The University of South Bohemia in České Budějovice

Faculty of Science

Optimization of Single-molecule FRET Experiment to observe the Conformational Dynamics in Sec Translocon During Protein Transport.

Bachelor thesis

Lorin Lehner

Advisor: Mgr. Tomáš Fessl, Ph.D.

České Budějovice 2023

Lehner, L., 2023: Optimization of Single-molecule FRET Experiment to observe the Conformational Dynamics in Sec Translocon During Protein Transport. Bc. Thesis, in English, - 28 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic

Annotation

In this thesis, single molecule experiments were optimized for the detection of conformational changes of the lateral gate in translocon. Fluorescence Correlation Spectroscopy was performed to find the best suited acceptor dye in single molecule Förster Resonance Energy Transfer (smFRET). The smFRET data was then analyzed by the open-source python notebook FRETbursts and the different fluorescence states were classified by two different methods utilizing Hidden Markov Modeling.

Declaration

I declare that I am the author of this qualification thesis and that in writing it I have used the sources and literature displayed in the list of used sources only.

České Budějovice, 14.12.2023

Student's signature

Table of Contents

1. Introduction

Fluorescence is a powerful tool to observe biological processes on a molecular level. Employing techniques like single molecule Förster Resonance Energy Transfer (smFRET) or Fluorescence Correlation Spectroscopy (FCS), it is possible to observe conformational changes of individual macromolecules, such as proteins, via the labeling of the respective sample with fluorescent dyes [1], [2]. Here we focus on optimization the experimental conditions to extract rates of conformational dynamics in SecYEG translocon, conserved nano-machine responsible for transport of proteins accross inner membrane of bacteria. These conformational dynamics can be quantified via photon-by-photon hidden Markov modeling (H2MM) analysis, even if they appear on a timescale of micro- to milliseconds [3]. Recently, two different approaches for multi-dimensional H2MM methods were introduced, as an open-source toolkit that works in combination with FRETbursts, an already established open-source toolkit [3], [4], [5], [6]. The first method classifies FRET states based on FRET efficiency and FRET stoichiometry, the second, newer method employes fluorescence lifetime as a third dimension for the analysis. While the reliability of all these tools is already established, to our knowledge the two H2MM approaches have not been directly compared. In this thesis, the transitions between the conformational and photophysical states in translocon dynamics will be attempted to be classified and quantified, and the resulting kinetics will be directly compared. Furthermore, selection of the ideal acceptor dye (with red absorption) will be guided by an FCS experiment in an attempt to classify fluorescent dyes based on their tendency to blink or photobleach.

When performing a FRET experiment, the right choice of the fluorescent dyes serving as FRET pair is a success-determining step. When selecting a suitable acceptor dye for the FRET pair, here we focus on the red part of the spectrum. This is motivated by high autofluorescence of polar lipid extract from *Escherichia coli* in shorter wavelength. Besides the consideration for excitation and emission ranges, crucial limiting factors hereby are the sensitivity to environmental factors and photostability, since photophysical effects, such as protein induced quenching, photoblinking and -bleaching, can temporarily or permanently inhibit the molecules fluorescent properties. For example, photoblinking is most commonly described as a result of a photochemical reaction, that can temporarily alter the fluorescent state of the molecule, caused by frequent transitions to the triplet state [7]. To test the photophysical properties of candidates for FRET acceptor, specifically their propensity to undergo triplet transition, Fluorescence Correlation Spectroscopy (FCS), Förster Resonance Energy Transfer

(FRET) and hidden Markov Modeling (HMM) methods were applied. For better control over the FRET pair stoichiometry a pulsed interleaved excitation laser (PIE) setup was used. Following is a theoretical introduction of each of these methods.

1.1 Photophysics: Fluorescence, Photoblinking and Photobleaching

A fluorescent dye will undergo a transition from the ground to the excited state upon absorption of a photon of suitable wavelength. The depopulation of the excited state in our system is preferentially happening either via fluorescence-emission of a photon with energy corresponding to the energy difference between the excited and ground state or FRET. FRET can be used as a tool to detect conformational dynamics of biomacromolecules covering the time-range between nanoseconds and minutes (details below in section 1.3). Extraction of rates of interconversion between different conformational states relies on detection of dwell times in each conformational state.

Photophysical effects such as photoblinking or photobleaching may temporarily or permanently inhibit fluorescence [7] and bias/shorten the dwell time of the conformational state. There must be a differentiation between these two effects. Photoblinking is a reversible process caused by fluctuations between fluorescent states and dark (non-fluorescent) states, which occur on a scale of microseconds to minutes. Due to its disruptive nature, as mentioned above, photoblinking can obscure the interpretation of dwell times. On the other hand, photobleaching is an irreversible effect and it introduces another kind of bias by reducing our capability to detect long-lasting dwell times. Bleaching is caused by two main mechanisms. The first is an oxygen-dependent pathway, in which molecular oxygen reacts with the reactive excited fluorophore [7]. This effect can be minimized by the addition of oxygen scavenging species [7]. The second mechanism is via the transition from a singlet state into the triplet state in which the molecule possesses strongly increased reactivity. As a consequence of the triplet state, permanent modifications within the molecule may occur [7]. This results in the permanent loss of fluorescence. Compared to other photochemical events, photobleaching is a less likely event. The same molecule needs to undergo multiple excitation processes for photobleaching to be of relevant concern. Photostability is a measurement of how many excitation periods any fluorophore can complete before being destroyed. Short survival times only count a few hundred cycles, long survival times can include several million excitation cycles. High resolution experiments require dyes with long survival times as photodegraded material results in a reduced signal-to-noise ratio [7]. Thus, photodegrading effects are

generally tried to be minimized. Diagram 1 below shows the possible energy pathways within a fluorescent molecule.

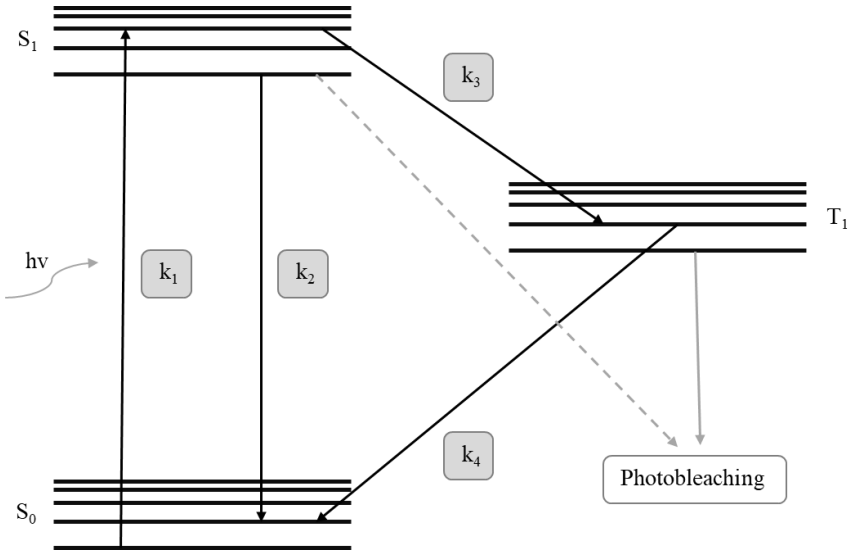

Diagram 1: Schematic of the energy pathways within a fluorescent molecule for the first excited states. Higher excited states work in a similar fashion but are linked with even stronger reactive species. Energy is supplied via photons of energy hv. S_0 describes the ground state and S_1 the first excited state. k_1 represents the excitation into the first excited state, and k_2 is the inverse, which results in the release of a photon and therefore fluorescence. k_3 represents the pathway into the first triplet state T_1. From here, it is still possible to return to the ground state (via k_4), however, photobleaching is a possible occurrence. Lastly, the dashed arrow represents other photobleaching reactions such as the reaction with molecular oxygen.

