Supplementary Information for
Directionality of light absorption and emission in representative fluorescent proteins

Jitka Myšková¹, Olga Rybakova¹,², Jiří Brynda¹,³, Petro Khoroshyy¹, Alexey Bondar¹,², Josef Lazar¹,²*

*Josef Lazar, Institute of Organic Chemistry and Biochemistry CAS, Flemingovo nám. 2, 16610 Prague 6, Czech Republic, phone: +420 723 415 295
**Email: lazar@uochb.cas.cz

This PDF file includes:

- Supplementary text
- Figures S1 to S6
- Tables S1 to S4

-
Supplementary Information Text
Supplementary Methods: Calculating Ratios of Extinction Coefficients ($\varepsilon_\parallel / \varepsilon_\perp$).

Based on Lambert-Beer’s law we can write for light polarized parallel and perpendicular to the long axis of crystal:

$$A_\parallel = \log \frac{I_{0\parallel}}{I_\parallel} = \varepsilon_\parallel \cdot c \cdot l \quad A_\perp = \log \frac{I_{0\perp}}{I_\perp} = \varepsilon_\perp \cdot c \cdot l$$  \hspace{1cm} (1, 2)

The corresponding crystal transmittances can then be described as

$$T_\parallel = \frac{I_\parallel}{I_{0\parallel}} = 10^{-\varepsilon_\parallel \cdot c \cdot l} \quad T_\perp = \frac{I_\perp}{I_{0\perp}} = 10^{-\varepsilon_\perp \cdot c \cdot l}$$

The above relationships were used to determine the values of extinction coefficients $\varepsilon_\parallel, \varepsilon_\perp$. However, in order to make our TDM direction determinations as accurate as possible, we wanted to avoid introducing unnecessary experimental errors (from measurements of crystal thickness and protein concentration) Therefore, for determining TDM directions, we used ratios of extinction coefficients $\varepsilon_\parallel, \varepsilon_\perp$, requiring only measurements of light transmission:

$$\frac{\varepsilon_\parallel}{\varepsilon_\perp} = \frac{\log T_\parallel}{\log T_\perp} = \frac{\log \frac{I_\parallel}{I_{0\parallel}}}{\log \frac{I_\perp}{I_{0\perp}}}$$  \hspace{1cm} (3)

Eq. 3 was used to calculate the values of $\varepsilon_\parallel / \varepsilon_\perp$ for measurements with excitation wavelengths of 545 and 593 nm, whose polarization purity was better than 99.8%. For excitation with 405 nm, 458 nm and 488 nm light, we amended our calculations to include contributions of polarizations perpendicular to the one desired, in the following manner.

If the light polarized parallel to the long axis of crystal contains a fraction $(1 - x_\parallel)$ of perpendicular polarization, we can write:

$$I_\parallel = x_\parallel I_0 10^{-\varepsilon_\parallel \cdot c \cdot l} + (1 - x_\parallel) I_0 10^{-\varepsilon_\perp \cdot c \cdot l}$$  \hspace{1cm} (4)

Similarly, for light polarized perpendicular to the long axis of crystal containing a fraction $(1 - x_\perp)$ of parallel polarization, we write:

$$I_\perp = x_\perp I_0 10^{-\varepsilon_\perp \cdot c \cdot l} + (1 - x_\perp) I_0 10^{-\varepsilon_\parallel \cdot c \cdot l}$$  \hspace{1cm} (5)

