The University of South Bohemia in České Budějovice

Faculty of Science

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

Bachelor thesis

Lorin Lehner

Advisor: Mgr. Tomáš Fessl, Ph.D.

České Budějovice 2024

Lehner, L., 2024: Optimization of Single-molecule FRET Experiments to observe the Conformational Dynamics in Sec Translocon During Protein Transport. Bc. Thesis, in English, - 37 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, 9.05.2024

Student's signature

Table of Contents

1. Introduction

Understanding biological processes has always been a major goal in human development. As the fundamental understanding of biology has improved across the millennia, the focus in recent history has shifted more and more onto understanding the extraordinary details of processes of individual molecules on a nanoscale. While genetic material (DNA, RNA) has largely been understood, proteins form a much more complex group of macromolecules, given they act as motors, ion pumps, cross-membrane channels and much more. Additionally, each specialized functionality consists of unique protein machinery and unique mechanisms. While classical structural techniques, such as x-ray crystallography or electron microscopy, can provide insight on different conformational states as a snapshot in time, fluorescence-based techniques allow the observation of the change of conformational states in a dynamic fashion. Understanding the complex nature of biomacromolecules requires understanding the changes in conformational states, as many macromolecules inherently behave dynamically or are altered upon binding of a ligand. [1] The analyzed protein in this thesis is the SecYEG translocon, a cross-membrane channel protein complex in the bacteria *Escherichia coli* that transports pre-proteins across the inner membrane [2] [3].

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

The SecYEG complex is a membrane protein complex in the bacteria E. coli, which is responsible for the transport of polypeptides across, or into, the membranes lipid bilayer. The protein forms a channel conducting preproteins across the membrane with an opening to membrane called a lateral gate. Intrinsic dynamics of the lateral gate, i.e. opening and closing is key for the mechanism of protein transport. This process is allosterically modulated by the nucleotide state of its cytosolic counterpart - ATPase SecA [4] [5].

The conformational dynamics of individual molecular machines (here: the opening and closing of the translocon channel) can only be accessed by single molecule fluorescence experiments. In this thesis, the focus lies on the optimization of extraction of rates from the point of view of experimental conditions as well as data analysis. Flurescence of individual molecules has some specifics including blinking and bleaching. Unfortunatelly, blinking and some other photophysical phenomena often happens on the same timescale as conformational

dynamics of molecular machines, hence removing photobleaching and blinking is of utmost importance.

In this thesis, we optimize experimental conditions and data analysis for extraction of interconversion rates in an artificial mutant of the complex, the SecYEG-prlA4. The variation "SecYEG-prlA4" describes a mutant of the SecY protein which allows the translocation of defective preproteins or such that are of missing or non-functional signal sequences. This mutant is a result of two mutations in the SecY gene "that result in the amino acid substitutions F286Y and I408N in transmembrane segment 7 and 10, respectively". What further separates the prlA4 variant from the wildtype is its increased translocation efficiency [5].

This work focuses on the technical aspects of data analysis, specifically on extraction of rates of interconversion between conformational states. Based on our recent results it seems plausible that altered conformational dynamics of the translocon caused by the mutations, is a major cause for the observed phenotype. More specifically, the increased translocation efficiency of the prlA4 mutant is linked with a prolonged residence time of the lateral gate in the "open state", which subsequently allows for more efficient diffusion of transport proteins through the pore [4] [5] [6].

Observing this process is done by labeling the lateral gate with fluorophores and performing a single molecule förster resonance energy transfer (smFRET) experiment [10]. To achieve the highest possible temporal resolution, and to remove potential bias caused by overlapping timescales of dye photophysics and conformational dynamics of the SecYEG complex, the choice of the right donor and acceptor dye is crucial. For the interpretation of conformational dynamics on a temporal scale, hidden Markov modeling methods applied photon-by-photon were employed [7] [8] [9].

The desire to experiment on smaller and faster scales has motivated the need for more precise methods of measurement and analysis. Especially the field of fluorescence imaging and spectroscopy has undergone unprecedented growth [2]. Fluorescence is a powerful tool to observe biological processes on a molecular level. Employing techniques like smFRET or fluorescence correlation spectroscopy (FCS), it is possible to observe conformational changes of individual macromolecules via the labeling of the respective sample with fluorescent dyes [10], [11]. Here the focus lies on the optimization of the experimental conditions to extract rates of conformational dynamics in SecYEG translocon, a conserved nano-machine responsible for transport of proteins across the inner membrane of bacteria. These

conformational dynamics can be quantified via photon-by-photon hidden Markov modeling (H^2MM) analysis, even if they appear on a timescale of micro- to milliseconds. Recently, two different approaches for multi-dimensional H^2MM methods were introduced, as an open-source toolkit that works in combination with FRETbursts, an already established open-source toolkit [7] [8] [9] [12]. 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 H^2MM 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, the selection of the ideal acceptor dye (with red absorption) will be guided by an FCS experiment to classify fluorescent dyes based on their tendency to blink or photobleach.

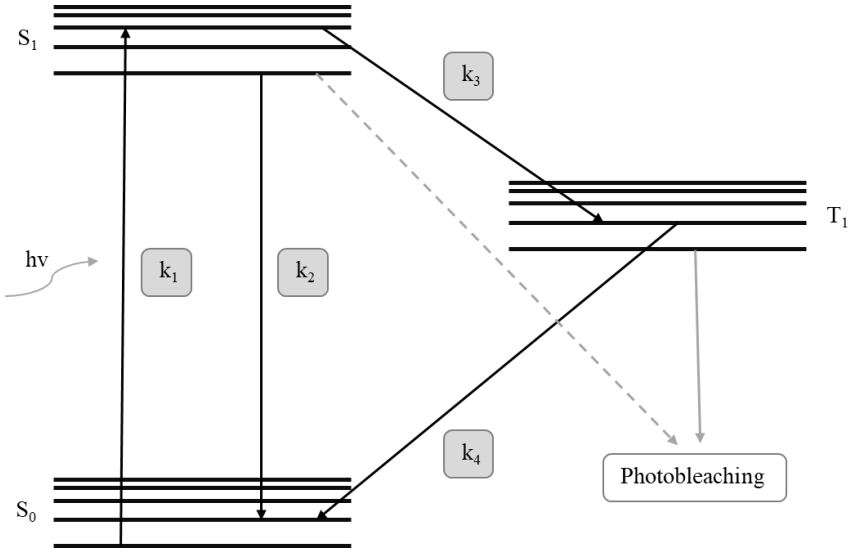
When performing a FRET experiment, the right choice of the fluorescent dyes serving as FRET pair is a success-determining step. When selecting a suitable acceptor dye for the FRET pair, here the focus lies on the red part of the spectrum. This is motivated by high autofluorescence of polar lipid extract from *E. coli* at a 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 [13]. To test the photophysical properties of candidates for FRET acceptor (specifically their propensity to undergo triplet transition) FCS, 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.

1.2 Photophysics: Fluorescence, Photoblinking and Photobleaching

A fluorescent dye will undergo a transition from the ground state to the excited state upon absorption of a photon of suitable wavelength. The depopulation of the excited state in the 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). The 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 [13] and bias/shorten the dwell time of the conformational state. 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, photoblinking can obscure the interpretation of dwell times. On the other hand, photobleaching is an irreversible effect and it introduces a different kind of bias by reducing the capability to detect long-lasting dwell times. Photobleaching is caused by two main mechanisms. The first is an oxygen-dependent pathway, in which molecular oxygen reacts with the reactive excited fluorophore. This effect can be minimized by the addition of oxygen scavenging species. The second mechanism is via the transition from a singlet state into the triplet state in which the molecules reactivity is strongly increased. As a consequence of the triplet state, permanent modifications within the molecule may occur. This results in the permanent loss of fluorescence. Photobleaching is less likely to occur than other photochemical events. 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 [13]. Thus, photodegrading effects are generally tried to be minimized. Diagram 1 below shows the possible energy pathways within a fluorescent molecule.



Schematic 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. The dashed arrow represents other photobleaching reactions such as the reaction with molecular oxygen.