1.2 Fluorescence Correlation Spectroscopy

Fluorescence Correlation Spectroscopy (FCS) is an analytical method that allows the detection of freely diffusing molecules in solution. Only molecules entering the confocal volume of the microscope can be observed. FCS can provide insight to the diffusion dynamics (diffusion coefficient) of a molecule and/or its photophysical properties, such as triplet lifetime and amplitude. FCS provides excellent temporal resolution as it calculates an autocorrelation function, that measures the correlation of a signal with itself in future events. This way it can detect any changes in fluorescence signal over a wide range of time, spanning from picoseconds to minutes. The autocorrelation function G(τ) is described in Equation 1.

$$G(\tau) = \frac{\langle F(t) \cdot F(t+\tau) \rangle}{\langle F(t) \rangle^2} \qquad \textit{Equation 1}$$

Where F(t) represent intensity fluctuations occurring by the e.g. fluorophore entering or leaving the confocal volume (translational diffusion), triplet transitions, FRET, etc. τ is the lagtime and the brackets < > represent the temporal average [8].

1.3 Förster Resonance Energy Transfer (FRET)

FRET is a highly sensitive method capable of the observation of changes in single molecules, be they conformational or dynamic, as well as molecular interactions. FRET is therefore well equipped for the analysis of proteins and related biological processes such as protein folding or the opening and closing of cell-membrane channels or molecular gates [1], [2]. FRET enables a precise measurement of relative distances on the scale of nanometers. Since macromolecules in solution are prone to undergo conformational changes, single molecule FRET (smFRET) specifically, also provides insight of these dynamic changes on a temporal scale.

An excited fluorescent dye, acting as a Donor, can either fall back down to the ground state leading to the emission of a photon and therefore, fluorescence. If, however, a FRET partner, an acceptor dye, is present in close vicinity, the donor will transfer some of its energy to the acceptor, exciting it in turn. This energy transfer is radiationless. The acceptor molecule then returns to its ground state, emitting a photon itself. The efficiency of energy transfer depends on the relative distance between the two fluorophores, making the emission intensity of the donor and the relative distance indirectly proportional. The distance relation between donor and acceptor can be described by Equation 2 [2].

$$k_t = \left(\frac{1}{\tau_D}\right)\left(\frac{R_0}{r_{DA}}\right)^6 \qquad \textit{Equation 2}$$

According to this equation, the transfer rate constant kt is proportional to the sixth power of the distance between the donor and acceptor molecule (r_{DA}). The term R_0 represents Forster distane. R_0 is a characteristic distance of each FRET pair, at which the efficency of energy transfer is exactly 50%. It depends on the spectral overlap between the emission spectrum of donor and absorption spectrum of acceptor, permitivity of the environment, quantum yield of donor and molar absorptivity of the acceptor. The term $1/\tau_D$ describes the rate of deactivation from the excited state via all available pathways. Hence, measuring the lifetime of the donor allows an exact measurement of the distance. Furthermore, measuring the intensity of both dye's emission will yield the same result (see the discribtion of PIE below).

1.4 Pulsed Interleaved Excitation

Pulsed Interleaved Excitation (PIE) uses two lasers that excite both fluorophores in an alternating fashion and is often used in FRET experiments. The alternation occurs on a timescale of nanoseconds. A short downtime between each pulse ensures that crosstalk can be

avoided. Since both, the donor and acceptor dye are directly excited by their respective laser, and all emitted photons are detected by separate channels, we can distinguish between active FRET pairs and inactive donor and acceptor molecules. Further, due to the fast pulsing, excellent temporal resolution can be achieved. PIE directly enables the calculation of FRET efficiency and photon stoichiometry (see Equation 3and 4) [9]. If no acceptor is present, the ratio goes to S = 1 ("100 % donor"), while the absence of a donor would lead to S = 0 ("0 % donor").

$$S = \frac{n_{A-FRET} + n_D}{n_{A-FRET} + n_D + n_{A-direct}} \qquad \textit{Equation 3}$$

Here, n_A-FRET are photons collected in the acceptor channel without the photons emitted after FRET, n_D are those collected in the Donor channel and n_A-direct include FRET-emitted photons.

$$E = \frac{n_A}{n_A + \gamma n_D} \qquad \textit{Equation 4}$$

Similarly, n_D and n_A represent photons collected by the donor and acceptor channel respectively. This equation differs from the conventional definition of FRET efficiency due to the nature of PIE, where the acceptor die is also directly excited, not only through the FRET interaction [9].

1.5 The Translocon: opening and closing of the lateral gate in the SecYEG complex

The SecYEG complex is a membrane protein complex in bacteria (here: specifically, *E. coli*), which is responsible for the transport of polypeptides across, or into, the membranes lipid bilayer. The protein forms a channel (lateral gate), whose opening and closing is key for the mechanism of protein transport. This process is modulated by the nucleotide state of its cytosolic counterpart - ATPase SecA. Observing this process is done by labeling the lateral gate with fluorophores and performing a smFRET experiment [10]. To achieve the highest possible temporal resolution, the choice of the right donor and acceptor dye is important. For the interpretation of conformational dynamics on a temporal scale, hidden Markov modeling methods in the form of a novel software package were employed.

1.6 FRETbursts and mpH2MM

FRETbursts is an open-source toolkit specializing in the analysis of smFRET datasets. The software offers high customizability while offering a reliable method of analysis. It is hosted on Github and run in Python, with multiple notebooks and templates that are optimized to be executed via Jupyter Notebooks. It is therefore easily accessible and understandable, regardless of prior coding knowledge. It is capable of analyzing selected sets of raw data produced by multiple photon streams, allows the estimation of background rates, and the selection of photon bursts based on their signal-to-background ratio. Further, burst sizes can be specified, such that only certain, significant events will be processed [4]. Photon bursts refer to short time periods, usually on the scale of milliseconds, in which one molecule in the confocal volume undergoes multiple excitation-emission steps which leads to a period of high photon count rate [3]. An extension to this program is the photon-by-photon Hidden Markov Modeling (H2MM) analysis tool. This tool allows the quantification of FRET dynamics in single biomolecules on a sub-millisecond timescale. Using the multi-parameter H2MM (mpH2MM) variant introduces a way to differentiate between conformational changes and photophysical transitions within FRET dynamics [5]. This approach is limited by the timescale of microseconds (referred to as “approach 1” or “classical approach”). Recently, a new method was published, which takes the fluorescence lifetime as an extra dimension for classification of states. This allows the observation of faster dynamic transitions on a timescale closer to nanoseconds. Not only should the extra dimension make classification of states more precise, but it also allows for a control of FRET states calculated from donor fluorescence lifetimes. Because there is usually not enough photons in a microsecond to fit fluorescence lifetime using the conventional approaches (Maximum Likelihood Estimate), Authors employed this “divisor approach” to estimate fluorescence lifetime on this fast timescale [3].

2. Aims of Work

In this work, single molecule experiments will be optimized for the detection of conformational changes of the lateral gate in the Sec translocon. First, the best candidates for the FRET acceptor dye will be determined, based on the low propensity of triplet transitions. Here, Atto643, Atto 647N and Cy5 will be tested using FCS and classified according to their triplet portion of the FCS curve.

Subsequently, two photon-by-photon Hidden Markov Modeling methods will be directly compared. The first method classifies fluorescent states based on the FRET efficiency and Stoichiometry and the second method adds fluorescent lifetimes as an extra dimension in the classification process.

3. Methods

3.1 FCS measurements

In the FCS, 3 different, red dyes were analyzed. Specifically, Atto 643, Atto 647N and Cy5 were tested at a concentration of roughly 1nM and a volume of 50 µl. The samples were diluted in TKM buffer (20 mM TRIS, 50 mM KCl, 2 mM $MgCl_2$; pH 7.5). Each measurement was repeated three times.

The data acquisition was done in a confocal microscope with a custom-built laser setup [10]. The microscopes focus was set to be 20 µm above the coverslips glass interface, ensuring that the confocal volume is fully within solution, no external forces (adhesion, surface tension) would be present, and observed molecules would be free in solution.