By combining equations 4 and 5, we can express the individual extinction coefficients:

$$\varepsilon_\perp = -\frac{1}{c \cdot l} \log \left( \frac{x_\parallel - 1}{x_\perp + x_\parallel - 1} \right) = -\frac{1}{c \cdot l} \log \left( \frac{x_\parallel T_\parallel + x_\perp T_\perp}{x_\perp + x_\parallel - 1} \right)$$  \hspace{1cm} (6)

$$\varepsilon_\parallel = -\frac{1}{c \cdot l} \log \left( \frac{x_\perp I_0 - (1 - x_\parallel) I_0}{x_\parallel + x_\perp - 1} \right) = -\frac{1}{c \cdot l} \log \left( \frac{(x_\parallel - 1) T_\parallel + x_\perp T_\perp}{x_\parallel + x_\perp - 1} \right)$$  \hspace{1cm} (7)
Therefore, if we know the polarization purity ($x_∥, x_⊥$; Table S2), we can calculate $\varepsilon_∥/\varepsilon_⊥$ as follows:

$$\frac{\varepsilon_∥}{\varepsilon_⊥} = \frac{\log \left( \frac{(x_∥ - 1) T_⊥ + x_⊥ T_∥}{x_∥ + x_⊥ - 1} \right)}{\log \left( \frac{(x_⊥ - 1) T_∥ + x_∥ T_⊥}{x_∥ + x_⊥ - 1} \right)}$$

(8)

For light of pure polarization, expression 8 becomes equivalent to expression 3.
Supplementary Methods: Mathematical modeling of light absorption by FP crystals.

At the core of the mathematical model is a simple $\cos^2$ relationship describing the rate of light absorption as a function of the angle between the xTDM vector and the electric field vector of the incoming light. The model takes into account (i) the distinct molecular orientations present within a crystallographic unit cell, (ii) orientation of this unit cell within a crystal, as well as (iii) orientation of the crystal within the laboratory coordinate frame. Here we illustrate the crucial steps of our approach on crystals of the P2₁2₁2₁ space group, as implemented in the computational software Mathematica. For other space groups, calculations were carried out analogously.

After storing the list of PDB coordinates of the fluorophore atoms in a variable (fluorAtoms), the fluorophore plane was defined by calculating its normal direction (fluoNormal) using singular value decomposition:

\[
\{U, S, V\} = \text{SingularValueDecomposition}[Y = # - \text{Mean}@# &\text{fluorAtoms}[\text{Transpose}]]; \\
\text{fluoNormal} = \text{Normalize}[U[[;;, 3]]];
\]

Coordinates of the aromatic ring centers (ring1center, ring2center) were calculated by averaging the coordinates of the member atoms of each aromatic ring. A vector describing the direction of the line connecting the two ring centers was then calculated as:

\[
\text{centerLineApprox} = \text{ring1center} - \text{ring2center};
\]

Since fluorophore atoms do not lie exactly within a plane, neither does this vector. Its projection into the fluorophore plane (centerLine) was therefore calculated and subsequently used to define the angle $\tau$ describing the TDM directions:

\[
\text{centerLine} = \text{Normalize}[	ext{Cross}[\text{Cross}[\text{centerLineApprox}, \text{fluoNormal}], \text{fluoNormal}]]];
\]

Coordinates of a unit-sized TDM vector within the fluorophore plane were then defined as a function of an angle $\tau$, describing a rotation of the line connecting the centers of the aromatic rings (centerLine) around the fluorophore plane normal (fluoNormal):

\[
\text{pdbTDM}[\tau_] := \text{RotationMatrix}[\tau, \text{fluoNormal}].\text{centerLine};
\]

TDM vectors for distinct molecules within a crystallographic unit cells can be calculated by applying crystallographic symmetry operations to the coordinates of the TDM for one of the molecules. The symmetry rotations for the P2₁2₁2₁ space group are defined by the matrices:

\[
sym1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad sym2 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad sym3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad sym4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.
\]

The coordinates of the TDMs (in the crystal coordinate system) for the four molecules within a P2₁2₁2₁ unit cell were therefore defined as:

\[
\text{pdbTDM1}[\tau_] := \text{sym1.pdbTDM}[\tau]; \\
\text{pdbTDM2}[\tau_] := \text{sym2.pdbTDM}[\tau]; \\
\text{pdbTDM3}[\tau_] := \text{sym3.pdbTDM}[\tau]; \\
\text{pdbTDM4}[\tau_] := \text{sym4.pdbTDM}[\tau];
\]

