $(M+3H)^{3+}$ ionization of the peptide fragment.

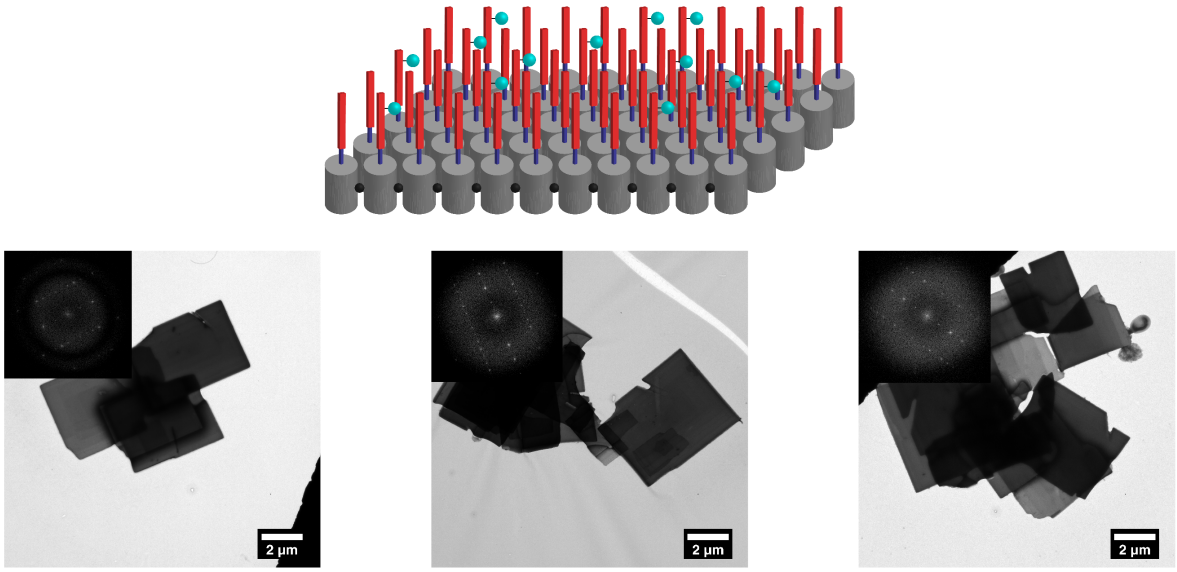


Figure S6. TEM characterization of enzymatically labeled RIDC3-ybbR arrays. A cartoon representation of enzymatically modified RIDC3-ybbR array is shown above with representative TEM micrographs after TAMRA-CoA, GFP-CoA, and co-labeling treatments.

