

smCellFRET

Documentation

Release 1.3

December 2020, Javitch Lab

Columbia University / RFMH

1. Overview

The *smCellFRET* analysis pipeline was built to streamline the data analysis of large single molecule FRET (smFRET) data acquired by Live-Cell TIRF microscopy. It was developed in cooperation with the Blanchard lab and utilizes functions of the single molecule analysis platform *SPARTAN* [1]. Before employing the pipeline, smFRET imaging data needs to be pre-analyzed using the open source particle tracking software *u-track* [2]. Tracking results acquired with *u-track* will then serve as the main input for the analysis pipeline. The software also uses functions published in the context of transient mobility analysis, called "divide-and-conquer moment scaling spectrum" (*DC-MSS*) [3], to detect diffusion segments in individual single molecule tracks and classify them as either free, confined, directional or immobile diffusion. The results of the tracking, diffusion and FRET data analysis are stored in a single structure variable and can be evaluated via the accompanying graphical user interface.

The *smCellFRET* package is written in *MATLAB* and provides tools to:

- extract time traces for the donor and acceptor's position, intensity, and FRET efficiency
- filter time traces and retain only a subset of useful donor-acceptor traces
- visualize individual molecule tracking data with their corresponding FRET traces
- generate FRET distribution histograms and
- analyze track lifetimes

In this documentation an example data set is used to guide the user through the data analysis workflow.

2. License

smCellFRET is a software platform to analyze large single molecule FRET data acquired by Live-Cell Imaging Microscopy.

Copyright (c) 2020, Javitch Lab - Columbia University

smCellFRET is a software platform to analyze large single molecule FRET data sets acquired by Live-Cell Imaging Microscopy.

This software will be made freely available by Columbia University for non-commercial research use. The software will be hosted on the Columbia Technology Ventures web site, where a license can be downloaded and submitted. After approval for research use, the software can be downloaded from the site. In addition, a link to a GitHub site will be provided and users of the software who make modifications will be encouraged to do so using the GitHub site so that others can potentially avail themselves of additions and improvements. We will monitor these changes and may choose to incorporate enhancements into new versions that will be made available through the same mechanism. Commercial use is encouraged and licensing and fees will be negotiated with CU Technology Ventures.

3. System Requirements

3.1 Hardware Requirements

The program was developed and tested under the following configuration:

Operating System: Windows Server 2012 Standard

Processor: Intel(R) Xenon(R) CPU E5-2660 v4 @ 2GHz 2.00GHz (2 processors)

Installed Memory (RAM): 256GB

System type: 64-bit OS, x64-based processor

3.2 Software Requirements

MATLAB: Version 2019a (64-bit). The installation requires the following toolboxes: Image Processing, Parallel Computing, Statistics and Machine Learning and Curve Fitting.

Functions / programs that are used by the *smCellFRET* pipeline:

MathWorks File Exchange: distance.m, uigetfile_n_dir, (tiffread: version 3.0), msdAnalyzer

SPARTAN (MATLAB): Version 3.7 (<https://www.scottblanchardlab.com/spartan-download>)

DC-MSS (MATLAB): <https://github.com/kjaqaman/DC-MSS>

Additional Software

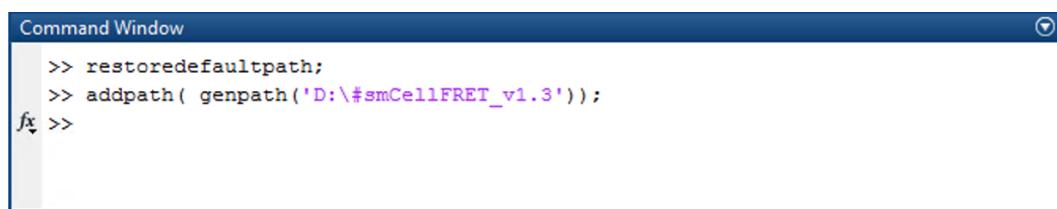
u-track (MATLAB): Version 2.2 (<https://github.com/DanuserLab/u-track>)

track Mate (ImageJ): <https://imagej.net/TrackMate>

4. Software

4.1 Installation

- Download the latest version of the *smCellFRET* software from the *Columbia Technology Ventures* website ([link will be provided in the future](#)) and extract all files from the zip file folder.
- Install *MATLAB* 2019a with all required toolboxes
- Add the extracted folder #*smCellFRET* to the *MATLAB* search path using the commands as shown in **Figure 4.1**. In this example the package #*smCellFRET* is installed in root directory D:\.
- The *smCellFret* Package includes: *smCellFRET*, *Spartan*, *DC-MSS*, *MFE* functions
- **Note:** Adding the # as a leading symbol to *smCellFRET*, as shown in Figure 4.1, is required.
- Typical Install time: Matlab (30min), *smCellFret* Package (5min)



The image shows a screenshot of a MATLAB Command Window titled "Command Window". It contains the following text:
>> restoredefaultpath;
>> addpath(genpath('D:\#smCellFRET_v1.3'));

Figure 4.1 Statements to configure the *MATLAB* search path.

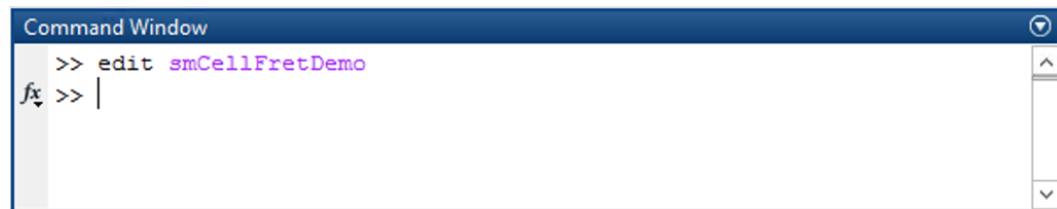
Running both statements will reset the *MATLAB* search path to its default-installed state and add all folders to the search path necessary to run the *smCellFRET* pipeline. Run the *path* command to view all the folders on the *MATLAB* search path. The *downloads* folder contains the complete packages of the *Spartan* Software v3.7, the *DC-MSS* transient motion analysis software and shared functions from the MathWorks File Exchange library.



Caution: *u-track* and *smCellFRET* folders cannot be added to the search path at the same time due to naming conflicts between both software packages.

4.2 Running the *smCellFretDemo* script

Open the demo script in the *MATLAB* Editor by entering the following command in the *MATLAB Command Window* as shown in **Figure 4.2**.



The image shows a screenshot of a MATLAB Command Window titled "Command Window". It contains the following text:
>> edit smCellFretDemo

Figure 4.2 Starting the *smCellFretDemo*-Script.

The demo script was written to guide the user systematically through the *smCellFret* pipeline using the *MATLAB* editor. Each code section in the editor should be executed step by step using the *Run and Advance*

button on the Editor toolbar (see **Figure 4.3**, Box 1). In the first code section, the user is asked to enter the installation directory (see **Figure 4.3**, Box 2). Afterwards, the demo script can be executed further. It is strongly recommended to read the corresponding paragraphs in the documentation where indicated in each code section to familiarize with the basic program routines. Before executing the next code section, also please wait until the *Busy Indicator* (see **Figure 4.4**, red box) in the lower left corner of the *MATLAB Command Window* disappears. The total runtime of the demo script on a Dell R7910 workstation is about 13 minutes.