In order to find the TDM coordinates in the laboratory coordinate system, the orientation of the crystallographic unit cell within the crystal was taken in account. In all the P2₁2₁2₁ crystals we investigated, the crystallographic axes B, C coincided with the diagonal directions of the crystal.
cross-section. Therefore, in crystals positioned flat on the glass cover slip, the crystallographic unit cell was rotated with respect to the xy plane of the microscope by an angle ($\gamma$). If we define the long axis of the crystal (which in our $P2_12_12_1$ crystals coincides with the crystallographic axis A) to be the laboratory axis $x$, we can describe the TDM orientations in the laboratory coordinate system:

\[
\text{roll}[\gamma_] := \text{RotationMatrix}[\gamma, \{1, 0, 0\}];
\]

\[
\text{TDM1}[\tau_, \gamma_] := \text{roll}[\gamma].\text{pdbTDM1}[\tau];
\]

\[
\text{TDM2}[\tau_, \gamma_] := \text{roll}[\gamma].\text{pdbTDM2}[\tau];
\]

\[
\text{TDM3}[\tau_, \gamma_] := \text{roll}[\gamma].\text{pdbTDM3}[\tau];
\]

\[
\text{TDM4}[\tau_, \gamma_] := \text{roll}[\gamma].\text{pdbTDM4}[\tau];
\]

For light propagating along the laboratory z-axis (vertically), the electric field vector (elVector) can be defined as a function of an angle (polAngle) with respect to the $x$-axis (coinciding with the crystal long axis):

\[
elVector[\text{polAngle}_] := \{\text{Cos}[\text{polAngle}], \text{Sin}[\text{polAngle}], 0\};
\]

Combining the above functions allows expressing the absorption rate of a crystal as a function of three variables: $\tau$ (TDM orientation within the fluorophore plane), $\gamma$ (crystal roll angle), and polAngle (direction of excitation light polarization):

\[
\text{absorptionRate}[\tau_, \gamma_, \text{polAngle}_] :=
(\text{elVector}[\text{polAngle}].\text{TDM1}[\tau, \gamma])^2 + (\text{elVector}[\text{polAngle}].\text{TDM2}[\tau, \gamma])^2 + (\text{elVector}[\text{polAngle}].\text{TDM3}[\tau, \gamma])^2 + (\text{elVector}[\text{polAngle}].\text{TDM4}[\tau, \gamma])^2;
\]

In principle, the TDM orientation within the fluorophore ($\tau$) can then be calculated from known values of $\gamma$, polAngle, and absorptionRate. However, the absorption rate in this equation is expressed in units of an unknown size. To bypass this issue, we solved for $\tau$ an equation using the unitless ratio (or logratio) of absorption rates of light polarized parallel and perpendicular to the long axis of the crystal (polAngle equal to 0 and $\pi$/2, respectively):

\[
\text{Solve}[\text{absorptionRate}[\tau_, \gamma_, 0]/\text{absorptionRate}[\tau_, \gamma_, \pi/2] == \text{observedAbsRateRatio}, \tau]
\]

This approach invariably yielded two values of $\tau$ (Fig 2, Table 1), of which one (designated as $\tau_1$) was consistent with results of our measurements of crystals of various tilts along their long axis (angles $\gamma$), while the other solution (designated as $\tau_2$) varied with crystal tilt, and therefore was not consistent with the physical concept of the transition dipole moment.

Apart from calculating the TDM orientation (angle $\tau$), the above approach also allowed us to generate plots of logratios of extinction coefficients for various TDM orientations ($\tau$) (Fig. S3) and crystal orientations ($\gamma$) (Fig. S4), for example:

\[
\gamma = 68.03/180 \pi; \ \tau = ;
\]

Plot[absorptionRate[$\tau$, $\gamma$, 0]/absorptionRate[$\tau$, $\gamma$, $\pi$/2], { $\tau$, -$\pi$/2, $\pi$/2}]

or

\[
\gamma = ; \ \tau = 3.6/180 \pi ;
\]