3.2 FRET sample and measurement

The FRET Sample preparation and all FRET measurements were performed by Crossley et al. [10] and is described in full detail within their article. The Sample consist of SecYEG-PrlA4 apo from *E. coli* with lateral gate labeling via ATTO 565 as Donor and ATTO 643 as Acceptor dye. The SecYEG complex features two unique Cysteine residues, denoted A103C and V353C, in the SecY, at which the labeling of the complex occurred. The labeling of each residue with each dye was random. The PrlA4 variant was prepared via site-directed mutagenesis. The labeled SecYEG complex was then introduced into proteoliposomes with diameter of 100 nm [10].

SmFRET data was acquired on a confocal microscope with a pulsed interleaved excitation (PIE) regime [10]. Samples were measured in a bovine serum albumin (BSA) coated 8-well sample chamber at a concentration of 30 pM within a TKM buffer solution. 1 mM aged Trolox was added to reduce photoblinking and photobleaching [10] [11].

3.3 Analysis via FRETbursts and H2MM

Analysis was performed by the FRETbrusts and H2MM python library packagess [4], [5], [6] in JupyterLab 3.6.3 and Spyder 5.4.3 within a Python 3.8.18 environment. The user interface code was created via multiple pre-made templates and altered to fit the specific needs of this work. A full transcript of the implemented and altered user-interface code can be found in the appendix at the end. The code's application in its principle is explained in Diagram 2 below.

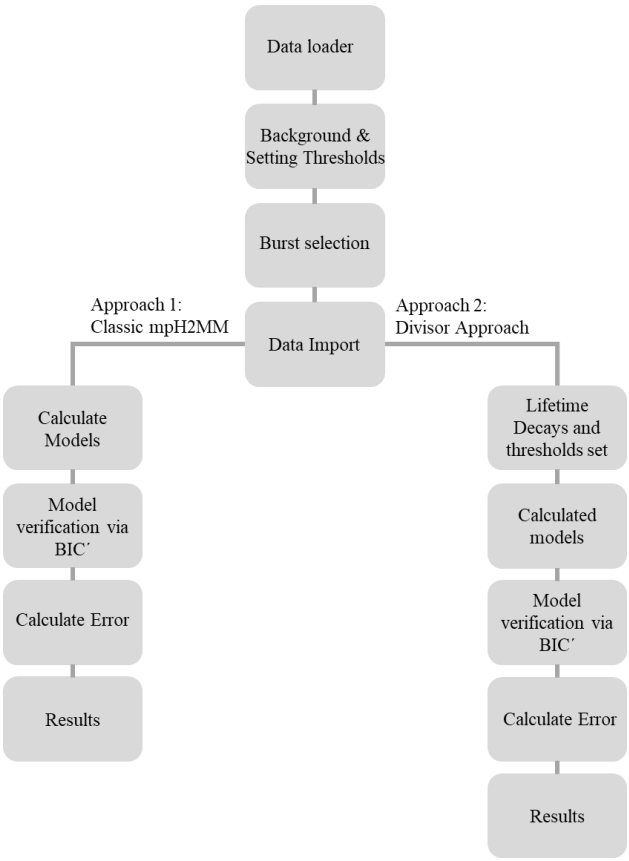

Diagram 2: Overview of the process done by the program. The altered dataset after burst selection for the FRET ES plot was stored as a new object and then used in all further calculations.

Burts-searches were implemented with a minimum threshold of 6 times higher than the background in all channels, and a minimum burst size of 50 photons, in order to ensure satisfactory signal-to-background and signal-to-noise ratio of data considered for further analysis.

For the classical approach to H2MM, models were selected by implementation of the modified Bayes Information Criterion (BIC') as convergence criteria. A fitting model size was chosen, and the results were plotted into an Efficiency-Stoichiometry scatter plot (ES scatter plot) with precise information of each states position, size, and errors, as well as transition rates between each state being summarized in form of a table. All these results can be found in Section 4 in Figure 7 and Table 2 and 4. Errors were defined as standard deviations calculated via bootstrapping.

For the Divisor approach, the FRET selected data was first analyzed for its lifetime decays and appropriate thresholds for the instrument response function (IRF) of each stream were set. Next, HMM models were calculated again using BIC' as convergence criteria. Again, the

results were plotted into an ES scatter plot (Figure 7). State positions, size and transition rates were calculated, and errors of all parameters were estimated via bootstrapping (Table 2 and 4). Lastly, the distribution of fluorescent lifetimes per state were plotted in a histogram (Figure 9).

4. Results

4.1 FCS Results

The autocorrelation functions resulting from the FCS experiments were fitted with the assumption, that only one diffusing species was present, and that triplet state transitions occurred in all cases. Table 1 summarizes the necessary parameters. The parameters were calculated via the fit. τ_{Trip} gives information on the lifetime of the triplet state. Comparing the three dyes, there is a significantly longer triplet lifetime in Cy5 (τ_{Trip} = 12 µs) than in both Atto dyes (τ_{Trip} = 0.5 µs). This difference can also be seen in the graphical comparison (Figure 1). Furthermore, the triplet state amplitude (T) directly measures the propensity of the system to undergo transitions into the triplet state. ATTO 643 and ATTO 647 have the same propensity to form triplet (within the experimental error; ~0.24), while Cy5, a well established and frequently used red fluorophore, spends significant portion of time in a triplet state (T=0.46). This property could heavily bias the determination of conformational dwell times/ interconversion rates.

Table 1: FCS parameters.

Parameter	**Atto643**		**Atto647N**		**Cy5**	
	Value	Error	Value	Error	Value	Error
T []	0.248	0.045	0.235	0.071	0.46	0.024
τ_{Trip}[µs]	0.5	0	0.5	0	12	0

Table 1 and Figure 1 show the FCS results in comparison. It can be seen that Atto 643 (Figure 1A) and Atto 647N (Figure 1B) appear more photostable than the Cy5 (Figure 1C), where a significant triplet portion can be seen. Furthermore, Atto 643 shows an overall higher correlation than Atto647N and Cy5 which seem to be similar.

