

FRET: Data analysis and protocol preparation

Chenyu Jin

Experimentalphysik I, Universität Bayreuth, Bayreuth, Germany.
September 3, 2021

In the practice course, we are to measure the FRET efficiency in different cells with different methods. This is a guide for the data analysis and protocol preparation.

1 Sensitize Emission

In this part, we take for each cell three images:

- **Donor**, excited with donor excitation wavelength, and collect signal in donor emission channel;
- **s.e.** (sensitized emission), excited with donor excitation wavelength, and collect signal in acceptor emission channel;
- **Acceptor**, excited with acceptor excitation wavelength, and collect signal in acceptor emission channel;

Then we calculate FRET efficiency based on these three images and some correction factors.

This method, including the data analysis is not as straight-forward as it seems. There are many things to be pay attention to. Following is the explanation of some pitfalls and a recommended protocol of data analysis.

1.1 Dark current, background fluorescence and shading

Even if the signal-to-noise ratio on an image is high enough, we still need to clean up the signal before further analysis. In our case, what you will find out is, even if a ROI without any cell is

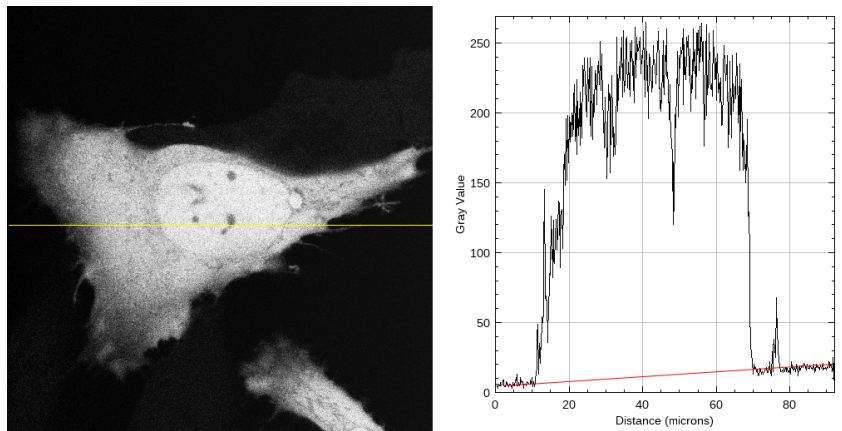


Figure 1: An example image of the cell, and the intensity profile along the yellow line, showing the non-zero background and the shading (the inclination of the base line).

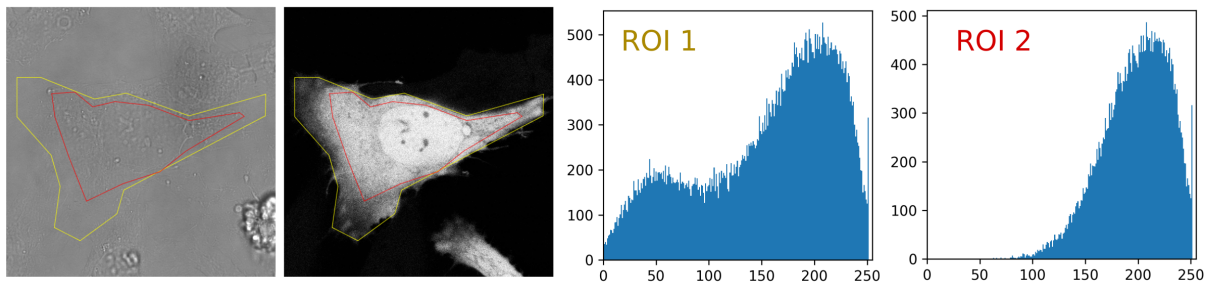


Figure 2: Histogram of intensity when different ROIs are selected.

selected, the intensity measured by our detector is still not zero. This background is from (a) the dark current of the detector and (b) the background fluorescence on the slide. It adds to the real signal, and has to be substrated from the measured signal.

Another error is induced by the shading. Here it is mostly caused by the inhomogeneous illumination. But in our case the effect brought by shading is minor so we will not correct shading in our analysis.

1.2 Selection of the region of interest (ROI)

Fluorescence is not homogeneous over the whole image. In our experiment, the membrane is the brightest, the cell body might also be fluorescent, and the surrounding should be dark. In a homogeneous region, the intensity distribution of pixels should be Gaussian (Poisson if too dark). Ideally, we should select only the “bright” homogeneous region as the ROI. Figure 2 shows what happens when some darker part is also included in the ROI.

