bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

A Simple Method to Dual Site-Specifically Label a Protein Using Tryptophan Auxotrophic *Escherichia coli*

Ti Wu, and Simpson Joseph*

Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California, United States of America.

* Corresponding author
E-mail: sjoseph@ucsd.edu

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

37 # Abstract

38 Site-specifically labeling proteins with multiple dyes or molecular moieties is an important
39 yet non-trivial task for many research, such as when using Föster resonance energy transfer
40 (FRET) to study dynamics of protein conformational change. Many strategies have been devised,
41 but usually done on a case-by-case basis. Expanded genetic code provided a general platform to
42 incorporate non-canonical amino acids (ncAA), which can also enable multiple site-specific
43 labeling, but it's technically complicated and not suitable for some applications. Here we present
44 a streamlined method that could enable dual site-specific protein labeling by using a tryptophan
45 auxotroph of *Escherichia coli* to incorporate a naturally found tryptophan analog, 5-
46 hydroxytryptophan into a recombinant protein. As a demonstration, we incorporated 5-
47 hydroxytryptophan into *E. coli* release factor 1 (RF1), a protein known to possess two different
48 conformations, and site-specifically attached two different fluorophores, one on 5-
49 hydroxytryptophan and another on a cysteine residue. This method is simple, generally
50 applicable, efficient, and can serve as an alternative way for researchers who want to install an
51 additional labeling site in their proteins.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

52 # Introduction

53 Föster or fluorescence resonance energy transfer (FRET)-based method is one of the most
54 powerful and commonly used technique to understand the dynamics of conformational change in
55 proteins [1–3]. FRET is a phenomenon that an excited "donor" fluorescent molecule transfers
56 energy to an "acceptor" fluorescent molecule via a long-range non-radiative dipole-dipole
57 coupling mechanism. The efficiency of energy transfer (E) between two fluorescent molecules
58 depends on the separation distance (r) between donor and acceptor molecules with a relationship
59 of inverse 6th-power law ($E = 1/[1 + ((r/R_o)^6)]$, where R_o is the Föster distance of this pair of
60 donor and acceptor). Hence the measurement of the change in FRET efficiency can be used to
61 reveal the dynamic structural information of proteins if two fluorescent dyes are carefully
62 installed so the difference in the separation distance between two dyes can represent the
63 conformational change of the protein of interest.

64 One of the technical hurdles for implementing FRET experiments is that it's non-trivial and
65 sometimes even tricky to attach two or more different fluorescent dyes in a site-specific manner
66 to a protein [4]. Classic site-specific labeling reactions most commonly utilize thiol-targeting
67 functional groups, such as iodoacetamides and maleimides, amine-targeting functional groups,
68 such as isothiocyanates, activated esters, sulfonyl chlorides, etc., and a few alcohol-targeting
69 reagents. One practical concern is that thiol, amine, and alcohol groups are commonly present in
70 multiple positions in a protein, which means that to achieve multiple site-specific labeling,
71 extensive mutation and engineering may be required. Another strategy is to introduce non-
72 canonical amino acids (ncAA) that possess bioconjugatable side chain into proteins [5–7], so
73 they are capable to perform a wider arrays of bioconjugation reactions [8], such as click

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

74 chemistry, tetrazine ligation, etc. Recent advances in expanded genetic code provide a platform
75 to incorporate non-canonical amino acids into proteins [9,10]. In short, this technology would
76 assign a codon, usually one of the stop codons, to the ncAA of interest, then find and engineer an
77 orthogonal pair of transfer RNA (tRNA) that can recognize the assigned codon and the
78 corresponding tRNA synthetase that will only catalyze the ligation reaction between that specific
79 ncAA and tRNA. While extremely powerful and versatile, this technology requires some
80 specially engineered organisms and chemical components, and might not be suitable for some
81 research projects, such as monitoring the conformational changes of RF1, which directly
82 compete with the orthogonal tRNA for the stop codon.

83 Here we present a streamlined method that could enable dual site-specific protein labeling by
84 incorporating a common tryptophan analog, 5-hydroxytryptophan, into a recombinant protein.
85 This method utilizes a tryptophan auxotrophic strain of *E. coli*, which, when supplied with 5-
86 hydroxytryptophan in a minimal growth media, can readily use them for protein synthesis.
87 Combining with 5-hydroxytryptophan targeting bioconjugation chemistry and thiol-targeting
88 maleimide dye, we can achieve dual site-specific labeling in a fascicle manner using the standard
89 recombinant protein expression protocol.

90 As a model system to test our approach, we used *E. coli* release factor 1 (RF1). Class I
91 release factor proteins, including RF1 and RF2 in bacteria and eRF1 in eukaryotes, are
92 responsible for recognizing stop codon on mRNA at the A site of the ribosome and catalyzing
93 the peptidyl-tRNA hydrolysis and the release of the newly synthesized polypeptide from the
94 ribosome. It's known that RF1 has two vastly different conformational states, "open" and
95 "closed" [11–18]. Previously we have used transition metal ion FRET to study the dynamics of

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

96 this conformation change and the role it plays during translation termination [3]. Here we will
97 use RF1 to demonstrate direct site-specific dual labeling with two different fluorescent dyes,
98 which may open up new opportunities for studying the structural dynamics of RF1 during stop
99 codon recognition. More importantly, this simplified method can be used to dual site-specifically
100 label any protein with fluorescent dyes, biotin, or other moieties.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

101 # Materials and Methods

102 ## Chemicals, buffers, and bacterial strains

103 L-tryptophan, 5-aminofluorescein (FLA), and tetramethylrhodamine-5-maleimide (TMR)
104 were purchased from Sigma-Aldrich. L-5-hydroxytryptophan (Acros) was purchased from Fisher
105 Scientific. M63 minimal media were prepared using premixed M63 Medium Broth powder
106 (VMR). Spectroscopic experiments were carried out in a buffer of 50 mM K-HEPES (pH 7.5)
107 and 300 mM NaCl. The Trp auxotroph *Escherichia coli* BL21 (λDE3)/NK7402 was a gift by Dr.
108 A. Rod Merrill (University of Guelph, Canada).

109 ## Mutant RF1 expression and purification

110 Mutant RF1 was produced by site-directed mutagenesis (QuikChange, Stratagene). Starting
111 from a cysteine-free RF1 (C51S, C201S, C257S) gene in pPROEx-HTc vector (Invitrogen), we
112 first subcloned the gene into pBAD LIC 8A vector (Addgene #37501), and then introduced
113 single-tryptophan mutation (W55H), followed by the single-cysteine mutation (C257). RF1
114 mutant proteins were purified by nickel-affinity chromatography, and concentrated using a 10-
115 kDa MWCO spin column (Amicon Ultra-15, EMD Millipore). Purified proteins were then
116 quantitated by the Bradford assay, flash-frozen, and stored at -80 °C.