1.3 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 $G(\tau)$ 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 $F(t)$ represents intensity fluctuations occurring by e.g. the fluorophore entering or leaving the confocal volume (translational diffusion), triplet transitions, FRET, etc. τ is the lagtime and the brackets $\langle\ \rangle$ represent the temporal average [14].

In this Thesis, the focus lies on the identification of Triplet states, or other mechanisms inhibiting fluorescence, of the acceptor dye candidates. The triplet state can be identified in an FCS as a region of fast decay rates on a microsecond time scale [15].

1.4 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 [10], [11]. 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, causing an indirect proportionality of the emission intensity of the donor and the relative distance. The distance relation between donor and acceptor can be described by *Equation 2* [11].

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

According to this equation, the transfer rate constant k_t is proportional to the sixth power of the distance between the donor and acceptor molecule (r_{DA}). The term R_0 represents Forster distance. R_0 is a characteristic distance of each FRET pair, at which the efficiency of energy transfer is exactly 50%. It depends on the spectral overlap between the emission spectrum of donor and absorption spectrum of acceptor, the permittivity of the environment, the quantum yield of donor, and the 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 dyes emissions will yield the same result (see the description of PIE below).

1.5 Pulsed Interleaved Excitation

Pulsed Interleaved Excitation (PIE), also often referred to as nanosecond alternating laser exctitation (nsALEX), 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 *Equations 3* and *4*) [16]. The Stoichiometry is defined as the ratio of photons emitted by the donor and acceptor after excitation of only the donor, and the sum of all photons emitted after the excitation of both the donor and acceptor directly. If no acceptor photons are present, the ratio goes to S = 1 ("donor only"), while the absence of a donor would lead to S = 0 ("acceptor only").

$$S = \frac{n_{DexAem} + n_{DexDem}}{n_{DexAem} + n_{DexDem} + n_{AexAem}} \qquad \textit{Equation 3}$$

Here, n_{DexAem} ("Donor excitation, Acceptor emission") is the number of photons collected in the acceptor channel as a result of the FRET interaction, n_{DexDem} ("Donor excitation, Donor emission") are those collected in the Donor channel and n_{AexAem} ("Acceptor excitation, Acceptor emission") describes the number of photons collected in the acceptor channel after direct excitation of the acceptor dye.

$$E = \frac{n_{DexAem}}{n_{DexDem} + n_{DexAem}} \qquad \textit{Equation 4}$$

Similarly, when calculating the FRET efficiency n_{DexDem} and n_{DexAem} represent photons collected by the donor and acceptor channel respectively [16]. Please note that the naming convention has been changed from the paper to fit consistently with the naming convention used in this thesis.

1.6 FRETbursts and mpH^{2}MM

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 [12]. 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. A "burst" therefor describes an ensemble of photons which are detected by single-photon avalanche diode (SPAD) detectors in the experimental setup. On a technical level, bursts are periods in the photon stream in which the rate of photons detected is above a minimum threshold rate. This threshold rate can be manually adjusted by either specifying the number of consecutive photons *m*, or by accounting for the background rate and choosing a rate threshold that is *F* times larger than the background rate [12] [8]. An extension to this program is the photon-by-photon Hidden Markov Modeling (H^2MM) analysis tool which employs a maximum likelihood estimation (MLE) algorithm [9]. This tool allows the quantification of FRET dynamics in single biomolecules on a sub-millisecond timescale. Using the multi-parameter H^2MM (mpH^2MM) variant introduces a way to differentiate between conformational changes and photophysical transitions within FRET dynamics by incorporating additional photon streams (DexDem, DexAem, AexAem streams) as multiple parameters [7]. This approach is limited by the timescale of microseconds (in this paper referred to as "approach 1" or "classical mpH^2MM approach"). Recently, a new method was published, which takes the fluorescence lifetime as an extra dimension for the 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 are usually not enough photons in a microsecond to fit fluorescence lifetime using the conventional approaches (MLE via mpH^2MM), the low photon arrival probability leads to poor optimizations, meaning that the dataset used in regular mpH^2MM performs poorly for analysis including the fluorescent lifetimes. Thus, Authors employed this "divisor approach" to estimate fluorescence lifetime on this fast timescale [8]. A divisor in this context is a threshold of nanosecond time bins, where time bins before and after the divisor are treated as having separate parameter value respectively. One divisor allows for two parameter values: before and after the divisor. A number of *n* divisors allow for *n+1* parameter values.

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, the classic approach, classifies fluorescent states based on the FRET efficiency and Stoichiometry and the second method, the divisor approach, 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 1 nM 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 at 20 minutes each.

The data acquisition was performed on a confocal microscope with a custom-built laser setup [17]. Excitation of each dye sample was done by a 640 nm diode laser at average power of 70 µW. The laser beams reflected from a dichroic mirror into a water-immersion objective. 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. The emitted photons were collected by the same objective before passing through the dichroic mirror. These photons were then focused through a 100 µm pinhole. After passing a filter, which removed residual light not of interest, the emission photons were collected by the detector.

Triplet states were identified as distortions of the FCS curve on a microsecond timescale [15]. After testing of all 3 dyes for their photostability, Atto 643 was chosen as the best available acceptor dye in the following FRET experiment.

3.2 FRET sample and measurement

The SecYEG protein sample preparation and all FRET measurements were performed as described by Crossley et al. [17]. The Sample consist of SecYEG-PrlA4 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 consisting of *E. coli* polar extract with diameter of 100 nm [17].

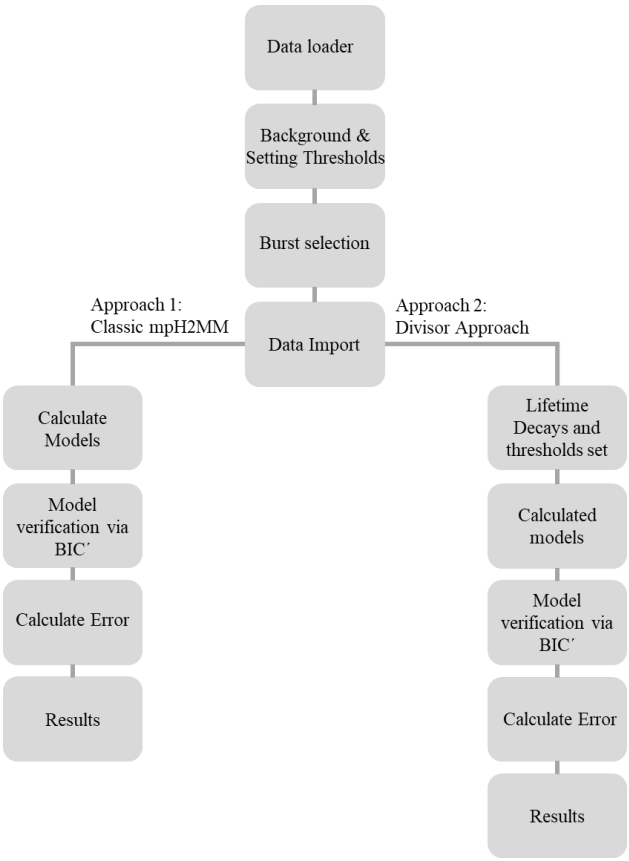
SmFRET data was acquired on a custom-built confocal microscope with a pulsed interleaved excitation (PIE) regime [17]. Samples were measured in a bovine serum albumin (BSA) passivated 8-well sample chamber at a concentration of 30 pM within a TKM buffer solution. 1 mM aged Trolox was added to reduce photoblinking and photobleaching [17] [18].

Excitation of the labeled protein complex was done by 560 nm (donor exciting laser) and 640 nm (acceptor exciting laser) diode lasers which were pulsed in an alternating fashion at a frequency of combined 40 MHz. The average power of the donor laser was 75 µW and 35 µW for the acceptor laser. The lasers were first combined by a dichroic mirror and afterwards coupled to an optical fiber. The laser beams were collimated and afterwards reflected from a polychroic beam splitter into a water-immersion objective. The focus was set 10 µm above the surface of the coverslip ensuring that no surface effects would interfere with the measurement.