1.3 Image processing protocol

The images can of course be processed with Python or Matlab. However, since we only need to process tens of images. I recommend Fiji (advanced version of ImageJ).

During the experiment, we have taken four images for each cell: **Donor**, **s.e.**, **Acceptor** and **BF** (bright field). They are not yet the **D**, **S** and **A** that we can feed to the equation to calculate the FRET efficiency. Some pre-processing is necessary:

- Split frames. Convert **Donor**, **s.e.**, **Acceptor** to 32-bits.
- Select a cell-free region on **BF**, measure the mean intensity on **Donor**, **s.e.**, **Acceptor**. These are the corresponding background values.
- If we only need a mean value, here we can draw ROI over the bright part, measure the mean intensity, minus the background values and feed the corrected intensity to the equations.
- To do image calculation, use “Process – Math” to substrate each image with the corresponding background value.
- The background substration will lead to some very small, even negative intensity values here and there on the image. You can hit “H” and see from the histogram. Later in image division these small values will cause problems, especially as denominator. We hence need a threshold to get rid of these small values. Open “Image – Adjust – Threshold”, set the lower threshold (upper bar) to around 2 or 3, click “Apply” and choose “Set to NaN”. Now all the pixels with a intensity value lower than the threshold will have a value “NaN”, which does not participate in calculation.

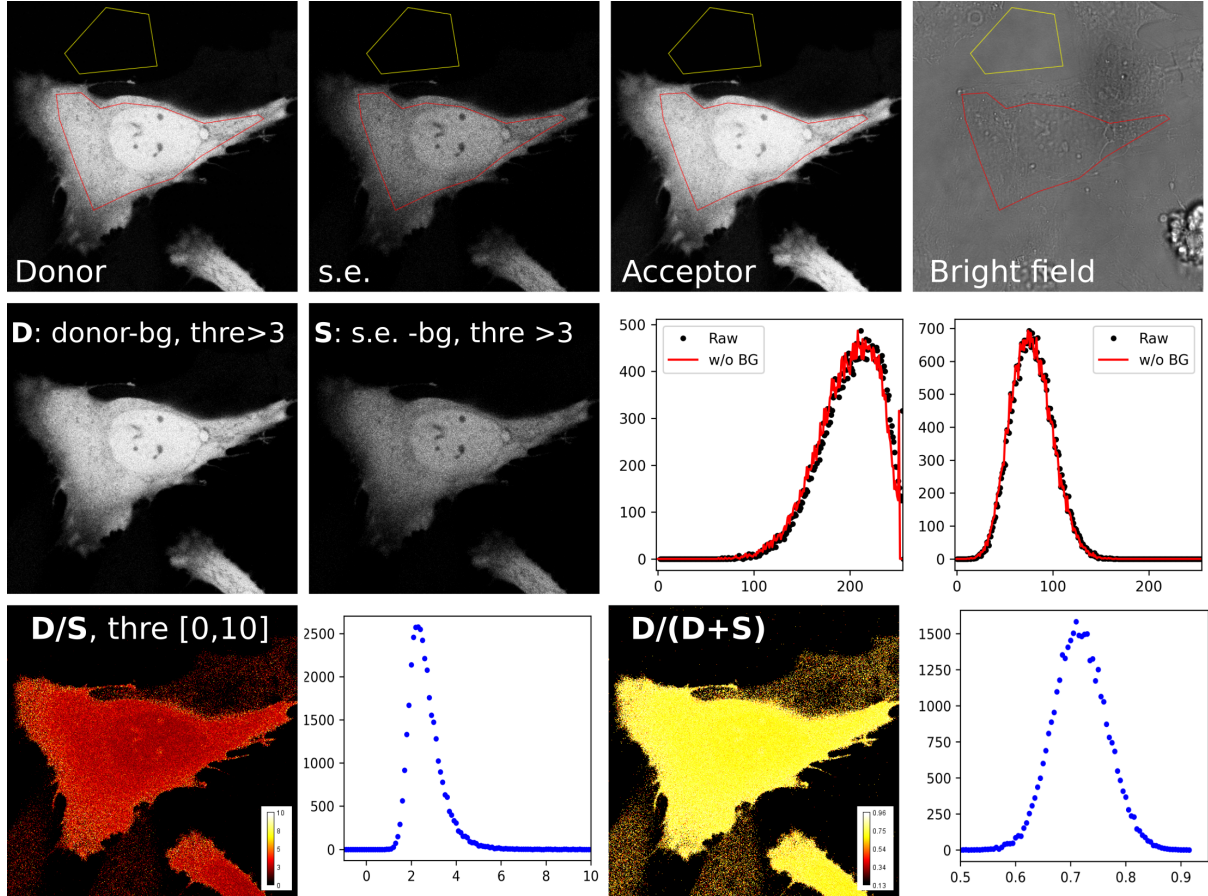


Figure 3: Steps of data processing, and the resulting histograms inside the red ROI. Before the image division, the **D** and **S** images are thresholded between [3, 250].