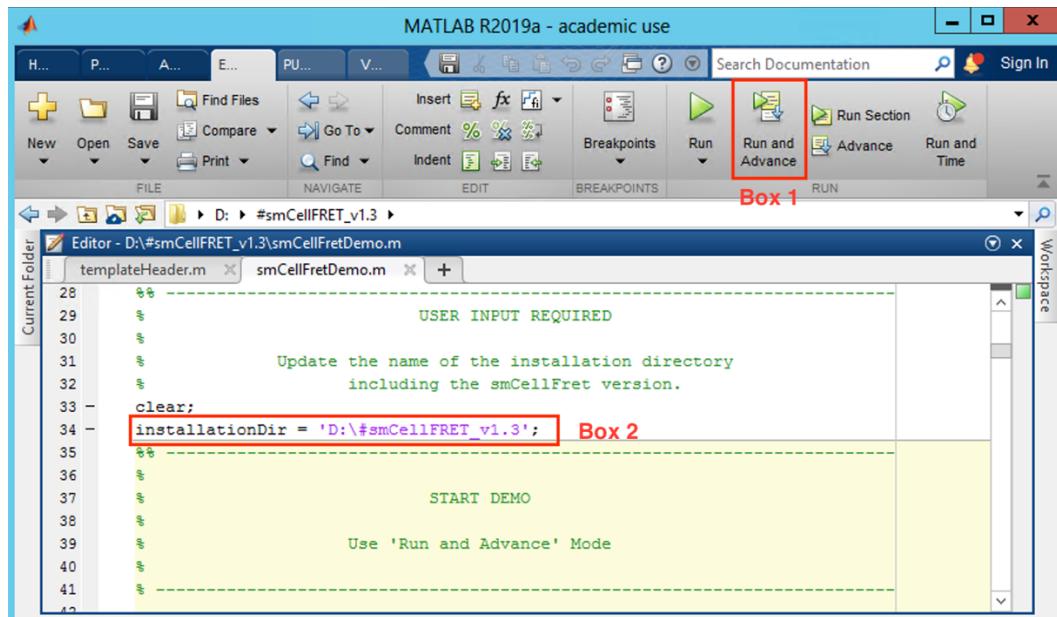


Figure 4.3 Section 1 of the *smCellFretDemo*-Script after opening in the *MATLAB-Editor*. Box 1: The *Run and Advance* button executes the code in the selected section and progresses to the next section. Box2: The user is asked to enter the installation directory here.

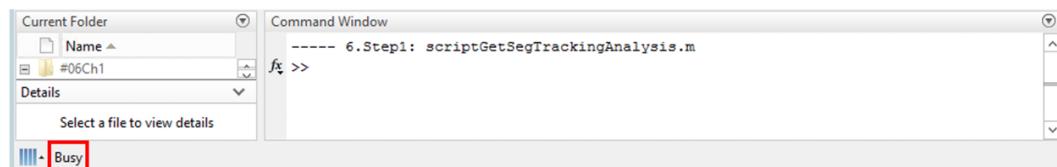


Figure 4.4 Wait with the execution of the next code section in the demo script until the *Busy* indicator (see red box) disappears.



Attention: When running the demo script, the user is prompted to enter file names at several places.

Since *smCellFret* is designed to analyze several data sets at the same time, the input dialog box expects file names until the user presses the *Cancel* button. In the demo only **one dataset** is provided, therefore the user has to press *Cancel*, after entering a filename once, to leave the input dialog box again. The sequence of events for entering a file name is indicated in the title bar of the input dialog box (see **Figure 4.5**, Title). Please always read the instructions in the title bar before selecting a file name, a directory or a pair of directories. In the given example shown in **Figure 4.5**, the user has to highlight the file name *40msG3mW8oPd100G_o6.stk* (Step 1) and press the *Open* button (Step 2). After the selection is complete, click the *Cancel* button to quit the dialog box (Step 3).

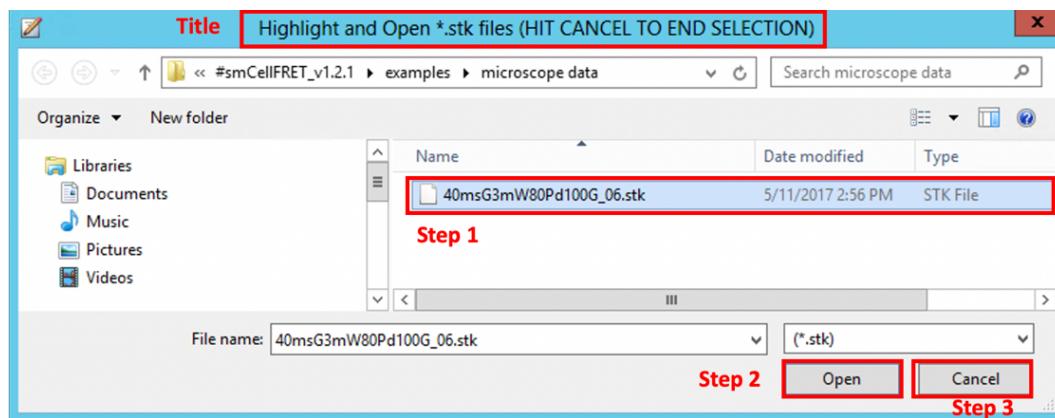


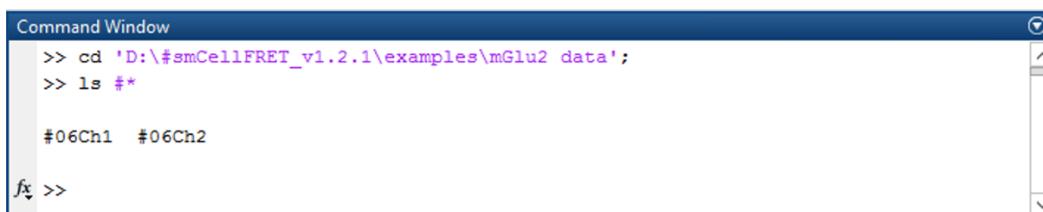
Figure 4.5 Sequence of input events for entering a file name.

5. Documentation

5.1 Project File Folder Structure

The file folder organization of a single molecule project needs to be named in a consistent way so that the software can pool data from different projects together for statistical analysis. Each *smCellFRET* experiment in a project folder consists of two separate folders with the naming format: #%%Ch1 for the acceptor data and #%%Ch2 for the donor data. %% is a 2-digit field reserved for the experiment number.

After the demo script has been executed successfully as described in section 4.2, the organization of the data in a single-molecule project can be displayed. By changing the current folder in *MATLAB* to the project folder **D:\examples\mGlu2 data** (assuming the *smCellFRET* software is installed on the *D*: drive) and running the list command, the two experiment folders 6 are listed as shown in **Figure 5.1**.

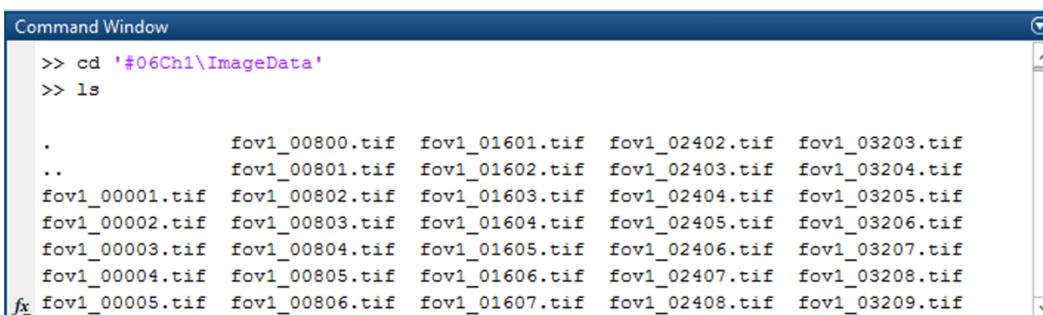


```
Command Window
>> cd 'D:\#smCellFRET_v1.2.1\examples\mGlu2 data';
>> ls #*
#06Ch1  #06Ch2

fx >>
```

Figure 5.1 The data of a particular *smCellFRET* experiment (here experiment number 6) are stored in an acceptor (#06Ch1) and a donor (#06Ch2) file folder.