Collection of the light emitted by the sample was done by the same water-immersion objective mentioned above. Passing through the polychroic beam splitter, the light was focused onto a 100 µm pinhole by the use of an achromatic doublet lens. Another achromatic lens after the pinhole collimated the collected light before it was separated by wavelength into the donor and acceptor emission channels using a dichroic mirror. Filters were used to remove residual excitation wavelength light in each channel and the filtered emitted light was focused onto single photon avalanche diode detectors. [17]

3.3 Analysis via FRETbursts and H^2MM

Analysis was performed by the FRETbrusts and H^2MM python library packages [7] [9] [12] 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.



Schematic 2: Overview of the process done by the program. First, the collected data was loaded into the software. The dataset was then corrected for background and parameters for burst detection were defined. Selected bursts underwent data analysis using the mpH2MM approach. In the step "calculate Models", Approach 1 allows direct analysis of the corrected dataset via the in HMM deployed MLE algorithm while Approach 2 requires the user to define the number of divisors and the calculation of fluorescent lifetimes before the HMM algorithm can calculate models of possible states observed. In both cases, these models are then verified and the model best describing the number of FRET states is selected. Lastly before the results, errors in the established probabilities are calculated. Please note that the 2 approaches are shown here to run in parallel. They are represented in this way to establish a comparable workflow between the two approaches which also allows for direct comparison of the results of each approach in Section 4.2: FRETbursts and mpH^{2}MM Results. The code, however, first cycles through the entirety of the Classic mpH2MM approach before repeating the process for the divisor approach.

Burts-search was implemented with a minimum threshold of burst intensity to be 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 mpH^{2}MM, 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, see Figure 7) with precise information of each states position, its error and population

size, as well as transition rates. The Results are summarized in Section 4 in Figure 7 and Tables 2 and 4. Errors were defined as standard deviations calculated via the bootstrap method. This method aims to estimate the distribution of the entire population (including the variance) by defining sample distributions of size n, where N values of a given sample distribution are randomly selected and analyzed (including multiple reselections of the same value – resampling with replacement). To compute variance, this process is repeated M times, usually in scope of hundreds to thousands of repetitions, forming a robust and reliable estimation of the standard deviation [19].

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 (Tables 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. Table 1 summarizes the necessary parameters. The parameters were calculated via the fit. T_{blink} gives information on the lifetime of the triplet state. Comparing the three dyes, a significantly longer/higher blinking rate is observed in Cy5 (τ_{blink} = 12 µs) than in both Atto dyes (τ_{blink} = 0.5 µs). This difference can also be seen in the graphical comparison (Figure 1). While for the Atto dyes, this effect is most likely the result of the triplet state, in Cy5 this effect may be caused by photo induced isomerization [20]. Furthermore, the triplet state amplitude (T) directly measures the propensity of the system to undergo transitions into a triplet or photoisomerized 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 a significant portion of time the photoisomerized state (T =0.46). This property could heavily bias the determination of conformational dwell times/ interconversion rates.

Table 1: FCS parameters. Cps are the photon counts per secont, T is the Triplet state amplitude and τ_{blink} are lifetimes of the triplet states/isomers.

Parameter	**Atto643**		**Atto647N**		**Cy5**	
	Value	Error	Value	Error	Value	Error
Cps [cnts/s]	18773	------	10969	------	5124	------
T []	**0.248**	**0.045**	**0.235**	**0.071**	**0.46**	**0.024**
τ blink[µs]	**0.5**	**------**	**0.5**	**------**	**12**	**------**

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) as seen by the portions of fast decay rates at the microsecond region. Figure 1C shows a larger region of fast decay at the 100-microsecond range, indicating that photoisomerization is the more common mechanism than triplet state. Furthermore, Atto 643 shows an overall higher correlation than Atto647N and Cy5 which seem to be similar.

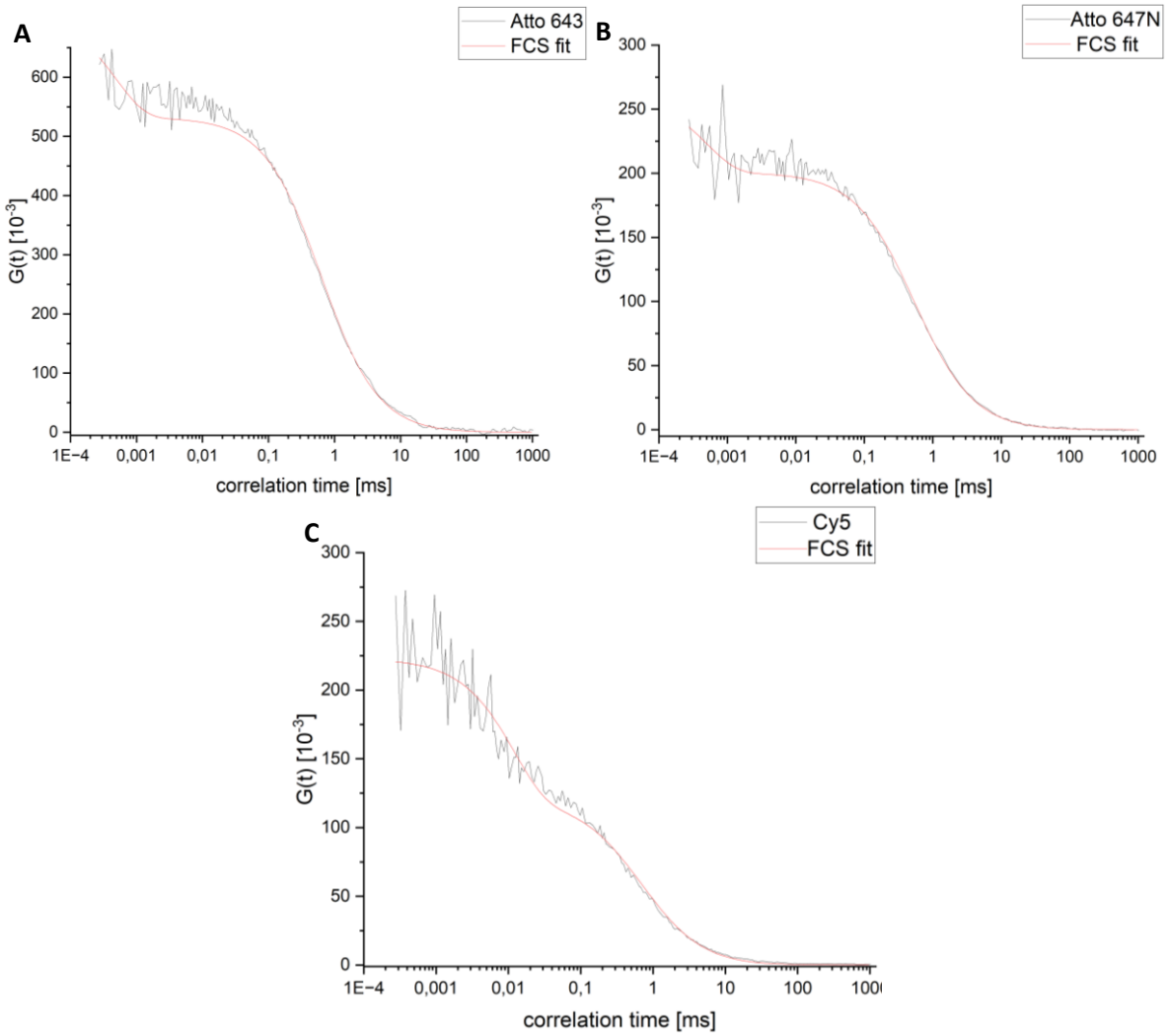


Figure 1: FCS results for all three samples. On the x-axis the correlation time is shown on a logarithmic scale. The y-axis shows G(t), which is the autocorrelation function defined as the correlation of a signal at time t with itself at various lagtimes τ. (A) shows the results for Atto 643, (B) for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

Since both Atto dyes display higher photostability than Cy5, a decision had to be made between them. Ultimately, the choice for either was not only influenced by their photostability but also other factors, such as binding specificity. Atto 643 was found to be the best suitable acceptor dye for the FRET experiments performed afterwards.