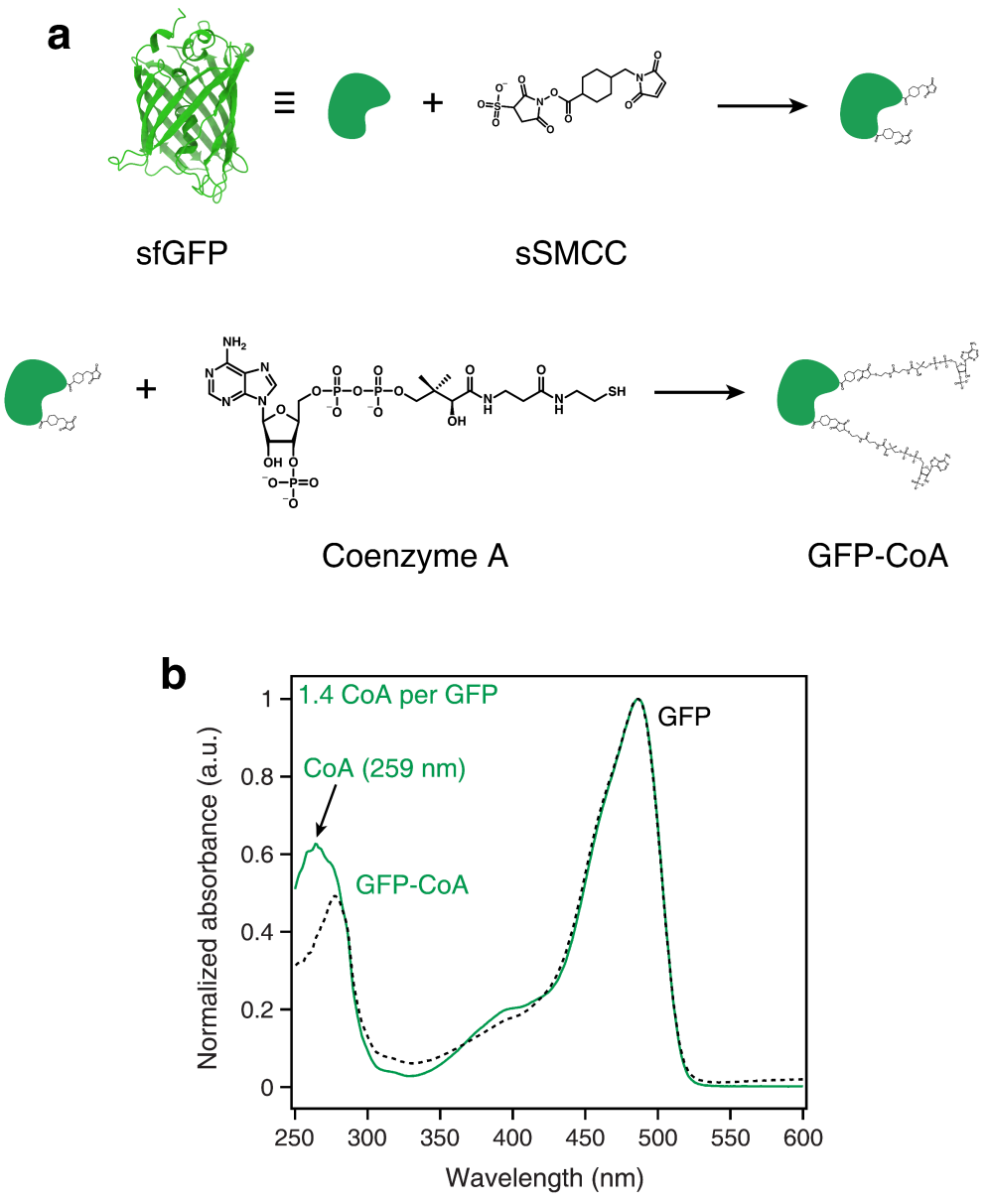


Figure S7. Generation and characterization of GFP-CoA. (a) Scheme for generation of GFP-CoA. (b) UV-vis characterization of GFP and GFP-CoA, shown normalized to the absorption maximum of GFP at 485 nm. The increase in absorbance at 259 nm is indicative of CoA conjugation onto GFP.