Each of the folders #06Ch1 and #06Ch2 has a subfolder named *ImageData*, where the donor or acceptor movie images are saved as individual TIF files. Changing from the current project folder **mGlu2 data** to the experiment folder **#06Ch1/ImageData** and running the list command shows a series of TIFF Images acquired in the acceptor emission channel of the microscope (see **Figure 5.2**) The corresponding donor image data of are saved in **#06Ch2/ImageData**.



```
Command Window
>> cd '#06Ch1\ImageData'
>> ls

.
..
fov1_00800.tif  fov1_01601.tif  fov1_02402.tif  fov1_03203.tif
fov1_00801.tif  fov1_01602.tif  fov1_02403.tif  fov1_03204.tif
fov1_00001.tif  fov1_00802.tif  fov1_01603.tif  fov1_02404.tif  fov1_03205.tif
fov1_00002.tif  fov1_00803.tif  fov1_01604.tif  fov1_02405.tif  fov1_03206.tif
fov1_00003.tif  fov1_00804.tif  fov1_01605.tif  fov1_02406.tif  fov1_03207.tif
fov1_00004.tif  fov1_00805.tif  fov1_01606.tif  fov1_02407.tif  fov1_03208.tif
fov1_00005.tif  fov1_00806.tif  fov1_01607.tif  fov1_02408.tif  fov1_03209.tif

fx
```

Figure 5.2 Acceptor image data are saved in the *ImageData* directory in TIF file format.

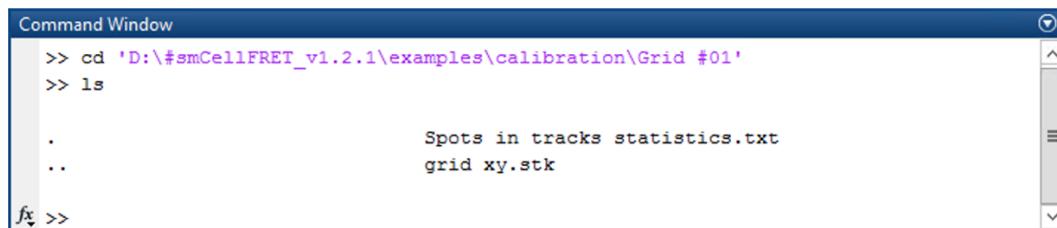
! The format of the image file name is specified by 3 fields: a 5-letter string **'fov1'** followed by a **5-digit image number** and the **'.tif'** file extension. All image names need to follow this convention.

5.2 General Data Analysis Workflow

5.2.1 Image registration between the donor and acceptor emission channel

To monitor FRET efficiencies of diffusing donor - acceptor labeled proteins in cells single molecule fluorescence needs to be co-tracked both in the acceptor and donor emission channel of the microscope. In the *smCellFRET* software co-tracking of the anticorrelated donor – acceptor signal is handled by primarily

tracking the acceptor signal while the position of the attenuated donor signal is calculated by a local weighted mean transformation *tform_lwm* which maps any position in the acceptor channel to the donor channel. The transformation function is specific for the imaging system in use and was measured by tracking a set of control points in both emission channels as described [4]. An example set of control points can be found in the folder **calibration / Grid #01** under examples (see **Figure 5.3**).



```
Command Window
>> cd 'D:\#smCellFRET_v1.2.1\examples\calibration\Grid #01'
>> ls
.
..
Spots in tracks statistics.txt
grid xy.stk
fx >>
```

Figure 5.3 An image series of measured control points is saved in the movie file *grid xy.stk*. The corresponding control point coordinates are saved in the file *Spots in tracks statistics.txt*.

The file '*grid xy.stk*' is a stack of TIFF images which can be played as a movie using the image processing software ImageJ. The movie shows how a pair of control points translate stepwise across the acceptor (left control point) and donor emission channel (right control point) thus forming a grid of 26 rows and 24 columns in each emission channel. The location of each control point was measured using the ImageJ plugin TrackMate [5] and saved in the same folder under the name '*Spots in tracks statistics.txt*'. The geometric transformation function is than calculated by the function *imageRegistrationTrackMate.m* which requires '*Spots in tracks statistics.txt*' as input file. The function *imageRegistrationTrackMate.m* is part of *scriptGetFretTraces.m* used downstream (Step 5.2.4) in the analysis pipeline.

5.2.2 Acquire *smCellFRET* movies and save movie data

Acquire a movie of *smCellFRET* images using a calibrated image splitting system (see Step 5.2.1) which projects the donor and acceptor emission signal side by side onto the camera sensor (see **Figure 5.4**). In order to track single donor and acceptor fluorescence spots using *u-track* the movie file needs to be first converted into a sequence of individual tiff images. Second, the two spectrally separated regions in each tiff image must be split and saved as a series of numbered images in the acceptor (#%%Ch1/ImageData) and donor (#%%Ch2/ImageData) image directories (see section 5.1). The *smCellFRET* movies, saved as Metamorph *.stk files in the example folder **D:\examples\mGluz2 data**, can be converted and split using the script '*scriptStk2Tiff.m*'.

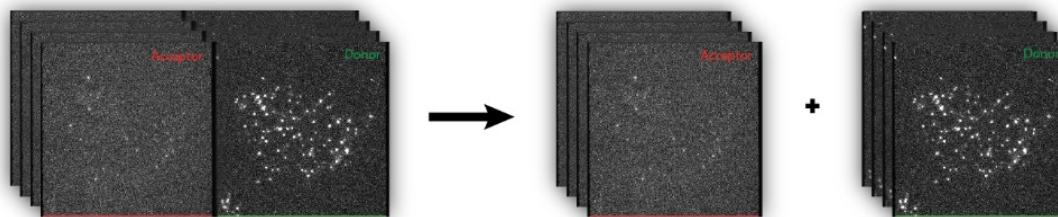


Figure 5.4 A *smCellFret* movie is converted first into a series of single frames. Secondly, the two spectrally different areas of the acceptor and donor are divided into two separate image series and stored in the corresponding acceptor and donor image directories.

5.2.3 Process Tiff Images using *u-track* to extract single molecule tracks

Before running *u-track* [2] it is recommended to tune parameters for particle detection and particle tracking to increase tracking performance. Parameter tuning should be done separately for the donor and acceptor

image data. After parameter tuning run *u-track* on both the acceptor and donor image stacks. Use the region of interest (ROI) option in *scriptDetectGeneral.m* to limit tracking to the cell area only. Spatial coordinates of the ROI data need to be provided as an ascii text file and can be generated by tracing the boundary region of the cell area. Since the boundary is poorly defined in *smCellFRET* images, the ROI coordinates were collected from the projected acceptor image calculated from the raw Metamorph *.stk file using the imageJ [6] plugin ZProject [7] with projection type 'Standard Deviation' or 'Max Intensity' (**Figure 5.5**). The acceptor ROI text file *STD_*.txt* must be saved in both the experiment folder #%%Ch1 for the acceptor data and in the experiment folder #%%Ch2 for the donor data.

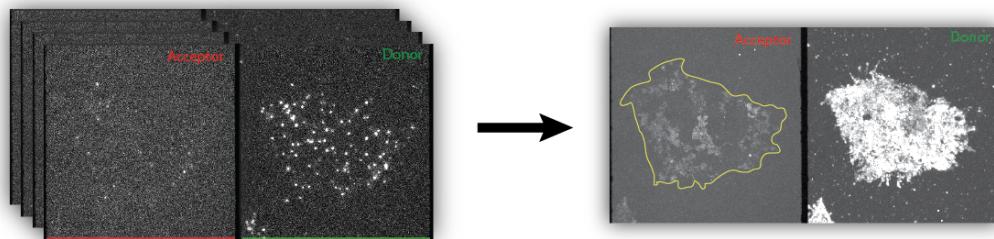


Figure 5.5 A Region of Interest (ROI) can be generated by tracing the cell boundary (yellow line) in the projected cell image on the right. The ROI text file is saved under the file name *STD_40msG3mW80Pd100G_06.txt* in the folder *examples\mGluz data\#06Ch1*.

After tracking finishes the *u-track* produces two output variables: the structure array *detectionMaxProject.mat* with the spot detection results and the structure array *Tracking.mat* which contains the individual particle tracks. The donor and acceptor tracking results must be saved in the experiment folder #%%Ch1 for the acceptor data and in the experiment folder #%%Ch2 for the donor data. At the end of this step the experiment folder for the donor and acceptor data should contain the data files as shown in **Figure 5.6**.