117 ## Expression of RF1 protein with 5-hydroxytryptophan (5-HW)

118 The 5-hydroxytryptophan-incorporated RF1 protein was expressed and purified using the Trp
119 auxotrophic strain as reported [19] with several adjustments. Electrocompetent *E. coli* BL21
120 (λDE3)/NK7402 Trp auxotropic cells were prepared [20] and stored at -80 °C. The competent
121 cells were transformed with pBAD plasmids containing the desired RF1 mutant gene by

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

122 electroporation and grown overnight at 37°C on LB/Ampicillin (Amp) plates. Next day, each
123 plate was scraped into 5 mL of Super Optimal Broth (SOB) with ampicillin and 2% glucose
124 (from sterile filtered 20% Glucose solution) and incubated at 37°C for 1 h. The 5 mL culture was
125 then transferred to a 4 L flask containing 1 L M63 minimal medium supplemented with 2.0%
126 glucose, 100 µg/mL ampicillin, 0.25M L-Trp, and 0.4% glycerol. This culture was grown to 0.5-
127 0.7 OD600 at 37°C, after which the cells were pelleted by centrifugation. The cell pellet was
128 then washed twice with 500 mL of M63 medium supplemented with 0.2% glycerol to remove all
129 traces of residual L-Trp. The cell pellet was then resuspended into the original volume of M63
130 media containing 0.6% glycerol and 100 µg/mL ampicillin and grown for a further 20 min to
131 deplete any residual tryptophan in the culture. Subsequently, the tryptophan analogues (D,L-
132 forms) (Sigma, St. Louis, MO) were added to the minimal medium at a final concentration of 0.5
133 mM, and the cells were induced with 1% arabinose (pBAD). The culture was allowed to grow
134 for 3h at 37°C, and the cells were harvested by centrifugation. Proteins were purified by nickel-
135 affinity chromatography and HiTrap Q HP anion exchange chromatography, and concentrated
136 using a 10-kDa MWCO spin column (Amicon Ultra-15, EMD Millipore). Purified proteins were
137 then quantitated by the Bradford assay, flash-frozen, and stored at -80 °C.

138 Labeling of RF1 mutants

139 For cysteine labeling, 100 µL RF1 mutants (40 µM final concentration) in labeling buffer [50
140 mM K-HEPES (pH 7.5) and 300 mM NaCl] was incubated with 20-fold excess (1 mM final
141 concentration) of tetramethylrhodamine-5-maleimide (TMR) (Invitrogen) at room temperature in
142 the dark for 2–4 h. Bioconjugation of 5-hydroxytryptophan with 5-aminofluorescein (FLA)
143 (Sigma-Aldrich) was carried out as reported [21] with several adjustments. 100 µL RF1 mutants

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

144 (40 μM final concentration) in labeling buffer [50 mM K-HEPES (pH 7.5) and 300 mM NaCl]
145 was incubated with 100-fold excess (4 mM final concentration) of FLA and 5 equivalent
146 ferricyanide (0.2 mM final concentration) at room temperature in the dark for 2–4 h.

147 The excess dye was removed by dialyzing against protein storage buffer [50 mM K-HEPES
148 (pH 7.5) and 100 mM NaCl] in the dark overnight. Proteins were further purified by HiTrap Q
149 HP anion exchange chromatography, and concentrated using a 10-kDa MWCO spin column
150 (Amicon Ultra-15, EMD Millipore). Purified proteins were then quantitated by the Bradford
151 assay, flash-frozen, and stored at -80 °C.

152 Staining, imaging and fluorescence spectroscopy

153 SDS-PAGE were stained with Coomassie Brilliant Blue R-250 following the standard
154 protocol [22], and stained gels were scanned and digitalized by Epson Perfection 2450 Photo
155 Flatbed Scanner. Gel containing fluorescent protein samples were scanned with Typhoon FLA
156 9500 imager (GE Healthcare), using 473 nm blue LD laser/LBP (510LP) emission filter for FLA
157 and 532 nm green SHG laser/BPG1 (570DF20) emission filter for TMR. Raw images were
158 analyzed and processed using ImageJ software.

159 Fluorescence spectroscopy was performed with Jasco FP-8500 Series Fluorometers. 0.1 μM
160 of protein samples were excited at 310 (2.5) nm for 5-hydroxytryptophan, 492 (2.5) nm for FLA,
161 and 544 (2.5) nm for TMR and scanning for a range of emission wavelength at 0.5 nm step. Data
162 were analyzed and plotted using GraphPad Prism software.

163

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

164 # Results

165 ## General schema of the protocol

166 Our goal is to establish a recombinant protein expression protocol to introduce additional
167 bioconjugation reaction sites via incorporation of non-canonical amino acids that can be easily
168 implemented by laboratories equipped with standard molecular biology setup. To minimize
169 technical complexity, residue-specific incorporation of non-canonical amino acids into proteins
170 using amino acid auxotrophs was chosen [6]. To be specific, we chose to use the tryptophan
171 auxotroph *Escherichia coli* strain BL21(λDE3)/NK7402 to incorporate tryptophan analogs [19],
172 because the general low occurrence of tryptophan in proteins could make this method more
173 feasible, and there are several well-known analogs and corresponding bioconjugation tools
174 available. Briefly, the method is as follows: the Trp auxotrophic *E. coli* is transformed with an
175 expression plasmid with the gene of interest tightly regulated by the pBAD promoter. The
176 transformed cells are first grown in minimal media supplied with tryptophan. Once the cells
177 grow to the desired density, the cells are spun down and washed to remove free tryptophan
178 molecules in the media, and then resuspended with fresh media supplied with a tryptophan
179 analog of choice and arabinose as the inducer. Finally, the cells are harvested, and the
180 recombinant protein purified via fractionation and/or chromatographic methods. The tryptophan
181 analog-incorporated protein could then be labeled with dyes using various bioconjugation
182 methods (Figure 1).

183 **Figure 1. Schema of incorporation of tryptophan analogs into recombinant protein using**
184 **Trp auxotrophic *E. coli* for dual site-specific labeling. Transformed cells are first cultured**
185 **in minimal medium supplied with regular tryptophan until the desired cell density. After a**

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

186 **few rounds of washing to remove free tryptophan, cells are resuspended in a new growth**
187 **medium supplied with the tryptophan analog along with the inducer molecule for protein**
188 **over-expression. Recombinant protein can then be harvested, analyzed, and site-specifically**
189 **labeled with compatible bioconjugation reactions.**

190 Design of single cysteine single tryptophan RF1 (scswRF1)

191 Our model protein is *E. coli* RF1, which shows two distinctive conformations–open and
192 closed (Figure 2A). A crystal structure of *E. coli* RF1 complexed with PrmC methyltransferase
193 (PDB code: 2B3T) served as the template for the closed conformation [23], and a cryo-EM
194 structure of *E. coli* RF1 in the translation termination complex (PDB code: 6OSK) was the
195 template for the open conformation [17]. To find two labeling sites whose separation distance
196 can reflect the conformational change, it's natural to use GGQ motif and anti-codon PXT motif
197 as reference points and look for potential sites in the domains they are located in, namely,
198 Domain III and Domain II, respectively. Fortunately, out of three cysteine sites and two
199 tryptophan sites present in wild-type *E. coli* RF1, one of the cysteine (C257) is located in
200 Domain III and one of the tryptophan (W144) is located in Domain II. Based on the model, the
201 estimated separation distance of these two residues are 24Å and 49Å in closed and open states,
202 respectively (Figure 2B and 2C).