4.2 FRETbursts and H^{2}MM 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.

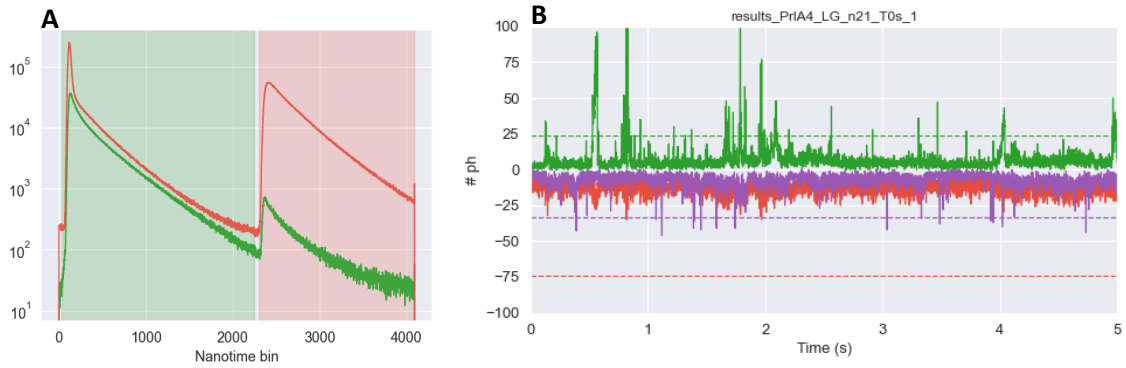


Figure 2: (A) Alternation histogram for nanosecond - Alternating Laser Excitation Time-Correlated Single Photon Counting (ns-ALEX TCSPC) measurements. Photons detected by the donor channel are green, those detected by the acceptor channel are in red. The shaded areas define the donor and acceptor periods. Detector periods were set manually to fit the dataset. (B) 5 second interval of the complete timetrace which functions as visual quality control of the measurement. 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. FRET interaction can be expected.

The next step, as established by Schematic 2 is the review of background noise and, following, the setting of thresholds for burst selection. The estimated background rates are shown in Figure 3. Figure 3A shows an exponential fit (lines) on a set of datapoints representing individual timestamps for all observed photon streams. Figure 3B shows background rates of individual photon streams and the total (black) as a function of time. Both figures serve as a quality control, estimating the influence of background noise on the overall quality of the measurement.

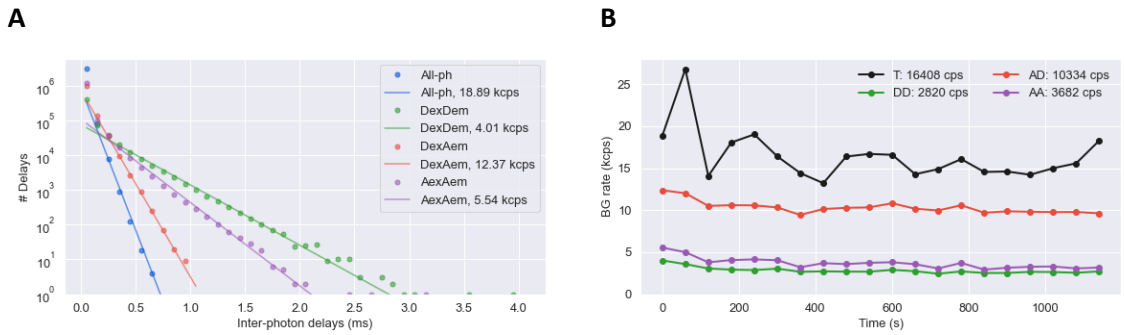


Figure 3: (A) Background rate histogram. Inter-photon delays fitted with an exponential function for different photon streams. Experimental distribution of inter-photon delays (dots) and their corresponding fits (lines). Shown are 4 fits: the sum of all photons is in blue, donor excitation with donor 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.

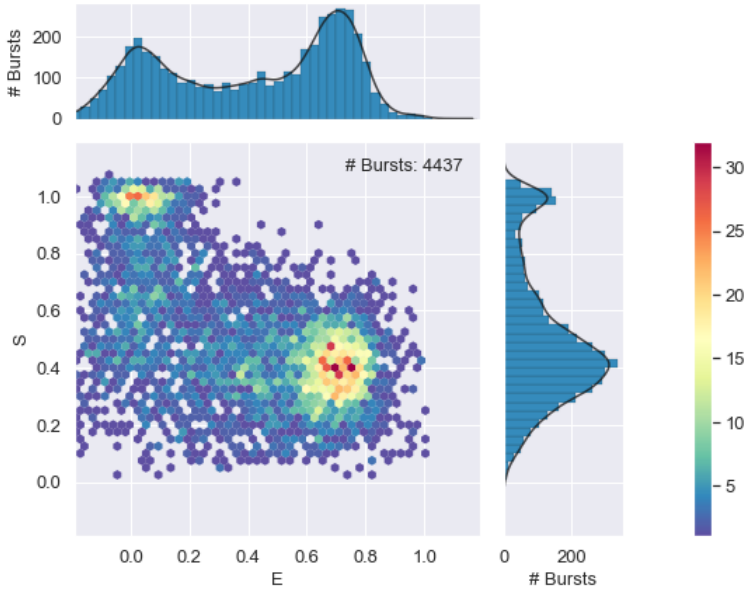


Figure 4: E-S Histogram showing FRET interaction between the donor and acceptor. The colorcoded scale (blue to red) gives information on the amount of bursts at any given position in the 2D histogram as indicated on the right. In the upper left (S > 0.8, E < 0.2) is donor only; down and to the right is acceptor only (S < 0.2). In the middle (E = 0.7 - 0.8; 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 mpH^{2}MM methods were applied to the FRET data, as indicated in Schematic 2. 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 for the divisor approach. The derived lifetimes can be used as corroboration that the low FRET state should have higher donor fluorescence lifetime than the 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.

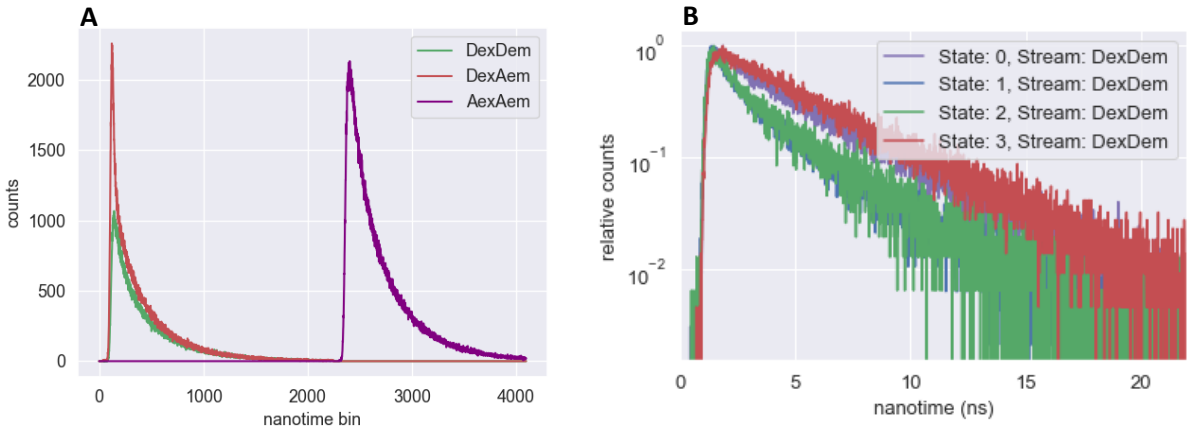


Figure 5: (A) Alternation modulo. The information gained from this graph was used to set the IRF thresholds for each of the three channels. In principle, this graph is a derivation of Figure 2A. (B) shows the lifetime decay for each state in the DexDem photon stream. It is a species verification of states by matching the lifetime decays to the states. Faster decays indicate a de-excitation mechanism, be it a FRET interaction or photophysical effects. State 0 (purple) describes the donor only state. State 1(blue, behind State 2 in green) describes the donor photons experiencing FRET interaction, hence loosing energy, which is displayed as a faster decay.

