

Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

**Investigating the Precision of Dynamic DNA Origami Platforms with
Multi-Color Single-Molecule FRET Methods and using them to
Benchmark Deep-Neural Networks Analysis Approaches**



Pooyeh Asadiatouei

aus

Savadkoh, Iran

2024

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Don C. Lamb betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 08.10.2024

Pooyeh Asadiatouei

Dissertation eingereicht am: 09.10.2024

1. Gutachter Prof. Don C. Lamb, Ph.D.

2. Gutachter Prof. Philip Tinnefeld, Ph.D.

Mündliche Prüfung am: 28.10.2024

Abstract

Many fundamental biological processes and interactions can be successfully explored by applying Förster resonance energy transfer (FRET) techniques which have become crucial tools in molecular biology and biophysics. By strategically incorporating suitable dye pairs into different positions of biomolecules, we can resolve both inter- and intra-molecular distances, as well as dynamic interactions with sub-nanometer resolution. Such level of detail is essential for understanding the molecular machinery of life, as it allows us to observe interactions and conformational changes that are often invisible when using traditional methods. FRET is particularly powerful due to its extreme sensitivity to the distance between the centers of molecules since it is related to the 6th power of the distance, so it is one of the most effective tools for detecting small changes. Unlike conventional ensemble measurements, which average out molecular behaviors across an ensemble, single-molecule FRET (smFRET) enables the detection of subtle variations such as conformational and functional heterogeneities within a sample. These underlying variations provide critical insights into the complex behaviors of biomolecular systems. The ability to observe such differences with the appropriate statistics, can uncover hidden states, transient interactions and rare events that play key roles in biological functions.

Traditionally, FRET experiments were focused on two-color systems, where a donor and acceptor dye pair report on a single distance. However, recent advancements have pushed the boundaries of FRET applications by introducing three-color FRET assays. Incorporating three labels into the system allows for the simultaneous measurement of multiple distances, creating a more comprehensive three-dimensional vision of the molecular interactions and structures. This multi-color approach not only helps track individual distances between each dye pair but also reveals correlations between them, offering a richer understanding of the molecular dynamics under study. Although three-color FRET assays can complicate both the experimental design and data analysis, these challenges have been addressed. Thanks to significant advancements in instrumentation and the development of sophisticated analysis software, the once intimidating task of analyzing three-color FRET data has become more manageable. The technological innovations in detection sensitivity, data processing and automation have transformed what was once a time-consuming and user-prone process into a streamlined, reliable workflow. As a result, researchers can now perform complex three-color FRET experiments with reduced analysis times, removing barriers that previously might have discouraged the widespread adoption of such techniques.

The combination of solution-based and surface-based smFRET assays offers a versatile approach for studying biomolecular systems. Solution-based assays are ideal for capturing the fast dynamic behavior of molecules in their native states and not bound to the surface, while surface-immobilization assays allow for long-term observations of molecular kinetics by fixing molecules on a chamber surface area. By employing both techniques, we can gain a comprehensive understanding of the sample, taking advantage of the different time scales and environmental conditions each method provides. This dual approach enables the examination of individual biomolecules behavior in ways that might be impossible using just one of the techniques.

A particularly exciting development in the field of smFRET is its combination with DNA origami nanotechnology. DNA origami structures allow for the precise design and construction of nanoscale samples, and have revolutionized many areas of single-molecule biophysics. When complemented with smFRET, DNA origami provides a versatile platform for positioning fluorophores with sub-nanometer precision, enabling extreme control over the spatial arrangement of dyes. This level of

control facilitates the study of a wide range of fundamental questions, from unraveling the structural features and distances, and molecular functions to probing the single fluorophores behavior in dynamic molecular changes and interactions.

In this thesis, both two- and three-color smFRET assays were employed to benchmark a powerful and versatile analysis tool based on deep learning. The primary goal was to significantly shorten the data analysis time and eliminate user bias during the analysis process, which is a common issue in manual data interpretation. The development of this tool has reduced the time required for analysis from weeks to minutes, revolutionizing how data is processed and visualized, and open a whole new world of possible experiments. Moreover, the experiments conducted on various L-shaped DNA origami nanostructures, which were integral to the development of these smFRET assays, uncovered a number of interesting characteristics related to the behavior of fluorophores and labeling strategies. These findings have prompted us to further investigate the precise control and manipulation of fluorophore positioning and transient binding kinetics of single-stranded DNA in the context of the DNA origami structure. A comprehensive set of experiments was executed to explore how accurately and consistently one can arrange and control the transient binding kinetics of DNA single strands. These studies revealed a number of factors that can interfere with the behavior and kinetics of the system, including the type of fluorophores used and the specific labeling or binding positions. By understanding and characterizing such factors, we can now predict and control the performance of FRET-based assays more reliably.

Zusammenfassung

Viele grundlegende biologische Prozesse und Interaktionen können erfolgreich durch die Anwendung von Förster-Resonanzenergietransfer (FRET)-Techniken erforscht werden, die zu wichtigen Werkzeugen in der Molekularbiologie und Biophysik geworden sind. Durch das strategische Einbringen geeigneter Farbstoffpaare an verschiedenen Positionen von Biomolekülen können sowohl inter- als auch intramolekulare Abstände sowie dynamische Interaktionen mit subnanometrigener Auflösung erfasst werden. Ein solches Detailniveau ist entscheidend für das Verständnis der molekularen Maschinerie des Lebens, da es uns ermöglicht, Interaktionen und konformationelle Veränderungen zu beobachten, die bei herkömmlichen Methoden oft unsichtbar bleiben. FRET ist besonders leistungsfähig aufgrund seiner extremen Empfindlichkeit gegenüber dem Abstand zwischen den Molekülzentren, da dieser im Zusammenhang mit der sechsten Potenz des Abstands steht, was es zu einem der effektivsten Werkzeuge zur Erkennung kleinster Veränderungen macht. Im Gegensatz zu herkömmlichen Ensemblemessungen, die das Verhalten von Molekülen über ein Ensemble hinweg mitteln, ermöglicht FRET auf Einzelmolekülebene (smFRET) die Erkennung subtiler Unterschiede wie konformationelle und funktionelle Heterogenitäten innerhalb einer Probe. Diese zugrundeliegenden Variationen bieten wichtige Einblicke in die komplexen Verhaltensweisen biomolekularer Systeme. Die Fähigkeit, solche Unterschiede mit den entsprechenden statistischen Methoden zu beobachten, kann verborgene Zustände, flüchtige Interaktionen und seltene Ereignisse aufdecken, die eine Schlüsselrolle bei biologischen Funktionen spielen.

Traditionell konzentrierten sich FRET-Experimente auf Zwei-Farben-Systeme, bei denen ein Donor-Akzeptor-Farbstoffpaar eine einzige Distanz misst. Jüngste Fortschritte haben jedoch die Grenzen der FRET-Anwendungen erweitert, indem Dreifarbf-FRET-Assays eingeführt wurden. Das Einfügen von drei Markierungen in das System ermöglicht die gleichzeitige Messung mehrerer Abstände, wodurch eine umfassendere dreidimensionale Sicht auf die molekularen Interaktionen und Strukturen entsteht. Dieser mehrfarbige Ansatz hilft nicht nur, einzelne Abstände zwischen den Farbstoffpaaren zu verfolgen, sondern zeigt auch Korrelationen zwischen ihnen auf, was zu einem tieferen Verständnis der untersuchten molekularen Dynamiken führt. Obwohl Dreifarbf-FRET-Assays das experimentelle Design und die Datenanalyse komplizieren können, wurden diese Herausforderungen bereits bewältigt. Dank bedeutender Fortschritte in der Instrumentierung und der Entwicklung ausgeklügelter Analysesoftware ist die einst einschüchternde Aufgabe, Dreifarbf-FRET-Daten zu analysieren, besser handhabbar geworden. Technologische Innovationen in der Detektionsempfindlichkeit, der Datenverarbeitung und der Automatisierung haben den vormals zeitaufwendigen und anfälligen Prozess in einen optimierten und zuverlässigen Workflow verwandelt. Infolgedessen können Forscher nun komplexe Dreifarbf-FRET-Experimente mit verkürzten Analysezeiten durchführen, wodurch Hindernisse beseitigt wurden, die früher die weitverbreitete Anwendung solcher Techniken behindern konnten.

Die Kombination von lösungsbasierten und oberflächenbasierten smFRET-Assays bietet einen vielseitigen Ansatz zur Untersuchung biomolekularer Systeme. Lösungsbasierte Assays sind ideal, um das schnelle dynamische Verhalten von Molekülen in ihrem natürlichen Zustand, also nicht an einer Oberfläche gebunden, zu erfassen, während oberflächenimmobilisierte Assays langfristige Beobachtungen der molekularen Kinetik ermöglichen, indem die Moleküle auf einer Kammeroberfläche fixiert werden. Durch den Einsatz beider Techniken können wir ein umfassendes Verständnis der Probe gewinnen, indem wir die unterschiedlichen Zeitskalen und Umweltbedingungen nutzen, die jede Methode bietet. Dieser duale Ansatz ermöglicht die

Untersuchung des Verhaltens einzelner Biomoleküle auf eine Weise, die mit nur einer der Techniken möglicherweise unmöglich wäre.

Eine besonders spannende Entwicklung im Bereich der smFRET ist die Kombination mit der DNA-Origami-Nanotechnologie. DNA-Origami-Strukturen ermöglichen das präzise Design und den Bau von nanoskaligen Proben und haben viele Bereiche der Einzelmolekül-Biophysik revolutioniert. In Kombination mit smFRET bietet DNA-Origami eine vielseitige Plattform zur Positionierung von Fluorophoren mit subnanometergenauer Präzision, was eine extreme Kontrolle über die räumliche Anordnung der Farbstoffe ermöglicht. Dieses Maß an Kontrolle erleichtert die Untersuchung einer Vielzahl grundlegender Fragen, von der Aufklärung struktureller Merkmale und Abständen, molekularer Funktionen bis hin zur Untersuchung des Verhaltens einzelner Fluorophore bei dynamischen molekularen Veränderungen und Interaktionen.

In dieser Dissertation wurden sowohl Zwei- als auch Dreifarbst-FRET-Assays eingesetzt, um ein leistungsstarkes und vielseitiges Analysetool auf Basis von Deep Learning zu testen. Das Hauptziel war es, die Datenanalysezeit erheblich zu verkürzen und Nutzerbias während des Analyseprozesses zu eliminieren, was bei der manuellen Dateninterpretation ein häufiges Problem darstellt. Die Entwicklung dieses Tools hat die für die Analyse benötigte Zeit von Wochen auf Minuten reduziert, revolutioniert die Datenverarbeitung und -visualisierung und eröffnet eine ganz neue Welt möglicher Experimente. Darüber hinaus brachten die Experimente an verschiedenen L-förmigen DNA-Origami-Nanostrukturen, die wesentlich zur Entwicklung dieser smFRET-Assays beigetragen haben, eine Reihe interessanter Eigenschaften im Zusammenhang mit dem Verhalten von Fluorophoren und Markierungsstrategien zutage. Diese Ergebnisse haben uns dazu veranlasst, die präzise Kontrolle und Manipulation der Positionierung von Fluorophoren und die transiente Bindungskinetik von einzelsträngiger DNA im Kontext der DNA-Origami-Struktur weiter zu untersuchen. Ein umfassender Satz von Experimenten wurde durchgeführt, um zu erforschen, wie genau und konsistent man die transiente Bindungskinetik von DNA-Einzelsträngen arrangieren und kontrollieren kann. Diese Studien haben eine Reihe von Faktoren aufgezeigt, die das Verhalten und die Kinetik des Systems beeinflussen können, einschließlich der verwendeten Fluorophore und der spezifischen Markierungs- oder Bindungspositionen. Durch das Verständnis und die Charakterisierung dieser Faktoren können wir nun die Leistung FRET-basierter Assays zuverlässiger vorhersagen und steuern.

List of Abbreviations

ALEX: Alternating Laser EXcitation

bp: Base Pair

BSA: Bovine Serum Albumin

DNA: Deoxyribonucleic Acid

dsDNA: Double-stranded DNA

EDTA: Ethylenediaminetetraacetic Acid

FRET: Förster Resonance Energy Transfer

msALEX: Millisecond Alternating Laser EXcitation

μsALEX: Microsecond Alternating Laser EXcitation

NA: Not Available

nsALEX: Nanosecond Alternating Laser EXcitation

oTIRF: Objective-type Total Internal Reflection Fluorescence

PBS: Phosphate Buffered Saline

pTIRF: Prism-type Total Internal Reflection Fluorescence

ssDNA: Single-stranded DNA

smFRET: Single-molecule Förster Resonance Energy Transfer

TCSPC: Time-correlated single photon counting

List of Figures

Figure 1. Schematic view of different single-molecule assays for FRET measurements.

Figure 2.1. A typical Jablonski diagram showing different energy levels and molecular processes occurring between absorption and emission of photons.

Figure 2.2. A schematic of fluorescence anisotropy measurement using separate detectors for parallel and perpendicular light orientations.

Figure 2.3. A typical Jablonski diagram for showing the phenomenon of Förster resonance energy transfer between a donor molecule in the excited state and an acceptor molecule in the ground state.

Figure 2.4. A schematic showing how orientation factor between a donor and acceptor is calculated.

Figure 2.5. Principle of the DNA origami structure assembly.

Figure 3.1. Schematic of the home-built confocal single-molecule three-color MFD-PIE setup used for the solution-based FRET measurements in this work.

Figure 3.2. Schematic of burst search methods for three-color experiments provided with the PAM software.

Figure 3.3. Schematic of three-color FRET.

Figure 3.4. Representative plots of different parameters obtained in a MFD-PIE experiment and analysis obtained from measuring version 1 of a two-color, two-state L-shaped DNA origami structures labeled with Cy3B on a tether and Atto647N on the origami structure.