Plot[absorptionRate[$\tau$, $\gamma$, 0]/absorptionRate[$\tau$, $\gamma$, $\pi$/2], { $\gamma$, -40/180$\pi$, 40/180$\pi$}]


Fig. S1. Microscopy setups. (A) Setup for measurements of linear dichroism (polarization resolved light absorption); (B) Setup for measurements of fluorescence polarization.
Fig. S2. Atoms used to define the fluorophore plane. (A) mTurquoise2; (B) eGFP, green form of mEos4b; (C) mCherry; (D) red form of mEos4b.
Fig. S3. Mathematical modeling predictions of log₂(ε∥/ε⊥) and log₂(F∥/F⊥) values for various TDM orientations (angle τ) within the fluorophore plane. Experimentally determined values of log₂(ε∥/ε⊥) and log₂(F∥/F⊥) and the corresponding TDM orientations (τ₁, τ₂) are indicated. (A) mTurquoise2; (B) eGFP (pH 7.5); (C) eGFP (pH 3.8); (D) mCherry; (E) mEos4b (green form); and (F) mEos4b (red form)
Fig. S4. Distinguishing between \( \tau \) values by observations of linear dichroism of crystals tilted along their long axis. (A) mTurquoise2; (B) eGFP; (C) mCherry; (D) mEos4b (green form); (E) mEos4b (red form). Points: experimental data from individual crystals. Curves: predictions for the two \( \tau \) values (\( \tau_1, \tau_2 \)) consistent with data obtained from crystals oriented horizontally.
Fig. S5. Fluorescence anisotropy of FP solutions. (A) Fluorescence anisotropy as a function of excitation wavelength for the investigated FPs. Points represent mean values, error bars show 95% confidence intervals. Values of the angle between the xTDM and mTDM ($\beta$) corresponding to the observed fluorescence anisotropy are indicated on the right y-axis. (B) Correlation between values of the angle between the xTDM and mTDM determined in FP solutions ($\beta$) and in FP crystals ($\Delta\tau_1$; difference between $\tau_1$ values determined for an xTDM and a corresponding mTDM). A linear relationship is apparent (indicated by a dashed line).
Fig. S6. Fluorophore bend in the green form of mEos4b. A stereo view of the structures of the fluorophore of eGFP (green) and of the green form of mEos4b (orange)
Table S1. Diffraction data and structure refinement statistics.

|                      | mTurquoise2 | eGFP pH7.5 | eGFP pH3.8* |
|----------------------|-------------|------------|-------------|
| **Data collection statistics** |             |            |             |
| Wavelength (Å)       | 1.5419      | 1.5419     | 1.5419      |
| Space group          | $P2_12_12_1$| $P2_12_12_1$| $P2_12_12_1$|
| Cell parameters (Å, o) | 51.15 61.40 68.60 | 50.940 62.12 68.86 | 50.97 62.15 68.79 |
| Resolution range (Å) | 50.0 - 1.85 (1.90 - 1.85) | 50.00 - 1.63 (1.67 - 1.63) | 50.00 - 1.55 (1.59 - 1.55) |
| Number of unique reflections | 18997 (1375) | 24211 (644) | 61255 (4560) |
| Multiplicity         | 12.6 (6.8)  | 5.9 (1.7)  | 5.8 (2.5)   |
| Completeness (%)     | 99.9 (99.2) | 86.0 (31.6) | 99.9 (99.9) |
| $R_{merge}^a$        | 5.2 (190.6) | 7.7 (162.3) | 5.5 (96.6)  |
| $CC_{1/2}(\%)$       | 100 (49.6)  | 99.9 (28.5) | 99.9 (37.3) |
| Average I/σ(I)       | 29.5 (1.0)  | 15.3 (0.55) | 16.9 (1.0)  |
| Wilson B (Å²)        | 38.6        | 26.2       | 26.1        |
| **Refinement statistics** |             |            |             |
| Resolution range (Å) | 39.27 - 1.85 (1.90 - 1.85) | 39.39 - 1.65 (1.69 - 1.65) | 46.12 - 1.55 (1.59 - 1.55) |
| No. of reflections in working set | 18046 (1299) | 22868 (1055) | 30780 (2250) |
| No. of reflections in the test set | 950 (68) | 1204 (56) | 1620 (118) |
| $R_{work}$ value (%)$^b$ | 18.4 (34.6) | 17.8 (35.8) | 18.6 (26.2)) |
| $R_{free}$ value (%)$^c$ | 21.8 (35.2) | 21.6 (30.1) | 21.3 (27.0) |
| RMSD bond length (Å) | 0.02 | 0.014 | 0.013 |
| RMSD angle (°)       | 1.7 | 1.9 | 1.7 |
| Mean ADP value (Å²)  | 42.4 | 24.2 | 23.1 |
| **Ramachandran plot statistics** |             |            |             |
| Residues in favored regions | 98.6 % | 98.2 % | 98.6 % |
| Residues in allowed regions | 0.0 % | 0.5 % | 0.0 % |
| PDB code             | 6YLN        | 6YLQ       | 6YLP        |
## Data collection statistics