For each mpH^{2}MM 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 that would reasonably explain the full dataset. Generally, a model which falls below the threshold BIC' < 0.005 is considered reliable. Here, however, only models with 6 states or more fulfill this requirement. The BIC' is considered unreliable in this case and the model best describing the number of states must be determined based on reasonableness. After testing models accounting for 3, 4 or 5 states, it was concluded that the dataset was best described with 4 states (an analysis of models with more or less than 4 states can be found in section 4.3). 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.

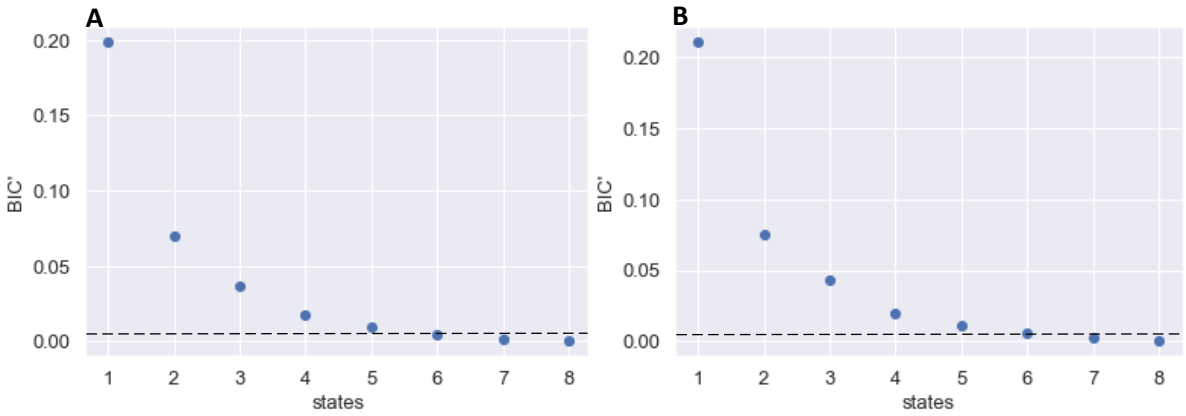


Figure 6: BIC' for 8 modeling attempts. (A) shows the result of the classic mpH2MM approach and (B) for the divisor approach. Valid models should satisfy the condition BIC' < 0.005 (indicated by the black dashed line). Here, models describing 6 states or more fulfill this condition. Nevertheless, the 4-state model was chosen as the most trustworthy one in both cases even though it does not satisfy the 0.005 threshold. 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, hence, losing all statistical meaning. More will be discussed in Section 4.3.

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 mpH^{2}MM 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.

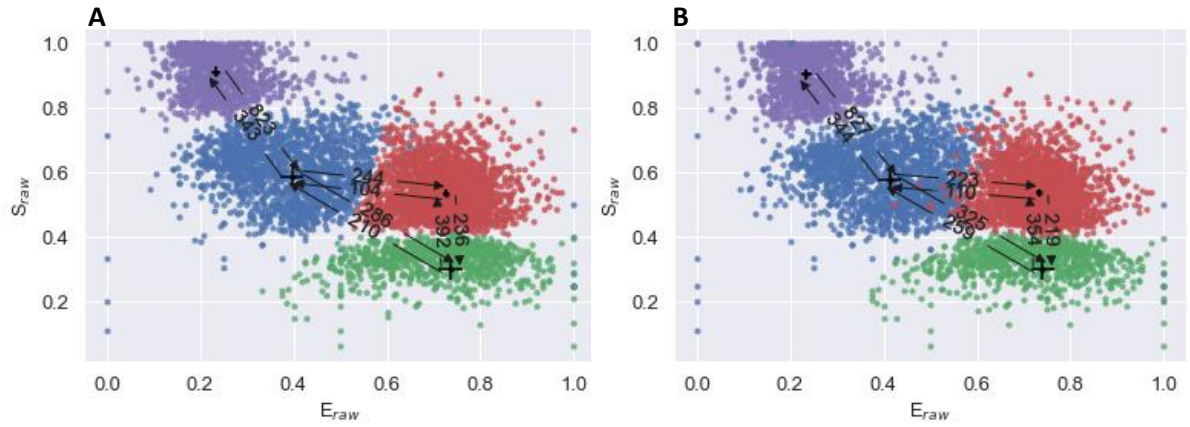


Figure 7: Comparison of the results from the classical H2MM and divisor based H2MM. ES scatter plots are derived from the classic approach (A) and from the divisor approach (B). Consecutive photons with the same state are considered as a single dwell. Donor only population is in purple, acceptor only 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. This result suggests 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 standard deviations obtained by the bootstrap method. 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 mpH²MM 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

Table 4 summarizes the transition rates which are already graphically indicated in Figure 7 (arrows). The errors here were also calculated via the bootstrap method. It can be seen that the transition rates derived by both approaches are similar. The divisor approach yielded 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.

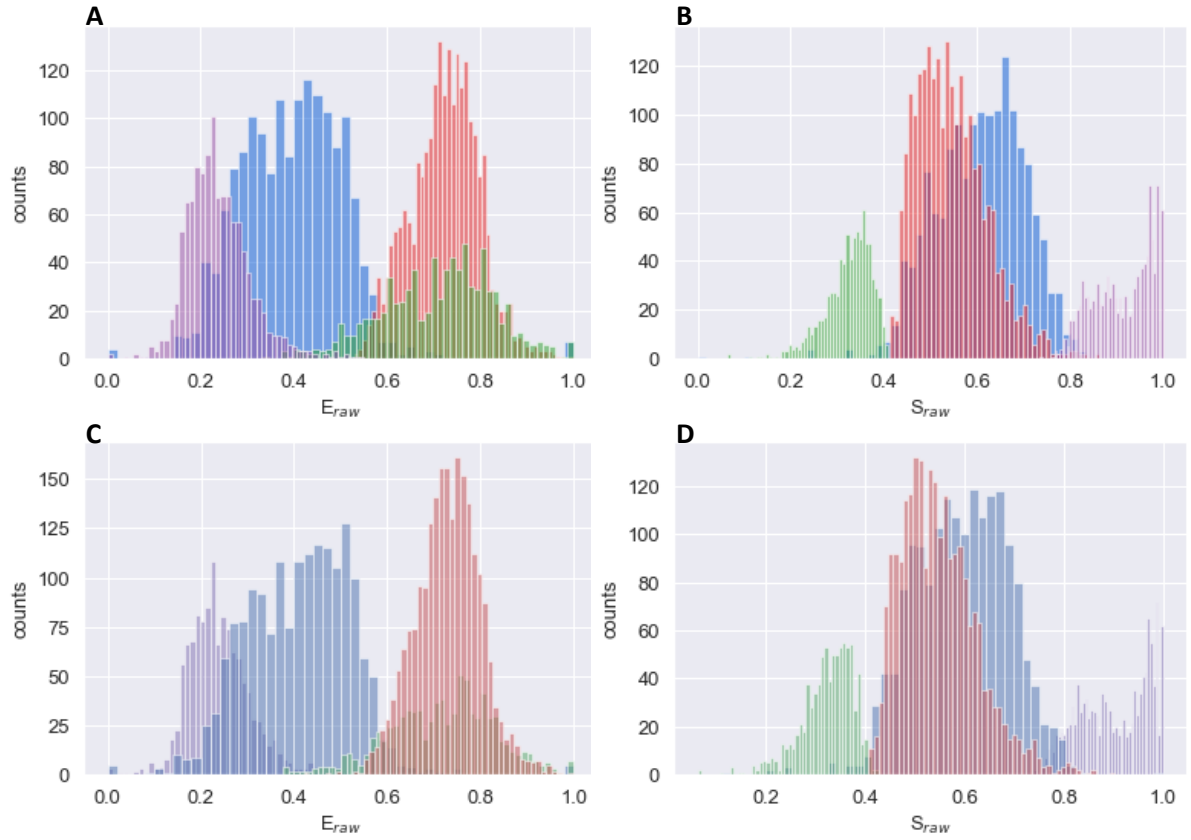


Figure 8: Population histogram for the classic approach (A and B) and for the divisor approach (C and D). The y-axis shows the photon count. The x-axis shows the FRET efficiency (A, C) and the photon Stoichiometry (B, D). The E histogram and S histogram correspond to Figure 7 with the color coding being 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.