Figure 3.5. Schematic showing the propagation of a light beam from a glass medium to water with corresponding change of angle from the normal as a result of different refractive indices at the interface between the two media.

Figure 3.6. The design and assembly of the prism used as both a medium to create the evanescent field and a chamber to hold the single molecules of interest after surface passivation.

Figure 3.7. Schematic smTIRF setup with five excitation lasers and four detection channels.

Figure 4.1. Benchmarking Deep-LASI software using L-shaped DNA origami structures.

Figure 4.2. Deep-LASI software overview.

Figure 4.3. Using Deep-LASI software to analyze three-color three-state smFRET data after TIRF microscopy.

Figure 4.4. Schematic overview of the L-shaped DNA origami structures used in this work.

Figure 4.5. Single-molecule TIRF results for the two- and three-color, two-state V1 L-shaped DNA origami structures.

Figure A.1. Single-molecule TIRF results for the three- and two-color, two-state V1 L-shaped DNA origami structures containing an IR dye.

Contents

| | |
|---|-----|
| Abstract..... | i |
| Zusammenfassung | iii |
| List of Abbreviations | v |
| List of Figures | vi |
| Contents..... | vii |
| 1. Introduction | 1 |
| 2. Theoretical Background | 4 |
| 2.1. Fluorescence | 4 |
| 2.1.1. Characteristics of Fluorescence | 5 |
| 2.1.2. Properties of Fluorophores | 6 |
| 2.1.3. Fluorescence Anisotropy..... | 7 |
| 2.2. Förster Resonance Energy Transfer | 9 |
| 2.2.1. Single-molecule FRET | 11 |
| 2.2.2. Fluorescence Microscopy..... | 12 |
| 2.3. DNA Origami Platforms as a Standard Platform used in Multi-color FRET Studies | 13 |
| 3. Single-molecule FRET measurements..... | 16 |
| 3.1. SmFRET Measurements in Solution | 16 |
| 3.1.1. Confocal Microscopy..... | 16 |
| 3.1.2. Three-color MFD-PIE Microscope | 17 |
| 3.1.3. Pulsed-interleaved Excitation (PIE)..... | 19 |
| 3.1.4. Sample Preparation for smFRET Experiments on MFD-PIE Setup | 19 |
| 3.1.5. Data Analysis for smFRET in Solution..... | 19 |
| 3.2. SmFRET Measurements on the Surface..... | 24 |
| 3.2.1. Total Internal Reflection Fluorescence Microscopy (TIRFM) | 24 |
| 3.2.2. Prism-type TIRF Setup..... | 25 |
| 3.2.3. Surface-passivation of Prisms to Immobilize Single Molecules | 27 |
| 3.2.4. SmTIRF Experiments Using Alternating Laser EXcitation (ALEX)..... | 28 |
| 3.2.5. Data Analysis for smFRET on the Surface | 29 |
| 4. Publications Overview..... | 32 |
| 4.1. Deep-LASI: deep-learning assisted, single-molecule imaging analysis of multi-color DNA origami structures..... | 32 |
| 4.1.1. Motivation..... | 32 |
| 4.1.2. Contribution..... | 33 |

| | |
|---|-----|
| 4.1.3. Key Results | 33 |
| 4.1.4. Discussion..... | 35 |
| 4.2. Deep-LASI, single-molecule data analysis software..... | 36 |
| 4.2.1. Motivation..... | 36 |
| 4.2.2. Contribution..... | 37 |
| 4.2.3. Key Results | 37 |
| 4.2.4. Discussion..... | 40 |
| 4.3. Distance and Kinetic Tunability of Dynamic DNA Origami Structures Examined at the Single-Molecule Level | 41 |
| 4.3.1. Motivation..... | 41 |
| 4.3.2. Contribution..... | 42 |
| 4.3.3. Key Results | 42 |
| 4.3.4. Discussion..... | 45 |
| 5. Conclusion and Outlook..... | 46 |
| Appendix 1: Single-molecule FRET Measurements on DNA Origami Structures Labeled with an IR dye | 48 |
| Appendix 2: Appended Publications..... | 53 |
| Paper 1- Deep-LASI: deep-learning assisted, single-molecule imaging analysis of multi-color DNA origami structures..... | 54 |
| Paper 2- Deep-LASI, single-molecule data analysis software | 119 |
| Paper 3- Distance and Kinetic Tunability of Dynamic DNA Origami Structures Examined at the Single-Molecule Level | 134 |
| Acknowledgments..... | 219 |
| Bibliography | 221 |

1. Introduction

Fluorescence spectroscopy and imaging assays are fast, non-invasive and quantitative techniques, which are ideal for exploring biomolecules interactions in cellular environments and solutions. They help unravel the structural and functional features of biomolecules such as DNA, RNA and proteins [1, 2]. Among all microscopy methods, fluorescence resonance energy transfer (FRET) is a powerful tool to study interactions, distances and dynamics of biological systems at the nanoscale (typically in the range of 2-10 nm) [3, 4]. FRET occurs through a non-radiative transfer of energy from a donor fluorophore to an acceptor fluorophore with matching photophysical properties when they are in close proximity. The efficiency of such energy transfer is extremely sensitive to the distance between the donor and acceptor molecules to the extent that FRET is known as a molecular ruler, making it an excellent tool for studying molecular interactions, conformational changes and dynamics [5, 6].

The way we approach and address biological systems and questions has been revolutionized by single-molecule spectroscopy. In contrast to ensemble measurements, which result in an average behavior of the whole system under study, single molecule FRET (smFRET) extends the capabilities of FRET to observe individual molecules and study their evolution in real time [7-9]. They empower us to discover individual events from a heterogeneous ensemble and to identify, sort and compare subpopulations. Hence, smFRET enables the detection of heterogeneities that are often hidden in conventional ensemble experiments [10, 11]. SmFRET experiments performed both on freely diffusing molecules in solution and surface-immobilized molecules offer a thorough and detailed view of molecular mechanisms providing information about crucial processes such as protein folding, enzyme activity and nucleic acid dynamics [3].

Solution-based smFRET measurements are usually performed using confocal microscopy whereas surface-based experiments are typically conducted using total internal reflection fluorescence (TIRF) microscopy. Both techniques take the advantage of alternating laser excitation (ALEX) [3, 12]. Measuring with ALEX mode, fluorophores get excited and detected in separate channels alternatively, so one can sort their emissions spectrally and understand the photophysical properties of individual dyes [13, 14]. ALEX is usually performed with different time scales for the solution- and surface-based techniques. In solution-based assays microsecond ALEX (μ sALEX) or nanosecond ALEX (nsALEX, PIE) is used, whereas in surface-based methods millisecond ALEX (msALEX) is the most popular approach. So, various parameters such as distances, fluorescence lifetime, anisotropy and kinetics are determined by combining the solution- and surface-based methods that lie within nanoseconds to minutes [12, 15] ([Figure 1](#)). Usually the conformational and distance changes are studied using two-color FRET. However, adding more labels to the same system and performing multi-color FRET empower us to follow the interactions of multiple binding partners and multiple distances at the same time on the same molecule [4, 16].

Combining the versatility and sensitivity of smFRET techniques with the programmability of DNA origami nanostructures made a major breakthrough in our attempts to understand the nanoscale world. DNA origami techniques work through utilizing the self-assembly properties of DNA strands base-pairing to design and construct precisely defined and complicated nanostructures [17]. The assembly of DNA origami structures is based upon using a long, single-stranded DNA known as the scaffold, and complementary short oligonucleotides called staples, which are designed such that the complementary sequences make the scaffold fold into the desired complex shapes and structures with nanometer precision. Offering such predictable designing, DNA origami technology provides a precise platform for designing the arrangement of molecular components including fluorophores, proteins

and DNA-based units such as pointers in a defined spatial positioning. The functional moieties on DNA origami structures are utilized in various assays including sensing [18], catalysis [19] and fabrication of nanodevices [20, 21].

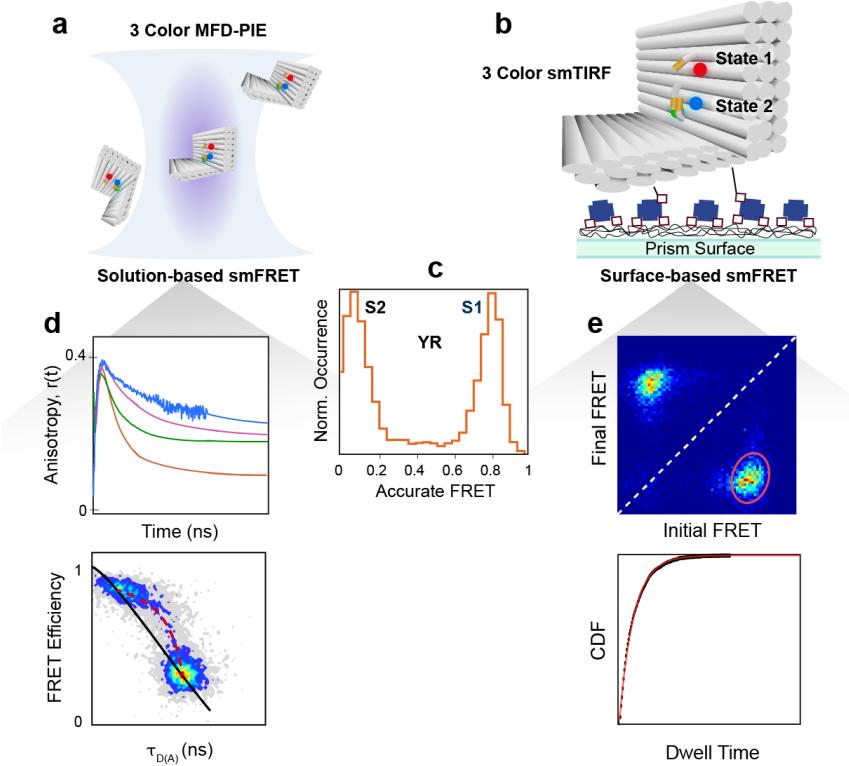


Figure 1. Schematic view of different single-molecule assays for FRET measurements. (a) SmFRET measurement of molecules freely diffusing in solution passing through the confocal volume of excitation lasers. (b) SmFRET experiments with immobilized molecules on a passivated and functionalized surface. (c) Both techniques result in FRET efficiency histograms consisting of the values from each measured single molecule with the conformational states being detected. (d) As a result of fast ALEX schemes used in the confocal method, the fluorescence lifetime and anisotropy of every fluorophore is measured. By plotting the anisotropy and FRET efficiency as a function of lifetime, we can acquire detailed information about measured molecules. (e) Analyzing the underlying kinetics in millisecond time scale is possible with smTIRF experiments. With the transition density plots showing the transitions between any two existing states and choosing the desired cluster of transitions, the distribution of the corresponding state's dwell time is plotted which we can fit to the proper model.

Combining smFRET methods with DNA origami nanotechnology providing flexible designing via exact base-pairing and labeling possibilities offers the precise spatial and temporal investigation of DNA origami assemblies with the sensitivity of single-molecule FRET measurements. This powerful combination enables us to investigate complex molecular interactions and develop sophisticated nanoscale devices used in various research fields such as structural biophysics, molecular machines, biosensing and nanotechnology.

The main content of this thesis starts with the theoretical overview of fluorescence in chapter 2. The basic principles and characteristics of fluorescence emission and some fluorophores' features, the concept of fluorescence anisotropy and Förster resonance energy transfer are described. The theoretical section also covers the basics of single-molecule assays and microscopy techniques. In the end of chapter 2, we focus on DNA origami structures from their assembly to their applications in multi-color FRET studies. Chapter 3 focuses on single-molecule FRET microscopy and measurements both in solution and on surface. Schematics of the both microscopes are depicted and various optical

components are mentioned in detail. The chapter also covers some experimental and analysis approaches including the description of FRET correction factors. The concept of TIRF microscopy with both oTIRF and pTIRF assays are also introduced.

Chapter 4 gives an overview of the three publications that were prepared throughout this work. The first article ‘Deep-LASI: deep-learning assisted, single-molecule imaging analysis of multi-color DNA origami structures’ introduces an analyzing tool based on deep-learning neural networks which helps with analyzing images from various surface-based methods. The software was benchmarked with a lot of static and dynamic experimental data taken from various L-shaped DNA origami structures. The second publication ‘Deep-LASI, single-molecule data analysis software’ introduces the MATLAB-based software package that makes the analysis of multi-color single-molecule data more straightforward by providing user-friendly graphical interfaces and tabs. The summarized functionalities of the software and the online documentation corresponding to different experimental designs help the user have a more clear data analysis pathway from the raw images to the final graphs and plots.

The third publication was initiated after measuring various two- and three-color static and dynamic DNA origami structures having different number of FRET states. Combining the results from the MFD-PIE and TIRF data implied some interesting insights about the dynamics of the origami structures and how they might be affected by the applied fluorophores in different regions. Therefore we designed more assemblies with pre-designed structural variations to check our assumptions, and included some control measurements. We believe this article would add many interesting details to the single-molecule FRET and DNA nanotechnology communities regarding to structural and experimental designs.

The thesis also includes an appending chapter summarizing smFRET data of DNA origami structures labeled with an IR fluorophore. The initial motivation was to check if we could improve the quality of our single-molecule time traces by replacing the blue fluorophore with an IR one. However, the resulting data did not show the desired improvement, so we went on with the structures containing the blue fluorophore. Since the IR data collection included upgrading our smTIRF setup from a three-color to a four-color microscope as a part of this PhD, the corresponding data are summarized on this extra chapter. Appendix 2 provides all the three published or submitted PhD themes in scientific journals.