|                         | mCherry | mEos4b* |
|-------------------------|---------|---------|
| **Wavelength (Å)**      | 1.5419  | 1.5419  |
| **Space group**         | C2      | P2₁2₁2₁ |
| **Cell parameters (Å, °)** |    |      |
|                         | 107.28  | 42.92   | 85.77   | 38.50  | 70.47 | 90.57  |
|                         | 90.00   | 128.50  | 90.00   | 90.00  | 90.00 | 90.00  |
| **Resolution range (Å)** | 50.0 - 1.60 | 50.0 - 1.55 |
|                         | 1.59 - 1.55 | 1.63 - 1.55 |
| **Number of unique reflections** | 39100 (2425) | 61581 (1747) |
| **Multiplicity**        | 6.2 (4.4) | 2.5 (1.2) |
| **Completeness (%)**    | 96.0 (81.8) | 89.0 (33.9) |
| **R<sub>merge</sub>**   | 6.6 (122.3) | 6.5 (32.4) |
| **CC(1/2) (%)**         | 99.9 (50.6) | 99.7 (74.9) |
| **Average I/σ(I)**      | 14.8 (1.1) | 9.2 (1.0) |
| **Wilson B (Å<sup>2</sup>)** | 29.35 | 21.0 |

## Refinement statistics

|                         | mCherry | mEos4b* |
|-------------------------|---------|---------|
| **Resolution range (Å)** | 67.12 - 1.60 | 35.44 - 1.55 |
|                         | 1.64 - 1.60 | 1.59 - 1.55 |
| **No. of reflection in working set** | 37956 (2342) | 33203 (1404) |
| **No. of reflection in the test set** | 1173 (72) | 1027 (43) |
| **R<sub>work</sub> value (%)**<sup>b</sup> | 21.1 (33.0) | 18.5 (24.9) |
| **R<sub>free</sub> value (%)**<sup>c</sup> | 21.6 (39.4) | 21.4 (33.2) |
| **RMSD bond length (Å)** | 0.014 | 0.014 |
| **RMSD angle (°)**      | 1.7 | 1.8 |
| **Mean ADP value (Å<sup>2</sup>)** | 28.1 | 21.1 |

## Ramachandran plot statistics

|                         | mCherry | mEos4b* |
|-------------------------|---------|---------|
| **Residues in favored regions** | 98.6 % | 99.1 % |
| **Residues in allowed regions** | 0.0 % | 0.5 % |
| **PDB code**            | 6YLM    | 6YLS    |

Data in parentheses refer to the highest-resolution shell for data collection statistic.

<sup>a</sup> Marks data sets where Friedel's Pairs were not merged.