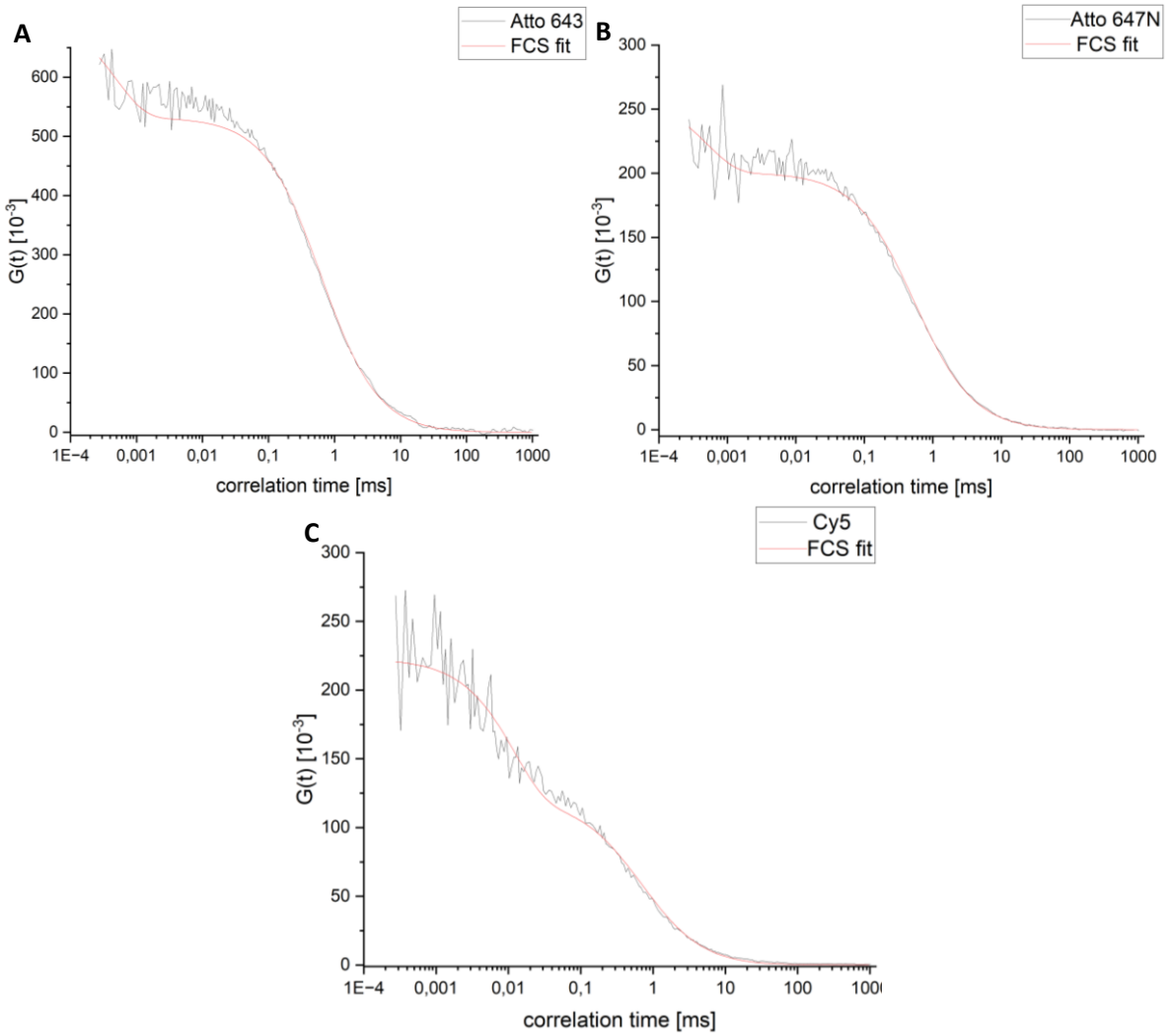

Figure 1: FCS results for all three samples. On the x-axis the correlation tie is shown on a logarithmic scale. The y-axis shows G(t), which is the autocorrelation function defined as the correlation of a signal at time t with itself at various lagtimes τ. (A) shows the results for Atto 643, (B) for Atto 647N and (C) for Cy5. In each case, the raw data correlation data is shown in black and the fit in red. For each fit, the presence of a triplet state was assumed.

4.2 FRETbursts and H2MM Results

In the FRET experiment, laser alternation periods for the PIE laser setup were first checked. Donor and Acceptor excitation periods were assigned accordingly (shaded areas in Figure 2A). The Donor and acceptor excitation time ranges were set to match the experimental settings. Furthermore, the experiments timetrace was visually reviewed for the presence of aggregates. Figure 2B shows that singular bursts of photons are distinguishable. It is concluded that the sample is within a correct dilution such that smFRET analysis can be performed.

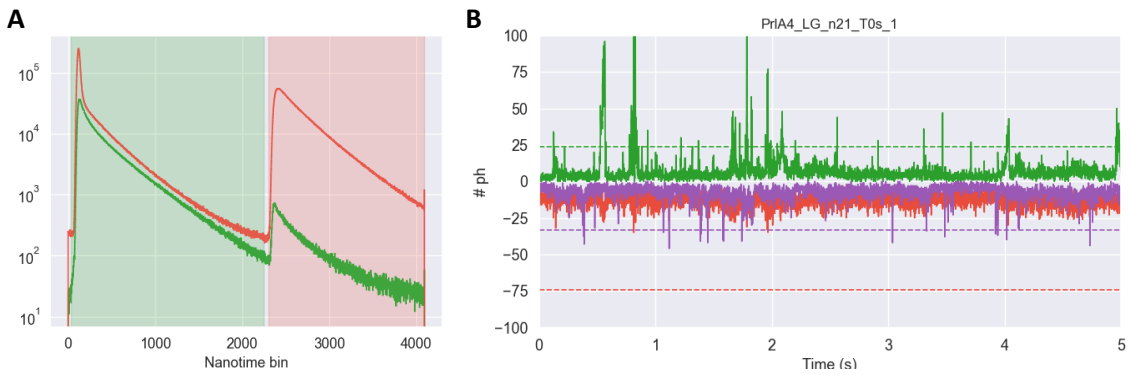

Figure 2: (A) Alternation histogram for nanosecond - Alternating Laser Excitation Time-Correlated Single Photon Counting (ns-ALEX TCSPC) measurements. Photons detected by the donor channel are green, those detected by the acceptor channel are in red. The shaded areas define the donor and acceptor periods. Detector periods were set manually to fit the dataset. (B) 5 second interval of the complete timetrace. Donor photons are in green, Photons resulting from the FRET interaction are in purple, and acceptor photons in red. Singular peaks, and therefor singular molecules within the confocal volume, can be distinguished. Regular FRET interaction can be expected.

The background was then calculated and corrected for. The results can be seen in Figure 3.

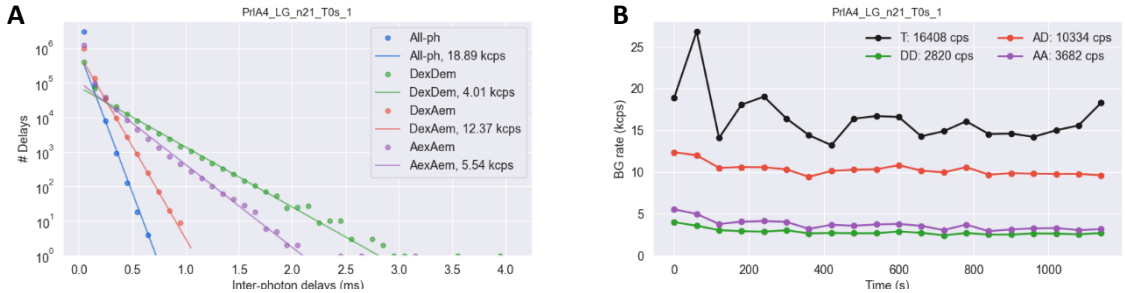

Figure 3: (A) Background rate histogram. Inter-photon delays fitted with an exponential function for different photon streams. Experimental distribution of inter-photon delays (dots) and their corresponding fits (lines). Shown are 4 fits: the sum of all photons is in blue, donor excitation with donor emission in green, abbreviated as DexDem (Donor excitation Donor Emission), donor excitation with Acceptor emission in orange/red, abbreviated as DexAem (Donor excitation Acceptor emission), and lastly, Acceptor excitation with Acceptor emission is in purple (AexAem – Acceptor excitation Acceptor emission). (B) Background rates as a function of time of different photon streams. Color-coding is the same as in (A) with the total being black instead of blue. Each datapoint represents a 54 second window.

Burst selection was done such that the minimum burst size was 50 photons, and a threshold was set to 6 times above the background. The resulting ES plot is shown in Figure 4. With stoichiometry close to 0.5, higher E-values (~0.8) correspond to a closed lateral gate and lower efficiency values (~0.4) correspond to an open later gate of the translocon. Higher S-values in low FRET regions (S~1, E~0) correspond with a donor only population and low S-values (S~0) to an acceptor only population. The most populated states visible are the donor only (top and left) and high FRET (middle and to the right) states.

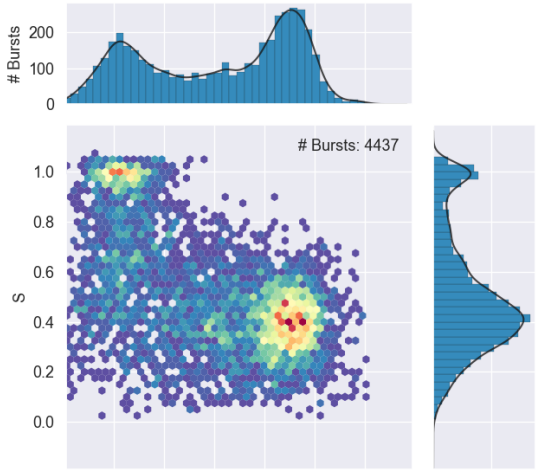

Figure 4: E-S Histogram showing FRET interaction between the donor and acceptor. In the upper left (S > 0.8, E < 0.2) is donor only; down and to the right is acceptor only (S < 0.2). In the middle (E = 0.7-8; S ~ 0.4), high FRET interaction is shown. Low Fret interaction is shown at S ~ 0.4 and E < 0.4. Selected datapoints shown here were used for further analysis in the HMM methods.