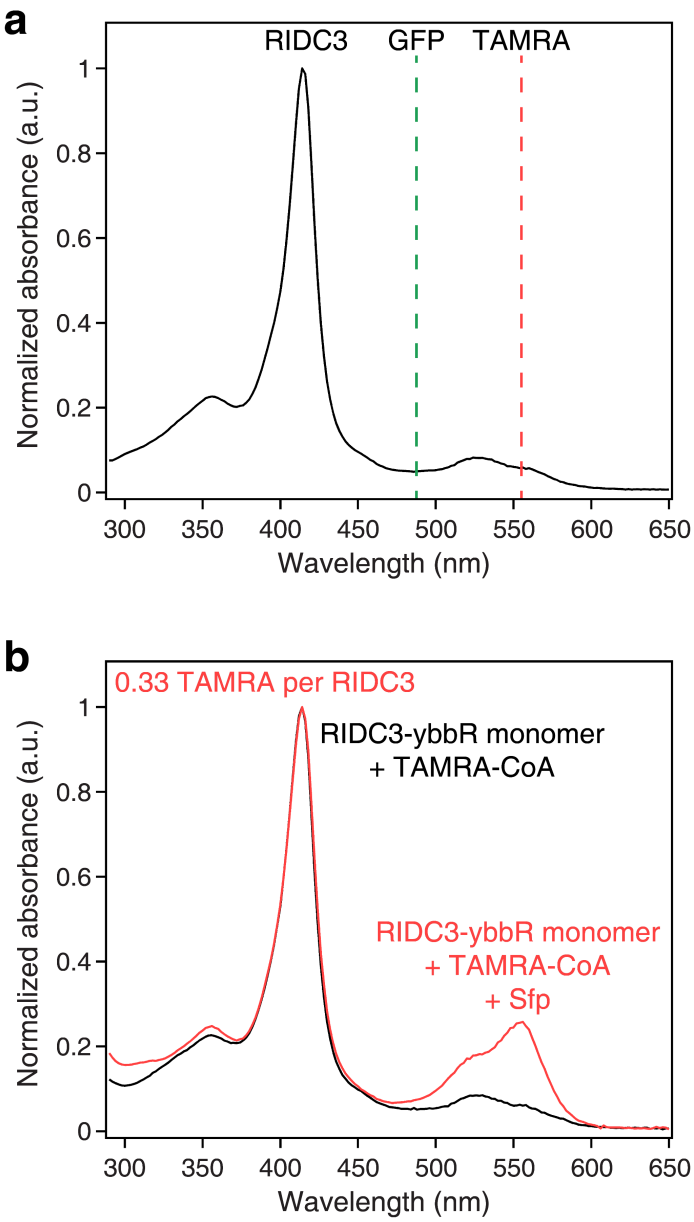
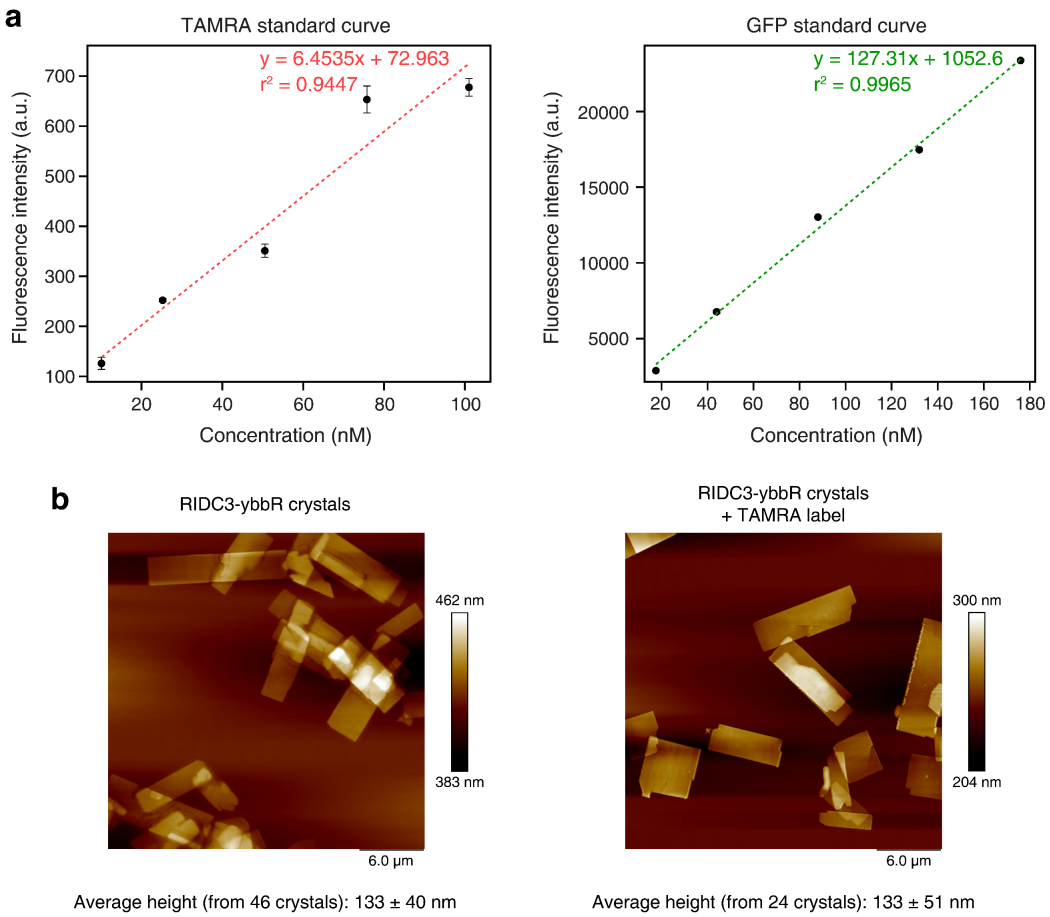