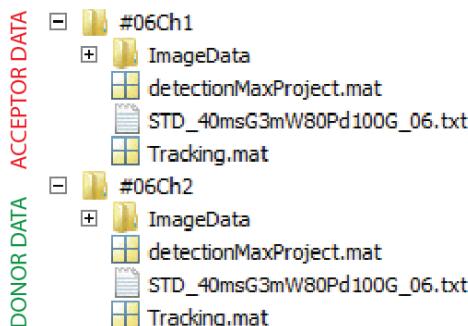


Figure 5.6 Files saved in the two experiment folders #06Ch1 and #06Ch2 after tracking is completed successfully.

5.2.4 Run scriptGetFretTraces to extract fluorescence time traces from single molecule tracks.

The script extracts donor and acceptor fluorescence time traces from the *u-track* output files *detectionMaxProject.mat* and *Tracking.mat*. Running the script prompts the user to select the calibration file ('Spots in tracks statistics.txt', see section 5.2.1) which contains the coordinates of the control points required to infer the transformation function. If the user clicks *Cancel* the program will continue without using a transformation function (no mapping) and prompts the user to select pairs of #%%Ch1 – #%%Ch2 experiment folders. After all desired pairs of folders are selected, click *Cancel* to close the File Explorer and continue running the *MATLAB* script. Not providing donor-acceptor pairs of experiment folders will lead to errors. The script outputs various files (files with extension *.traces, see **Figure 5.7**) most of which are used to

track down possible errors in the pipeline. The final results are saved in a structure array called `fretTraces.mat` in the `#%%Ch2` experiment folder (see **Figure 5.7**, highlighted in red)

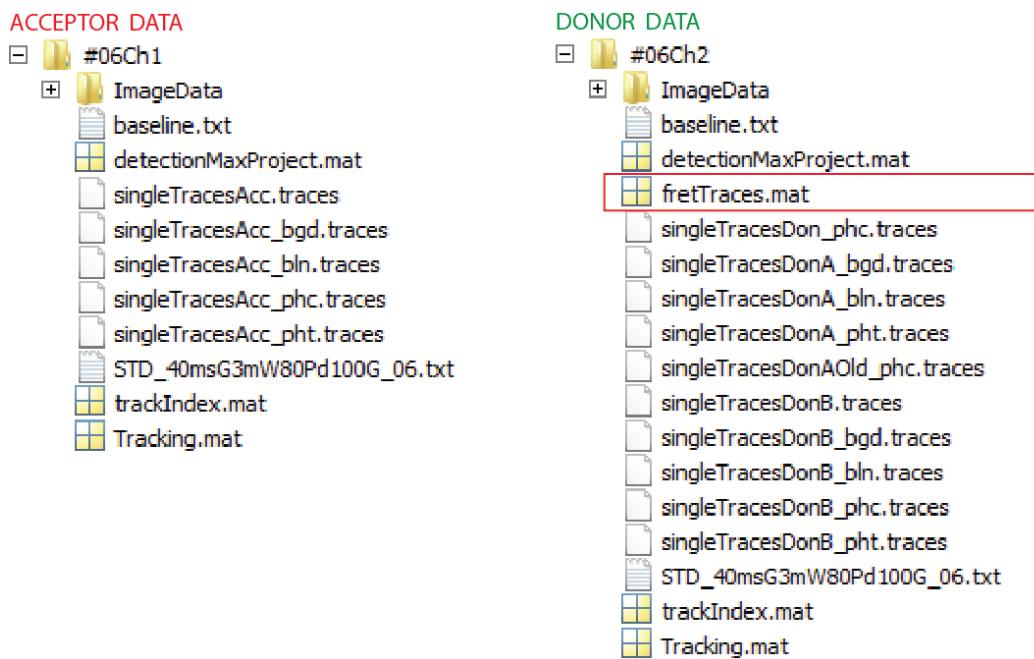


Figure 5.7 Data files in the two folders of experiment #06 after running `scriptGetFrettraces`. The main results are saved in the structure array `fretTraces.mat` (highlighted in red).

5.2.5 Filter multiple data sets to discard unwanted donor and acceptor traces

The program `scriptFretFilter.m` allows the user to select a combination of data filters and apply them to multiple `*_phc.traces` files. (see paragraph 5.2.4). The following filter types are available:

- **Signal to Background Ratio of the Acceptor:** Time averaged ratio of the acceptor intensity of the diffraction limited single molecule spot and its bordering background intensity. Time-averaging is performed over the range of the acceptor track time. All traces t that have a mean signal to background ratio s in the interval $S = \{ s \mid s_{min} < s < s_{max} \}$ are selected.
- **Average Crosstalk Value:** Ratio of the time averaged acceptor intensity of the diffraction limited single molecule spot and its corresponding time averaged donor intensity. The lower limit crt_{min} of the crosstalk value can be calculated as the ratio of the acceptor intensity and the donor intensity after acceptor photobleaching and estimates the spectral bleed-through of the donor signal into the acceptor emission channel [8]. Applying this filter removes primarily traces of tracked particles with intensities in the order of spectral crosstalk. In general all traces t that have a crosstalk value crt in the interval $Crt = \{ crt \mid crt_{min} < crt < crt_{max} \}$ are selected.
- **Acceptor Intensity:** All traces t that have a time averaged acceptor intensity int_{acc} in the interval $Int_{acc} = \{ int_{acc} \mid int_{acc}^{min} < int_{acc} < int_{acc}^{max} \}$ are selected. Time-averaging is performed over the range of the acceptor track time.
- **Donor Intensity:** All traces t that have a time averaged donor intensity int_{don} in the interval $Int_{don} = \{ int_{don} \mid int_{don}^{min} < int_{don} < int_{don}^{max} \}$ are selected. Time-averaging is performed over the range of the donor track time.

- **Total Intensity:** The total intensity is the sum of the donor and acceptor intensity. All traces t that have a time averaged total intensity int in the interval $Int_{tot} = \{ int_{tot} | int_{tot}^{min} < int_{tot} < int_{tot}^{max} \}$ are selected. Time-averaging is performed over the range of the donor track time.
- **Average FRET Value:** All traces t that have a time averaged FRET efficiency $e = int_{acc}/(int_{don} + int_{acc})$ in the interval $E = \{ e | e_{min} < e < e_{max} \}$ are selected. Time-averaging is performed over the range of the acceptor track time.
- **Acceptor Lifetime:** All traces t with a track lifetime lt in time interval $LT = \{ lt | lt_{min} < lt < lt_{max} \}$ are selected.
- **Multiple Event Filter:** If a trajectory θ_i that belongs to a series of trajectories $X = \{\theta_k^X, k = 1 \dots n | n > 25\}$ reappears at the same location (i.e. centroid distance $d(\theta_i|X) < 1.5 \text{ pixels}$) AND the mean track lifetime of $X < t_{limit}$ (unit of t_{limit} is s) then the analysis returns: logical 0 (θ_i is rejected). Otherwise the filter returns logical 1 (θ_i is selected).

Filters are selected by setting the field `[property]FilterOn` of the variable `fltPrm` to logical true (cf. header of the MATLAB script `scriptFretFilter.m`). All selected filters are linked with AND logic, which means that a track t must pass all filter criteria to be accepted.