2. Theoretical Background

2.1. Fluorescence

The phenomena of light being emitted from a substance is called luminescence, which occurs when electrons in electronically excited states return to ground states by emitting photons. A specific type of luminescence known as photoluminescence, happens when the excitation of electrons occurs with absorption of a photon. The exact phenomenon depends on the nature of the excited state, and can happen in three ways, i.e. fluorescence, phosphorescence and delayed fluorescence. Fluorescence happens when the excited electron is paired to a second electron in the ground-state orbital (opposite spins, singlet state). Returning to the ground state is then spin-allowed and occurs by emission of a photon typically within a couple nanoseconds. The average time duration of the excited state is referred to as the fluorescence lifetime (τ). Phosphorescence happens upon returning of the excited electron from a triplet state to the singlet ground states. In such cases, the transitions to the ground state are spin-forbidden, and the emission rates are slower. Phosphorescence lifetimes are typically in the range of milliseconds to seconds [22, 23]. Delayed fluorescence occurs after thermal excitation of the triplet state back to the excited singlet state. The singlet excited molecules generated in this way from triplet states can then decay radiatively to the electronic ground state [24].

The various molecular processes that occur between the absorption and emission of light are usually understood and schematically illustrated by a Jablonski diagram. A rather simple version of it is shown in [Figure 2.1](#). The scheme shows different ground and excited electronic states. The letter 'S' on diagram is used to depict singlet electronic states and the superscripts are used to show ground, first or second states of increasing energy. At each electronic level, fluorophores might have different rotational and vibrational energy levels, which are shown by numbers on the right side of each electronic level. The letter 'T' on diagram is used to depict a triplet electronic state. Other processes like colloidal quenching, energy transfer to a neighboring molecule and solvent interactions make the molecule lose the excitation energy and return to the ground state without emitting photons [25].

The transitions between different states happen within 10^{-15} s which is a very short time range for any significant nuclei displacement. This statement is known as Born–Oppenheimer principle, which suggests that due to the large masses of nuclei in a molecule, the coordinates of nuclei will not be affected by electronic, vibrational or rotational transitions. As shown in [Figure 2.1](#), internal conversion is the fast process of relaxing from a higher excited state to the lowest or first excited singlet state without emitting of a photon and could result in heating the surroundings. Since any excitation ending up at higher vibrational states is followed quickly by the internal conversion process that happens faster than fluorescence, this transition makes the peak position of fluorescence emission independent of the excitation wavelength. This process is referred to as Kasha's rule. Intersystem crossing of the excited molecule from a singlet to a triplet state can also prohibit the fluorescence emission that involves a change in the relationship between the spins of the electrons between the excited state and partly occupied molecular orbitals [25, 26].

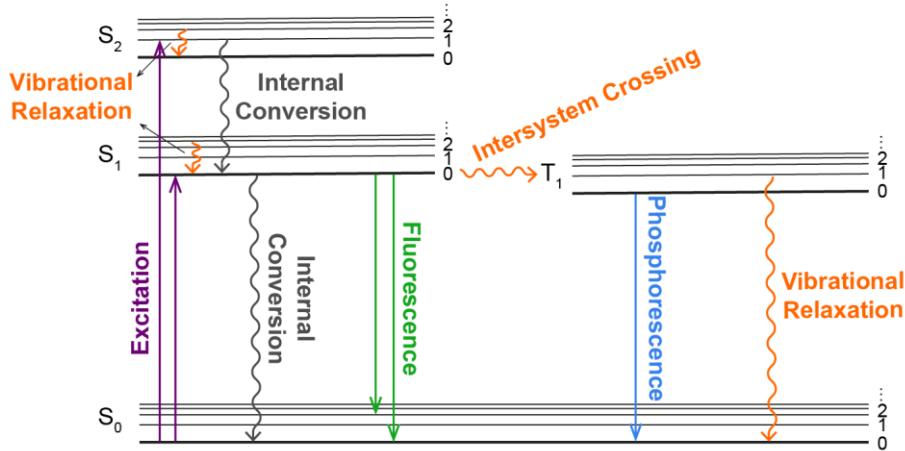


Figure 2.1. A typical Jablonski diagram showing different energy levels and molecular processes occurring between absorption and emission of photons. The horizontal lines show electronic and vibrational states labeled with letters S or T on their left side for singlet and triplet states respectively. The numbers on the right side of each state show the increasing vibrational energy levels. Solid vertical arrows show radiative processes such as excitation, fluorescence and phosphorescence. Wavy arrows depict non-radiative processes like intersystem crossing and internal conversion.

2.1.1. Characteristics of Fluorescence

An important observation of fluorescence is that the emission energy is usually smaller than the excitation energy resulting in a spectral shift of fluorescence emission spectrum to longer wavelengths than the excitation spectrum. This phenomenon was characterized by Stokes and is known as the Stokes shift [27]. There are a couple contributing factors that shift the emission spectrum to longer wavelengths. One factor is internal conversion as mentioned in the previous section, which releases a portion of the absorbed energy before emission. Fluorescence can also happen after transitioning to higher vibrational levels of the ground state, and then some further energy gets lost with relaxing to the lowest vibrational state of S_0 . According to the Franck-Condon principle the most probable transitions between different vibrational levels of ground and excited states are those with the highest overlap between the vibrational wavefunctions of the electronic ground state and the ones of the excited state, and transitions between vibrational states can be explained by wavefunctions overlap [28]. In addition, excited molecules can transfer a portion of their energy to the surroundings as loss of energy in collisions with other molecules, energy transfer, solvent relaxing effects and excited state reactions. Practically, accurate measurements of a fluorescence excitation or emission spectrum require the sample to be dilute enough so that only a negligible fraction of the incident or emitted light is lost [25, 29].

Since the same S_0 - S_1 transitions (to and from) are involved for excitation and emission processes and the vibrational sublevels of S_0 and S_1 are rather similar in spacing, the fluorescence emission spectrum of a molecule is usually almost a mirror image of its absorption spectrum (of course exceptions exist for some fluorophores). In other words, assuming that the Born–Oppenheimer approximation holds (meaning that the electronic and nuclear motions are separated) and the vibrational modes are harmonic with the same frequencies in the ground and excited electronic states, then the energies of the allowed vibronic transitions in the absorption and emission spectra will be symmetrically located

on the two opposite sides of the corresponding transition energy. This feature of fluorescence is known as the mirror-image law [25].

Considering all the competing factors for dissipating energy in addition to fluorescence emission, fluorescence yields (defined as the ratio of the number of emitted photons to the absorbed photons) are often much less than unity. Measurements of fluorescence yields thus can provide information on whether a chromophore attached to a macromolecule is accessible to quenchers dissolved in the surrounding solution or attached to a different site on the same measured molecule [25].

2.1.2. Properties of Fluorophores

Fluorescence lifetime (τ), quantum yield (ϕ) and brightness (ϵ) are three of the most important characteristics of a fluorophore to be considered before use. The lifetime value gives a very accurate vision about the environment of the molecule related to its potential interactions and information about its emission. If the excited state decays only through fluorescence, its population would decrease exponentially with time and the time constant of the decay would be τ_{fl} (lifetime related to fluorescence radiation). However, as discussed earlier, other relaxing mechanisms depopulate the excited states which consequently decrease the excited state lifetime. The total rate constant from all the possible phenomena includes the fluorescence emission rate constant as k_{fl} , formation of triplet states (intersystem crossing) with rate constant k_{isc} , non-radiative decay to the ground state (internal conversion) with the rate constant k_{ic} , transfer of energy to other molecules (resonance energy transfer, k_{rt}) and electron transfer (k_{et}).

The total rate constant of decay followed by all of the mentioned phenomena that happen in parallel would be the sum of rate constants of every individual process (Equation 2.1).

$$k_{tot} = k_{fl} + k_{isc} + k_{ic} + k_{rt} + k_{et} + \dots = k_{fl} + k_{nr} \quad (2.1)$$

We can also simplify equation 2.1 by summarizing all non-radiative decay rate constants in one term shown as k_{nr} . The measured lifetime of a fluorophore is then given by the inverse of the total rate constant which includes all the decaying processes as shown on equation 2.2.

$$\tau = \frac{1}{k_{tot}} = \frac{1}{k_{fl} + k_{nr}} \quad (2.2)$$

The other important parameter, 'quantum yield' is defined as the ratio of the number of emitted photons to the number of absorbed photons (equation 2.3), which can describe the efficiency of photochemical processes. In other words, it is the fraction of the excited molecules that decay upon fluorescence emission.

$$\phi = \frac{\text{emitted photons}}{\text{absorbed photons}} = \frac{k_{fl}}{k_{tot}} = \frac{\tau_{tot}}{\tau_{fl}} \quad (2.3)$$

Some external factors such as solvent, temperature and excitation wavelength might impact the quantum yield of a fluorophore [30].

Fluorescence quenching refers to a decrease in fluorescence emission intensity because of the advent of a nonradiative decay from the excited state. There are a diverse set of phenomena that result in fluorescence quenching [31]. Collisional quenching occurs when fluorophores in their excited-state collide with quenchers in their vicinity, which is referred to as dynamic quenching. The most typical

example for such quencher is molecular oxygen. Static quenching however, happens when fluorophores in their ground state form a non-fluorescent complex with a potential quencher. The main feature of static quenching is that while the absorption spectra of the fluorophore undergoes some changes upon binding to the quencher, the fluorescence lifetime remains unchanged, because the lifetime still comes from the free fluorophores and not when they are quenched by complexation. Since dynamic quenching happens during the excited-state lifetime of the fluorophore, it results in a decrease in the measured lifetime. Hence, dynamic quenching gives faster and more reliable detection than the static quenching. Anyhow to realize which type of quenching is dominant, both the spectrum and the lifetime of the fluorophore have to be obtained [25, 32].

The brightness of a single molecule can give similar information as the fluorescence quantum yield that is crucial for understanding the photophysical properties for individual multi-color systems in inhomogeneous environments. Molecular brightness is defined as the detected fluorescence rate per molecule, which reports on the ability of a fluorophore to emit light via photoexcitation. In another words, it is the number of recorded counts per unit time per molecule. It is usually measured in kilocounts per second for each molecule, and can be obtained by dividing the fluorescence emission by the number of emitters [33, 34]. The molecular brightness of a fluorophore can be calculated using equation 2.4.

$$\varepsilon = k \cdot \sigma \cdot \emptyset \cdot W(0) \quad (2.4)$$

Where k stands for the detection efficiency, σ is the absorption cross section at the excitation wavelength, \emptyset is the quantum yield and $W(0)$ is the intensity of exciting laser at the center of the point-spread function. So, unlike the quantum yield, the brightness is not a molecular property of the fluorophore, but depends on the precise optical conditions including the light intensity, light collection efficiency of the instrument, and the counting efficiency of the detector. Molecular brightness might change upon molecular interactions like energy transfer phenomena [22, 35].

2.1.3. Fluorescence Anisotropy

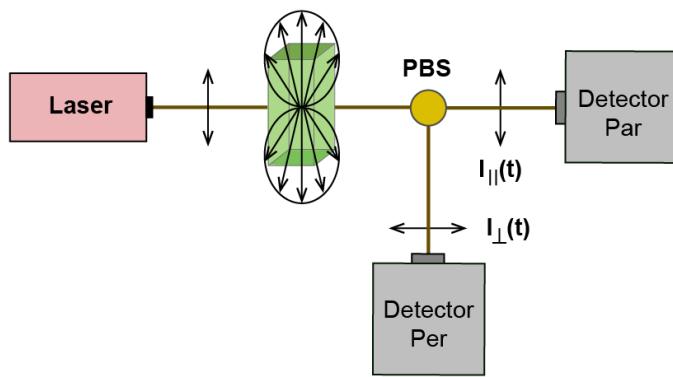


Figure 2.2. A schematic of fluorescence anisotropy measurement using separate detectors for parallel and perpendicular light orientations. After exciting the sample with polarized light, the emission, which is parallel or perpendicular to the excitation light is measured.

PBS: polarizing beam splitter, Par: parallel, Per: perpendicular.

Measuring molecular orientations and rotational mobility and studying the processes that affect them open another important dimension of fluorescence information. Anisotropy is an important parameter for understanding rotational motions of fluorophores. When a sample with randomly oriented molecules gets excited with polarized light (Figure 2.2), the fluorophores that have their absorption transition dipole aligning with the electric vector of the incident light have a higher probability for excitation than the molecules with other orientations of their transition dipole

moments. Such photoselection and unequal excitation of molecules results in a partially polarized fluorescence emission. Emission from the transition moments that are parallel to the incident light will occur more frequently than other orientations. In practice, fluorescence anisotropy (r) is determined using the parallel and perpendicular polarized emission intensities as equation 2.5 shows:

$$r = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + 2I_{\perp}} \quad (2.5)$$

In equation 2.5, ' r ' is the fluorescence anisotropy and I_{\parallel} and I_{\perp} are the fluorescence intensity of the parallel and perpendicular polarized emission after the sample has been excited with polarized light. The denominator is the total emission intensity along all the orientations. The fluorescence polarization can be defined with a similar equation to anisotropy as shown in equation 2.6.

$$P = \frac{I_{\parallel} - I_{\perp}}{I_{\parallel} + I_{\perp}} \quad (2.6)$$

Anisotropy or polarization can be used based on the existing questions to be answered. However, we mostly use the anisotropy.