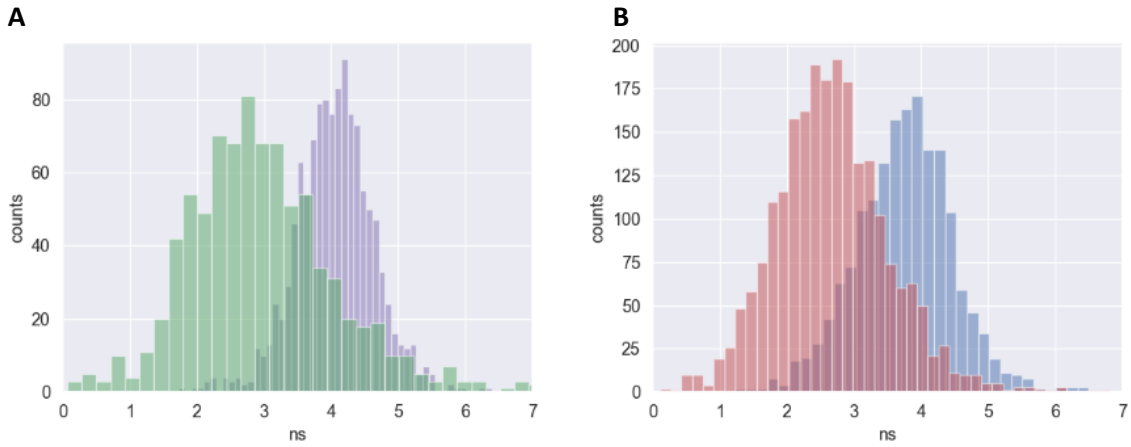


Figure 9: Fluorescence lifetimes derived by the divisor approach. The y-axis shows counts, the x-axis shows the fluorescent lifetimes in nanoseconds. (A) shows only the donor (purple) and acceptor (green) state, while (B) shows both FRET states (low FRET in blue; high FRET in red).

The fluorescent lifetimes allow the calculation of FRET efficiencies via the *Equation 5* [FRET-based stoichiometry in living cells]

$$E = 1 - \frac{\tau_{DA}}{\tau_D} \qquad \textit{Equation 5}$$

where τ_D is the mean fluorescence lifetime of the Donor and τ_{DA} is the mean fluorescence lifetime of the donor in presence of the acceptor (hence FRET interaction).

With τ_D = 4.15 ns and τ_{DA} = 2.41, E = 0.412, which broadley corresponds to the E value for the high FRET state as derived from Figure 7 (E = 0.413). The low FRET state is broadly resembled by an E value of E = 0.16.

4.3 State models of different size

Since no clear threshold in the BIC' can be identified, which would allow a simple estimation of the number of states purely on that, other models wer analyzed as well. Here we compared 3, 4 and five states, showing the strength of lifetimes as an extra dimension in the analysis, while they have a physical meaning and they could guide us to identify whether the algorithm correctly classified and grouped dwells and states. Following will be an analysis according to the divisor approach where the fitted models account for 3 and 5 states.

4.3.1 The 3 State model

After the application of the BIC', as described in Figure 6, a new ES scatter plot was calculated, similarly to Figure 7. Figure 10 below shows this resulting ES scatter plot according to a 3 state model. Here, the algorithm split the data in three states more or less

resembling donor only, very broad FRET distribution and acceptor only (Fig 10a). We don't think this classification is plausible, mostly because the lifetime analysis of the acceptor only population closely resembles the FRET distribution (compare Figure 10B to 10D), indicating that they are not well separated.

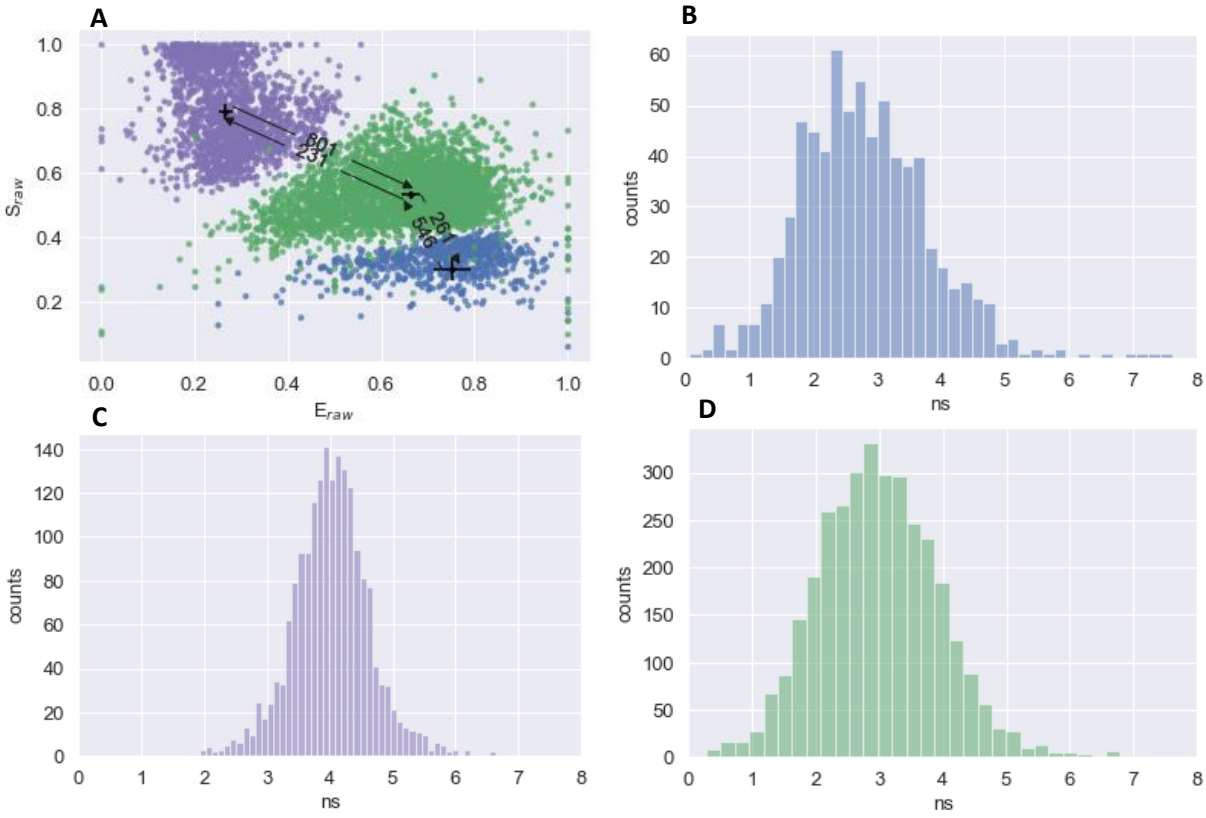


Figure 10: ES scatter plot according to a 3-state model including transition rates between the states. Donor only population is in purple, acceptor only in blue. The FRET state is indicated as the green population. Lifetimes for the DexDem stream are given in the histograms (B-D). (B)shows the acceptor state, (C) the Donor State and (D) the FRET state.

In general, comparing to the 4-state model in Figure 7, it appears that significant amounts of photons previously described in the low FRET state have now been assigned with the Donor only state, portions of the two FRET states seem to have merged together and another portion of the high FRET state appears to have merged with the acceptor only state.

4.3.2 The 5 State Model

Similarly to the 3-state model, the resulting model didn't work to our satisfaction. Figure 11 shows the results.