The result images are **D**, **S** and **A**. Use “Process – Math” or “Process – Image calculator” to calculate the “SE” and “FRET efficiency” images according to the equations. Choose a lookup table for the result FRET efficiency image (“Image – Lookup Tables”, e.g., “Red Hot”).

The FRET efficiency image now shows the FRET efficiency of each pixel. Please note that at each pixel it should be a number between 0 and 1. The whole image is actually a 2D plot. It is only meaningful when you add the calibration bar from “Analyze – Tools”. Don’t forget to add the scale bar as well.

Then please select on the FRET efficiency image the same ROI that you use to measure the mean intensity, and compare the mean FRET efficiency with what you calculated with the mean intensity.

2 Donor emission after the acceptor bleaching

There is not lot of data analysis to do in this part. In your protocol, please include:

- A typical image showing the ROIs.
- Plot of fluorescence intensity to time in donor and s.e. channel.
- Calculate the FRET efficiency with

$$E = 1 - \frac{D_{CY,pre}}{D_{CY,post}}$$

3 Fluorescence Life-Time Measurement (FLIM)

In this part, we have measured the distribution of time between the excitation pulse and the emission photon. By fitting the histogram with exponential distribution we will get the life-time of the fluorescence.

There is only one excitation wavelength: 470 nm. Signals are collected in both donor and acceptor channels. Hence for the three types of cells:

- C cells: donor signal in channel 1, leaked-through donor signal in channel 2.
- Y cells: cross-excited leak-through signal in channel 1, cross-excited signal in channel 2.
- CY cells: donor signal in channel 1, FRET signal in channel 2, mixed signal in channel 1+2. Note that only the dominating signals are listed here.

The first task of data analysis is single exponential fit to the histograms (declining branch). The fitting result will be the life-time of the corresponding fluorescence.

The second task is double-exponential fit to the histogram of the mixed signals (CY cells, channel 1+2): $f(t) = A_1 \cdot \exp(-t/\tau_1) + A_2 \cdot \exp(-t/\tau_2)$. There are 4 parameters to be found out with this fitting. Fitting with so many parameters will not give a meaningful result. The most famous example here is a quota by John von Neumann: “With four parameters I can fit an elephant, and with five I can make him wiggle his trunk”.

What we can do is to go back to the original purpose of fitting: to find out a reasonable function that is most close to the data. Here is one possible approach:

- Inherit A_1 from the fitting result of channel 1.
- Make A_2/A_1 the parameter to be found out. Take an initial guess from the fitting result of channel 2.
- Try a few numbers around the initial guess. Each time fit the data with only τ_1 and τ_2 with the fitting parameters. Calculate the abbreviation $\sum (y_{\text{data}} - y_{\text{fit}})^2$
- Find the A_2/A_1 at which the abbreviation is the smallest.

The third task is to fit the full data range. You only need to do it with one dataset, preferentially one from CY/channel 2, CY/channel 1+2, or Y/channel 2, because the signal-to-noise ratio is high.

The detected signal is the convolution (Deutsch: Faltung) of the real signal and the Instrument Response Function (IRF). So what you need is the convolution function (numpy.convolve, scipy.signal.convolve, or conv in Matlab). Due to the deadtime of the detector t_{tot} after the laser pulse, the measured signal $f(t)$ and the real signal $e(t)$ should be:

$$f(t) = \text{IRF} * e(t),$$
$$e(t) = \begin{cases} 0 & (t < t_{\text{tot}}) \\ A \cdot \exp(-t/\tau) & (t > t_{\text{tot}}) \end{cases}$$

A possible approach:

- Guess a t_{tot} from the measured signal (beginning of the rise, peak position, etc.).
- Take the result of the single exponential fit as the initial guess. You can fix A and aim for optimize τ .

- Try a few numbers around the initial guess. Calculate the abbreviation (as in the second task) after convolution, and find out the best fit.

Please do this with two types of IRFs: the one given by the TCSPC program, and a Gaussian curve deduced from the rising branch of the histogram.

Please include plots of your fitting result. Also, you are very encouraged to use other approaches, e.g., error function.

The FRET efficiency is then estimated with

$$E = 1 - \frac{\tau_{\text{donor in C}}}{\tau_{\text{donor in CY}}}$$