Supporting Information

for

Crystallographic Evidence For Chromophore Isomerisation In 3-
Fluorotyrosyl-Green Fluorescent Protein

Jae Hyun Bae, Prajna Paramita Pal, Luis Moroder, Robert Huber
and Nediljko Budisa*

[A] Fermentation, protein expression and purification
[B] Spectroscopic and mass spectrometric analyses
[C] “Enhanced Cyan Fluorescent Protein” ECFP
[D] Fluorescence pH titration of chromophore
[E] Crystal structure analysis data
[F] Mesomeric structures
[G] References

*Dr. Nediljko Budisa
Max-Planck-Institut für Biochemie
Abt. Strukturforschung & AG Bioorganische Chemie
Am Klopferspitz 18a,
D-82152 Martinsried bei München,
Germany
telephone: +49-89-8578-2661 (office) or 2683 (lab),
facsimile: +49-89-8578-3516
e-mail budisa@biochem.mpg.de
[A] Fermentation, protein expressions and purification

**Chemicals, bacterial strains, and media for the growth.** The amino acids L-Tyrosine and L-3-fluorotyrosine were from Sigma. All other chemicals were purchased from Sigma or Aldrich unless stated otherwise. For all bioincorporation experiments we used the tyrosine deficient *E. coli* strain AT2471 (λ-, tyrA4, relA1, spoT1, thi-) which was a generous gift from Prof. Barbara Bachmann from Yale University, New Haven. Fermentation experiments were performed in New Minimal Medium (NMM) as described elsewhere.

**Growth, culture conditions, expression and purification of labelled proteins.** The incorporation experiments were performed using cultures grown in NMM in the presence of 100mg/l ampicillin and 70mg/l kanamycin and 0.03 mM tyrosine as optimal limiting concentrations of the native substrate at 30 °C. This medium provides enough "healthy" cells before depletion of tyrosine as a native substrate, in the mid-logarithmic phase of growth (OD600 about 0.5 - 0.6) - an optimal point where the non-canonical substrate is added and target protein expression gets switched on by addition of Isopropyl-β-thio-galactoside (IPTG). The culture fermentation after the induction of protein synthesis was 4 hours or overnight at 30 °C.

*E. coli* strain AT2471 transformed with two plasmids in NMM at 30 °C. T5 promotor-based expression systems (Qiagen) inducible with IPTG were used for protein expressions. The protein expression host *E. coli* AT2471 was routinely co-transformed with two plasmids: ampicillin resistant pQE60-EGFP harbouring “enhanced” GFP (Phe64Leu/Ser65Thr) gene sequence (BD Biosciences) which is Nt-His-Tagged (Qiagen) and kanamycin resistant pREP4 containing a repressor gene lacIq. The expression of 3-fluorotyrosyl-EGFP in a soluble form gives similar yields to those of the parent protein for both proteins (10-30 mg/L). Native, and 3-fluorotyrosyl-EGFP were purified through two columns (i) Ni-NTA Agarose (Qiagen) via imidazole gradient (0 – 100 mM) in 100 mM Na-phosphate buffer pH 8.0 and 0.5 M NaCl, and (ii) phenyl-sepharose (Pharmacia) with ammonium sulphate gradient 20 – 0% in 20 mM TrisCl pH 8.0 and 1 mM EDTA. The purity of the proteins was analysed by SDS/PAGE, HPLC and electrospray mass analysis.

[B] Spectroscopic and mass spectrometric analyses

**UV absorbance.** For UV-absorption spectra of the proteins in Phospate Buffered Saline (PBS: 115 mM NaCl, 8 mM KH2PO4, 16 mM Na2HPO4, pH 7.3) were recorded on a Perkin-Elmer Lambda 17 UV/VIS spectrophotometer. Molar extinction coefficients ($\varepsilon_M$) measured at room temperature (20 °C) for native and fluorinated protein in the Tyr+Trp absorption region were determined as described elsewhere.
while chromophore extinction coefficients were derived by normalisation to these values. When compared with EGFP ($\lambda_{\text{max}} = 277$ nm; $\varepsilon_M = 19887 \pm 754$ M$^{-1}$cm$^{-1}$; $\lambda_{\text{max}} = 488$ nm: $\varepsilon_M = 35870 \pm 1316$ M$^{-1}$cm$^{-1}$ ratio: chromophore/(Trp+Tyr) = 1.80) the intensity of absorption profiles of (3-F)Tyr-EGFP is almost unchanged while both maxima are blue shifted for 4 nm ($\lambda_{\text{max}} = 273$ nm: $\varepsilon_M = 20341 \pm 987$ M$^{-1}$cm$^{-1}$; $\lambda_{\text{max}} = 484$ nm: $\varepsilon_M = 35610 \pm 1115$ M$^{-1}$cm$^{-1}$) ratio: chromophore/(Trp+Tyr) = 1.70).