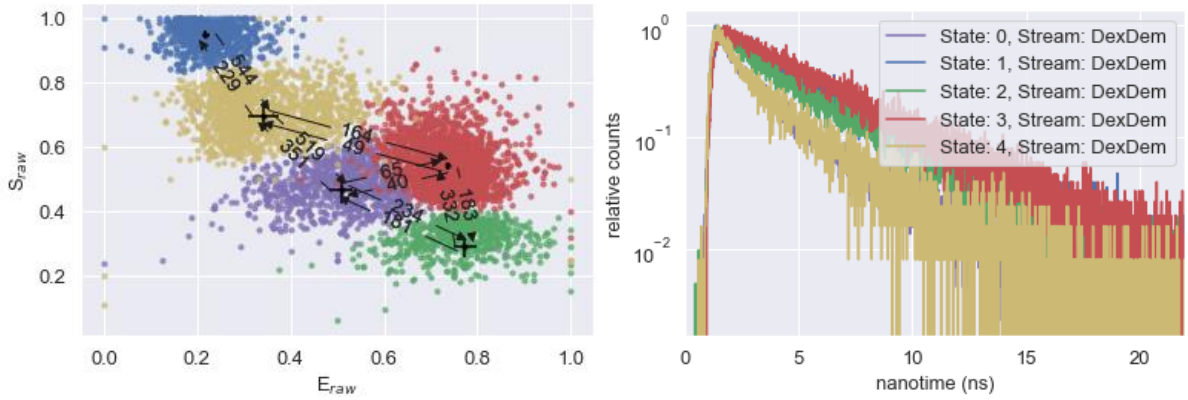


Figure 11: ES scatter plot, including transition rates, of the 5-state model. Blue and green populations describe the donor and acceptor only populations respectively. Yellow, purple and red populations describe FRET or Pseudo-FRET states, which will be denoted as FRET1 (red), FRET2 (purple) and FRET3 (yellow).

Here, it appears that a new state has been formed from bursts previously (in the 4 state model) affiliated with the donor only and the low FRET state (yellow in Figure 12). This state can be described as a pseudo-state, accounting for bursts capturing the transitions between the donor only state and the FRET states but are misinterpreted because of photoblinking of the acceptor, leading to the formation of this dynamic state. the yellow state describes a “bridge“, similar to what is shown in the paper of Kong et al. [21] in Figure 2 (as a simluation) and Figure 4 (based on experimental data).

.

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 an 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. Dyes containing NHS-esters are generally known to react easily with amino groups of proteins and other biomolecules and are therefore unspecific in their interactions. However, Atto 643, also containing a NHS-ester group, shows a lower probability of unspecific binding (compared to Atto 647N) due to its high hydrophilicity, making it the best suitable choice [22] [23]. Cy5 seems to spend significantly longer times in, what was first assumed, a triplet state. Comparing these results to Windengren and Schwille [20] suggests, however, that the triplet state formation is relatively inefficient. Instead, it seems more common for a sterically non-hindered Cy5 molecule to undergo a photo-induced isomerization from the (fluorescing) trans conformation into (significantly less fluorescing) cis conformation. Nevertheless, due to the higher probability of Cy5 to lose its fluorescent property, it is considered a less suitable choice for our purposes.

In the FCS experiments, Atto 643 showed its low tendency to form Triplet states. Nevertheless, the H^2MM 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 [13] [18], was present during the measurement. In the results of this Thesis, the as such identified triplet states occur as fast decay rates in the microsecond timescale (Figure 1). The triplet state is indicated by distortions of the FCS curve on a timescale microseconds, as described by Widengren, Mets and Rigler [15], supporting this interpretation. Despite the addition of aged Trolox, photoblinking and -bleaching events still occurred and had influence on the quality of the smFRET data. 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 and antibleaching reagents in the future.

In the comparison of the H^2MM models, one crucial step was the determination of the number of states. The BIC', implemented to guide the selection of a model best describing the data, contained only little information of value as models with more than four states - even though they achieve better statistical fits in the BIC' - encounter the risk of being overfitted in an

attempt to achieve the least statistical error. The algorithm starts splitting states with artificially sharp separating borders, which indicates that the division is no longer physiological and statistically meaningful. In an attempt to gain more meaningful information, another selection criterion, the integrated complete likelihood (ICL), was first applied on the dataset. The ICL provides an extremum-based decision on the most likely state model by showing a minimum value for the most likely model [7]. This, however, also did not provide information of value as no minimum was shown for models with reasonable amounts of states (as shown in Figure A1 in the appendix). The source of this error for both, the BIC' and ICL, remains unknown. Hence, a decision could only be made based on reasonableness and plausibility of the state models as shown in Section 4.3. It was concluded that 4 states are present. Comparing this decision to the results of Crossely et al [17], which share the experimental setup and the analyzed protein complex in question, SecYEG, reinforces this conclusion as they also present the 4-state model. Furthermore, one would intuitively also expect a 4-state system, as the primary function of the SecYEG complex is the translocation of proteins via opening and closing of its transport channel. A donor only and acceptor only population is expected as the labeling process may label individual protein complexes incompletely, eg. only one of the dyes is bound to the complex. Another reason for the existence of Donor only/acceptor only states is the loss of fluorescence by one of the two, leading to a transition from a FRET state into the "only" state. The FRET states, low FRET and high FRET, describe the SecYEG translocon in its open and closed state respectively. A visual comparison of Figure 4 to the simulations and results of Kong et al. [21] already indicates some amount of photobleaching of the acceptor. It is plausible that models of 5 or more states identify bleaching events as separate states.

The comparison of both H^2MM methods yielded overall only minor deviations. A clearly "more accurate" method cannot be decided. Harris et al. [8], [7] already demonstrate the reliability of their analyzing tools. It can be concluded that both methods are reliable and viable options. However, the divisor approach offers additional insight on the lifetime of FRET states, which will have other relevant applications, such as the determination of kinetics faster than tens of microseconds.

6. Literature and References

[1] Lerner, E., Cordes, T., Ingargiola, A., Alhadid, Y., Chung, S., Michalet, X., Weiss, S. (2018). Toward dynamic structural biology: Two decades of single-molecule Förster resonance energy transfer. *Science (New York, N.Y.)*, *359*(6373), eaan1133. https://doi.org/10.1126/science.aan1133

[2] Selvin, P.R., Ha, T. (2008). Single-molecule techniques: a laboratory manual.*Cold Spring Harbor Laboratory Press*

[3] Flechsig, H., Mikhailov, A. S. (2019). Simple mechanics of protein machines. *Journal of the Royal Society Interface*, *16*(155), 20190244. https://doi.org/10.1098/rsif.2019.0244

[4] Crossley, J. A., Allen, W. J., Watkins, D. W., Sabir, T., Radford, S. E., Tuma, R., Collinson, I., Fessl, T. (2024). Dynamic coupling of fast channel gating with slow ATP-turnover underpins protein transport through the Sec translocon. *The EMBO journal*, *43*(1), 1–13. https://doi.org/10.1038/s44318-023-00004-1

[5] van der Wolk, J. P., Fekkes, P., Boorsma, A., Huie, J. L., Silhavy, T. J., Driessen, A. J. (1998). PrlA4 prevents the rejection of signal sequence defective preproteins by stabilizing the SecA-SecY interaction during the initiation of translocation. *The EMBO journal*, *17*(13), 3631–3639. https://doi.org/10.1093/emboj/17.13.3631

[6] Adams, H., Scotti, P. A., Luirink, J., Tommassen, J. (2002). Defective translocation of a signal sequence mutant in a prlA4 suppressor strain of Escherichia coli. *European Journal of Biochemistry*, *269*(22), 5572-5580. https://doi.org/10.1046/j.1432-1033.2002.03263.x

[7] Harris, P. D., Narducci, A., Gebhardt, C., Cordes, T., Weiss, S., Lerner, E. (2022). Multi-parameter photon-by-photon hidden Markov modeling. *Nature communications*, *13*(1), 1000. https://doi.org/10.1038/s41467-022-28632-x

[8] *Harris, P.D., Lerner, E. (2022). Identification and quantification of within-burst dynamics in singly labeled single-molecule fluorescence lifetime experiments. Biophysical Reports, 2. DOI: 10.1016/j.bpr.2022.100071*