Running this script will open the file explorer and prompts the user to select pairs of #%%Ch1 – #%%Ch2 experiment folders. After all desired pairs of folders are selected, click *Cancel* to close the File Explorer and continue running the MATLAB script. The script outputs two new workspace files, *TrackingPst.mat* with the filtered tracking data (*Acceptor file directory*, see **Figure 5.8**) and *fretTracesPst.mat* with the filtered FRET data (*Donor file directory*, see **Figure 5.8**)

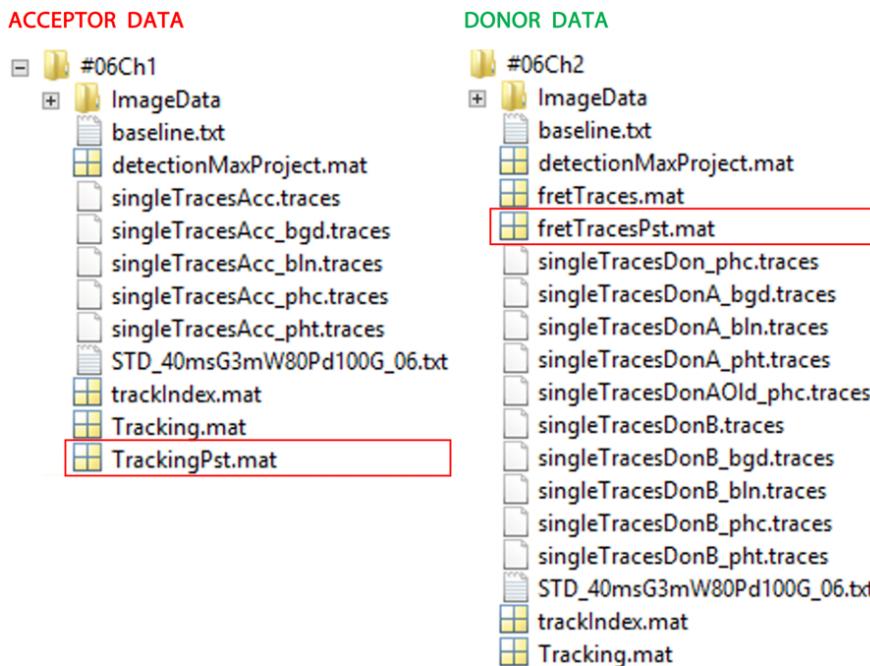


Figure 5.8 Data files in experiment #06 after running `scriptFretFilter.m`. Filtered data are saved in the structure array *fretTracesPst.mat* (marked red in the directory Ch2). The u-track result is updated as *TrackingPst.mat* (marked red in the directory Ch1).

5.2.6 Transient Motion Analysis of Acceptor Tracks

Motion analysis of acceptor tracks is performed to detect diffusion segments in individual tracks and classify them as free, confined, directed and immobile diffusion. The analysis utilizes functions published in the transient mobility analysis framework, termed “divide-and-conquer moment scaling spectrum” (*DC-MSS*) [3] and can be performed in *smCellFRET* by running the following scripts in sequence from the *MATLAB* command window:

Step 1. *scriptGetSegTrackingAnalysis.m*: The script applies the function *basicTransientDiffusionAnalysisV1* (part of *DC-MSS*) to tracking data processed with the *smCellFret* pipeline. The *DC-MSS* function detects potential diffusion segments in a track and performs a Moment Scaling Spectrum (MSS) analysis for these segments. Running the script will open the file explorer and prompt the user to select a list of *Acceptor File Directories* (#%%Ch1 directories). Each selected directory must contain the post filtered data file *TrackingPst.mat* (see **Figure 5.8**). The script will save a data file *results.mat* (Classification results and MSS data for each track segment) and a figure file *figCell*.Figure* (showing the color-coded diffusion segments) in the respective *Acceptor File Directory* (see **Figure 5.9**).

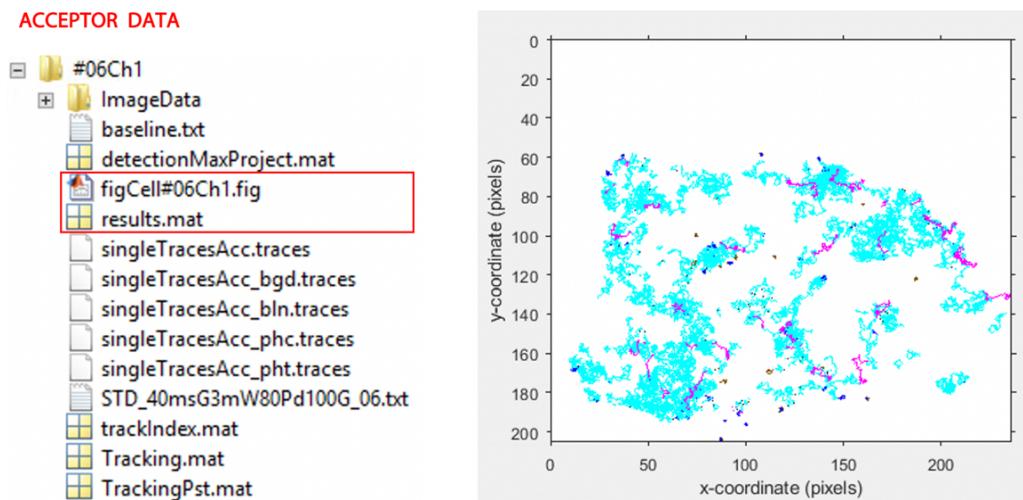


Figure 5.9 Left: Data and image files in experiment #06 generated by the DC-MSS routine *scriptGetSegTrackingAnalysis.m* are marked red. Right: Color coded track segments in a single cell indicate different diffusion modes (brown: immobile, blue: confined diffusion, cyan: normal diffusion, magenta: super diffusion, black: unclassified).

Step 2. *scriptGetSegResults.m*: Running this script will open the file explorer and prompt the user to select a list of *Acceptor File Directories* (#%%Ch1 directories). Each selected directory must contain the data file *results.mat* (see **Figure 5.9**) from the previous analysis (Step 1). The script converts the structure array *results.mat* into a matrix *SegResultsFinal.mat* and performs a statistical analysis on the four segment populations: free, confined, directed and immobile diffusion. The results of the statistical analysis are saved in the data file *SegStats.mat* (see **Figure 5.10**). If multiple experiment folders are selected the script vertically concatenates the matrix *SegStats.mat* from individual experiments and saves the final results in the data file *CombinedStats.mat* (see **Figure 5.10**) located in the parent folder of the last experiment folder. For a detailed description of *SegResultsFinal.mat*, *SegStats.mat* and *CombinedStats.mat* review the function header of *scriptGetSegResults.m*.

Step 3. *scriptDiff2Frettraces.m*: In the last step of the motion analysis track segments which are classified as either free, confined, directed or immobile diffusion will be linked to the *smCellFRET* data. Running the script opens the file explorer and prompts the user to select pairs of *Acceptor and Donor File Directories* (#Ch1 and #Ch2 directories). Selected acceptor directories must contain data files of type *SegResultsFinal.mat*, while donor directories must contain data files of type *fretTracesPst.mat*. Updated *smCellFRET* data containing motion dynamics which are linked to FRET Traces are saved as data files of type *fretTracesDif.mat* (see **Figure 5.10**) and can be displayed in *cellFretViewtraces* (see **Figure 5.11**).

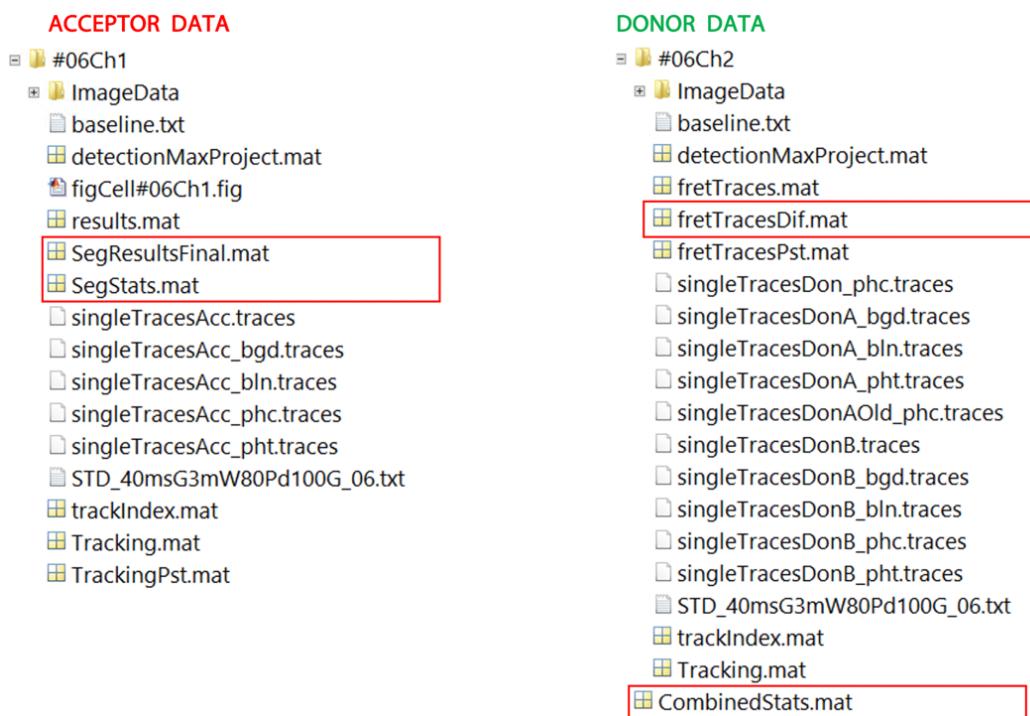


Figure 5.10 When *scriptGetSegResults.m* is executed, it performs a statistical analysis of the diffusion data and stores the results in the Ch1 folder of experiment #06 (files are marked red). If the script is executed with multiple experiments, the additional variable *combinedStats.mat* is stored in the Ch2 folder. The *scriptDiff2Frettraces.m* links the diffusion data with the *smCellFRET* data and saves the final result as *fretTracesDif.mat* in the Ch2 folder.