During its time in the excited state, fluorophores can make angular displacements through rotational diffusion, which depolarizes the emission light and the anisotropy value decreases. Measuring anisotropy gives valuable information regarding the rotational freedom, environment and interactions of fluorophores. The Perrin equation is used to describe the rotational diffusion of spherical rotors, which is normally given as equation 2.7 where r_0 is the fundamental or time-zero anisotropy, τ is the fluorescence lifetime and ρ is the rotational correlation time, which is the time span for the anisotropy to decay to $1/e$ of its initial value.

$$r = \frac{r_0}{1 + \tau/\rho} \quad (2.7)$$

The excitation of molecules with polarized light results in selective excitation of molecules of a particular orientation. The anisotropy then depends on the distribution of the angle (Θ) between the absorption dipole moment of a molecule and the polarization of the incident light as formulated in equation 2.8 with $\langle \dots \rangle$ meaning the average value.

$$r = \frac{3\langle \cos^2 \Theta \rangle - 1}{2} \quad (2.8)$$

The excitation of fluorophores in different orientations results in $\cos^2 \Theta = 3/5$, and accordingly, a maximum theoretical value of anisotropy would be 0.4. For systems with more than one excited state, the initial value could be higher than 0.4 but such systems are more complicated to study through anisotropy analysis. Right after an excitation pulse and as the molecule rotates within its excitation lifetime, the anisotropy will decay to zero with useful information about the shape and flexibility or freedom extracted from the decay kinetics [22].

Monitoring the steady-state and time-resolved fluorescence anisotropy of individual fluorophores helps characterize the rotational properties of fluorophores and the labeled biomolecules, which provides insights about complex formation and possible undesired interactions of the dye with the biomolecule regions. Anisotropy analyses are also used to assess the uncertainties in conformational orientation factors and determine the inter-dye distances more accurately [3, 12, 36].

2.2. Förster Resonance Energy Transfer

Following the inter-molecular distances reporting on different regions of biomolecules is critical to unravel their conformations and interactions. Resonance energy transfer (RET) is a popular tool to investigate the structure and dynamics of a wide range of biomolecules due to its sensitivity to sub-nanometer scale, which is compatible with the size of important biomolecules like proteins. When two fluorophores are in close proximity with matching spectroscopic properties, the excited energy of the donor molecule (D) can be transferred to an acceptor molecule (A) in its ground state. This energy transfer, which happens without a photon, is the result of dipole-dipole interactions between the donor and acceptor fluorophores and happens faster than the fluorescence emission from donor. It is characterized by measuring the acceptor emission followed by exciting the donor molecules. This energy interaction depends on the magnitude of the two dipoles, the distance between the two centers of charge, the two angles of dipoles with respect to the line which connects the two centers of dipoles and the spectral overlap between the emission spectrum of the donor and absorption spectrum of the acceptor. Figure 2.3 shows a Jablonski diagram for the schematic process of resonance energy transfer [37].

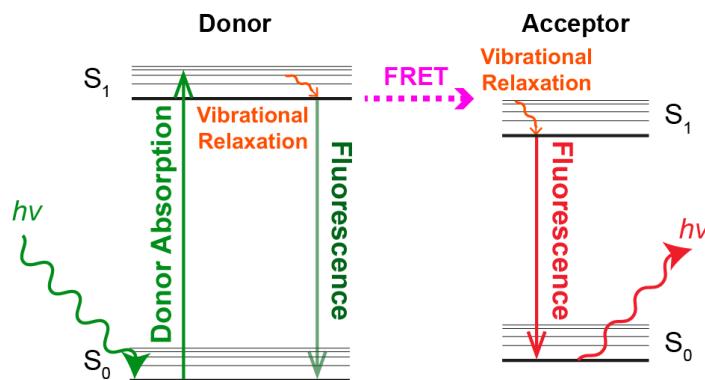


Figure 2.3. A typical Jablonski diagram for showing the phenomenon of Förster resonance energy transfer between a donor molecule in the excited state and an acceptor molecule in the ground state. When proper circumstances for the energy transfer regarding to the distance, angle and spectroscopic properties of the dye pair are met, the excited energy can be transferred via a non-radiative, dipole-dipole interaction.

Since Theodor Förster was the scientist who correctly characterized the energy transfer of a donor-acceptor pair in terms of the distance and spectroscopic properties of the two fluorophores, the described energy transfer is widely known as Förster resonance energy transfer (FRET). Another expression for this type of energy transfer is 'fluorescence resonance energy transfer' but has become less favored because this interaction does not involve emission of photons, so the use of this term could be misleading from the true identity of the FRET process [25].

The energy transfer rate between a single donor and acceptor is given using equation 2.9.

$$k_{rt} = \frac{Q_D \kappa^2}{\tau_D R^6} \left(\frac{9000(\ln 10)}{128\pi^5 N n^4} \right) \int_0^\infty F_D(\lambda) \varepsilon_A(\lambda) \lambda^4 d\lambda \quad (2.9)$$

where Q_D is the donor quantum efficiency in the absence of the acceptor, κ^2 is the orientation factor, R is the distance between the centers of the dye molecules, N is Avogadro's number, n is the refractive index of the medium, $F_D(\lambda)$ is the corrected and normalized fluorescence intensity of the donor, $\varepsilon_A(\lambda)$ is the molar extinction coefficient of the acceptor. The orientation factor (κ^2) describes the relative orientation of the transition dipoles and can range between 0 and 4 going from

perpendicularly oriented to head-to-tail parallel transition dipoles, and is usually assumed to have the average value of 2/3, which is a suitable approximation for donors and acceptors with free rotational diffusion. Equation 2.10 is used to determine κ^2 , and [Figure 2.4](#) depicts the important parameters.

$$K^2 = (\cos \Theta_T - 3 \cos \Theta_D \cos \Theta_A)^2 \quad (2.10)$$

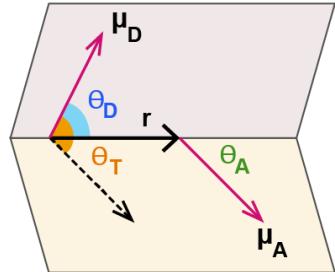


Figure 2.4. A schematic showing how orientation factor between a donor and acceptor is calculated. μ_D and μ_A are the donor and acceptor transition dipoles. θ_D and θ_A are the angles between donor and acceptor transition moments and the vector r connecting their centers. θ_T is the angle between the transition moments of the donor and acceptor.

The integral term in equation 2.9 is called the overlap integral, usually shown as $J(\lambda)$, and reports on the magnitude of spectral overlap between the donor emission and acceptor absorption spectra. The factors contributing to the overlap integral come from the overlapping region of the normalized emission spectrum of the donor and absorption spectrum of the acceptor, which is a representative of the resonance energy transfer spectrum.

Considering that the fluorescence lifetime of the donor in the absence of the acceptor is τ , the center-to-center distance of the two molecules is R , and the distance in which 50 % of the energy decay happens through the energy transfer is R_0 (known as Förster radius, typically around 20-60 Å), the rate constant for resonance energy transfer (k_{rt}) can also be written as:

$$k_{rt} = \frac{1}{\tau_D} \left(\frac{R}{R_0} \right)^6 \quad (2.11)$$

Although determining the intermolecular distances from the rate of energy transfer gives accurate results, a couple factors might affect the precision of distance calculations. To begin with, the orientation factor used in equation 2.9 to determine the transfer rate is usually an assumed value and if fluorophores have limited rotational freedom, it could add uncertainty to the distance calculation. To avoid this averaging, one can try to calculate the value of κ^2 for the system under study through anisotropy experiments. However, this approach might add extra complications especially to the analysis steps. Also, a refractive index of 1.4 is usually considered as a good compromise between the buffer and protein [6] might not be that exact depending on the molecules and structures under study. The overall effect of the mentioned uncertainties is however rather small and the complicated ways of determining the mentioned parameters might be a bigger challenge. Calculating the Förster radius (R_0), derived from equations 2.9 and 2.11, gives a helpful and widely accepted approach for determining the intermolecular distances [10]. After inserting all the known constants in equation 2.9, the more simple formula for calculation of R_0 in Å (Angstrom) would be as follows:

$$R_0 = 0.211(\kappa^2 n^{-4} Q_D J(\lambda))^{-6} \quad (2.11)$$

Since energy transfer rates are usually higher than the emission rates of fluorophores, FRET happens efficiently during the lifetime of the donor. FRET efficiency is a widely used term that reports on the ratio of the photons being emitted from the acceptor to the total number of photons that had been absorbed by the donor. This definition can also be written as the ratio of energy transfer rate to the total rates of donor lifetime decay and the transfer rate (equation 2.12).

$$E = \frac{k_{rt}}{\tau_D^{-1} + k_{rt}} \quad (2.12)$$

Replacing the distances from equation 2.11 gives the very helpful relation between the FRET efficiency and the distance between a donor and acceptor pair as shown on equation 2.13. This high dependence of energy transfer on the distance (to the power of 6th) between the centers of the fluorophores made FRET known as a molecular ruler, which is a very powerful tool for measuring the short distances between any FRET pair in close vicinity. It is especially sensitive to any distance changes when the distance of the two fluorophores is close to the Förster radius of the dye pair.

$$E = \frac{R_0^6}{R_0^6 + R^6} \quad (2.13)$$

By measuring the FRET efficiency using the donor lifetime in the absence and presence of the acceptor or calculating the efficiency from the fluorescence intensities of the donor and acceptor before and after the photobleaching step of the donor, we can calculate the distance between the two fluorophores and determine how they might change over time, which is crucial in studying dynamic systems. However, there are some considerations when doing such calculations in terms of having the true or accurate value for the FRET efficiency because the efficiency we obtain from measurements could be affected by various factors that we need to correct for. Every method used for determining the FRET efficiency and distances have specific correction factors that should be taken into account for a successful assessment of the nanometer range distances in the system under study.

2.2.1. Single-molecule FRET

During the past decade, great progress has been made in the field of single-molecule fluorescence detection, enabling us to study physical, chemical and biological properties of molecules [38]. Förster resonance energy transfer as a widely used method for studying biomolecular structures and dynamics benefited a lot from such achievements for the both in vitro and in vivo studies. FRET is widely carried out at both the ensemble and single-molecule levels. Conventional ensemble measurements complicate our understanding of the detailed properties of a system because of averaging over the parameters obtained from a whole population of molecules that could be significantly heterogeneous. There might be short-lived intermediates or sub-populations of molecules that simply have different behavior than some other molecules under study which will all be masked during the averaging over the whole system.

Since 30 years ago, single-molecule FRET techniques have been empowering us with molecular insights due to their high spatial resolution (2-10 nm) together with a range of accessible time scales from nanoseconds to minutes [6, 39, 40]. These features make smFRET an ideal tool in the research fields of dynamic structural biology. Measurements of freely diffusing molecules in and out of a confocal excitation volume and counting the emitted photons using point detectors, offers sub-nanosecond time resolution and is widely used to unravel fast dynamics [3]. To investigate conformational dynamics of single molecules for an extended time (seconds to minutes), labeled molecules are usually immobilized on passivated glass surface, and imaged using cameras [11].

Other advantages of single-molecule FRET techniques include: very low sample consumption (usually in the range of pico-molar concentrations), high sensitivity, local probing of molecules or different

regions of biomolecules, understanding the photochemical properties of different fluorescent labels when exciting and detecting them alternatively, following the conformational changes with the sequence and order or movements of biomolecules, high sample purification because of focusing on the study of single molecules of interest, observing rare or transient states and the compatibility of this technique to many experimental assays carried out both on the surface and in solution [41, 42].

2.2.2. Fluorescence Microscopy

Fluorescence microscopy is the technique that aims to reveal only the objects of interest in an otherwise black background and is one of the major characterization assays widely used in the life sciences, and has become a method with high sensitivity and resolution (spectral, spatial, temporal and order). Most researchers practicing microscopy for cell and molecular biology have difficulty with understanding the underlying photophysical phenomena and parameters used in fluorescence methods. Moreover, fluorescence microscopy is in a state of rapid evolution, with new developments in assays, probes and equipment, which make our experiments more informative [43]. All fluorescence assays are based on illuminating samples with specific excitation wavelengths followed by counting the emitted photons or creating images from the field of view [44]. Many different types of fluorescence microscopy assays are widely in use which are mentioned briefly in this section.

In epifluorescence microscopy, which is the most popular method, observation and excitation of the fluorescence occur from the same side of the specimen. The resulting fluorescence is then focused on the detector using the same objective used for excitation [45]. Confocal microscopy approaches use a focused spot of light as the confocal volume and remove out-of-focus emission by incorporating a pinhole in the image plane [46]. In contrast to epifluorescence microscopes, confocal microscopy excludes secondary fluorescence in regions outside of the focal plane [45]. Single- and multiple-beam scanning confocal microscopes scan regions of interest, typically in a raster pattern or pinhole array respectively, to offer a wide range of imaging rates [47].

Multiphoton fluorescence excitation microscopy is an optical sectioning method that uses high powers of lasers. The energy density is then large enough at the focal spot resulting in a 3D resolution at depths far exceeding those achievable by single-photon excitation microscopy. Optical sections are acquired by raster scanning the specimen in the x-y plane, and a full three-dimensional image is composed by scanning the specimen at sequential z positions. Interestingly, photodamage and photobleaching are confined in the focal plane because the very fast laser's long wavelength emission (~800 nm) results in less damage, and the multiphoton absorption mostly transpires in the focal plane. Hence, it is a widely-used technique for studying biological tissues deep within living organisms. Multiphoton assays offer high-resolution imaging deep within intact specimen by using non-linear excitation and overcoming the scattering of light in heterogeneous environment of tissues [45, 48-50].

Total internal reflection fluorescence (TIRF) microscopy was developed to reduce the problems caused by the background by using a thin evanescent field. It is a very useful technique for monitoring events at the plasma membrane of cells [51, 52]. TIRF helps with directly monitoring biomolecules even at single-molecule levels (this method will be described in more details in chapter 3) [53]. Fluorescence correlation spectroscopy (FCS) is another versatile technique that can be exploited at the single-molecule level. This technique became especially powerful after being combined with confocal detection [54]. FCS analyzes very small fluctuations in the fluorescence emission of small complexes to study underlying inter- and intramolecular dynamics. Analyzing the autocorrelation function on the

intensity fluctuations over time can reveal an average retention time that particles spend in the confocal volume, and since the size of the confocal volume is known, the diffusion rate can be determined [55, 56].