Figure S8. UV-vis characterization of RIDC3-ybbR monomers. (a) UV-vis characterization of RIDC3 with absorption maxima for GFP (green, 485 nm) and TAMRA (red, 554 nm) shown as dashed lines. (b) UV-vis characterization of RIDC3-ybbR monomers and TAMRA-CoA in the presence and absence of Sfp.



Average height (from 46 crystals): 133 ± 40 nm

Average height (from 24 crystals): 133 ± 51 nm

c

	+TAMRA label	+GFP label	+both labels	control
Protein concentration (µM)	16.5	22	24	21
Surface-accessible proteins (nM)	623.7	831.6	907.2	793.8
Dye concentration (nM)				
TAMRA fluorescence	63	0.06	84.8	5.4
GFP fluorescence	1.3	30.1	32.1	3.03
% labeling				
Total protein				
TAMRA fluorescence	0.38	0	0.35	0.03
GFP fluorescence	0	0.14	0.13	0.01
Surface-accessible proteins				
TAMRA fluorescence	10.1	0	9.3	0.68
GFP fluorescence	0.2	3.6	3.5	0.38

Figure S9. Quantification of fluorescent labeling of RIDC3-ybbR crystals. (a)Standard curves plotting dye concentration vs. fluorescence. (b) Representative AFM images of RIDC3-ybbR crystals before and after enzymatic labeling with TAMRA-CoA. The average heights of measured crystals are listed below. (c) Table showing protein and dye concentrations calculated based on UV-vis (for protein) or using the standard curves in (a) for the dye. Percentage of surface labeling is calculated based off the concentration of surface-accessible proteins available for modification. Labeling values calculated for the control sample represent background fluorescence and/or non-specific binding of dye molecules to the protein arrays.

	A	B	C
Protein concentration (μM)	22.6	23.7	25.3
Surface-accessible proteins (nM)	854	896	956
Dye concentration (nM)			
GFP fluorescence	68.42	2.99	2.12
% labeling			
Total protein			
GFP fluorescence	0.3	0.013	0.0084
Surface-accessible proteins			
GFP fluorescence	8	0.33	0.22

	A	B	C
RIDC3-CoA	+	+	+
^{ybbR}GFP	+	+	–
Sfp	+	–	–

Figure S10. Quantification of enzymatically labeled RIDC3-CoA arrays. A table for quantification of GFP labeling using fluorescence data. Labeling values calculated for control samples (B and C) represent background fluorescence or non-specific binding of GFP to the protein arrays. Ex/Em = 485/510 nm, n=3 for fluorescence measurements.