5.2.7 Run the GUI *cellFretViewtraces* to visualize individual molecule tracks and intensity time traces

The program *cellFretViewtraces* (see **Figure 5.11**) displays the motion of individual donor-acceptor-labeled molecules, their tracks with their corresponding intensity and FRET time traces and allows the user to manually correct and sort traces into the three different categories.

Loading Traces Files: Select the data format from the drop-down menu in the *Traces Panel*. For unsorted traces select the format *fretTraces* else if traces were sorted previously select either *fretTraces_bestFret*, *fretTraces_allFret* or *fretTraces_noFret*. Click on the *SelectTracesFile* button to navigate with the file explorer to an experiment folder #Ch2, select a *fretTraces*.mat* (e.g. *fretTraces.mat*, *fretTracesPst.mat*, *fretTracesDif.mat*) file and open it. The first frame of the recorded acceptor (Ch1) and donor (Ch2) image stack is displayed in the image panel. Move the sliders in the *Image Processing Panel* to update and adjust the image

contrast before clicking the *LoadTraces* button. After loading the traces into the *MATLAB* workspace, the current *Molecule Number* and the traces ID for the donor and acceptor is shown in the traces panel.

Visualizing Tracks: Recorded single molecule image stacks are displayed in the image panel. The acceptor image is shown on the left (ch1) and the donor image is shown on the right (ch2). The currently selected FRET pair is circled in red (acceptor) and green (donor). Moving the *Frame Slider* in the *Tracking Panel* allows the user to move as a function of time through the image stack and observe particle localization and particle tracking (enable the radio button *Show Tracking*) for the current molecule. As an alternative to the *Frame Slider* the user can also scroll through the image sequence with the mouse wheel (*zoom in* or *zoom out* icons in the *Menu Toolbar* need to be disabled). Enabling the *Zoom to Molecule* radio button sets the zoom to a predefined magnification of 167% (can be edited manually) and zooms in on the current molecule. Enabling the *Hold* radio button preserves the current magnification when switching to the next FRET Pair. Another option to zoom in on the image is to enable *Zoom to Region* and set the image boundaries in the fields provided. The full trajectory of each FRET pair is plotted from track start to track end as a line plot in the lower left panel of the viewer.

Plotting Fluorescence and FRET Time Traces: The middle panel of the viewer shows three plots: The donor (green) and acceptor (red) fluorescence intensity time trace, the total intensity time trace (black) and the FRET time trace (blue) of the selected molecule. Elongated red (acceptor) and green (donor) bars below the fluorescence intensity time trace show the time window in which the molecule was tracked.

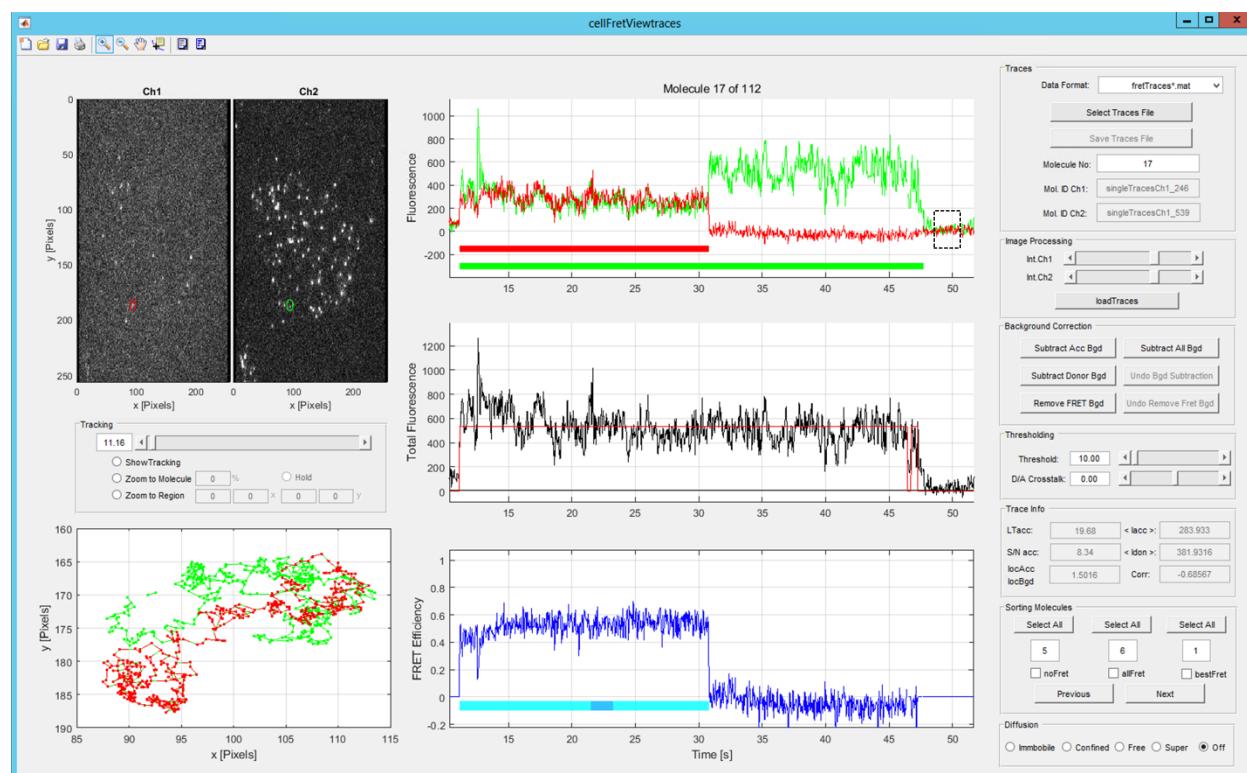


Figure 5.11 Single molecule tracks, intensity-time traces and the diffusion mode can be displayed with the graphical user interface *cellFretViewtraces*.

Background Correction Panel: The baseline of the donor and acceptor fluorescence intensity time trace can be manually set to zero if auto-background correction performed poorly. With zoom enabled (enable the *Zoom*

In icon in the upper left Menu Toolbar of the GUI) the axes range used for background correction is defined by a rectangle (see **Figure 5.11**) that is drawn by clicking and dragging the mouse cursor over the photobleached region of interest. After defining the axes range click either *SubtractAccBgd*, *SubtractDonorBgd* or *SubtractAllBgd* to bring the baseline fluorescence of the acceptor, or donor or both acceptor and donor down to zero. The action can be reverted by clicking *UndoBdgSubtraction*. Correcting the donor and acceptor fluorescence will also bring the baseline of the FRET time trace down to zero. Remaining noise in the FRET baseline can be further removed by zooming into the area of interest, e.g. the FRET baseline, and clicking on *RemoveFRETBgd*. This action can also be reverted by clicking *UndoRemoveFRETBgd*.

Thresholding Panel: The horizontal line shown in the plot of the total intensity time trace defines a cutoff threshold for low amplitude signals which are excluded from the calculation of FRET. For signal amplitudes below the threshold, FRET values are set to zero. The cutoff threshold can be changed manually by moving the threshold slider. The crosstalk value corrects the acceptor signal for leakage of the donor emission into the acceptor channel. If the D/A crosstalk slider is set to the correct value, the acceptor signal is zero after acceptor photobleaching.