203 **Figure 2. Structural schema of *E. coli* release factor 1 (RF1) in open and close**
204 **conformations. Structures are modified from published models for open (6osk) and close**
205 **(2b3t) conformation. (A) Alignment of RF1 in open and close conformations. (B)(C) Green**
206 **spheres indicate the tryptophan residue (W144) and red spheres indicate the cysteine**
207 **residue (C257) that will be labeled with fluorescent dyes. (D)(E) Residue-to-residue**

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

208 **distance map of the open and closed conformations of *E. coli* RF1 models. (D) The color**
209 **code shows the residue-to-residue difference in separation distance between open/closed**
210 **states of RF1. (E) The color code shows the residue-to-residue difference in FRET**
211 **efficiency between open/closed states of RF1.**

212 The optimality of the labeling sites were further examined by using two-dimensional residue-
213 to-residue maps [3,24]. The color-coded distance map shows the distance change of any residue
214 pair in closed and open states, which is calculated based on the relative location of α-carbons of
215 any two residues in the structural model (Figure 2D). The dark color region are those residual
216 pairs whose separation distance won't change a lot when the protein changes its conformation,
217 while the light color region shows the residual pairs which have significantly different separation
218 distances in closed and open states. The map clearly shows four folded domains, and the
219 C257/W144 pair lies close to the GGQ/PXT pair yet not on the same vertical and horizontal
220 lines, which means they won't be too close and likely affect the biological functions of GGQ and
221 PXT motifs. The distance map was further transformed into another 2-dimensional map showing
222 the estimated change of FRET efficiency based on fluorescein and tetramethylrhodamine (Figure
223 2E). The FRET efficiency map shows that C257/W144 pair may show a significant change in
224 FRET efficiency when labeled with fluorescein and tetramethylrhodamine.

225 Expression of 5-hydroxytryptophan-incorporated scswRF1

226 After trying a few tryptophan analogs, 5-hydroxytryptophan (5HW, Figure 3A) was chosen
227 as the main focus, because it could be readily incorporated by *E. coli* BL21(DE3)/NK7402 strain
228 with moderately good efficiency [19], has its own distinctive spectral properties [19,25–27]–it
229 and the proteins containing it absorb between 300 to 320 nm, and would emit at higher

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

230 wavelength region–and there are bioconjugation methods to target it specifically [21,28]. 5HW is
231 actually a natural occurring amino acid, known as the precursor of the neurotransmitter serotonin
232 [29]. Structurally, it differs from regular tryptophan molecule by one oxygen atom at position 5
233 on the indole ring (Figure 3A).

234 **Figure 3. Incorporation of 5-hydroxytryptophan into single-cysteine single-tryptophan**
235 **RF1. (A) Chemical structure of 5-hydroxytryptophan. (B) Coomassie-stained SDS-PAGE**
236 **showing purified 5HW-incorporated scswRF1 (5HW-RF1) and regular single-cysteine**
237 **single-tryptophan RF1 (Trp-RF1). (C) Fluorescence spectroscopy of 5HW-incorporated**
238 **RF1 (blue solid) and regular RF1 (black dotted) with excitation wavelength at 310 nm.**
239 **(D)(E) Deconvoluted LC-ESI-MS spectra of (D) 5HW-RF1 and (E) Trp-RF1, which shows**
240 **a 16 Da difference caused by regular tryptophan and 5-hydroxytryptophan.**

241 Following the above-mentioned protocol, we could successfully overexpress scswRF1
242 protein in moderately good yield, a few milligrams per liter liquid culture. The recombinant
243 scswRF1s with 5-hydroxytryptophan (5HW-RF1) or with regular tryptophan (Trp-RF1) were
244 further purified using affinity and ion exchange chromatography (Figure 3B). With fluorescence
245 spectroscopy 5HW-RF1 shows strong emission at 340 nm when excited at 310 nm while Trp-
246 RF1 does not. LC-ESI-TOF-mass spectrometry shows that the difference in intact protein masses
247 between Trp-RF1 and 5HW-RF1 is exactly 16, the atomic mass of an oxygen atom, indicating
248 the successful incorporation of one and only one 5-hydroxytryptophan into the recombinant
249 scswRF1 protein (Figure 3D-E).

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

250 ## Dual site-specific labeling of 5-hydroxytryptophan-incorporated
251 ## scswRF1

252 The scswRF1 is designed to be labeled with one fluorophore targeting the cysteine residue
253 and the other targeting the tryptophan analog. For 5-hydroxytryptophan site, 5-aminofluorescein
254 (FLA) could be covalently attached onto the indole ring under a mild oxidative condition [21],
255 while cysteine was labelled with tetramethylrhodamine (TMR)-maleimide (Figure 4A). To
256 determine the specificity of the labeling reactions, the scswRF1 was labeled with either FLA
257 (RF1-F) or TMR (RF1-T) and doubly labeled with both FLA and TMR (RF1-FT). The labeled
258 proteins were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and the gel
259 was scanned for fluorescence signals with a Typhoon scanner (excitation at 488 nm/emission at
260 525 nm for FLA; excitation at 532 nm/emission at 570 nm for TMR). We observed the FLA-
261 labeled and TMR-labeled proteins only at their corresponding fluorescence channels showing
262 that the proteins were successfully labeled with the individual dyes. The scswRF1 that was
263 reacted with both FLA and TMR dyes was responsive to both fluorescence channels showing
264 that the protein was conjugated to both dyes (Figure 4B).