Combining optical microscopy with FRET as a distance-sensitive interaction between electronic states of so-called donor and acceptor molecules, make the whole system a spectroscopic ruler that gives quantitative information about distances, interactions and kinetics [57]. The method is based on optical excitation of the donor molecule and interaction of optical transition dipoles with an acceptor molecule, which is able to fluoresce [58]. Microscopy experiments using FRET methods are the focus of this thesis, which are described in detail in the corresponding sections for theory and instrumentation.

Among all various types of fluorescence microscopy methods, FRET experiments are typically carried out using two types of wide-field and confocal configurations [45]. In wide-field methods, the sample is illuminated across a field of view with the most common approach being total internal reflection followed by imaging the surfaces on low-noise cameras [22]. In confocal approaches, the molecules are excited while freely diffusing through the small confocal volume of roughly one femtoliter. The emitted bursts of photons are detected using photon-counting detectors [12]. Advanced applications and development of fluorescence microscopy assays are continuously evolving the both techniques from experimental equipment such as detectors, to new computational and analysis approaches [59].

2.3. DNA Origami Platforms as a Standard Platform used in Multi-color FRET Studies

DNA origami nanostructures with two- or three-dimensional shapes offer many research-friendly features such as programmability [60, 61], self-assembly, flexibility [62], high specificity [63], biocompatibility [63], stability and easy tailoring compared to other biomolecules [64]. All these features have made DNA nanostructures well established and precise building blocks for various fields of life science such as drug delivery [65], molecular sensing and biosensing [18, 66-68], molecular machines [69], material synthesis [70], nanophotonics and plasmonics [71], and nanoelectronics [72]. DNA origami structures can be divided into two classes of static and dynamic nanostructures, and of course, the dynamic versions drew more attention and found more applications [32, 37].

Dynamic DNA nanostructures designed as nanoscale devices primarily function by undergoing controlled conformational rearrangement or motion. These dynamic assemblies have been used in a wide range of applications like therapeutics [73], DNA-based logic gates [74, 75] and robotics [76, 77] to name a few. Designing of the controlled motion mainly involves transient DNA strand displacement and application of external factors such as surrounding ions, effects of light and pH changes. Dynamic origami structures can also be made by combining self-assembly static components with reconfigurable modules, which undergo post-assembly structural changes such as, the DNA nano-box [78, 79], DNA tweezers [80, 81], DNA walkers [82, 83], and reconfigurable origamis [84, 85]. The conformational changes could be introduced by external stimuli such as applied fields or simply strand-displacement modules [86, 87].

DNA origami structures are usually made using a bottom-up strategy starting from a long, single-stranded scaffold and a set of short oligonucleotides arranging themselves through programmable complementary nucleotide bindings. One of the advantages of DNA origami assemblies is that we do not have to mix the scaffold and the small staple strands in a limited stoichiometric manner. Since the

hybridization happens specifically between the designed regions, any additional strands will be washed away during the purification steps [64]. The main consideration will only be carefully choosing the complementary nucleobases of the staples to reach to the final desired structure. [Figure 2.5](#) shows a simplified schematic of DNA origami assembly. During hybridization of the scaffold and staple strands, every nucleotide on the staple strand will bind to the right complementary sequence and thereby make the designed structure.

The shape of origami nanostructures is usually controlled by arranging DNA double helices into hexagonal or square cross sections, and curved structures are obtained by introducing intentional mismatches. Using DNA chemistry, each strand can be individually functionalized with a nanometer precision to have addressable chemicals like fluorophores, proteins, nanoparticles or any desired DNA motifs to design well-predictable interactions. All these features make DNA origami structures as suitable platforms for different fields of research [17, 88].

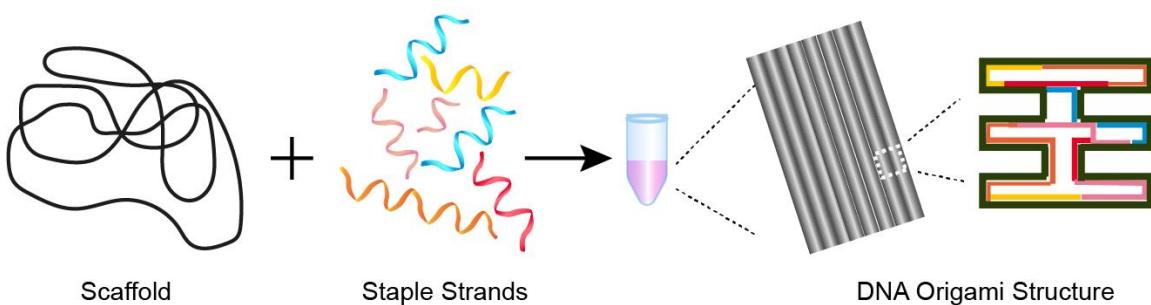


Figure 2.5. Principle of the DNA origami structure assembly. The long ssDNA, called the 'scaffold', is mixed with several hundreds of short ssDNA known as 'staples'. Upon the annealing process with programmed heating up and cooling down steps, the final 2D or 3D shape is obtained.

Single-molecule fluorescence methods in particular take advantage of the precise and tunable placement of fluorophores in the complex shapes of two- and three-dimensional DNA origamis. Due to the rather simple labeling approaches of DNA origami with organic dyes, researchers started to use the many functionalities of such systems to study energy transfer and FRET processes [88, 89]. In this regard and as early exemplary research works on origami capabilities in energy transfer, after dsDNA and 2D structures, 3D DNA origami structures were used as a channel to help transfer energy from a primary donor to an intermediate one and in the end to an acceptor or served as a breadboard to study the distance-dependent FRET within a range of 2.5 to 14 nanometers [90]. Directional multi-color energy transfer pathways were also studied at the single-molecule level starting from a blue fluorophore as the input dye, a yellow dye as an energy conductor or jumper in a close proximity to two other dyes (red and IR) as output acceptors [89-91]. DNA origami platform was used to model a DNA double-strand break as a form of DNA damage, and the end-joining reaction and DNA repair pathways were investigated using smFRET [92]. After such works, the smFRET field using DNA origami structures has been growing rapidly.

In addition, single-molecule FRET techniques have the capability of providing real-time structural and functional information of DNA origami structures with nanometer resolution, it can directly report on any change in the origami conformation and its time evolution. Additionally, this method reveals the structural heterogeneity among the nanostructures, rare events in mechanistic steps of dynamic nanostructures, and faulty or non-functional molecules. Single-molecule FRET has been increasingly used in diverse characterization of DNA nanostructures and DNA-based nanodevices, such as quality control, conformational transition, actuation and function. These characterizations are crucial for the optimization of the nanostructures and the devices constructed from them. Hence, the two fields of

DNA origami nanotechnology and single-molecule multi-color FRET have been affecting and helping each other while flourishing and advancing [37].

After the past couple decades of establishing DNA origami nanotechnology by works from different areas of research fields, the ability of developing DNA origami structures to be customized and adapted for addressing important fundamental questions via smFRET powerful techniques, made origamis trustworthy and robust candidates for such experimental designs. Recent applications of DNA origami nanostructures in smFRET studies are quit broad and advanced. Hence, only a couple of examples are summarized to show the broadness of their applications specifically in FRET studies.

Flexible DNA origami structures were designed to be curvature biosensors with smFRET readout [18]. Based on the angle of the curvature and the change in FRET efficiency as the signal, the particle size of various substances in the range of 50-300 nm was reported. In another recent application, dynamic DNA origami nanostructures were used to report inter-dye distance with nanometer precision between 4 and 100 nm to achieve the resolution offered by both smFRET and co-tracking of fluorophores, which push the distance range beyond what FRET approaches provide [93].

Researchers have shown another interesting potential of labeled DNA origami nanostructures using the orientation of organic dyes as donors and acceptors to manipulate the energy transfer between the covalently attached fluorophores. Such control comes with engineering the number of paired and unpaired nucleotides right next to the fluorophores. In this study, the same distance between the FRET pair shows different FRET efficiencies only by changing the orientation of the dyes [18, 93, 94].

Although the progress made in DNA origami nanotechnology already allows better designs for the detection of molecules of interest in the concentration range of nanomolar to picomolar, still some more challenges need to be addressed. For instance, improving the limit of detection to the sub-picomolar range would provide new opportunities for detecting a broader range of targets and for pushing the sensitivity of the already working targets to even lower concentrations. Improved or shorter response times of FRET sensors will also be an important accomplishment, especially for biomedical applications. Other properties that could to be improved is the stability of DNA origami platforms under physiological conditions [95].

3. Single-molecule FRET measurements

In this chapter, the principles, microscopy details and general data analysis steps for corresponding smFRET experiments are described. Single-molecule FRET measurements allow for investigation of hidden intermediate states and heterogeneities in biomolecules structure and function that can include a range of observed behaviors. SmFRET histograms reveal amplitudes of both fluorescence emission and FRET efficiency between the fluorophores. Using the advanced measurement approaches of smFRET for solution-based or surface-based assays, many details about the features, functions and kinetics of the system can be obtained. Since the two assays have different time-scale and modalities, combining their results gives a thorough understanding of biomolecular structure and dynamics extracted from hundreds or thousands of single molecules with high sensitivity and signal-to-noise ratio [3, 10, 40, 59].

3.1. SmFRET Measurements in Solution

Single-molecule FRET measurements can be done with freely diffusing molecules in solution using the burst analysis approach, where the emission of photon bursts from a dilute solution (usually 20-50 pM concentration) of fluorophores passing through the very small confocal volume is measured [10, 11]. The very low concentration range is needed to guarantee capturing one molecule at a time. When a molecule is diffusing through the confocal excitation volume (with femtoliter dimensions), a burst of fluorescence emission is produced and is measured by photon-counting detectors. As a result, the fluorescence intensities from individual molecules are used to determine the histogram of the FRET efficiency. The key idea in confocal microscopy is taking advantage of a pinhole that is placed in the detection path of the confocal microscope resulting in the blocking of out-of-focus light from outside the desired volume [4, 12].

3.1.1. Confocal Microscopy

Confocal microscopy technique was first developed by Marvin Minsky in 1955 with the pioneering idea of rejecting the out-of-focus light, achieved by the illumination- and detection-side pinhole apertures in the same conjugate plane to make them “confocal” [96]. In order to overcome some issues of measuring on wide-field microscopes where the whole sample area is illuminated by the excitation light, confocal microscopy was developed to increase the resolution of images by exciting only a very small region of the sample. The confocal volume is the in-focus volume within a sample, which is typically on the order of femtoliters. For confocal fluorescence measurements, a laser beam is focused in a fluorescent sample using a high numerical aperture (NA) objective [3]. Due to diffraction at the lenses, a zero-dimensional point-object will not be imaged as a point at the detector but instead, as a bright, round spot in the middle surrounded by a series of concentric bright and dark rings around it known as the Airy disk pattern. The response of an optical system to a luminescent point source is called the point spread function (PSF), which can usually be approximated by a Gaussian function.

The lateral resolution (r_{lateral}) in a confocal microscope is defined as the distance from the maximum of the airy disc to the first minimum, which is given by equation 3.1 when using an objective with a known numerical aperture (NA_{obj}). It also depends on the beam wavelength and the angle of the

incident beam entering the objective. Considering that in our experiments, objective and condenser are the same object, equation 3.1 is a common equation to calculate the lateral resolution.

$$r_{lateral} = 0.61 \frac{\lambda}{NA_{obj}} \quad (3.1)$$

where λ is the excitation wavelength. In practice, the lateral resolution of a technique is the minimum distance at which two adjacent point emitters can be distinguished and resolved.

There are various types of confocal microscopy techniques to measure molecules immobilized on a surface or freely diffusing molecules in a solution. For immobilized samples, scanning a surface is widely used by confocal microscopy. Another approach is to measure molecules in a solution and exciting them while they are diffusing through the confocal volume. The excitation is then focused on a very small and well-defined spot inside the solution. The latter measurement approach is what has been used in this thesis.

3.1.2. Three-color MFD-PIE Microscope

In this section, the microscopy setup that has been used to perform two- and three-color smFRET experiments in solution is described. Multi-parameter fluorescence detection provides time-resolved information of many intrinsic properties of fluorophores that can be studied by fluorescence [35]. After collecting the photons being emitted from the confocal volume, various parameters of the photons can be determined. This technique is especially powerful for FRET studies because the fluorescence properties of molecules can change over FRET interactions, especially in case of donor molecules. Using MFD, the selection of a subset of molecules from a mixture is possible enabling a species specific analysis. In MFD experiments, pulsed excitation sources are used and the detection path is split first according to polarization and then by spectral separation before being detected using time-correlated single-photon counting (TCSPC) devices. The repetition rate of laser pulses (with 10–100 ns resolution) together with the arrival time of the fluorescence photons with respect to the laser pulse are measured by TCSPC give an overall time resolution of picoseconds independent of how long the measurements run. All the information regarding to FRET efficiencies, fluorescence lifetime and anisotropy can be extracted from the detected photons coming from the same molecule [12, 97].

Figure 3.1 shows a schematic figure of the three-color MFD-PIE microscope used in our lab. Solution-based FRET experiments of freely diffusing single molecules are performed using pulsed-interleaved excitation (PIE) with three lasers of 485 nm, 560 nm and 640 nm (LDH-D-C-485, LDH-D-TA-560, LDH-D-C-640, PicoQuant, Berlin; Germany) at an alternation period of 16.67 MHz adjusted by the laser driver (Sepia II, PicoQuant) for both two- and three-color experiments. In the case of two-color samples, the excitation was done with two of the three laser lines at a time corresponding to the fluorophores present in the sample and keeping the third unused laser line off. The resulting fluorescence is collected by a 60x water immersion objective (Plan Apo IR 60x 1.27 WI, Nikon, Düsseldorf, Germany) and spectrally separated from the excitation beam by a polychroic mirror (zt405/488/561/633, AHF; Analysetechnik, Tübingen, Germany). The emission is first separated using a polarizing beamsplitter (Thorlabs, Dachau, Germany) and then spectrally filtered (AHF Analysetechnik, Tübingen, Germany) for the blue (ET525/50), yellow (ET607/36) and red (ET670/30) detection range. Photons are then collected using six photon counting avalanche photodiodes (2× COUNT-100B, LaserComponents, Olching, Germany; 4× SPCM- AQR-14, PerkinElmer, USA) and

registered by independent channels of the time-correlated single photon-counting hardware (TCSPC, HydraHarp400, PicoQuant). The detector signal is recorded using a home-written program in C# which is synchronized with the laser driver [4].