Traces Info Panel: This panel shows a few basic statistical measures of the trace quality for a given molecule.

- LT_{acc} : Acceptor fluorescence lifetime until photobleaching or termination of the tracker.
- $\langle I_{acc} \rangle$: Mean acceptor Fluorescence Intensity
- $\langle I_{don} \rangle$: Mean donor Fluorescence Intensity
- S/N_{acc} : The acceptor signal-to-noise ratio is defined as the ratio of the mean acceptor signal before photobleaching and the standard deviation of the background signal after photobleaching.
- loc_{acc}/loc_{bkgd} : Mean fluorescence intensity ratio of the diffraction limited single molecule spot and its bordering background intensity.
- *Corr*: The correlation between the donor and the acceptor is measured by the Pearson correlation coefficient over the range of the donor signal.

Sorting Molecules: Traces can be manually distributed into the three different bins by use of the checkboxes. Bin Names are arbitrary and can be used to create subpopulations of FRET traces. In general, the three categories are used as follows:

- **noFret**: Fret traces that show significant artifacts and are not useful for further analysis.
- **allFret**: Fret traces which show anticorrelation at any point of time.
- **bestFret**: Fret traces which show anticorrelation of the donor and acceptor signal at any point of time and at the end of the trace.

The number of traces in each bin is shown above each checkbox. Clicking the *NEXT* or *PREVIOUS* pushbuttons will display the next molecule. After sorting the data, save the results by pressing the 'Save Traces' button located in the *Traces Panel*. The selection will be saved as *fretTraces*_bestFret.mat*, *fretTraces*_allFret.mat* and *fretTraces*_noFret.mat* in the same directory where the unsorted data are saved (see **Figure 5.12**). In the corresponding text files, which have the same file naming, the selected 'trace' IDs are stored as integers.

Diffusion Panel: The Diffusion Panel is only activated if a motion analysis (see section 5.2.6) has been performed and data files of type *fretTracesDif.mat* have been loaded into the program memory. A color-coded bar (dark blue: immobile, blue: confined, cyan: free and magenta: super diffusion) superimposed on the time axis of the FRET efficiency diagram visualizes the detected diffusion segments in the current track and links them to the *smCellFRET* time trace. By default, the radio button is set to *off* so that all diffusion states in a track are visible. Switching the radio button to a different diffusion mode only shows the segments of the

selected mode in the FRET trace. Using this option, *smCellFRET* trace subpopulations of a specific diffusion mode can be generated.

DONOR DATA

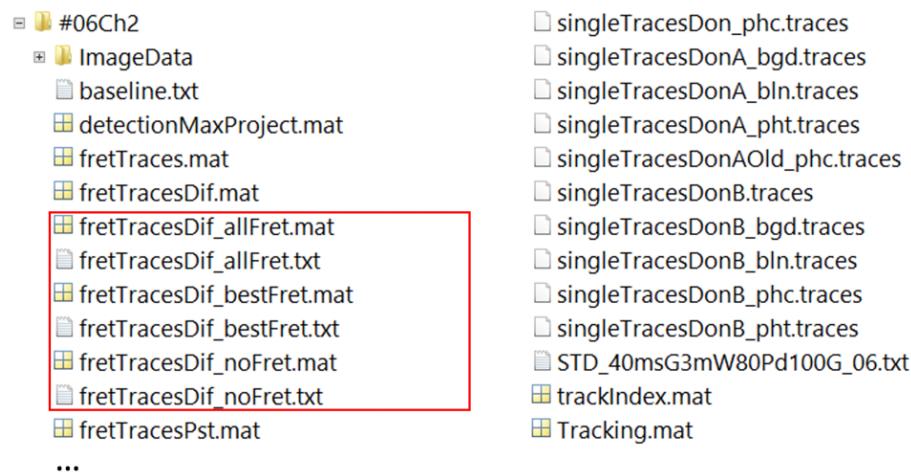


Figure 5.12 Saved data files in experiment #06 after manually sorting single molecule data into the three different bins *noFret*, *allFret* and *bestFret* using *cellFretViewtraces*.

5.2.8 Plot traces as contour plots and histograms to visualize their ensemble behavior.
Populations of FRET traces can be converted into FRET histograms which are useful to detect conformational states. In *smCellFret*, FRET histograms can be generated from subpopulations of traces which were spectrally corrected for donor bleed through and sorted in *cellFretViewtraces* (e.g. *fretTraces*_bestFret.mat* or *fretTraces*_allFret.mat* data files, see section 5.2.7). Histograms can be created from FRET traces that contain either all diffusion modes or from trace segments that contain only one specific diffusion mode (e.g. free diffusion).

Step 1. Compiling FRET traces into segmented traces with a single diffusion mode (optional): This step is only required for generating FRET histograms from trace segments belonging to a specific diffusion mode. Running the script *FretTraces2Seg.m* will open the file explorer and prompt the user to select a list of *Donor File Directories* (#Ch2 directories). Each selected directory must contain a data file of type '*fretTracesDif_bestFret.mat*' or '*fretTracesDif_allFret.mat*' which have the results of the motion analysis. The script then parses FRET traces with multiple diffusion modes per trace into a segmented trace with only one diffusion mode per trace and saves the results under the input file name using the following extensions: '**_SegImb.mat*' (segments with immobile diffusion modes), '**_SegCnf.mat*' (segments with confined diffusion modes), '**_SegFre.mat*' (segments with free diffusion modes) and '**_SegSpr.mat*' (segments with super diffusion modes).

Step 2. Add Fret traces from multiple experiment folders (optional): The script *AddFrettraces.m* combines data files of type '*fretTraces*.mat*' from several experiment folders into a single data file to perform an analysis of all files together. The new file is saved under the same name as the input file with the extension '_add.mat' in the last selected experiment folder. Analyzing all experiments together will not allow an error analysis based on the cell-to-cell variability.

Step 3. Post-synchronize Fret traces: In live cell single-molecule movies, particle tracks and FRET traces have different start and end times. To generate FRET distribution histograms, the individual fret traces must be synchronized with respect to their starting point. Running *scriptPostSyncTraces.m* will prompt the user to

select an input file e.g 'fretTracesDif_SegFre.mat' and synchronizes all its FRET traces to the start point of the movie. By default, the results are saved as a *Traces* object (for post-processing with the SPARTAN [1] software) under the same file name as the input file with the extension e.g. *fretTraces_SegFre_sync.traces*.

Note: By running the SPARTAN's *loadTraces.m* function from the *MATLAB* command window, a *.traces file can be loaded into the *MATLAB* workspace to display its variables.

Step 4. Filter traces by total intensity (optional): Running *scriptTotalFilter_traces.m* offers the possibility to include only FRET traces with total intensities in a predefined interval in the analysis. This step is particularly useful for cleaning up trace populations that have not yet been processed by other filters. The File Explorer prompts the user to select multiple **sync.traces* files to which the filter should be applied. Filtered traces are saved with the additional extension **sync_flt.traces* in the same directory as the input file.

Step 5. Generate FRET distribution histograms. Apparent FRET histograms can be created with *scriptContourPlot.m*. Running the script will open the file explorer and prompt the user to select a **sync.traces* File (located in #Ch2 directories). The script uses the SPARTAN [1] function *makeplot.m* and generates two ASCII data files which will be saved under the same file name as the input file with the additional extensions **_2dHst.txt* and **_1dHst.txt*. ASCII text files can be imported into a data analysis software (e.g. Origin) to generate 2D-FRET contour plots and 1D-FRET histogram distributions.

5.2.9 Correction of FRET Histograms

To convert apparent FRET values into real FRET efficiencies, a common workflow [9] was used, applying the following corrections to the data in this sequence: spectral crosstalk correction, direct excitation correction and a detection correction.

Step 1: Spectral crosstalk correction. The function *alphaCorrectCell.m* multiplies the donor intensities by a correction factor *alpha* and subtracts these adjusted intensities from the acceptor intensity. The *alpha* factor for spectral crosstalk takes into account the amount of donor fluorescence leakage into the acceptor emission channel. When executing the function, the user is prompted to select a list of trace files (e.g. **sync_flt.traces*) to be corrected. A second prompt asks the user to either confirm the default *alpha* factor or enter a new value before the corrections are applied. Corrected traces are saved in the same directory as the input file with the additional extension *a* for *alpha* (e.g. **sync_flt_a.traces*).