265 **Figure 4. Site-specific dual-labeling of 5HW-incorporated RF1 and spectroscopic analysis.**
266 **(A) Reaction schema of dual site-specific labeling on 5HW-incorporated RF1 protein. (B)**
267 **Fluorescence and Coomassie Blue-stained SDS-PAGE gel images of single-labeled (RF1-F**
268 **with fluorescein, RF1-T with tetramethylrhodamine) and double-labeled (RF1-FT) protein.**
269 **(C) Fluorescence spectroscopy of regular tryptophan (black dotted) and 5HW-**
270 **incorporated RF1 (blue solid) with excitation wavelength at 310 nm. (D) Fluorescence**
271 **spectroscopy of RF1-F (green solid line), RF1-T (red solid line) and RF1-FT (orange solid**

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

272 **line) with excitation wavelength at 492 nm. Sum of RF1-F and RF1-T spectra is plotted for**
273 **comparison (black dashed line). RF1-FT shows slightly lower signal at FLA emission peak**
274 **(520 nm) and slightly higher signal at TMR emission peak (575 nm) than the combined**
275 **signal of two single-labeled proteins.**

276 ## Fluorescence spectroscopic analysis of dye-labeled RF1

277 The dye-labeled scswRF1s also behaved as expected (Figure 4D). When excited at 492 nm,
278 the RF1 with only FLA conjugated to the 5HW site showed one emission peak at 520 nm (Figure
279 4D, green solid line), and the RF1 with only TMR conjugated to the cysteine site showed one
280 emission peak at 575 nm (Figure 4D, red solid line), while the double-labeled RF1 showed two
281 peaks corresponding to FLA and TMR in the emission spectrum (Figure 4D, orange solid line),
282 which confirm that the protein is labeled by both dyes.

283 With both single-labeled and double-labeled proteins in hand, we wished to see if there's
284 FRET in the presumably closed form of RF1 (Figure 4C). Comparing to the sum of the signals
285 from the two single-labeled proteins (Figure 4D, black dashed line), the FLA emission peak at
286 520 nm is slightly lower and TMR emission peak at 575 nm slightly higher in the spectrum of
287 RF1-FT, suggesting there is a small FRET in RF1 protein.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

288 # Discussions

289 Dual or multiple site-specific labeling of protein is useful for various kinds of biochemical
290 and biophysical research, yet it's not a trivial task. To overcome the limitation of direct
291 bioconjugation with amine- or thiol-reactive chemistry, scientists have developed many
292 strategies, such as labeling two fragments of a protein separately and then joining them together
293 into one protein, using technique such as native chemical ligation [30] or intein-mediated
294 ligations [31]. Recent advances in biorthogonal bioconjugation reactions and expanded genetic
295 code enables a more general strategy for site-specific labeling–first incorporate ncAA with a
296 chemical handle on the amino acid side chain, then perform bioconjugation reaction specific to
297 that chemical handle to attach fluorophores or other molecular moieties. Our method is also
298 leveraging the power of ncAA and bioconjugation reactions yet implemented by using
299 auxotrophic strain for ncAA incorporation for simplicity and efficiency. While not as multi-
300 purpose as expanded genetic code, our method could excel in many scenarios for people who
301 want to install an additional labeling site in their proteins.

302 Analysis of protein sequences have shown that tryptophan is the rarest amino acid in a
303 protein, on average only one are present in every one hundred amino acids in a protein sequence
304 [32]. This is a key advantage for site-specifically labeling a protein because a single tryptophan
305 at a unique position in a protein can be created with minimal changes to the protein's primary
306 sequence. Additionally, 5-hydroxytrytophan is an economic and commercially available
307 tryptophan analog and can be efficiently incorporated into over-expressed recombinant protein. It
308 has a distinctive fluorescence property compared to regular tryptophan, and can be employed as
309 FRET donor while using AEDANS as acceptor [19]. In the scswRF1 constructed here, it showed

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

310 a very broad emission range which could even serve as FRET donor to dyes such as fluorescein.
311 Fluorescence dye (e.g. fluorescein amine) or molecular moieties (e.g. 4-carboxydiazonium
312 (4CDZ)-biotin) can be attached onto it under ambient reaction condition, which makes it a useful
313 tool for site-specific labeling when it's incorporated into proteins. In principle, the same strategy
314 can work with other tryptophan analogs, for example 5-azidotryptophan, which can further
315 expand the applicable reactions for site-specific protein labeling.

316 While we successfully demonstrated that RF1 was site-specifically labeled with two different
317 dyes and have observed some FRET phenomenon, this current construct cannot provide any
318 further insights on the dynamic property of RF1. It may be because the position or the choice of
319 dyes are not optimal. Based on the theoretical calculation according to the residue-to-residue
320 map (Figure 2D-E), the separation distance change between open and closed state is 25Å and the
321 expected FRET efficiency change is up to 33%. However, the microenviroments inside the
322 protein and the equilibrium between open and closed or any other possible intermediate
323 conformations will affect the fluorescence properties of the dye and the observable signals.
324 Further optimization would be required if we want to study the exact conformational state of
325 release factor proteins in solution and further understand the reaction kinetics during translation.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

326 # Acknowledgments

327 We thank Dr. Krista Trappal for help in the early phase of this project. This work was
328 supported by the Department of Defense Army Research Office contract W911NF-13-1-0383.

329

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

330 # References

331 1. Selvin PR. Fluorescence resonance energy transfer. Methods in Enzymology. Elsevier; 1995.
332 pp. 300–334. doi:10.1016/0076-6879(95)46015-2

333 2. Selvin PR. The renaissance of fluorescence resonance energy transfer. Nat Struct Biol.
334 2000;7: 730–734. doi:10.1038/78948

335 3. Trappl K, Joseph S. Ribosome Induces a Closed to Open Conformational Change in Release
336 Factor 1. Journal of Molecular Biology. 2016;428: 1333–1344.
337 doi:10.1016/j.jmb.2016.01.021

338 4. Zhang G, Zheng S, Liu H, Chen PR. Illuminating biological processes through site-specific
339 protein labeling. Chem Soc Rev. 2015;44: 3405–3417. doi:10.1039/C4CS00393D

340 5. Hoesl MG, Budisa N. Recent advances in genetic code engineering in Escherichia coli.
341 Current Opinion in Biotechnology. 2012;23: 751–757. doi:10.1016/j.copbio.2011.12.027

342 6. Johnson JA, Lu YY, Van Deventer JA, Tirrell DA. Residue-specific incorporation of non-
343 canonical amino acids into proteins: recent developments and applications. Current Opinion
344 in Chemical Biology. 2010;14: 774–780. doi:10.1016/j.cbpa.2010.09.013

345 7. Liu CC, Schultz PG. Adding New Chemistries to the Genetic Code. Annu Rev Biochem.
346 2010;79: 413–444. doi:10.1146/annurev.biochem.052308.105824

347 8. McKay CS, Finn MG. Click Chemistry in Complex Mixtures: Bioorthogonal
348 Bioconjugation. Chemistry & Biology. 2014;21: 1075–1101.
349 doi:10.1016/j.chembiol.2014.09.002

350 9. Xiao H, Schultz PG. At the Interface of Chemical and Biological Synthesis: An Expanded
351 Genetic Code. Cold Spring Harb Perspect Biol. 2016;8: a023945.
352 doi:10.1101/cshperspect.a023945