To determine various parameters like fluorescence intensity, lifetime and anisotropy, emission photons arriving at the three detection channels are sorted according to the three excitation periods based on the photon arrival time. Photons are considered by the algorithm to belong to a molecular burst when the local count rate within a sliding window exceeds a given threshold. The threshold was described typically by a minimum total number of 50 photons having 30 consecutive photons within a time interval of 500 microseconds.

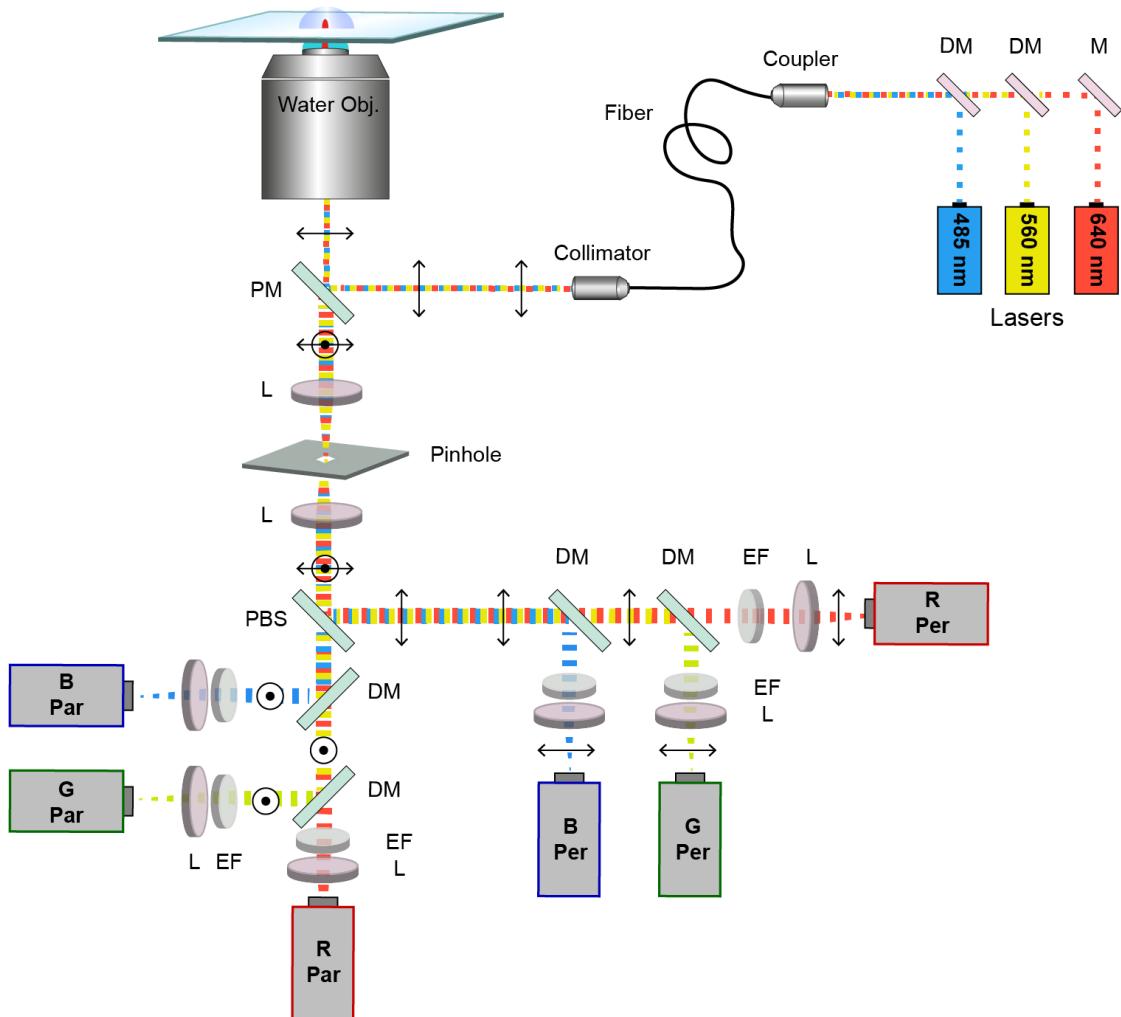


Figure 3.1. Schematic of the home-built confocal single-molecule three-color MFD-PIE setup used for the solution-based FRET measurements in this work. Three lasers are coupled into a single-mode fiber. The output of the fiber is collimated and focused on the sample above an objective. The fluorescence emission bursts are collected by the same objective and separated based on their polarization and spectral regions to be detected by the six APDs. The different abbreviations are as follows: M: mirror, DM: dichroic mirror, PM: polychroic mirror, L: achromatic lens, PBS: polarizing beam splitter, EF: emission filter.

3.1.3. Pulsed-interleaved Excitation (PIE)

Combining the advantages of MFD and PIE allows us to extract the maximum amount of information from smFRET experiments from the same experiment. Utilizing pulsed interleaved excitation (PIE) instead of millisecond or microsecond ALEX allows extraction of the fluorescence lifetime and time-resolved anisotropy information. Since the excitation source for each detected photon is distinguished using TCSPC detection, the presence of the acceptor can also be probed via the stoichiometry. Additional information is also available from MFD-PIE experiments that can be used to determine and correct for experimental artifacts like spectral crosstalks that may influence the FRET efficiency values [12, 15].

3.1.4. Sample Preparation for smFRET Experiments on MFD-PIE Setup

Two- or three-color smFRET experiments on MFD-PIE setup were performed using 8-well chamber slides (Nunc Lab-Tek, VWR) as the sample holder. The surface of each well was first passivated with 30 μ L of 1 mg/mL BSA in PBS for 10 min before putting the sample solution. The instrument response function (IRF) as the response timing adjustment of the setup was measured using a fluorescence-free solution like the commercial PBS (ThermoFisher Scientific) for 10 minutes. To determine the background level of measurements, the used buffer was measured also for 10 minutes. After measuring the IRF and background, each sample was prepared as a dilute solution of 30-50 pM concentration, 30 μ L of it was placed on a BSA-passivated chamber and measured for a couple hours (2 hours in case of two-color and 3 hours for three-color samples) to get suitable amount of statistics.

3.1.5. Data Analysis for smFRET in Solution

- *Burst Selection*

Burst analysis method is used to extract and examine single-molecule events from MFD-PIE experiments. While single molecules diffuse freely in the sample solution, they get excited upon illumination inside the lasers confocal volume resulting in bursts of fluorescence emission spreading in all directions with a part of it being gathered by the objective and guided to the detectors where photons are counted. By loading the raw data into the PAM software, we can observe the amplitude and timing of the bursts along the whole measurement time span as shown in [Figure 3.2](#) for a three-color experiment. The width or duration of the bursts show how slow single molecules diffuse through the confocal volume. For the example of big DNA origami structures, the width of the bursts will be bigger than small proteins or nucleic acids. One can use suitable thresholding to select the bursts generated by the molecules of interest for further analysis.

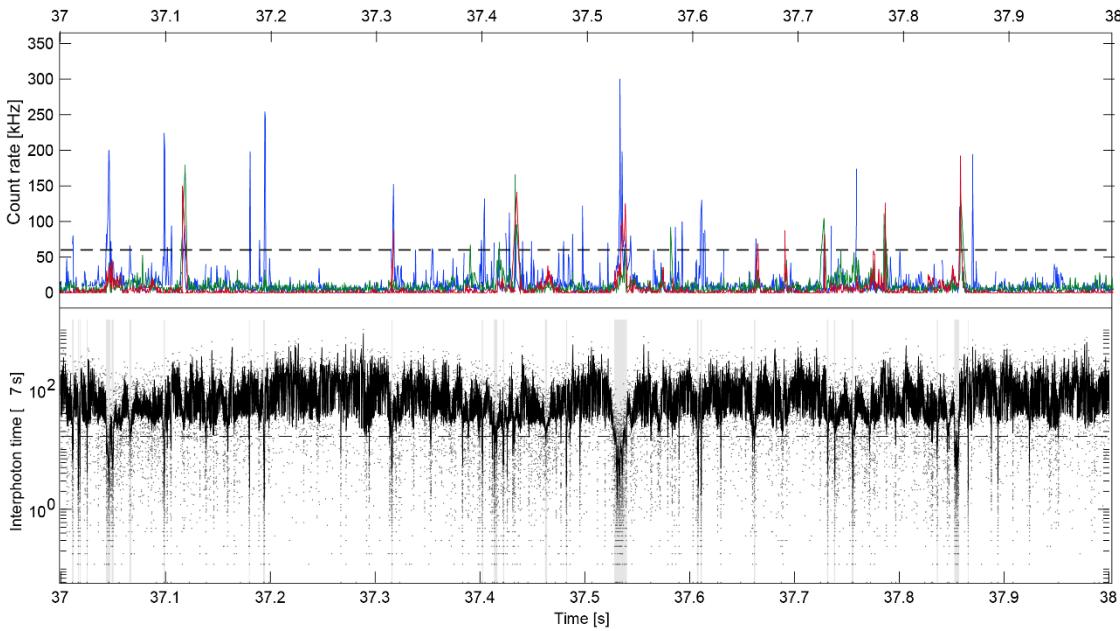


Figure 3.2. Schematic of burst search methods for three-color experiments provided with the PAM software. Top panel: sliding time window method. Bottom panel: inter-photon time method. Dashed lines represent the incorporated thresholds which are 30 kHz count rate on this example with the time window of 500 μ s. the inter-photon time was set to 30 μ s (minimum number of photons per burst was set to 100).

Throughout this work, the burst search method of all photon burst search (APBS) was used, which is the most general approach. Using this method, burst extraction occurs depending on the photons from all channels no matter how many labels are present. In other words, the bursts from molecules with all three labels as well as two and one labels will all be extracted; this approach helps with determining the FRET specific correction factors more accurately. When the number of detected photons exceeds the user defined threshold level, it will be considered as a burst and based on the width of the time window, the adjacent bursts on the left and right side of the main one will also be examined. If any of them also exceeds the threshold value, they will be combined to a burst. This approach is referred to as sliding time window [98].

- *Correction Factors for Quantitative Three-color smFRET*

Since the structure and function of biomolecules are highly dependent on the intermolecular distances and their changes over time, it is important to estimate them precisely to make correct assumptions about their conformational interactions. FRET is known as a spectroscopic ruler and makes it possible to determine distances at the sub-nanometer range, however some corrections on the obtained intensities are needed for this purpose. A huge advantage of using alternating laser excitation schemes such as PIE or ALEX is that the individual dye emissions are accessible and detected in corresponding detectors with separate time bins. This lets us determine FRET-specific correction factors even in the complicated case of a three-color experiment to achieve three-dimensional distance information [6, 12, 31, 35].

In the case of a three-color sample, the three fluorophores make the quantitative FRET measurements more complicated than a system consisting of one donor-acceptor pair. Here, we have three fluorophores measured at the same time: a donor with the highest excitation energy and an acceptor with a lowest energy together with a third dye that its excitation falls between the other two and acts

both as an acceptor for the donor molecule and as a donor for the acceptor molecule. Such three-color systems can be modeled and understood with three dyes as blue, green and red, making three FRET pairs of BG, BR and GR (the first letter stands for the donor and the second one for the acceptor). [Figure 3.3](#) shows a schematic of a three-color FRET system. For simplicity, we refer to the first donor dye (blue) as D, the second donor dye (green) as D' and the red dye as A.

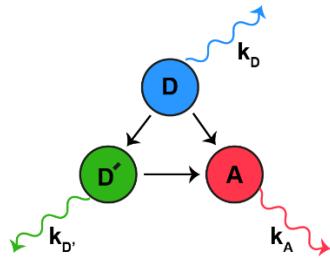


Figure 3.3. Schematic of three-color FRET. The fluorophores are shown as circles with colors corresponding to their excitation wavelength from the blue to red spectral regions. D stands for the first donor, D' for the fluorophore that behaves as both donor and acceptor, and A for the acceptor. The wavy arrows depict the fluorescence emission from each of the dyes and the solid arrows show the direction of energy transfer between molecules.

The apparent FRET efficiency between any of the two dyes can be calculated from the measured intensities. An alternating three-color excitation scheme is used to separate emissions after each excitation pulse meaning that we can measure emitted photons after direct excitation separate from the ones emitted after the energy transfer. These separate photon packages are used to calculate the amount of energy transfer. We start by calculating the energy transfer between the two acceptor molecules (the green and red dyes in the [figure 3.3](#)) after green excitation because we can treat it like a two-color experiment with one donor (D' in this case) and one acceptor. Equation 3.2 gives the FRET efficiency for the green-red dye pair using the intensities of the two different channels after green excitation.

$$E_{D'A} = \frac{I_{D'A}}{I_{D'A} + I_{D'D'}} \quad (3.2)$$

Here the letter I stands for the measured intensity and the two suffix letters in each term stand for the excitation and detection colors respectively. Hence, the nominator shows the intensity measured of the acceptor channel after excitation with the donor (D') laser, which would ideally be the acceptor intensity due to the FRET process. The denominator shows the sum of intensities from the both channels after D' excitation, hence the overall formula gives the FRET efficiency.