**Fluorescence.** Fluorescence spectra were recorded in PBS on a Perkin-Elmer spectrometer (LS50B) equipped with digital software. Protein samples (0.25 µM; slit 2.5 nm) were excited at 488 nm (wt-EGFP) or at 484 nm ((3-F)Tyr-EGFP) and the emission spectra were recorded in the 495–580 nm.

**Mass spectrometry.** The quantitative replacement of the 11 Tyr-side chains of parent EGFP by the (3-F)Tyr was routinely confirmed by electrospray mass spectrometric analyses (ESI-MS) performed with an ESI PE SCIEX API 165 (Perkin-Elmer) single quadrupole MS system. The spectra for detecting the products were collected with an auto-sampler Series 200 (Perkin-Elmer) at a rate of 10µl/min, an ion source high voltage of 4900 kV, an orifice voltage of 10 V, a dwell time of 0.4 ms per scan and a step size of 0.2 Da with the scan range of 20 to 50. Injection was regulated with the splitter. The injection volume of samples was 10µl dissolved in acetonitrile. The eluent was 100% acetonitrile (0.05% TFA).

[C] “Enhanced Cyan Fluorescent Protein” ECFP

This avGFP variant, known as “enhanced cyan fluorescent protein” ECFP Class VI, Ser65Thr/Tyr66Trp; ($\lambda_{\text{max,1}} = 434$ nm; $\varepsilon_{M1} = 24700 \pm 500$; $\lambda_{\text{max,2}} = 452$ nm; $\varepsilon_{M2} = 23600 \pm 700$ M$^{-1}$cm$^{-1}$; $\lambda_{\text{em,1}} = 476$ nm; $\lambda_{\text{em,2}} = 505$ nm) with an indole ring being an integral part of chromophore, exhibit a characteristic double humped blue-green emission. In this avGFP variant, an indole ring in the chromophore extents toward Val150 and Phe156 and its accommodation is favoured with other compensatory mutations: Phe64Leu/Asn146Ile/ Met153Thr/Val163Ala. It contains one Tyr residue less than EGFP since the Tyr66 in chromophore was replaced with Trp (total number of Tyr-residues in ECFP is 10). Full incorporation of (3-F)Tyr was confirmed by mass spectrometry: ECFP (expected mass: 27707.0 Da; experimental value: 27710.0 ± 3.0 Da ); 3-fluorotyrosyl-ECFP (expected mass: 27885.15 Da; experimental value: 27888.2 ± 4.5 Da ). As expected, global fluorination of ECFP does not influence spectral properties of this protein (3-fluorotyrosyl-ECFP: $\lambda_{\text{max,1}} = 434$ nm ($\varepsilon_{M1} = 23450 \pm 900$); $\lambda_{\text{max,2}} = 452$ nm; ($\varepsilon_{M2} = 23600 \pm 800$) M$^{-1}$cm$^{-1}$; $\lambda_{\text{em,1}} = 476$ nm; $\lambda_{\text{em,2}} = 505$ nm).

Therefore there is not even any marginal difference between ECFP and 3-fluorotyrosyl-ECFP, neither in the intensities of the absorption and emission nor in their maximum. This is a direct proof that observed spectral shifts in EGFP are not
due to the fluorination of surrounding Try-residues but rather due to the fluorination of Tyr66 in the chromophore.