353 10. de la Torre D, Chin JW. Reprogramming the genetic code. Nat Rev Genet. 2021;22: 169–
354 184. doi:10.1038/s41576-020-00307-7

355 11. Capecchi MR. Polypeptide chain termination in vitro: isolation of a release factor.
356 Proceedings of the National Academy of Sciences. 1967;58: 1144–1151.
357 doi:10.1073/pnas.58.3.1144

358 12. Rawat UBS, Zavialov AV, Sengupta J, Valle M, Grassucci RA, Linde J, et al. A cryo-
359 electron microscopic study of ribosome-bound termination factor RF2. Nature. 2003;421:
360 87–90. doi:10.1038/nature01224

361 13. Laurberg M, Asahara H, Korostelev A, Zhu J, Trakhanov S, Noller HF. Structural basis for
362 translation termination on the 70S ribosome. Nature. 2008;454: 852–857.
363 doi:10.1038/nature07115

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

364 14. Weixlbaumer A, Jin H, Neubauer C, Voorhees RM, Petry S, Kelley AC, et al. Insights into
365 Translational Termination from the Structure of RF2 Bound to the Ribosome. Science.
366 2008;322: 953–956. doi:10.1126/science.1164840

367 15. Korostelev A, Asahara H, Lancaster L, Laurberg M, Hirschi A, Zhu J, et al. Crystal structure
368 of a translation termination complex formed with release factor RF2. Proceedings of the
369 National Academy of Sciences. 2008;105: 19684–19689. doi:10.1073/pnas.0810953105

370 16. Brown A, Shao S, Murray J, Hegde RS, Ramakrishnan V. Structural basis for stop codon
371 recognition in eukaryotes. Nature. 2015;524: 493–496. doi:10.1038/nature14896

372 17. Fu Z, Indrisiunaite G, Kaledhonkar S, Shah B, Sun M, Chen B, et al. The structural basis for
373 release-factor activation during translation termination revealed by time-resolved cryogenic
374 electron microscopy. Nat Commun. 2019;10: 2579. doi:10.1038/s41467-019-10608-z

375 18. Svidritskiy E, Demo G, Loveland AB, Xu C, Korostelev AA. Extensive ribosome and RF2
376 rearrangements during translation termination. eLife. 2019;8: e46850.
377 doi:10.7554/eLife.46850

378 19. Mohammadi F, Prentice GA, Merrill AR. Protein–Protein Interaction Using Tryptophan
379 Analogues: Novel Spectroscopic Probes for Toxin–Elongation Factor-2 Interactions.
380 Biochemistry. 2001;40: 10273–10283. doi:10.1021/bi011035u

381 20. Green MR, Sambrook J. Transformation of Escherichia coli by Electroporation. Cold Spring
382 Harb Protoc. 2020;2020: pdb.prot101220. doi:10.1101/pdb.prot101220

383 21. Sarathi Addy P, Italia JS, Chatterjee A. An Oxidative Bioconjugation Strategy Targeted to a
384 Genetically Encoded 5-Hydroxytryptophan. ChemBioChem. 2018;19: 1375–1378.
385 doi:10.1002/cbic.201800111

386 22. Simpson RJ. Staining Proteins in Gels with Coomassie Blue. Cold Spring Harbor Protocols.
387 2007;2007: pdb.prot4719-pdb.prot4719. doi:10.1101/pdb.prot4719

388 23. Graille M, Heurgué-Hamard V, Champ S, Mora L, Scrima N, Ulryck N, et al. Molecular
389 Basis for Bacterial Class I Release Factor Methylation by PrmC. Molecular Cell. 2005;20:
390 917–927. doi:10.1016/j.molcel.2005.10.025

391 24. Yu X, Wu X, Bermejo GA, Brooks BR, Taraska JW. Accurate High-Throughput Structure
392 Mapping and Prediction with Transition Metal Ion FRET. Structure. 2013;21: 9–19.
393 doi:10.1016/j.str.2012.11.013

394 25. Hogue CWV, Rasquinha I, Szabo AG, MacManus JP. A new intrinsic fluorescent probe for
395 proteins Biosynthetic incorporation of 5-hydroxytryptophan into oncomodulin. FEBS
396 Letters. 1992;310: 269–272. doi:10.1016/0014-5793(92)81346-N

397 26. Ross JBA, RuSINOVA E, HUANGt YT, Laws WR, Hasselbacher CA. Spectral
398 enhancement of proteins: Biological incorporation and fluorescence characterization of 5-
399 hydroxytryptophan in bacteriophage A ci repressor. Proc Natl Acad Sci USA. 1992; 5.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

400 27. Alexander Ross JB, Szabo AG, Hogue CWV. Enhancement of protein spectra with
401 tryptophan analogs: Fluorescence spectroscopy of protein-protein and protein-nucleic acid
402 interactions. Methods in Enzymology. Elsevier; 1997. pp. 151–190. doi:10.1016/S0076-
403 6879(97)78010-8

404 28. Addy PS, Erickson SB, Italia JS, Chatterjee A. A Chemoselective Rapid Azo-Coupling
405 Reaction (CRACR) for Unclickable Bioconjugation. J Am Chem Soc. 2017;139: 11670–
406 11673. doi:10.1021/jacs.7b05125

407 29. Birdsall TC. 5-Hydroxytryptophan: a clinically-effective serotonin precursor. Altern Med
408 Rev. 1998;3: 271–280.

409 30. Lee TC, Moran CR, Cistrone PA, Dawson PE, Deniz AA. Site-Specific Three-Color
410 Labeling of α-Synuclein via Conjugation to Uniquely Reactive Cysteines during Assembly
411 by Native Chemical Ligation. Cell Chemical Biology. 2018;25: 797-801.e4.
412 doi:10.1016/j.chembiol.2018.03.009

413 31. Yang J-Y, Yang WY. Site-Specific Two-Color Protein Labeling for FRET Studies Using
414 Split Inteins. J Am Chem Soc. 2009;131: 11644–11645. doi:10.1021/ja9030215

415 32. Kozlowski LP. Proteome-pI: proteome isoelectric point database. Nucleic Acids Research.
416 2017;45: D1112–D1116. doi:10.1093/nar/gkw978