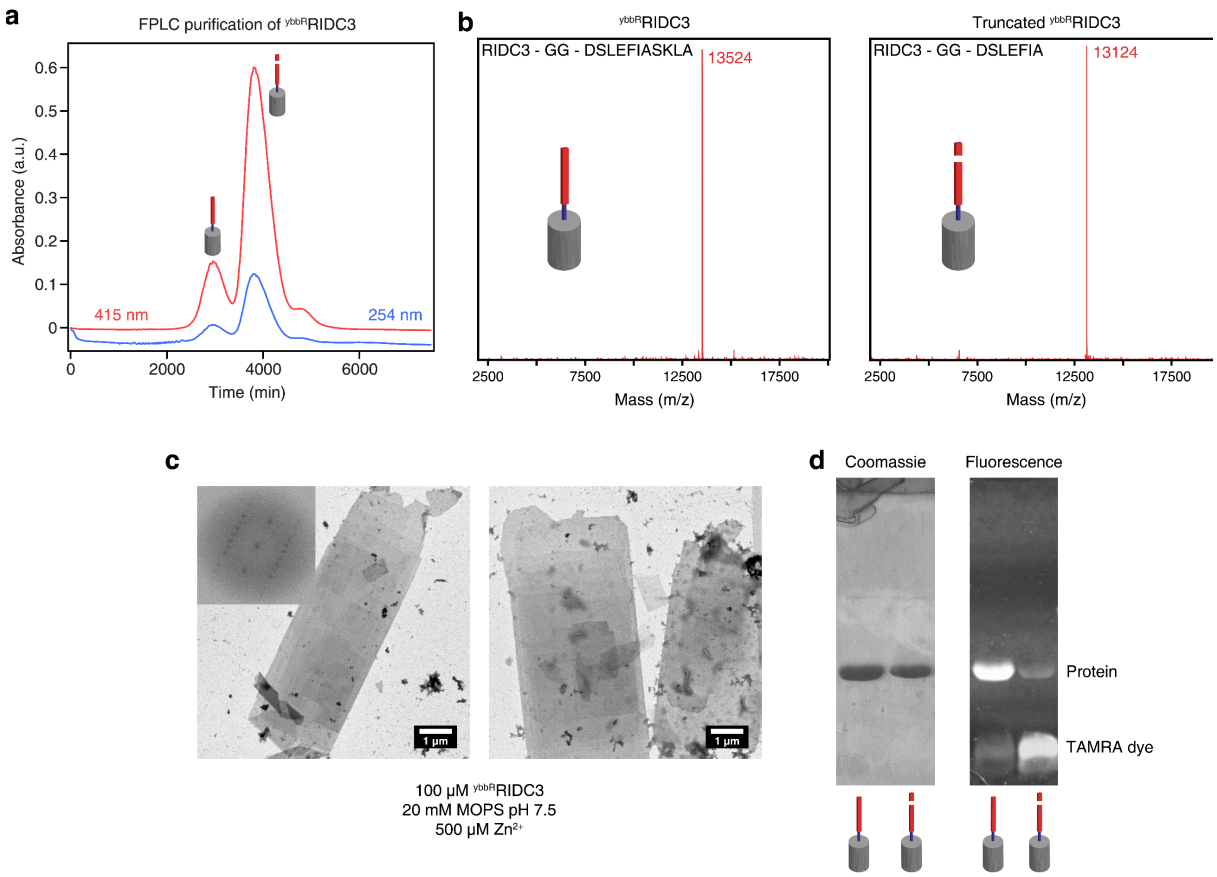
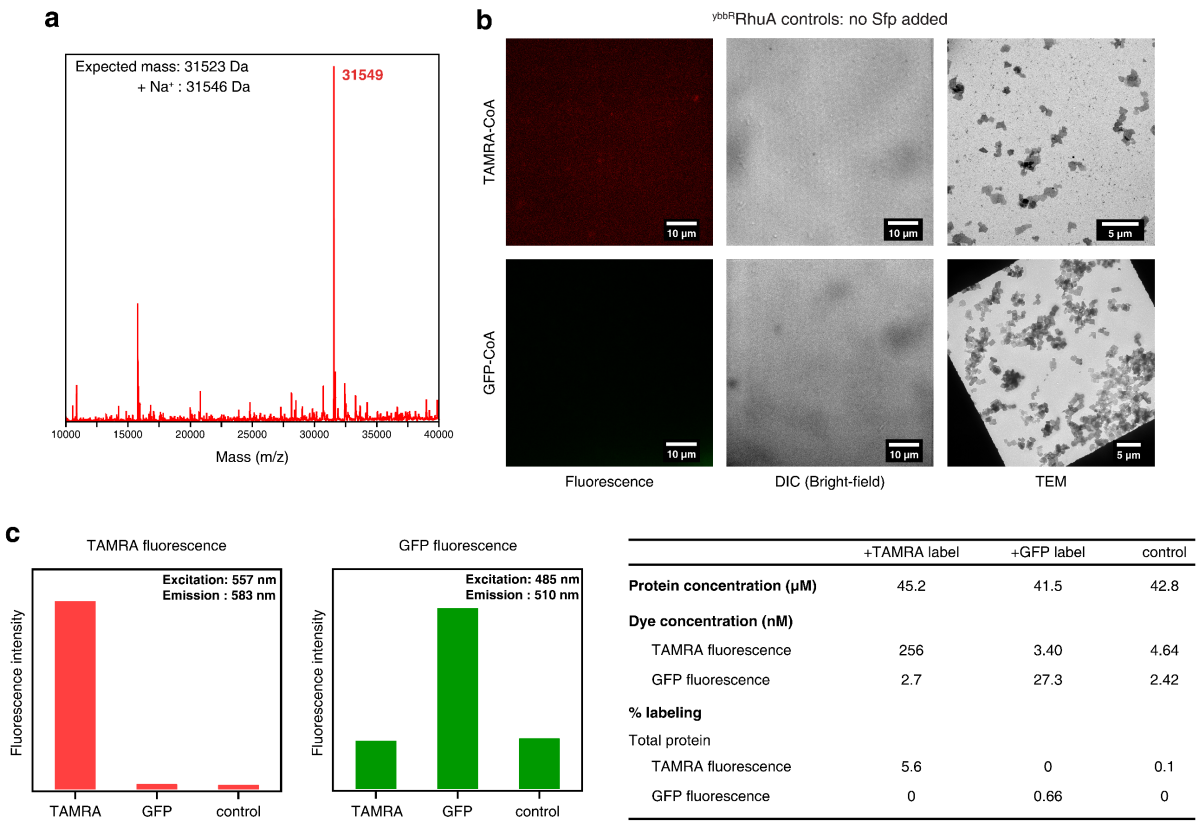