[D] Fluorescence pH titration of chromophore

Degassed and argon-saturated aqueous solutions contained 100 mM NaCl and 50 mM buffering substance: Na-citrate/HCl (pH 1.0); citric acid/Na$_2$HPO$_4$ (pH 3.0, 3.5); Na-acetate/acetic acid (pH 4.5, 5.0, 5.5); Na$_2$HPO$_4$/ NaH$_2$PO$_4$ (pH 6.0, 6.4, 6.8, 7.2, 7.6, 8.0); boric acid/NaOH (pH 9.0); Na-borate/NaOH (pH 8.5, 9.0, 9.5). Fluorescence spectra were recorded on a Perkin-Elmer spectrometer (LS50B) equipped with digital software. Protein samples (0.25 µM) were examined in the pH range from 1 to 11 in different buffers. Probes were excited at 484 and 488 nm respectively, (slit 2.5 nm) and the emission spectra were recorded at 510 and 514 nm, respectively.

![Figure S1](image_url)

**Figure S1.** pH titration profiles for wt-EGFP and 3-fluorotyrosyl-EGFP chromophores measured upon excitation at 484 and 492 nm respectively. Native EGFP has pKa 5.76 (±0.07) while this value for 3-fluorotyrosyl-EGFP is decreased by 0.45 pH unit pKa 5.31 (±0.05).
[E] Crystal structure analysis data

3-fluorotyrosyl-EGFP has been crystallised in 0.05M NaOAc, 0.1M Tris/HCl buffer pH 8.5, and 14% (w/v) Polyethylene glycol 4000 within 2 days. For the crystallisation, hanging drops were made of 1 µl of protein solution (34 µg/µl) and 1 µl of precipitant solution at 20°C. Drops were equilibrated against 0.2 ml of precipitant solution.

The protein structure was solved by the molecular replacement technique. As the platform for the molecular replacement, one of the GFP structures (EGFP, 1EMG) was used\(^4\). The space group of the crystal is primitive orthorhombic, P2\(_1\)2\(_1\)2\(_1\), same as previously solved structures. The data set of EGFP was collected with the resolution 2.1 Å. X-ray diffraction data for 3-fluorotyrosyl-EGFP crystal were collected on an X-ray image plate system (Mar Research, Hamburg, Germany) using CuK\(_{\alpha}\)-radiation generated by a Rigaku rotating anode at 5.4kW in a cryo-protected condition at 100 K. Crystal was transferred to its mother solution containing 20% (w/v) meso-Erythritol as a cryo-protectant and shock-frozen in a nitrogen stream.

Reflections were integrated with the program DENZO\(^4\), scaled and reduced using SCALE\(^4\). Relevant statistics are listed in Table S1. Model building and refinement were performed with CNS\(^5\). The initial model was refined alternating automatic minimisation protocols performed with CNS inspecting visual electron density map and manually adjusted using the program O\(^6\).

<table>
<thead>
<tr>
<th>Crystal (3-F)Tyr-EGFP</th>
<th>Source</th>
<th>CuK?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength, Å</td>
<td>1.54178</td>
<td></td>
</tr>
<tr>
<td>Resolution Range, Å</td>
<td>50-2.1</td>
<td></td>
</tr>
<tr>
<td>Completeness, %</td>
<td>98.8/96.9(2.18-2.10)</td>
<td></td>
</tr>
<tr>
<td>(R_{sym})(^1), %</td>
<td>8.4/25.4</td>
<td></td>
</tr>
<tr>
<td>No. Of protein atoms/waters</td>
<td>1820/61</td>
<td></td>
</tr>
<tr>
<td>(R_{factor})(^2)/((R_{free})(^3), %)</td>
<td>22.3/27.7</td>
<td></td>
</tr>
<tr>
<td>rms. bond lengths, Å</td>
<td>0.006665</td>
<td></td>
</tr>
<tr>
<td>rms. angles, Å</td>
<td>1.42684</td>
<td></td>
</tr>
<tr>
<td>Reflection used(I/Sigma &gt; 0)</td>
<td>13596</td>
<td></td>
</tr>
</tbody>
</table>

\(^1R_{sym} = \frac{\sigma(\langle I(h)\rangle) - \langle \sigma(I(h))\rangle}{\sigma(\langle I(h)\rangle)}; I(h)_i is the observed intensity of the ith measurement of reflection h, and I(h) the mean intensity of reflection h calculated after loading and scaling.

\(^2R_{factor} = \frac{\sigma(|F_{obs}| - |F_{calc}|)}{|F_{obs}|} \times 100\)

\(^3R_{free} was calculated randomly omitting 10% of the observed reflections from refinement and R factor calculation
[F] Mesomeric structures

Figure S2. Mesomeric structures of ionised 3-fluorotyrosyl-EGFP chromophores

[G] References