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

418

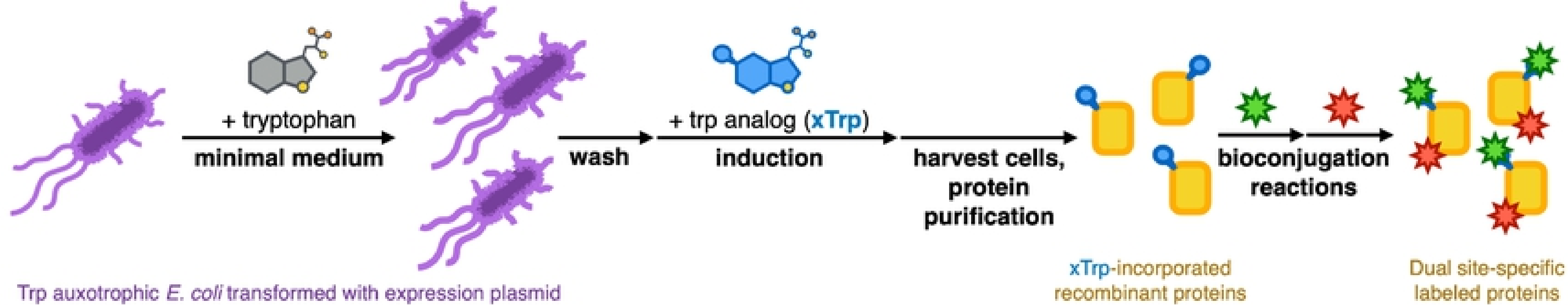


Figure 1

preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

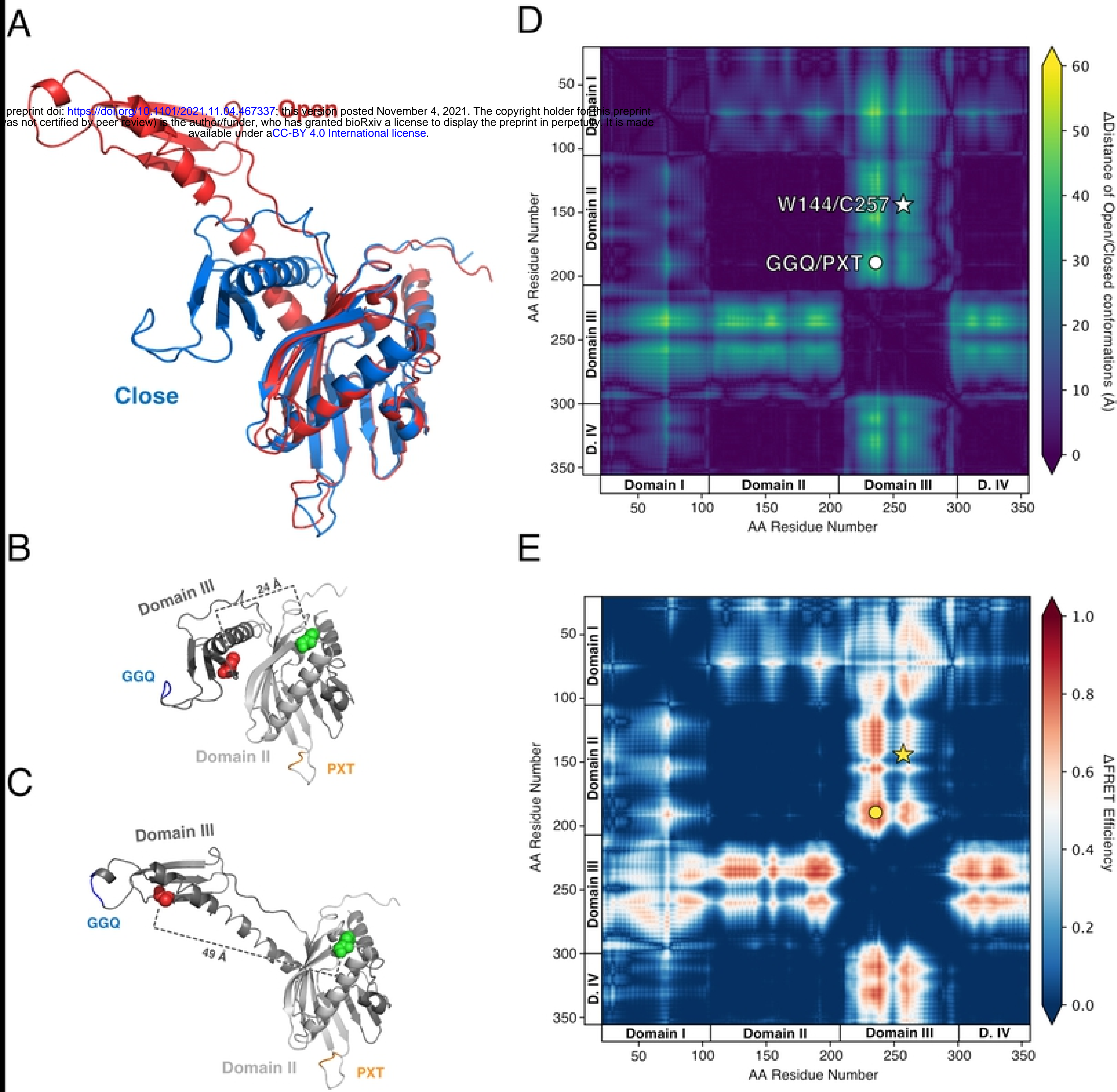
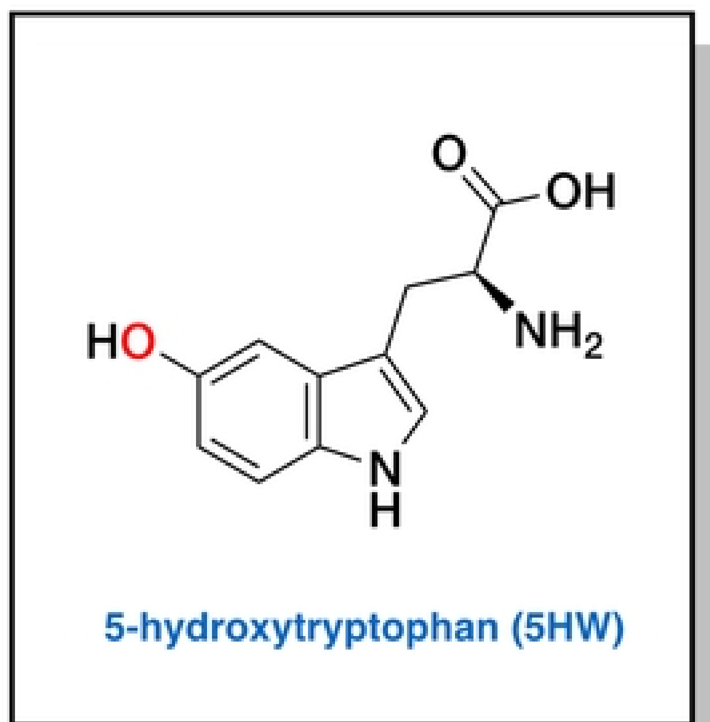