After blue excitation, all detection channels need to be analyzed, which makes the analysis more complicated to characterize because upon the excitation of the first dye, its energy can be transferred to both of the acceptors, and at the same time, the second donor (D') also gives energy away to the third dye. Considering these processes, we can write the following formulas to calculate the two-color FRET efficiency for BG and BR dye pairs.

$$E_{DD'} = \frac{I_{DD'}}{I_{DD}(1 - E_{D'A}) + I_{DD'}} \quad (3.3)$$

$$E_{DA} = \frac{I_{DA} - E_{D'A}(I_{DD'} + I_{DA})}{I_{DD} + I_{DA} - E_{D'A}(I_{DD} + I_{DD'} + I_{DA})} \quad (3.4)$$

As mentioned before, we can obtain the distances between the centers of the three fluorophores from the calculated FRET efficiencies. However, we should consider that a couple of important correction factors need to be applied to the results. Since we use separate detectors with different detection efficiencies for different wavelengths of emission, and different fluorophores with different quantum yields, such differences should be taken into account for a correct estimation of the absolute FRET efficiencies and calculated distances. Other correction factors are necessary to correct for the spectral overlap that the FRET pair has as one of the basic criteria for the energy transfer to happen. The corrections are usually done in a couple of steps and the resulting FRET efficiency values after each step is referred to with a specific term between the FRET community so it is worth mentioning the names that most researchers agree on [6]. The raw intensities that are measured by each detector are first corrected for the background counts and subtracting them to calculate FRET efficiencies like the ones in equations 3.2 - 4 will result in **proximity ratios**.

Next, we need to correct the proximity ratios from the “leaking” intensities due to direct excitation of the acceptor by the donor excitation (so not through the energy transfer), and donor emission that leaks into the acceptor channel. The first type of extra emission is called the direct excitation (de) and the latter is called the spectral crosstalk (ct). After applying these two correction factors to the proximity ratios, the resulting efficiency is known as **apparent FRET**. One last correction factor is still needed to take care of the different quantum yields of the dyes and different efficiencies of the detectors when dealing with non-equal wavelengths. This last correction factor is called the detection efficiency and is usually designated with the letter γ . After incorporating all these correction factors, we achieve the final value, which is known as **corrected FRET** efficiency. This corrected FRET efficiency can then be used to calculate the distances we are interested in.

Similar concept and approaches are needed for determining the correction factors for the both solution- and surface-based experiments. The procedure to calculate the correction factors are described in detail in section 3.2.5 for the surface measurements. In solution-based data, the spectral crosstalk of the donor and the direct excitation of the acceptor are specifically calculated from the donor-only and acceptor-only populations respectively. The gamma factor can also be calculated using the lifetime information, which are easily accessible with MFD-PIE method.

- *Fluorescence Lifetime and Anisotropy Analysis*

Beyond the spectral information needed for determining the FRET efficiencies, other valuable information is collected when using multi-parameter fluorescence detection (MFD) combined with PIE excitation: time-resolved fluorescence lifetime and anisotropy decays. A summary of the various types of information provided using MFD-PIE is summarized in [Figure 3.4](#) for a two-color, two-state L-shaped DNA origami structure. The origami is labeled with Cy3B on a pointer and Atto647N attached to the origami structure. The PAM software package offers environments to analyze such data molecule-wise or with ensembles or subensembles. The burst-wise fluorescence lifetime is determined using a maximum likelihood estimation where the expected value is calculated using the IRF and the contribution from background counts [12].

After loading the data into the software, the plot of apparent or corrected FRET efficiency versus stoichiometry (E-S plot) of donor and acceptor molecules can be obtained. [Figure 3.4a](#) shows the corrected E-S plot for the Cy3B-Atto647N dye pair. The lifetime curves can be fit with various

exponential models containing a different number of exponential terms. Since the energy transfer changes the lifetime of the donor molecule, we usually need two- or three-exponential components to find the correct lifetime values for the ensemble measurement on the donor channel, which contains all possible donor populations, for example donor-only, low-FRET or high-FRET populations. Figure 3.4b shows two FRET populations with the corresponding lifetime values for the donor (the donor-only population at $E = 0$ is not shown for simplicity). The pink line, known as the static FRET line, is made by the donor lifetime. Static populations captured during the bursts time intervals lie on this line. In the case that the sample dynamic changes happen during the burst duration, different averaging occurs leading to a deviation from the static line like the dashed black line on the graph showing the dynamic population between the two static states. Figure 3.4c shows the same plot for the lifetime of the acceptor after fitting the curve with a mono-exponential model. As shown in the plot, the lifetime of the acceptor here does not depend on the FRET efficiency but such data gives information about the environment or possible quenching of the acceptor.

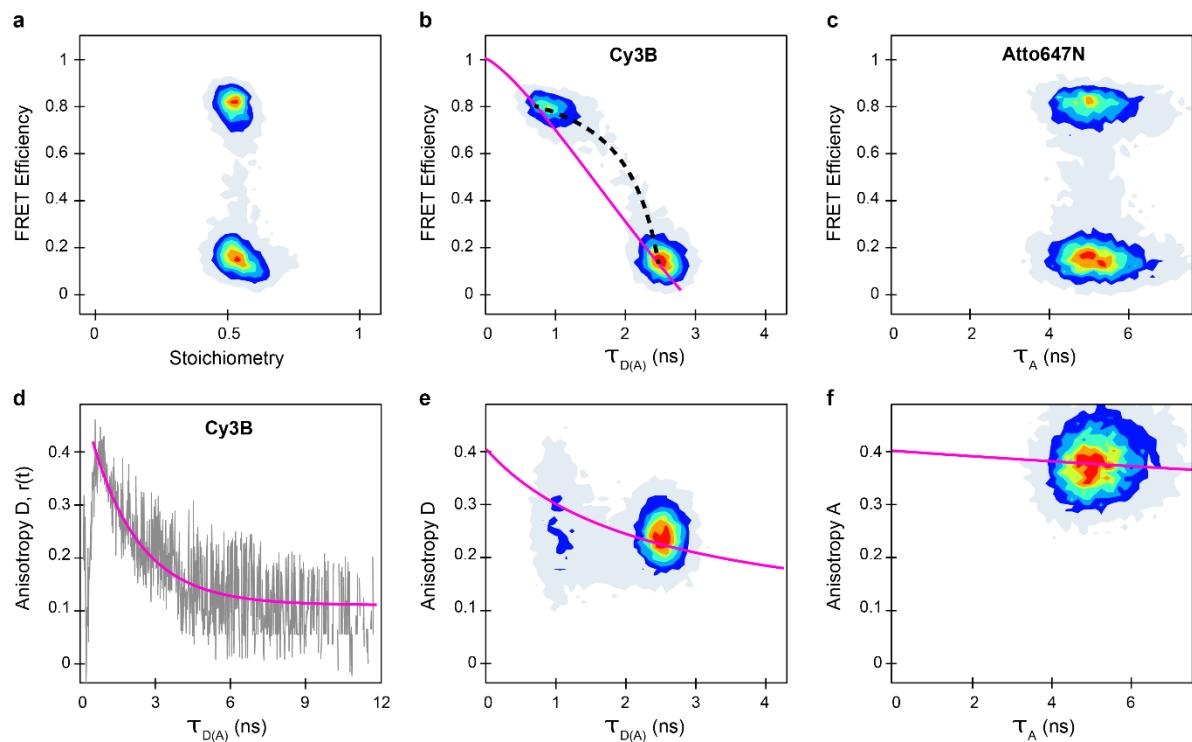


Figure 3.4. Representative plots of different parameters obtained in a MFD-PIE experiment and analysis obtained from measuring version 1 of a two-color, two-state L-shaped DNA origami structures labeled with Cy3B on a tether and Atto647N on the origami structure. The length of the complementary sequence for the two docking strands is 8 nucleotides including 1 mismatch. For details of the samples, see [99]. (a) The plot of FRET efficiency versus stoichiometry. (b) The E - τ (FRET efficiency versus fluorescence lifetime) plot of the donor in the presence of the acceptor. Two FRET efficiency populations and the corresponding fluorescence lifetime values for the donor are shown. (c) The E - τ plot for the acceptor shows the two FRET populations in efficiency that have two close values of acceptor lifetime. The shorter lifetime corresponds to interactions leading to the acceptor quenching. (d) Time-resolved anisotropy decay of Cy3B. The gray graph shows the real data, and the solid line depicts the fit to a one-component rotational model. (e) The plot of steady-state anisotropy versus the lifetime of the donor in the presence of acceptor. The rotational correlation time of the fluorophore is obtained as 1.09 ns by fitting the data to the Perrin equation (equation 2.7). (f) The graph of steady-state anisotropy versus the lifetime of the acceptor. The higher rotational correlation time for the acceptor as 4.44 ns indicates the more rigid status for the acceptor dye.

Fluorescence anisotropy gives insights about the fluorophores rotational freedom. In the case of a mono-exponential anisotropy decay, the relationship can be described by Perrin equation (equation 2.7). Time-resolved fluorescence anisotropy decays help with following the dynamics of the molecular processes such as rotational motion, ligand binding or following the movements of a tether that leads to emission depolarization [100, 101]. If there are simultaneous fast and slow motion components in the sample, we can distinguish between them using anisotropy decay and fitting it to the proper model [10]. [Figure 3.4d](#) shows the time-resolved anisotropy for Cy3B with the real data in gray and the fit in pink. [Figure 3.4e](#) and [f](#) show histograms of the steady-state anisotropy values versus the lifetime of donor and acceptor respectively. In this example, the rotational correlation time (ρ) for the acceptor (Atto647N) is longer than the donor (Cy3B), which indicates more rotational freedom for the donor fluorophore.

3.2. SmFRET Measurements on the Surface

3.2.1. Total Internal Reflection Fluorescence Microscopy (TIRFM)

A second approach allows the long-term observation of single molecules by immobilizing them on a surface and make them more phot-stable with some chemical tricks. Total internal reflection fluorescence (TIRF) microscopy is a great optical method that offers the excitation of fluorophores in an aqueous environment with a very thin electromagnetic region of about 100 nm depth known as the evanescent field. The depth of the evanescent field depends on the excitation wavelength and the incident illumination angle, so in the case of objective TIRF methods, the numerical aperture of the objective. The objective numerical aperture determines the amount of angular freedom we have for aligning the total internal reflection. As objective NA decreases, the ability to align a system for TIRF experiments becomes more difficult. There are a couple important advantages in using such a thin excitation layer: The background levels in TIRF methods are significantly lower than other fluorescence microscopy methods, so the signal to noise ratio is higher. In addition, there is no out-of-focus fluorophores that make low-quality images and the deeper parts of the sample will be kept away from any potential damages caused by the excitation light [102, 103].

When a light beam propagating in a medium like glass with a refractive index n_1 reaches an interface with a second medium like water with refractive index n_2 , the angle of the refracted beam relative to the normal (θ_2) is related to the angle of incidence beam (θ_1) by an equation known as Snell's law (equation 3.5). In the medium with the lower refractive index (like the water in the example here), the refracted beam bends away from the normal resulting in a larger angle between the beam and the normal at the surface as shown in [figure 3.5](#).

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad (3.5)$$

The critical angle is the angle of the incident beam at which the light propagates parallel to the border between the two media after reaching the interface, meaning $\Theta = 90^\circ$ with respect to the normal (equation 3.6).

$$\theta_c = \sin^{-1}(n_1/n_2) \quad (3.6)$$

In our example of the light propagating from glass ($n_1 = 1.55$) to water ($n_2 = 1.33$) and considering the 90° for θ_2 , the critical angle would be 59.3° . When the incident light hits the surface with an angle greater than the critical angle, the refracted beam will never exit the first medium because it gets totally reflected at the interface as shown on [figure 3.5](#). This phenomenon is known as total internal reflection (TIR) and is the basis of TIRF microscopy. As a result of TIR, a so-called evanescent wave forms at the interface and penetrates a finite depth into the second medium because the electric and magnetic fields of the incident light are continuous into the second medium. The evanescent field is the result of constructive interference between the incident and reflected light fields [25].

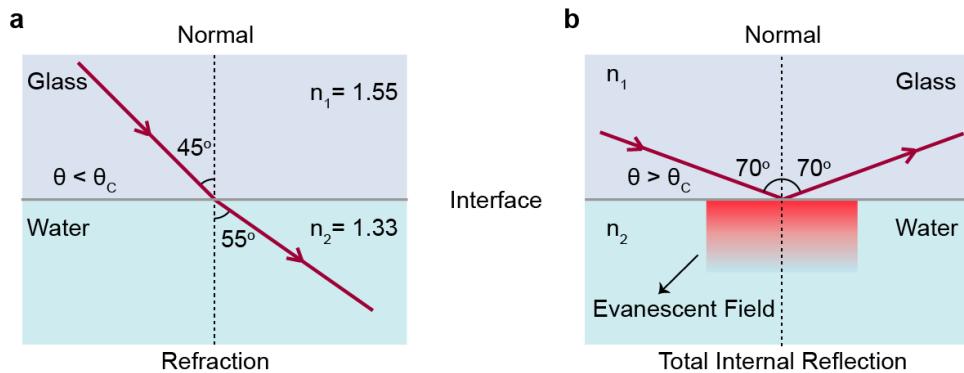


Figure 3.5. Schematic showing the propagation of a light beam from a glass medium to water with corresponding change of angle from the normal as a result of different refractive indices at the interface between the two media. (a) As the beam enters the medium with a lower refractive index, it propagates with a larger angle from the normal compared to the incident beam. (b) When the light reaches the interface at a greater angle than the critical angle, the beam becomes totally reflected at the interface and does not exit the first medium. However, an exponentially decaying evanescent field is formed in the second medium.