After background correction and burst selection, the two H2MM methods were applied to the FRET data.this data set was intentionally not corrected for the direct excitation of the acceptor and spectral cross-talk. Hence, the donor only population is not found exactly at E=0; S=1. Since the Divisor approach requires the specification of lifetime parameters, the lifetime decays were calculated as well. Figure 5 below shows the resulting fluorescence lifetime per state, the derived lifetimes can be used as corroboration that low FRET state should have higher donor fluorescence lifetime than high FRET state (Figure 5B). From Figure 5A the exact nanotimes of excitation pulses and corresponding instrument response function (IRF) were derived and fed into FRETbursts.

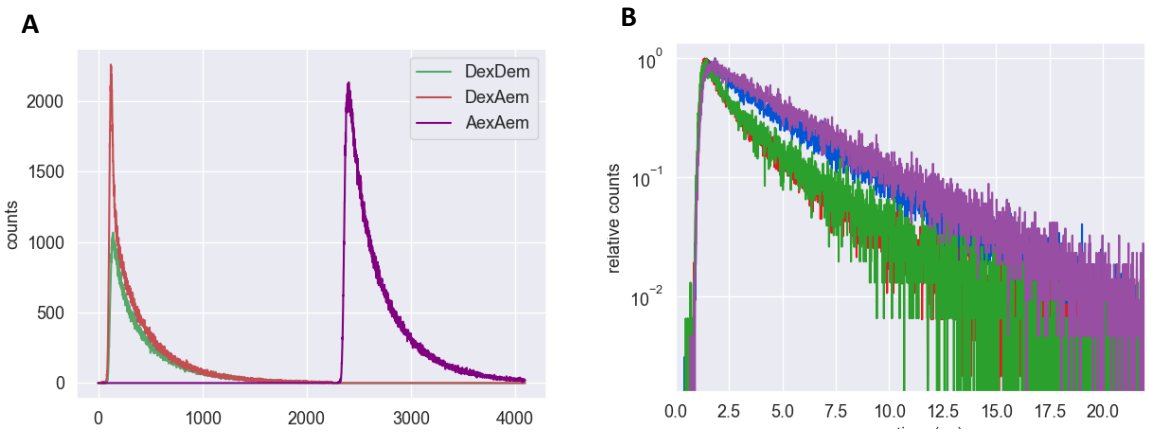

Figure 5: (A) Alternation modulo. The information gained from this graph was used to set the IRF thresholds for each of the three channels. In principle, this graph is a derivation of Figure 2A. (B) shows the lifetime decay for each photon stream. It is derived from the individual channels and allows a species verification of states.

For each H2MM method, 8 models in a range of 1 to 8 states in ascending order were calculated. Meaning model 1 accounted for 1 state and model 8 accounted for 8 different states.

Then the modified Bayes Information Criterion (BIC') was applied to determine the least amount of states which would reasonably explain the full dataset. BIC' suggests that the set in both approaches can be best described by four states. Two of these are conformational, meaning open and closed lateral gate, and two are photophysical, meaning the donor only and acceptor only states. Figure 6 shows the results of the BIC' for the classic (A) and the devisor (B) approach.

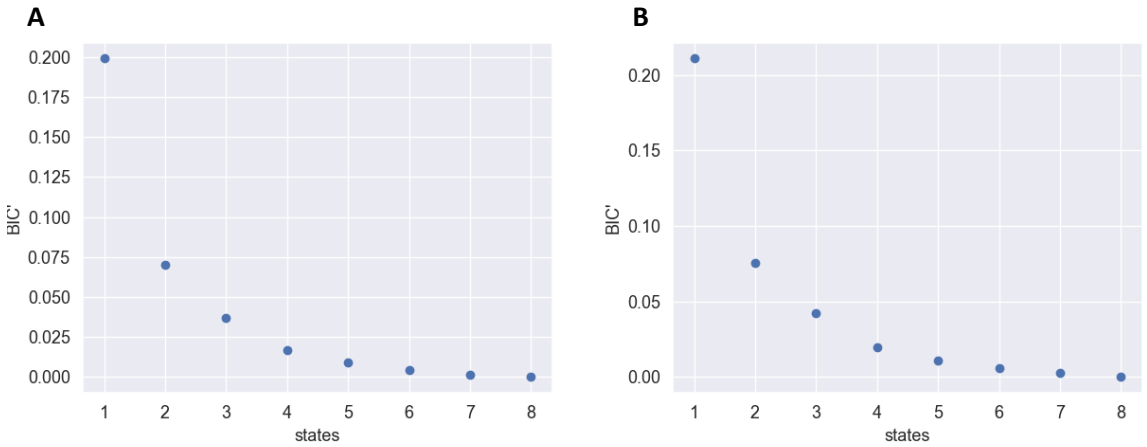

Figure 6: BIC' for 8 modeling attempts. (A) shows the result of the classic approach and (B) for the divisor approach. Valid models satisfy the condition BIC' < 0.005. The 4-state model was chosen as the most trustworthy one in both cases. Models with more states tend to overfit in order to achieve the least statistical error by separating bursts of the same state into two or more non-existing states.

Lastly, the established four states are shown in Figure 7. The transition rates, depicted as arrows, show the migration pathways between the states. The migration from the donor only population towards the low FRET dwell occurs most often. This event could be explained by photoblinking. This confirms again that even in high quality dyes with low percentage of triplet, blinking has to be considered as an important factor and needs to be accommodated for in the model. Not shown in this scatter plot are the transitions from the donor only population into the High FRET population and vice versa ($1.3 * 10^{-14}$ s^{-1} and $2.4 * 10^{-20}$ s-1 respectively) or the transitions from donor to acceptor directly ($1.5 * 10^{-73}$ s^{-1}). While these processes do happen, the probability of these transitions are low. Of particular interest is the transition into the donor only state. This transition indicates the loss of FRET interactions, whether it is due to blinking or bleaching cannot be differentiated at this point.

Both H2MM approaches yield fairly similar results (see Figure 7). An exact comparison can be found in Table 2 for the ES values, and Table 4, for the transition pathways.

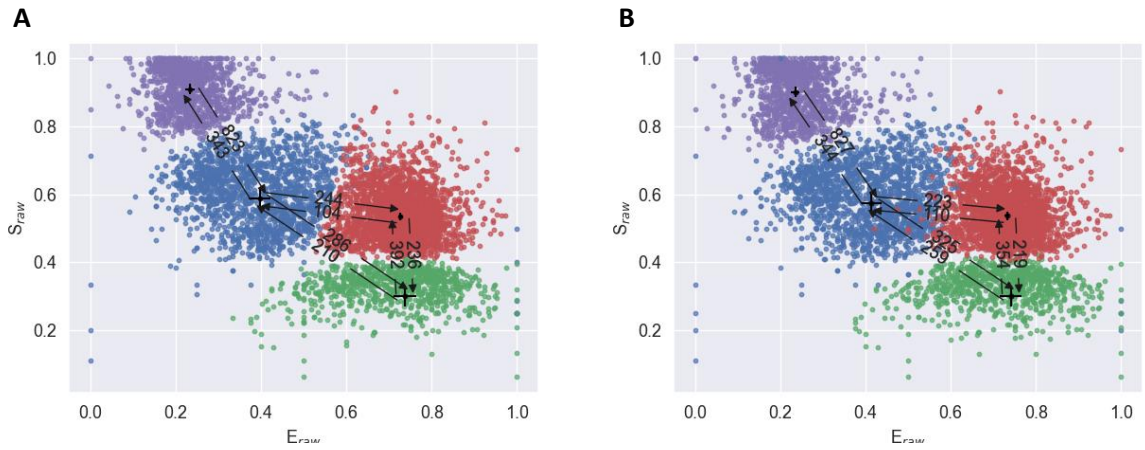

Figure 7: Comparison of the results from the classical H2MM and divisor based H2MM. ES scatter plots are derived from the classic approach (A) and from the divisor approach (B). Consecutive photons with the same state are considered as a single dwell. Donor only population is in purple, acceptor only in green. Low fret and high fret states are in blue and red respectively. Arrows show the transition from one dwell into the next with the corresponding number in units of "per second" (s^{-1}).