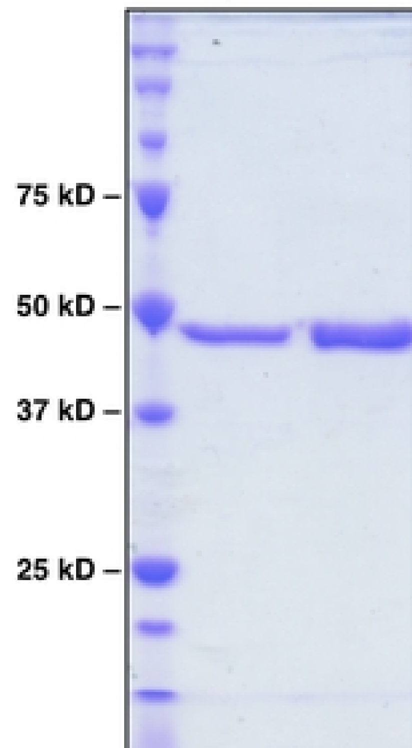
Figure 2

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.04.467337; this version posted November 4, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

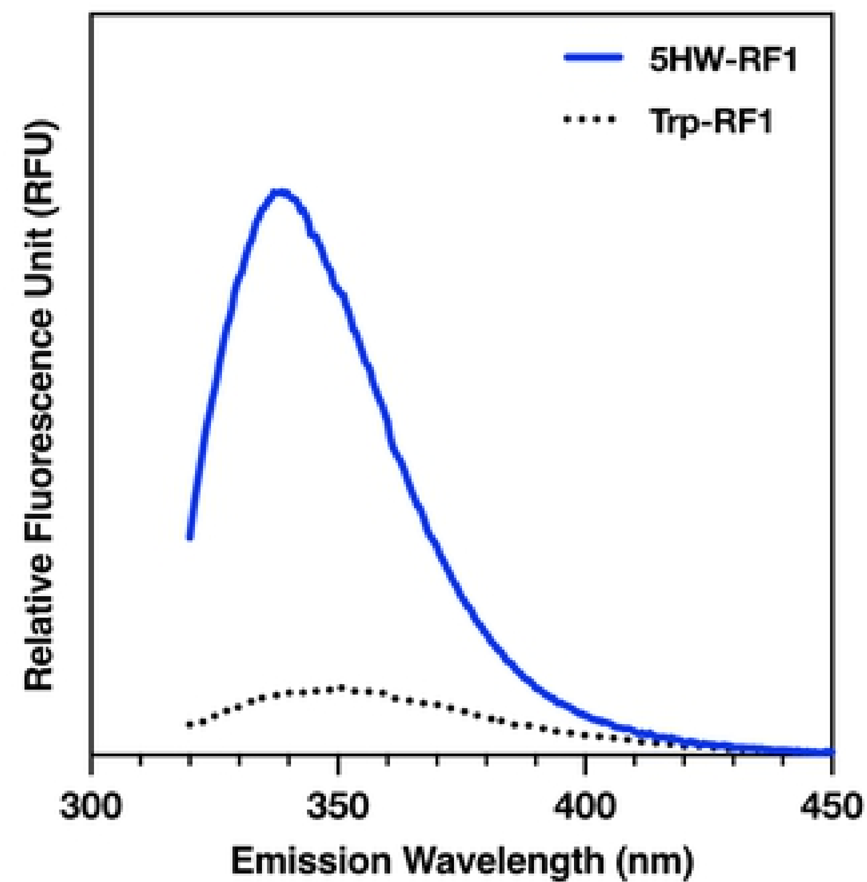
A



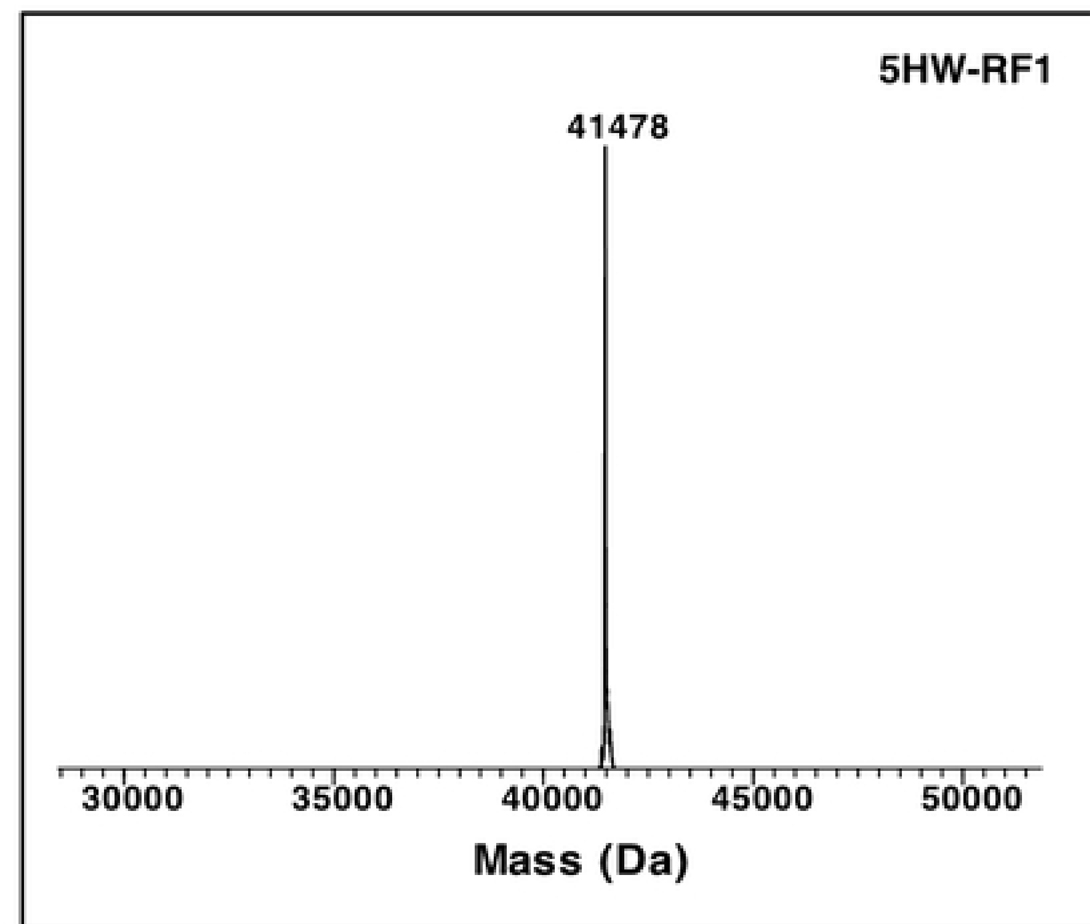
B



C



D



E

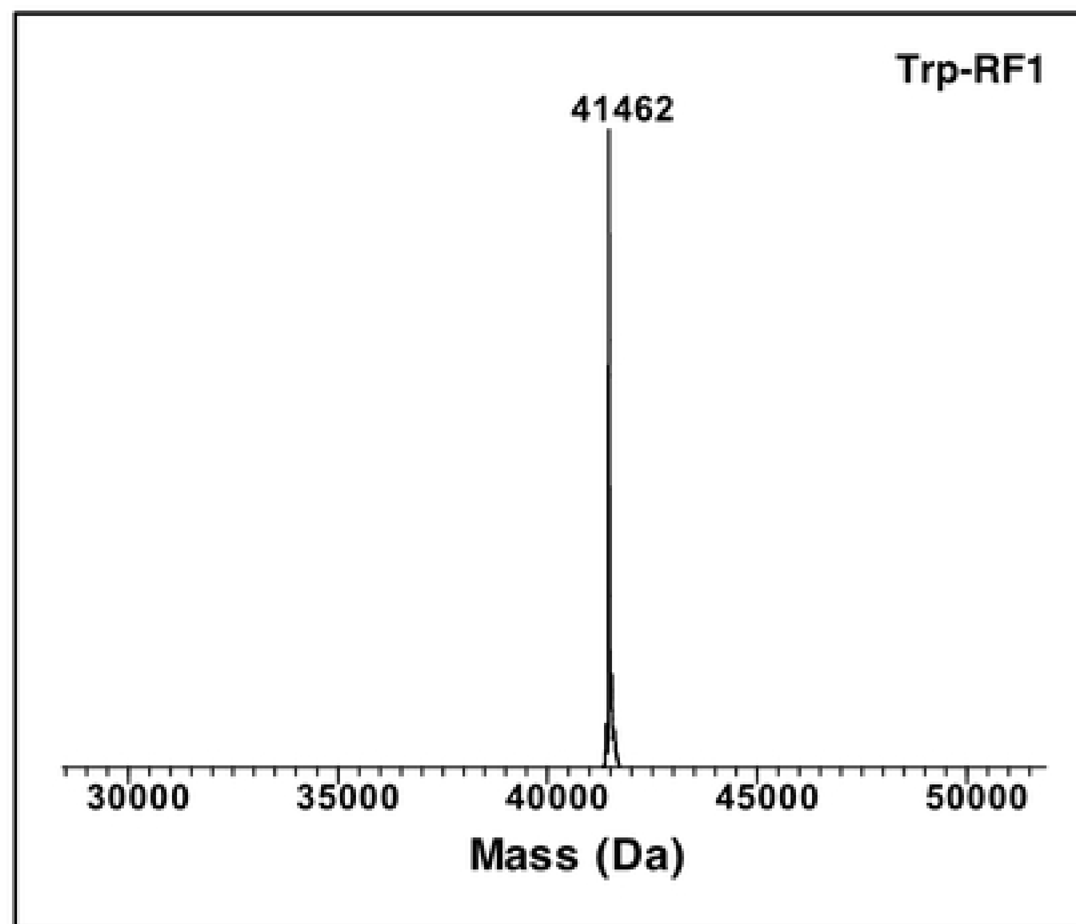


Figure 3

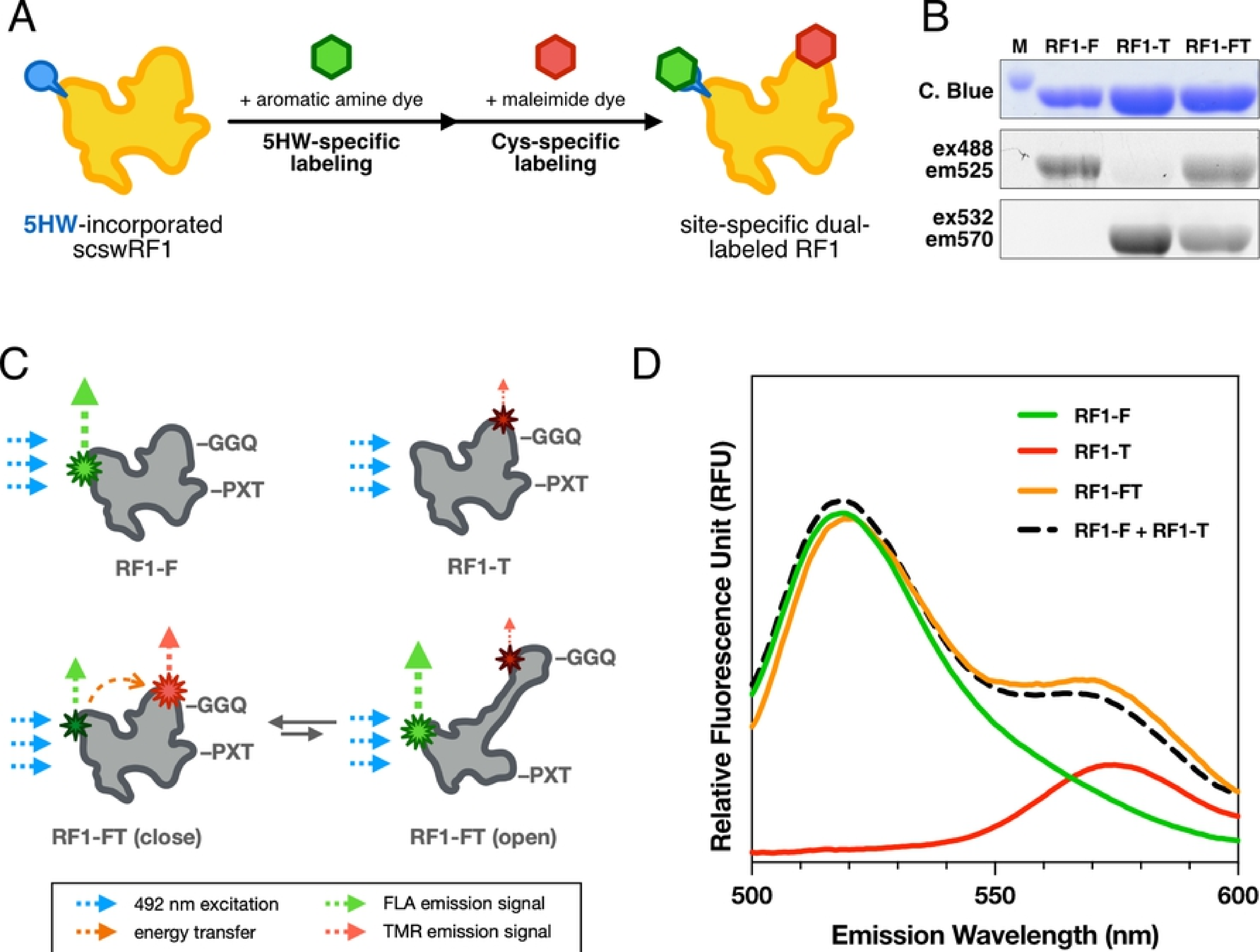


Figure 4