<sup>b</sup> R-value = ∥F<sub>o</sub> - |F<sub>c</sub>||/∥F<sub>o</sub>∥, where F<sub>o</sub> and |F<sub>c</sub>| are the observed and calculated structure factors, respectively.
Table S2. Polarization purity of the illuminating light.

| Wavelength | 0° (||) | 45° | 90° (⊥) | 135° | 180° (||) |
|------------|--------|-----|---------|------|-----------|
| 405 nm     | 99.7%  | 99.2% | 98.5%  | 98.8% | 99.5%     |
| 458 nm     | 99.8%  | 99.7% | 99.2%  | 99.5% | 99.8%     |
| 488 nm     | 99.8%  | 99.8% | 99.5%  | 99.5% | 99.8%     |
| 543 nm     | > 99.9% | 99.9% | 99.8%  | 99.8% | > 99.9%   |
| 594 nm     | > 99.9% | > 99.9% | > 99.9% | > 99.9% | > 99.9% |

Table S3. Molar extinction coefficients derived from FP crystals (in M⁻¹ cm⁻¹). The values listed have been rounded to the nearest hundred. N = 10 for all conditions listed. Information on the red form of mEos4b is not provided, as the concentration of the red form within mEos4b crystals could not be reliably ascertained.

| Fluorescent protein | Wavelength | ε₀ (mean ± 95% CI) | ε┴ (mean ± 95% CI) |
|---------------------|-------------|---------------------|---------------------|
| mTurquoise2         | 405 nm      | 23800 ± 3400        | 5500 ± 600          |
|                     | 458 nm      | 30300 ± 6800        | 8800 ± 1200         |
| eGFP, pH 7.5        | 405 nm      | 19700 ± 1800        | 4700 ± 800          |
|                     | 488 nm      | 28700 ± 7000        | 10200 ± 2000        |
| eGFP, pH 3.8        | 405 nm      | 26900 ± 3200        | 8100 ± 1600         |
|                     | 488 nm      | 7700 ± 5500         | 2100 ± 800          |
| mCherry             | 543 nm      | 1500 ± 500          | 1000 ± 400          |
|                     | 594 nm      | 2400 ± 700          | 1500 ± 500          |
| mEos4b              | 488 nm      | 30200 ± 700         | 1900 ± 400          |
Table S4: PDB coordinates of transition dipole moments.

| Fluorescent protein          | Wavelength   | PDB coordinates     |
|-----------------------------|--------------|---------------------|
| mTurquoise2 (6YN)           | 405 nm (xTDM)| -0.8502, 0.4115, 0.3283 |
|                             | 458 nm (xTDM)| -0.8482, 0.4091, 0.3365 |
|                             | 489 – 531 nm (mTDM) | -0.8509, 0.4123, 0.3255 |
| eGFP, pH 7.5 (6YLQ)         | 405 nm (xTDM)| -0.8586, 0.4102, -0.3076 |
|                             | 488 nm (xTDM)| -0.8457, 0.3922, -0.3619 |
|                             | 529 – 555 nm (mTDM) | -0.8642, 0.4191, -0.2783 |
| eGFP, pH 3.8 (6YLW)         | 405 nm (xTDM)| -0.8648, -0.4213, 0.27309 |
|                             | 488 nm (xTDM)| -0.8422, -0.3859, 0.3766 |
|                             | 529 – 555 nm (mTDM) | -0.8414, -0.3850, 0.3792 |
| mCherry (6YLW)              | 543 nm (xTDM)| 0.4878, -0.5450, -0.6820 |
|                             | 594 nm (xTDM)| 0.4628, -0.5466, -0.6979 |
|                             | 600 – 690 nm (mTDM) | 0.3240, -0.5469, -0.7720 |
| mEeos4b (green form, 6YLS)  | 488 nm (xTDM)| 0.9678, -0.0094, 0.25153 |
|                             | 529 – 555 nm (mTDM) | 0.9491, -0.0660, 0.3081 |
| mEeos4b (red form, 6YLS)    | 594 nm (xTDM)| 0.9221, -0.1258, 0.3659 |
|                             | 600 – 650 nm (mTDM) | 0.8746, -0.2049, 0.4395 |