The comparison of the E and S values together with their statistical variance are shown in Table 2. We proved that both methods yield similar results in both, the relative position of each state and the error (the "density") of each state. The variance between both methods is shown in Table 3, comparing the E values of each state and stoichiometry of each state.

Table 2: Exact values for the center of each state as shown in Figure 6. The abbreviations for each state are as follows: Donor only state (D, purple in Figure 7), low FRET state (LF, blue), high FRET State (HF, red), and acceptor-only state (A, green). The error columns represent the statistical deviations obtained by bootstrapping. In Figure 7, these are shown as crosses.

	Classic Approach				**Divisor Approach**			
	Efficiency	**Error**	**Stoichio-metry**	**Error**	**Efficiency**	**Error**	**Stoichio-metry**	**Error**
D	0.233	0.011	0.910	0.015	0.235	0.011	0.901	0.015
LF	0.397	0.025	0.587	0.031	0.413	0.025	0.575	0.030
HF	0.727	0.006	0.535	0.012	0.732	0.006	0.537	0.012
A	0.737	0.026	0.300	0.032	0.741	0.026	0.301	0.031

Table 3: comparison of Efficiency per state in both approaches and of stoichiometry per state in both approaches. The deviation is given in percent. As established, Donor only state is denoted as D, Acceptor only as A and the FRET states are denoted LF for low FRET and HF for high FRET state.

State	classic approach	divisor approach	comparison [%]
	Efficiency	Efficiency	
D	0.233	0.235	0.71
LF	0.397	0.413	3.90
HF	0.727	0.732	0.75
A	0.737	0.741	0.55
	Stoichiometry	Stoichiometry	
D	0.910	0.901	0.95
LF	0.587	0.575	2.21
HF	0.535	0.537	0.26
A	0.300	0.301	0.12

Table 3 shows that the low FRET states deviate the most in both efficiency (3.90%) and stoichiometry (2.21%) when comparing the classic approach to the devisor approach. The acceptor state on the other hand shows the most similarities in efficiency (0.55%) and stoichiometry (0.12%) when comparing the classic and divisor approach.

Table 4: Additional data for the comparison of the results from the classical H2MM based and divisor based H2MM. The abbreviations used in the transition pathways are as follows: Donor only state (D), low FRET state (LF), high FRET State (HF) and acceptor-only state (A).

	Classic Approach			Divisor Approach			Comparison
Pathway	**Transition rate [s^{-1}]**	**Error [s^{-1}]**	**Error [%]**	**Transition rate [s^{-1}]**	**Error [s^{-1}]**	**Error [%]**	**Transition rate difference[%]**
D→LF	823.3	116.5	14.2	827.3	138.8	16.8	0.5
LF→D	342.9	56.4	16.4	343.8	83.0	24.2	0.3
LF→HF	244.3	98.0	40.1	222.9	64.1	28.8	9.6
HF→LF	103.6	29.7	28.7	110.3	23.5	21.3	6.1
HF→A	236.0	97.8	41.4	218.6	91.8	42.0	8.0
A→HF	392.3	147.1	37.5	353.8	132.1	37.3	10.9
LF→A	285.6	94.7	33.1	325.5	88.6	27.2	12.3
A→LF	210.3	73.5	34.9	259.0	67.5	26.1	18.8

The transition rates derived by both approaches are similar (see Table 3). The divisor approach yields on average slightly smaller errors than the classical approach. The transition rate difference column shows the deviation of transition rates in the divisor approach from transition rates from the classic approach. The biggest difference is observed in the transition from the acceptor-only population into the Low FRET state, with an 18.8 % lower value of transitions per second in the classic approach compared to the divisor method. Transitions involving the Donor only state show the most similar results, where deviations are only 0.5 % / 0.3 %.

Figure 8 shows the population of photons in each state in dependance of the FRET efficiency (8A and 8C) and the stoichiometry (8B and 8D). In both cases, the FRET states are the most populated states. This corresponds to what was seen in the FRET ES plot (Figure 4) where the high FRET state was most populated.

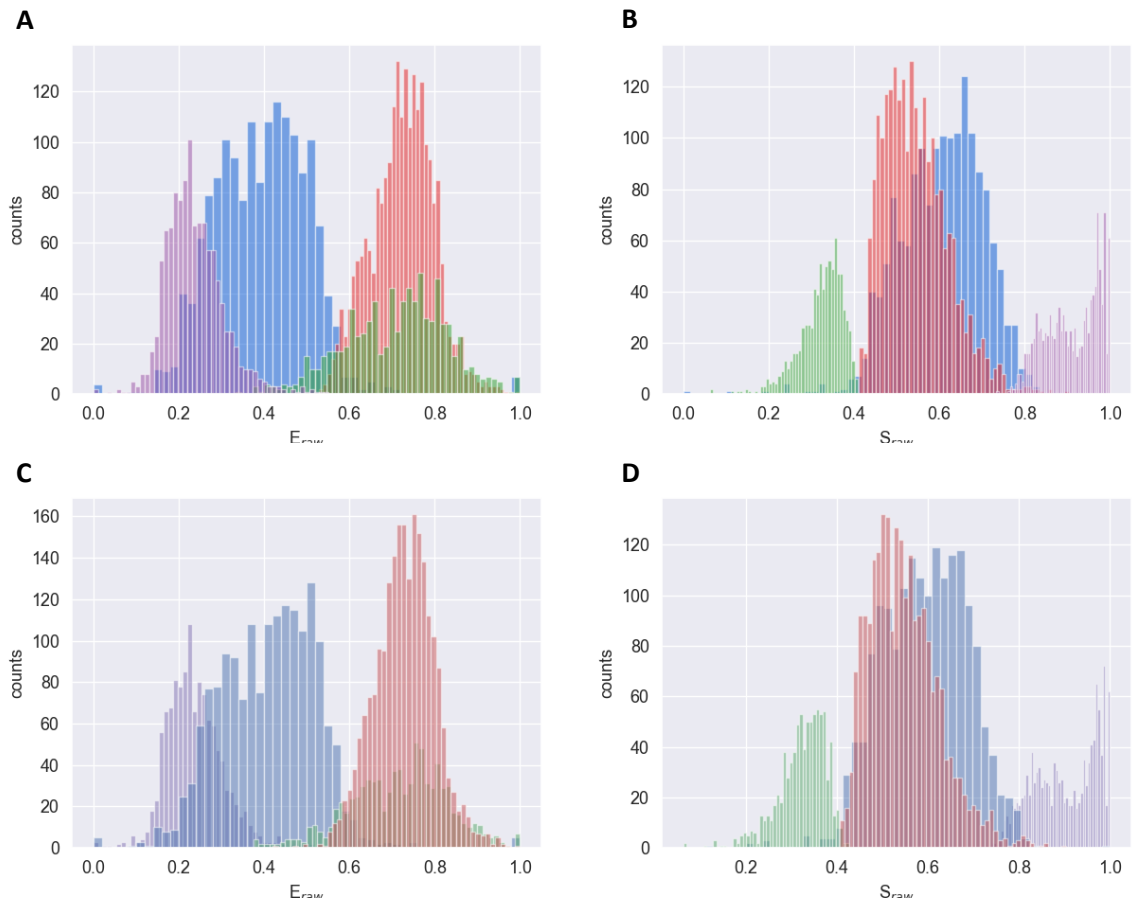

Figure 8: Population histogram for the classic approach (A and B) and for the divisor approach (C and D). The y-axis shows the photon count. The x-axis shows the FRET efficiency (A, C) and the photon Stoichiometry (B, D). The E histogram and S histogram correspond to Figure 7 with the color coding being identical to Figure 7 for each state. It can be seen in all cases that both FRET states are the most populated, followed by the donor-only and lastly by the acceptor-only state.

Lastly, the fluorescence lifetimes of all extracted states were derived by the divisor approach (Figure 9). The High FRET state has a lower average lifetime compared to the low FRET state by approximately 1 ns. It also shows that the donor-only state has a longer average lifetime than the acceptor-only state. This information is unique to the divisor approach. Since the FRET efficiency and stoichiometry are by default not corrected for cross-talk and direct acceptor excitation, the lifetimes do not match the exact FRET values.