[9] *Pirchi, M., Tsukanov, R., Khamis, R., Tomov, T.E., Berger, Y., Khara, D.C., Volkov, H., Haran, G., Nir, E. (2016). Photon-by-Photon Hidden Markov Model Analysis for Microsecond Single-Molecule FRET Kinetics. The Journal of Physical Chemistry. B, 120 51, 13065-13075 . DOI:10.1021/acs.jpcb.6b10726*

[10] *Sahoo, H. (2011). Förster resonance energy transfer - A spectroscopic nanoruler: Principle and applications. Journal of Photochemistry and Photobiology C-photochemistry Reviews, 12, 20-30. DOI: 10.1016/J.JPHOTOCHEMREV.2011.05.001*

[11] Robert M. Clegg.(2009). 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*. Volume 33. pp. 1-57. https://doi.org/10.1016/S0075-7535(08)00001-6

[12] Ingargiola A., Lerner E., Chung S., Weiss S., Michalet X.(2016).FRETBursts: An Open Source Toolkit for Analysis of Freely-Diffusing Single-Molecule FRET. *bioRXiv*. https://doi.org/10.1371/journal.pone.0160716

[13] Demchenko A. P. (2020). Photobleaching of organic fluorophores: quantitative characterization, mechanisms, protection. *Methods and Applications in Fluorescence*, *8*(2), 022001. https://doi.org/10.1088/2050-6120/ab7365

[14] Machán, R., Hof, M. Fluorescence Correlation Spectroscopy (FCS). PicoQuant GmbH.Available under: https://www.picoquant.com/images/uploads/page/files/17319/5_fcs.pdf [Accessed 14.12.2023]

[15] Widengren, J., Mets, Ü., Rigler, R. (1995). Fluorescence Correlation Spectroscopy of Triplet States in Solution: A Theoretical and Experimental Study. *The Journal of Physical Chemistry.* 99. 13368–13379. DOI: 10.1021/j100036a009

[16] *Ruettinger, S., Buschmann, V., Kraemer, B., Orthaus, S. (2013). FRET analysis with Pulsed Interleaved Excitation (PIE) using the MicroTime 200.PicoQuant GmbH.Berlin,Germany. Available under:* https://www.picoquant.com/?ACT=35&fid=125&d=7266&f=appnote_pie-fret.pdf

[17] *Crossley, J.A., Allen, W.J., Watkins, D.W., Sabir, T., Radford, S., Tuma, R., Collinson, I., Fessl, T. (2023). Energy landscape steering mediates dynamic coupling in ATP-driven protein translocation by the bacterial Sec machinery. bioRxiv. DOI: 10.1101/793943*

[18] Altman, R. B., Terry, D. S., Zhou, Z., Zheng, Q., Geggier, P., Kolster, R. A., Zhao, Y., Javitch, J. A., Warren, J. D., Blanchard, S. C. (2011). Cyanine fluorophore derivatives with enhanced photostability. *Nature methods*, *9*(1), 68–71. https://doi.org/10.1038/nmeth.1774

[19] König, S. L., Hadzic, M., Fiorini, E., Börner, R., Kowerko, D., Blanckenhorn, W. U., Sigel, R. K. (2013). BOBA FRET: Bootstrap-Based Analysis of Single-Molecule FRET Data. *PLoS ONE*, *8*(12). https://doi.org/10.1371/journal.pone.0084157

[20] Widengren, J., Schwille, P. (2000). Characterization of Photoinduced Isomerization and Back-Isomerization of the Cyanine Dye Cy5 by Fluorescence Correlation Spectroscopy. *Journal of Physical Chemistry A* 104. DOI: 10.1021/jp000059s.

[21] ATTO-Tec GmBH.(2024). Product Information: ATTO 647N.Retrieved from: https://www.atto-tec.com/ATTO-647N.html?language=en [Accessed: 08.05.2024]

[22] ATTO-Tec GmBH.(2024). Product Information: ATTO 643.Retrieved from: https://www.atto-tec.com/ATTO-643.html?language=en [Accessed: 08.05.2024]

[22] ATTO-Tec GmBH.(2024). Product Information: ATTO 565.Retrieved from: https://www.atto-tec.com/ATTO-565.html?language=en [Accessed: 08.05.2024]

[23] Kong, X., Nir, E., Hamadani, K., Weiss, S. (2007). Photobleaching pathways in single-molecule FRET experiments. *Journal of the American Chemical Society*, *129*(15), 4643–4654. https://doi.org/10.1021/ja068002s

7 Appendix

7.1 Supplemental material

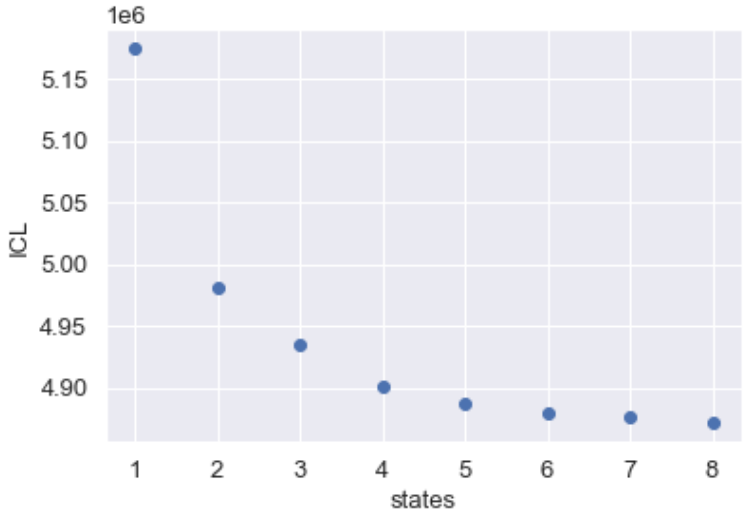


Figure A1: ICL plot attempting to guide an extremum-based decision on the number of states. It can be seen that no minimum is shown, hence, this ICL plot holds no information of value to the decision-making process.

7.2 User code

```
1.  #import of all necessary modules
2.  import numpy as np
3.  from matplotlib import pyplot as plt
4.  import fretbursts as frb
5.  import burstH2MM as hmm
6.  sns = frb.init_notebook()
7.  # load the data into the data object frbdata
8.  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

trans_std, E_std, S_std=bdata.models[3].bootstrap_eval()

print("E-array", "\n", Earray)
print("S-array", "\n", Sarray)
print("Transition rate-array", "\n", Transarray)

print("E st deviation", "\n", E_std)
print("S st deviation", "\n", S_std)
```

```
print("trans st deviation", "\n", trans_std)

# plot the dwell ES of the result
state_color = [{'color':'b'}, {'color':'r'}, {'color':'g'}, {'color':'m'}]
hmm.dwell_ES_scatter(bdata.models[3], state_kwargs=state_color )
# overlay with the main values,
hmm.scatter_ES(bdata.models[3], s=10, c="k", )
plt.errorbar(Earray, Sarray, xerr=E_std, yerr=S_std, ecolor='black', fmt='none')
hmm.trans_arrow_ES(bdata.models[3]);


plt.savefig('ES_scatter')
plt.savefig('ES_scatter.pdf')
hmm.dwell_E_hist(bdata.models[3], bins= 50, )
plt.savefig('E_hist')
plt.savefig('E_hist.pdf')
hmm.dwell_S_hist(bdata.models[3], bins= 50)
plt.savefig('dwell_S_hist.png')
plt.savefig('dwell_S_hist.pdf')
#Divisor approach


hmm.state_nanotime_hist(bdata.models[3], normalize=True)
plt.xlim([0,22])
plt.savefig('Nanotime_hist')
plt.savefig('Nanotime_hist.pdf')
hmm.raw_nanotime_hist(bdata)
plt.savefig('Nanotime_bins.png')
plt.savefig('Nanotime_bins.pdf')
#creates a file titled “output” in which the values corresponding
#to nanotime_hist in order: x-axis, DexDem, DexAem, AexAem
streams = bdata.ph_streams
stream_id = [np.argwhere([stream == psel for psel in bdata.ph_streams])[0,0] for stream in
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')
```