Step 2: Direct excitation correction. The function *deltaCorrectCell.m* multiplies the mean total trace intensities by a correction factor *delta* and subtracts these adjusted intensities from the acceptor intensities. This correction is based on the assumption that the direct acceptor excitation intensity can be approximated by the total intensity. Here the *delta* factor is a measure of the percentage of direct acceptor excitation with the donor laser. When executing the function, the user will be prompted to select a list of trace files to be corrected. Selected files need to be already crosstalk corrected (extension **_a.traces*). A second prompt asks the user to either confirm the default *delta* correction factor or enter a new value. Corrected traces are saved in the same directory as the input file with the additional extension *d* for *delta* (e.g. **sync_flt_a_d.traces*).

Step 3: Detection correction. The function *gammaCorrectCell.m* scales the donor intensities by a correction factor *gamma*. Gamma correction is necessary to adjust for differences in the donor - acceptor fluorescence quantum yields and fluorescence detection efficiencies. When executing the function, the user is prompted to select a list of trace files that are already corrected for spectral crosstalk and direct excitation (files with the extension **_a_d.traces*). A second prompt asks the user to either confirm the default *gamma* factor or enter a new value before the corrections are applied. Corrected traces are saved in the same directory as the input file with the additional extension *g* for *gamma* (e.g. **sync_flt_a_d_g.traces*).

Step 4: *Generate FRET distribution histograms.* Corrected FRET histograms can be created with scriptContourPlot.m (see previous section).

5.2.10 Lifetime Analysis

Single molecule track lifetime analysis can be performed on trajectory subpopulations which were selected using the FDT (freely diffusing trajectory: minimum track length ≥ 20 frames) or the NLT (non-limited lifetime trajectory) filters.

FDT Analysis:

Step 1: *Get lifetimes of trajectories with only one diffusion mode.* Running scriptDiffLT.m will open the file explorer and prompt the user to select a list of either #Ch1 (acceptor lifetime analysis) or a list of #Ch2 (donor lifetime analysis) directories. Each selected directory must contain the data file SegResultsFinal.mat (see 5.2.6 Transient Motion Analysis, Step 2). The script computes the number of diffusion modes per trajectories and extracts trajectories that have only one mode of diffusion during their entire lifetime. Each trajectory is then further classified as either entirely free diffusion, entirely confined diffusion, entirely directed diffusion or entirely immobile diffusion. Classification results will be saved in each directory under the file name diffFilter.mat. Single mode diffusion lifetimes (in units of frames) from multiple trajectories and experiments will be concatenated and exported as variables LtDataImob (lifetime data of immobile diffusion), LtDataConf (lifetime data of confined diffusion), LtDataFree (lifetime data of free diffusion), LtDataSupr (lifetime data of directed diffusion) to the MATLAB workspace. If no additional filtering is required, the variables can be used to generate histogram distributions and extract average lifetimes from exponential distributions using curve fitting routines (see also function header of scriptDiffLt.m).

Step 2: *Link single mode diffusion events with smCellFRET data files.* This step is only necessary if acceptor lifetimes need to be further filtered using fluorescence intensity-based filters. Running scriptDiffFilter.m opens the file explorer and prompts the user to select pairs of #Ch1 and #Ch2 directories. Selected #Ch1 directories must contain data files of type diffFilter.mat (see Step 1), while #Ch2 directories must contain data files of type fretTracesPst.mat (fretTraces data generated using the FDT filter condition). The script uses the results of the motion classification to extract from fretTracesPst.mat those traces associated with entirely immobile, entirely restricted, entirely free and entirely directed diffusion. The results will be saved under the filename fretTraces using the following extensions: *Imob.mat (FRET traces data associated with entirely immobile diffusion), *Conf.mat (FRET traces data associated with entirely confined diffusion), *Free.mat (FRET traces data associated with entirely free diffusion) and *Supr.mat (FRET traces data associated with entirely directed diffusion).

Step 3. *Add FretTraces*.mat files from multiple experiment folders:* Use scriptAddFrettraces.m to combine fretTraces*.mat files (with extensions *Imob.mat, *Conf.mat, *Free.mat or *Supr.mat) from multiple experiment folders into a single data file to perform an analysis of all files together. The new file is saved under the same name as the input file with the extension _add.mat in the last selected experiment folder.

Step 4. *Apply a total intensity filter to the fretTraces*_add.mat file:* The total intensity filter offers the possibility of including only FRET traces in the lifetime analysis with total intensities in a predefined interval. Running scriptTotalFilter.m opens the file explorer and prompts the user to select a fretTraces*.mat file. For the lifetime analysis of e.g. entirely freely diffusing molecules choose as input file fretTracesFree_add.mat. Applying the filter removes traces with average total fluorescence intensities outside the empirically determined interval [240; 785]. Filtered traces will be saved in the same directory and under the same name as the input file but

with extension `*_flt.mat`. A list of total intensity filtered lifetimes of e.g. entirely free diffusing molecules can now be obtained by loading the file `fretTracesFree_add_flt.mat` into the *MATLAB* workspace. Acceptor lifetimes (in units of frames) are stored in the structure variable `fretTraces.Ch1.traceMetadata.traceLen` and can be used to generate a histogram distribution. Average lifetimes can be determined from the exponential distribution using curve fitting routines.

NLT Analysis:

Step 1. Add *FretTracesPst.mat* files from multiple experiment folders: Use `scriptAddFrettraces.m` to combine `fretTracesPst.mat` files (fretTraces data generated using the NLT filter condition) from multiple experiment folders into a single data file to perform an analysis of all files together. The new file is saved under the same name as the input file with the extension `_add.mat` in the last selected experiment folder.

Step 2. Access lifetimes from the *fretTracesPst_add.mat* file: A list of measured fluorescence lifetimes can be obtained by loading the file `fretTracesPst_add.mat` into the *MATLAB* workspace. The acceptor lifetimes (in units of frames) are stored in the structure variable `fretTraces.Ch1.traceMetadata.traceLen` and can be used to generate a histogram distribution. Average lifetimes can be determined from the exponential distribution using curve fitting routines.

5.3 A note about data storage

If the user only wishes to use the `cellFretViewtraces.m` GUIs to sort and analyze data (without SPARTAN support), they can reduce their data storage usage by up to 80% by deleting all `*.traces` files, which is easily performed using the provided script `scriptDeleteTraces.m`. If the files need to be reproduced at any point, the user can run `scriptGetFretTraces.m` (as instructed in step 5.2.4, above).

6. References

1. Juette, M.F., et al., *Single-molecule imaging of non-equilibrium molecular ensembles on the millisecond timescale*. Nat Methods, 2016. **13**(4): p. 341-4.
2. Jaqaman, K., et al., *Robust single-particle tracking in live-cell time-lapse sequences*. Nat Methods, 2008. **5**(8): p. 695-702.
3. Vega, A.R., et al., *Multistep Track Segmentation and Motion Classification for Transient Mobility Analysis*. Biophys J, 2018. **114**(5): p. 1018-1025.
4. Churchman, L.S., et al., *Single molecule high-resolution colocalization of Cy3 and Cy5 attached to macromolecules measures intramolecular distances through time*. Proc Natl Acad Sci U S A, 2005. **102**(5): p. 1419-23.
5. Tinevez, J.Y., et al., *TrackMate: An open and extensible platform for single-particle tracking*. Methods, 2017. **115**: p. 80-90.
6. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis*. Nat Methods, 2012. **9**(7): p. 671-5.
7. Kelly, P. and H. Karten, *Reslice and ZProject plugin (Image -> Stacks -> Reslice and ZProject... commands)*. University of California, San Diego.
8. Hohng, S., C. Joo, and T. Ha, *Single-molecule three-color FRET*. Biophys J, 2004. **87**(2): p. 1328-37.
9. Hellenkamp, B., et al., *Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study*. Nat Methods, 2018. **15**(9): p. 669-676.