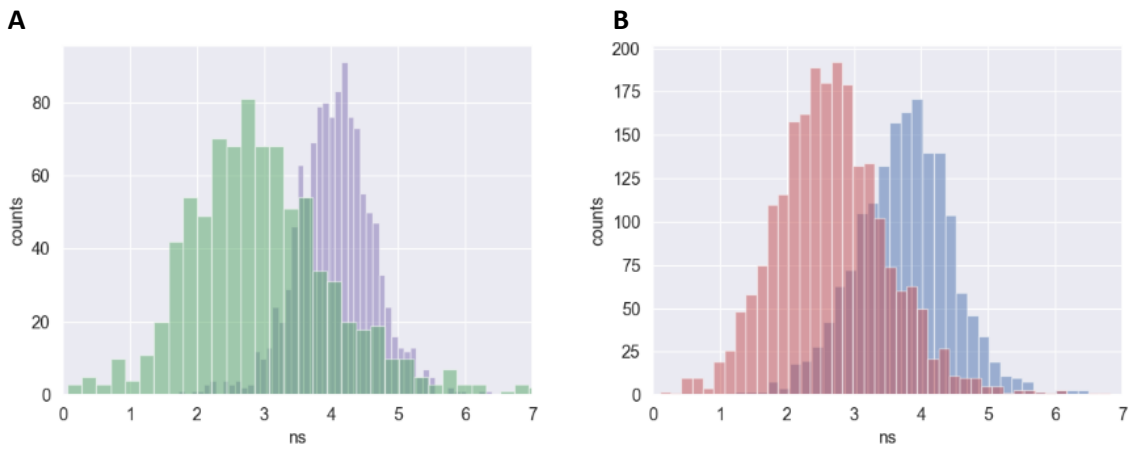

Figure 9: Fluorescence lifetimes derived by the divisor approach. The y-axis shows coutns, the x-axis shows the fluorescent lifetimes in nanoseconds. (A) shows only the donor (purple) and acceptor (green) state, while (B) shows both FRET states (low FRET in blue; high FRET in red).

5. Discussion and Conclusions

Single molecule experiments aimed to detect conformational dynamics of the lateral gate of translocon were optimized. Specifically, the best candidate for FRET as acceptor dye was determined based on their low propensity of triplet transition. The selected candidate was Atto 643, a modern dye with good photostability and good water solubility. Similarly low results for the triplet state abundance were seen in Atto 647N, which, however, is known to non-specifically interact with proteins, making it a less viable choice as FRET acceptor.

In the FCS experiments, Atto 643 showed its low tendency to form Triplet states. Nevertheless, the H2MM analysis still showed a high donor only population, indicating that a significant portion of acceptor molecules either blinked or bleached even though aged Trolox, which is known to increase photostability [7] [11], was present during the measurement. This suggest that there is still room for improvement regarding the photophysical properties of the acceptor dye. To improve these parameters, we are going to test other available antiblinking ad antibleaching reagents in the future.

In the comparison of the H2MM models, one crucial step was the determination of the number of states. Here, it was concluded that 4 states are present. Even though models with more states achieve better statistical fits, they encounter the risk of being overfitted in an attempt to achieve the least statistical error. With more than four states in our system, the algorithm starts splitting states with artificially sharp separating borders, which indicates that the division is no longer physiological and statistically meaningful.

Overall, the comparison of both H2MM methods yielded only minor deviations. A clearly "more accurate" method cannot be decided. Harris et al. [3], [5] already demonstrate the reliability of their analyzing tools. It can be concluded that both methods are reliable ad viable options. However, the divisor approach offers additional insight on the lifetime of FRET states, which will have other relevant applications, e.g. determination of kinetics faster than tens of microseconds.

6. Literature and References

[1] Harekrushna Sahoo, Förster resonance energy transfer – A spectroscopic nanoruler: Principle and applications, *Journal of Photochemistry and Photobiology C: Photochemistry Reviews,* vol. 12, no. 1, pp. 20-30, 2011.

[2] Robert M. Clegg, Chapter 1 Förster resonance energy transfer—FRET what is it, why do it, and how it's done, *Laboratory Techniques in Biochemistry and Molecular Biology,* vol. 33, pp. 1-57, 2009.

[3] E. L. Paul David Harris, Identification and quantification of within-burst dynamics in singly labeled single-molecule fluorescence lifetime experiments," *Biophysical Reports,* vol. 2, no. 3, 2022.

[4] Antonio Ingargiola et al., FRETBursts: An Open Source Toolkit for Analysis of Freely-Diffusing Single-Molecule FRET, *bioRxiv,* 2016.

[5] Harris Paul David et al., Multi-parameter photon-by-photon hidden Markov modeling, *Nature Communications,* 2022.

[6] Menahem Pirchi et al., Photon-by-Photon Hidden Markov Model Analysis for Microsecond Single-Molecule FRET Kinetics, *The Journal of Physical Chemistry B,* vol. 120, no. 51, pp. 13065-13075, 2016.

[7] Alexander P Demchenko, Photobleaching of organic fluorophores: quantitative characterization, mechanisms, protection," *IOP Publishing,* vol. 8, no. 2, 2020.

[8] M. H. Radek Macháň, Fluorescence Correlation Spectroscopy (FCS), PicoQuant GmbH, [Online]. Available: https://www.picoquant.com/images/uploads/page/files/17319/5_fcs.pdf. [Accessed 14 12 2023].

[9] Steffen Ruettingeret al., FRET analysis with Pulsed Interleaved Excitation (PIE) using the MicroTime 200, Berlin, Germany, 2013.

[10] Joel A. Crossley et al., Energy landscape steering mediates dynamic coupling in ATP-driven protein translocation by the bacterial Sec machinery, *bioRxiv,* 2023.

[11] Altman Roger B. et al., Cyanine fluorophore derivatives with enhanced photostability, *Nature Methods,* vol. 9, no. 1, pp. 68-71, 2012.

7. Appendix

```
#import of all necessary modules
import numpy as np
from matplotlib import pyplot as plt
import fretbursts as frb
import burstH2MM as hmm
sns = frb.init_notebook()
# load the data into the data object frbdata
filename = 'PrlA4_LG_n21_T0s_1.hdf5'
def apply_params(d, setup_params):
    d.leakage = setup_params['d_leakage']
    d.dir_ex = setup_params['direct_ex']
    d.gamma = setup_params['gamma']
    d.beta = setup_params['beta']
    d.add(D_ON=(setup_params['donor_ON'], setup_params['donor_OFF']),
          A_ON=(setup_params['acceptor_ON'], setup_params['acceptor_OFF']),
          det_donor_accept = (setup_params['donor_ch'], setup_params['acceptor_ch']))
    return d

#% apply correction parameters and calculate background/ check with
#alternation histogram fo validty of parameters
PIE_560_640 = {'name': 'PIE_560_640',
               'donor_ch' : 2,
               'donor_ON': 35,
               'donor_OFF': 2250,
               'donor_q': 2,
               'acceptor_ch': 3,
               'acceptor_ON': 2300,
               'acceptor_OFF': 4096,
               'acceptor_q': 1,
               'd_leakage':  0.16,
               'direct_ex': 0.15,
               'gamma': 0.85,
               'beta': 0.85}


d_raw = frb.loader.photon_hdf5(filename)
frbdata = apply_params(d_raw, PIE_560_640)
# plot the alternation histogram