Figure S11. Characterization of genetically modified ^{ybbR}RIDC3. (a) FPLC chromatogram showing separation of intact and truncated ^{ybbR}RIDC3 protein during purification. (b) Mass spectra of intact and truncated ^{ybbR}RIDC3 observed during purification. (c) TEM micrographs of 2D ^{ybbR}RIDC3 arrays formed at pH 7.5. FFT inset shows the same diffraction pattern as RIDC3 crystals. (d) SDS PAGE gels of ^{ybbR}RIDC3 and truncated ^{ybbR}RIDC3 proteins incubated with TAMRA-CoA and Sfp PPTase. Fluorescence imaging of the gel confirms that the truncated protein cannot be enzymatically modified. Ex/Em: 532/580 nm



	+TAMRA label	+GFP label	control
Protein concentration (µM)	45.2	41.5	42.8
Dye concentration (nM)			
TAMRA fluorescence	256	3.40	4.64
GFP fluorescence	2.7	27.3	2.42
% labeling			
Total protein			
TAMRA fluorescence	5.6	0	0.1
GFP fluorescence	0	0.66	0

Figure S12. Characterization of genetically modified ^{ybbR}RhuA. (a) ESI-MS characterization of ^{ybbR}RhuA after purification. No peptide degradation is observed. (b) Confocal microscopy and TEM characterization of ^{ybbR}RhuA control samples. In the absence of Sfp, no fluorescence is observed. ^{ybbR}RhuA crystals remain intact in the labeling conditions. (c) Fluorescence profiles for ^{ybbR}RhuA arrays after enzymatic modification with TAMRA-CoA or GFP-CoA. Control experiments were incubated with both modified-CoA substrates in the absence of Sfp. A table is shown on the right for quantification of enzymatic labeling using fluorescence data. Labeling values calculated for the control sample represent background fluorescence and/or non-specific binding of dye molecules to the protein arrays.