```

```
frb.bpl.plot_alternation_hist(frbdata)
plt.savefig('alternation_hist')
plt.savefig('alternation_hist.pdf')
# if the alternation period is correct, apply data
frb.loader.alex_apply_period(frbdata)
# calcualte the background rate
frbdata.calc_bg(frb.bg.exp_fit, F_bg=1.7, )
# plot bg parameters, to verify quality
frb.dplot(frbdata, frb.hist_bg)
plt.savefig('background_rate_hist')
plt.savefig('background_rate_hist.pdf')
# calcualte small section of timetrace
frb.dplot(frbdata, frb.timetrace)
plt.xlim(0,5)
plt.savefig('timetrace.png')
plt.savefig('timetrace.pdf')
frb.dplot(frbdata, frb.timetrace_bg)
plt.savefig('timetrace_bg.png')
plt.savefig('timetrace_bg.pdf')
# now perform burst search
np.float = float
frbdata.burst_search(m=10, F=6)
# make sure to set the appropriate thresholds of ALL size
# parameters to the particulars of your experiment
frbdata_sel = frbdata.select_bursts(frb.select_bursts.size, th1=50, th2=500, add_naa=False)
frb.alex_jointplot(frbdata_sel);
plt.savefig('FRET_burst.png')
plt.savefig('FRET_burst.pdf')
bdata = hmm.BurstData(frbdata_sel)
# calculate models
bdata.models.calc_models(to_state=3, max_state=8, conv_crit="BICp")
#given are 3 state selection methods: ICL, BIC and BIC'. we are
#interested in BIC'
hmm.BICp_plot(bdata.models)
plt.savefig('BICp.png')
plt.savefig('BICp.pdf')
# calculates exact ES values, transition matrix and
# std deviation for all 3
Transarray=bdata.models[3].trans
Earray=bdata.models[3].E
Sarray=bdata.models[3].S
```

```
81.
82.  trans_std, E_std, S_std=bdata.models[3].bootstrap_eval()
83.
84.  print("E-array", "\n", Earray)
85.  print("S-array", "\n", Sarray)
86.  print("Transition rate-array", "\n", Transarray)
87.
88.  print("E st deviation", "\n", E_std)
89.  print("S st deviation", "\n", S_std)
90.  print("trans st deviation", "\n", trans_std)
91.  # plot the dwell ES of the result
92.  state_color = [{'color':'b'}, {'color':'r'}, {'color':'g'}, {'color':'m'}]
93.  hmm.dwell_ES_scatter(bdata.models[3], state_kwargs=state_color )
94.  # overlay with the main values,
95.  hmm.scatter_ES(bdata.models[3], s=10, c="k", )
96.  plt.errorbar(Earray, Sarray, xerr=E_std, yerr=S_std, ecolor='black', fmt='none')
97.  hmm.trans_arrow_ES(bdata.models[3]);
98.
99.  plt.savefig('ES_scatter')
100. plt.savefig('ES_scatter.pdf')
101. hmm.dwell_E_hist(bdata.models[3], bins= 50, )
102. plt.savefig('E_hist')
103. plt.savefig('E_hist.pdf')
104. hmm.dwell_S_hist(bdata.models[3], bins= 50)
105. plt.savefig('dwell_S_hist.png')
106. plt.savefig('dwell_S_hist.pdf')
107. #Divisor approach
108.
109. hmm.state_nanotime_hist(bdata.models[3], normalize=True)
110. plt.xlim([0,22])
111. plt.savefig('Nanotime_hist')
112. plt.savefig('Nanotime_hist.pdf')
113. hmm.raw_nanotime_hist(bdata)
114. plt.savefig('Nanotime_bins.png')
115. plt.savefig('Nanotime_bins.pdf')
116. #creates a file titled “output” in which the values corresponding
117. #to nanotime_hist in order: x-axis, DexDem, DexAem, AexAem
118. streams = bdata.ph_streams
119. stream_id = [np.argwhere([stream == psel for psel in bdata.ph_streams])[0,0] for stream in
120. streams]
```

```
index = np.concatenate(bdata.models.index)
bc = [np.bincount(np.concatenate(bdata.nanos)[index==idx],
minlength=bdata.data.nanotimes_params[0]['tcspc_num_bins']) for idx in stream_id]
#print(bdata.data.nanotimes_params[0]['tcspc_num_bins'])
with open('output.txt', 'a') as file:
    for x in range(bdata.data.nanotimes_params[0]['tcspc_num_bins']):
        file.write(str(x) + ": ")
        for y in range (3):
            file.write(str(bc[y][x]) + " ")
        file.write("\n")

#selected datapoint from nanotime_hist as thresholds for the IRF
bdata.irf_thresh = np.array([144, 122, 2404,])
#creates divisor
div_name = bdata.auto_div(1)
# run H2MM analysis
bdata.div_models[div_name].calc_models(to_state=3, max_state=8, conv_crit="BICp")
#calculates exact ES values, transition matrix and std deviation
#for divisor data for all 3
Earray_div=bdata.div_models[div_name][3].E
Sarray_div=bdata.div_models[div_name][3].S
Transarray_div=bdata.div_models[div_name][3].trans

trans_std_div, E_std_div, S_std_div=bdata.div_models[div_name][3].bootstrap_eval()

print("Divisor_E-array", "\n", Earray_div)
print("Divisor_S-array", "\n", Sarray_div)
print("Divisor_trans-array", "\n", Transarray_div)

print("E st deviation", "\n", E_std_div)
print("S st deviation", "\n", S_std_div)
print("trans st deviation", "\n", trans_std_div)
# plot the dwell ES of the result
state_color = [{'color':'m'}, {'color':'b'}, {'color':'g'}, {'color':'r'}]
hmm.dwell_ES_scatter(bdata.div_models[div_name][3], state_kwargs=state_color ) #ax=ax,
states=[0,1,2,3,4,5,6,7] )
#ax.legend()
# overlay with the main values,
hmm.scatter_ES(bdata.div_models[div_name][3], s=10, c="k")
hmm.trans_arrow_ES(bdata.div_models[div_name][3]);
```

```
hmm.scatter_ES(bdata.div_models[div_name][3], s=10, c="k")
plt.errorbar(Earray_div, Sarray_div, xerr=E_std_div, yerr=S_std_div, ecolor='black', fmt='none' )
plt.savefig('ES_scatter_div_model.png')
plt.savefig('ES_scatter_div_model.pdf')
#E histogram
state_color = [{'color':'m'}, {'color':'b'}, {'color':'g'}, {'color':'r'}]
hmm.dwell_E_hist(bdata.div_models[div_name][3], bins= 50, states=[0,1,2,3],
state_kwargs=state_color)#, order_kwargs=[1,3,2,0])

plt.savefig('dwell_E_hist_div.png')
plt.savefig('dwell_E_hist_div.pdf')
#S histogram
state_color = [{'color':'m'}, {'color':'b'}, {'color':'g'}, {'color':'r'}]
hmm.dwell_S_hist(bdata.div_models[div_name][3], bins= 50, states=[0,1,2,3],
state_kwargs=state_color)
plt.savefig('dwell_S_hist_div.png')
plt.savefig('dwell_S_hist_div.pdf')
hmm.BICp_plot(bdata.div_models[div_name])
plt.savefig('BICp_div.png')
plt.savefig('BICp_div.pdf')
#tau histogram: lifetime of FRET states per state; for better
#visiblity split in 2 (following)
state_color = [{'color':'m'}, {'color':'b'}, {'color':'g'}, {'color':'r'}]
fig = plt.figure()
ax = fig.add_subplot(111)
hmm.dwell_tau_hist(bdata.div_models[div_name][3], bins=50, ax=ax, states=[0,1,2,3],
state_kwargs=state_color)

ax.set_xlim(0,8)
plt.savefig('dwell_tau_hist_all4.png')
plt.savefig('dwell_tau_hist_all4.pdf')
state_color = [{'color':'m'}, {'color':'g'}]
fig = plt.figure()
ax = fig.add_subplot(111)
hmm.dwell_tau_hist(bdata.div_models[div_name][3], bins=50, ax=ax, states=[0,2],
state_kwargs=state_color)

ax.set_xlim(0,7)
plt.savefig('dwell_Tau_hist_FRETinteract.png')
plt.savefig('dwel_Tau_hist_FRETinteract.pdf')
```

```
state_color = [{'color':'b'}, {'color':'r'}]
fig = plt.figure()
ax = fig.add_subplot(111)
hmm.dwell_tau_hist(bdata.div_models[div_name][3], bins=50, ax=ax, states=[1,3],
state_kwargs=state_color)

ax.set_xlim(0,7)
plt.savefig('dwell_tau_hist_DAonly.png')
plt.savefig('dwell_tau_hist_DAonly.pdf')


```