The energy of the evanescent field decays quickly after the interface with an exponential rate given in equation 3.7 with E_0 as the energy at the interface, z the distance from the interface and d the decay depth, which is defined in equation 3.8 with λ being the wavelength of the incident light. This thin evanescent field is still capable of exciting fluorophores that are attached or close to the surface [25].

$$E = E_0 e^{-z/d} \quad (3.7)$$

$$d = \frac{\lambda}{4\pi(n_1^2 \sin^2 \theta_1 - n_2^2)^{1/2}} \quad (3.8)$$

3.2.2. Prism-type TIRF Setup

Single-molecule TIRF microscopy is typically performed using one of three formats: waveguide-based [104], objective-type, and prism-type TIRF. These techniques take advantage of a thin evanescent field at the surface of the sample. In objective-based TIRF, the excitation light is projected through a high NA objective and the total internal reflection is achieved by adjusting the angle of off-axis beam exiting the objective to get a larger angle than the critical angle at the sample interface [105]. In waveguide-based TIRF the excitation light propagates in a planar optical waveguide. The evanescent field generated by waveguides is independent of the imaging pathway, enabling them to be used with any imaging objective on a standard microscope [106, 107]. In prism-type TIRF (pTIRF) in contrast to performing oTIRF with an objective that has a high numerical aperture (oTIRF), a prism is used to create the evanescent field and the emission gathered by the objective will be guided and focused to

the cameras. We have custom-designed and made the prisms from fused silica as the glass medium to have an angle of 72° on the incident surface as shown on [figure 3.6](#). The excitation laser enters the prism from the air along the normal and hits the interface of the prism and the aqueous solution at an angle greater than the critical angle and thus gets totally reflected inside the prism. The reflected light then exits the other side of the prism, which has an angle of about 45° so the light exits the prism at a 42° angle from the normal. The evanescent field penetrates the transparent solution below the lower prism surface where the molecules are immobilized.

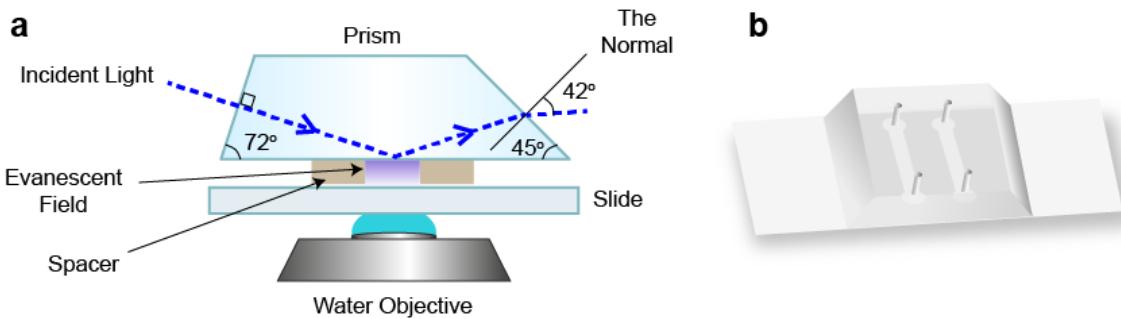


Figure 3.6. The design and assembly of the prism used as both a medium to create the evanescent field and a chamber to hold the single molecules of interest after surface passivation. (a) When the laser enters the prism medium from the right side across the normal, it hits the interface of prism and the buffer below and gets totally reflected towards the left side of the prism to exit in the end. (b) The schematic design of a prism attached to a coverslip. The prism has four holes and upon sandwiching a pre-cut nescofilm between the prism and coverslide, the two channels are prepared to insert the sample.

Many microscopy setups might not have the option of developing prism-type TIRF because it needs extra optical components installed above the microscope body. Hence, when the option is available, prism-type TIRF is preferred for three main reasons. First the excitation and emission paths are separated because the lasers are focused from above the prism surface and emission is gathered by the objective. In objective-type TIRF, both the excitation and emission lights are passing through the objective so extra dichroic mirrors have to be used before the objective to separate the two beam paths. OTIRF adds more complication to the setup alignment regarding to the excitation path. Another advantage is that, although surface passivation for the both cover slides and prisms takes the same amount of time and material, more experiments can be done using each prism because they have been designed with two or three separate channels, each one used for one whole experiment. The cover slides, on the other hand, support only one experiment run. Hence, from the time and cost points of view, passivating and using prisms is more efficient than cover slips.

The wide-field/TIRF setup used for the present thesis has the both options of objective- and prism-type TIRF depending on where the collimated excitation beam is placed. The schematic of the setup is shown in [figure 3.7](#). The alignment of the cameras are possible by mounting them on a custom-designed L-shaped block as shown in [Figure 3.8](#). The block allows for the fine adjustment of the camera's image capturing area with respect to other cameras by rotating the camera as desired. After making the final alignment, the camera can be fixed on the final position using the four screws on the holder. More details about the setup and its components can be found in the attached articles.

The laser-beam profiles at the bottom of the prism were determined by imaging a concentrated dye mixture of Atto488, Cy3B and Atto647N or Atto532 in the low micromolar range using a $10\times$ objective (Nikon, Plan Fluor air objective, NA = 0.30). The size of the fields-of-view on the different cameras were calibrated using a 1951 United States Air Force resolution test chart. Images were loaded into Fiji [108] to calculate the area of illumination. As the profile is an ellipse, the profile of the major and

minor axes were determined using the plot profile function in Fiji and fit to a Gaussian function. The area of the ellipse was calculated using:

$$A = \pi ab \quad (3.9)$$

where a and b represent the radius of the major and minor axes respectively. The major and minor axes were approximated by the width (σ) of the fitted Gaussian distribution.

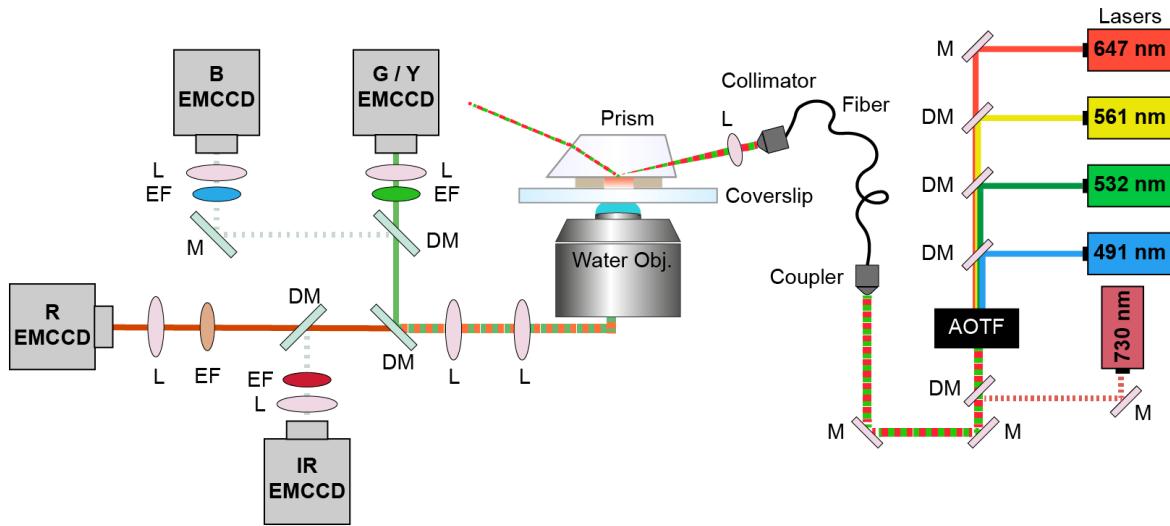


Figure 3.7. Schematic smTIRF setup with five excitation lasers and four detection channels. The setup is capable of both objective- and prism-TIRF techniques. All lasers (except for the IR laser) are coupled through an AOTF device for modulating the excitation with ALEX schemes. The IR laser is modulated separately. The final light enters a single-mode fiber and is guided to the sample either from the rear of the microscope body for objective-TIRF or to the top part of the body for prism-TIRF. The final fluorescence emission is spectrally divided using dichroic mirrors in the respective detection paths to reach the corresponding cameras after emission filters and focusing lenses.

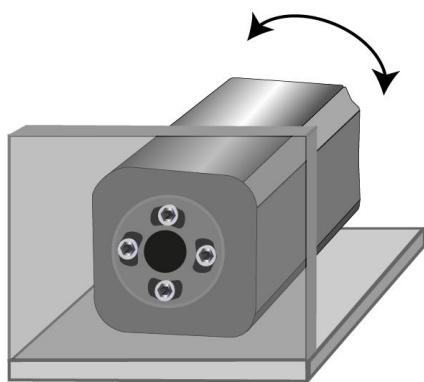


Figure 3.8. The schematic design of the L-shaped camera holder. The aperture of the camera is located on the hole designed on the vertical plane of the holder, which is shown transparent here for clarity. After aligning the detector with respect to other channels, the camera can be fixed to the holder using the four screws.

3.2.3. Surface-passivation of Prisms to Immobilize Single Molecules

Immobilizing the molecules on the surface is a common approach to increase the observation times to a couple minutes, which is extremely important for single-molecule TIRF microscopy as one wishes

to observe the dynamic behavior of a single molecule with time. However, immobilization is tricky as molecules might interact non-specifically with the surface causing negative impacts on the final results. To prevent such random interactions, the surfaces of the sample holder need to be passivated and, based on the assay to be performed, may even need to be encapsulated inside lipid vesicles.

There are a couple of known surface passivation and immobilization approaches including applying biotinylated BSA, poly(ethylene glycol) (PEG) conjugated to biotin, PEG with Ni^{2+} or Cu^{2+} ions and using click chemistry [109, 110]. The surface to be passivated must be thoroughly clean and then exposed to the passivation reagents. The presence of any impurities on the surface increase the background counts. Coating the surface with linear PEG is usually sufficient however, in the case of highly adsorbent proteins branched PEG are used to ensure the immobilization [111]. In this work, we applied polymer-coating for the prism surface (made of fused silica) to immobilize the molecules of interest and minimize any non-specific interactions with the surface.

The applied linear polymer is poly(ethylene glycol) (2-[methoxy (polyethyleneoxy) propyl]trimethoxysilane, mPEG, 6-9 PE units, #AB111226, abcr; Germany) and was mixed with an ~8 % fraction of biotinylated PEG (biotin-poly(ethylene glycol)-silane (biotin-PEG, MW3000, PG2-BNSL-3k, Nanocs, NY; USA) to provide an anchor for attaching the biotin-labeled molecules through high-affinity biotin-streptavidin binding with a K_D in picomolar range. The reaction between alkoxy silane and hydroxyl groups makes a strong conjugation of PEG to the surface of silica [111].

In order to clean the surface, prisms were first dipped in a 2 % solution of Hellmanex (Hellmanex™ III, Z805939-1EA, Merck, Germany) followed by 15 minutes sonication at room temperature. They were then washed three times with distilled water and then dipped in pure acetone (Aceton, 67-64-1, Merck, Germany) and sonicated for a 15 minutes. After rinsing with distilled water, the same sonicating procedure was repeated in ethanol (Ethanol, 64-17-5, Merck, Germany). The prisms were then rinsed with water again and dried using a strong flow of clean air. After placing them in a glass petridish, the surface was cleaned using a plasma cleaner (Diener Zepto plasma cleaner, Diener Electronic, Germany) for 20 min to remove any remaining organic impurities. A mixture of 350 μL PEG-silan and 25 μL PEG-biotin was made in 50 mL Toluene (Toluene 99.8%, 108-88-3, Merck, Germany) as the passivation buffer. The prisms were dipped in the mixture and placed in the oven (Heating cabinet VWR Dry-Line, 815095, VWR) for 4-5 hours at 55 degrees. Afterwards, the prisms were washed with ethanol twice and two rounds of 5-minute sonication was also done. After rinsing two times with distilled water, they were dried by blowing air. The passivated prisms are then directly assembled and kept in the freezer for usage in near future.

The assembly of clean and passivated prisms is done by sandwiching a piece of pre-cut nescofilm between the prism and a clean microscopy slide, which was also cleaned and passivated with the same mPEG polymer without adding any biotin. The cut on the nescofilm provides a channel for the sample molecules preventing them from leaking to other areas like the neighboring channel ([Figure 3.6b](#)). We then melt the nescofilm layer by putting the whole structure on a heater plate at a temperature around 100 °C in such a way that the slide is in contact with the hot plate. Upon cooling down, the molten layer acts as a glue for the two glass surfaces to keep them tightly together.

3.2.4. SmTIRF Experiments Using Alternating Laser EXcitation (ALEX)

FRET data analysis is rather involved because of various heterogeneities like inactive fluorophores or incomplete labeling of the structures of interest. Hence, if we excite the sample with only the donor

laser, we will lose important information about the acceptor. Especially for systems showing dynamic behaviors, it would be impossible to distinguish between a state transition or a fluorophore blinking event. In smTIRF measurements, usually alternating laser excitation (ALEX) with a millisecond timescale is used to get information about the photophysical state of individual fluorophores and the ability of molecule sorting over the resulting time traces. In this way, the fluorophores on the structure of interest gets directly excited and detected, and by adjusting the time gates between excitation pulses and data acquisition, the consecutive frames with their specific excitation laser lines are obtained. Another huge importance of using ALEX is that we get the photobleaching step of each fluorophore, which is necessary for determining accurate FRET efficiencies and the distances between the dye pairs by determining FRET specific correction factors.

3.2.5. Data Analysis for smFRET on the Surface

To obtain the single-molecule information from the immobilized molecules, a number of analysis steps are needed. A very detailed and precise molecule-wise data analysis is required to draw appropriate conclusions from the experiment. After mapping the detection channels with respect to a reference channel and loading the raw images into the software, the particles are located and the co-localized ones are detected. Each detected molecule in the smTIRF assay appears as a time trace with the intensity and FRET signal on the y axis. Measuring two or three labels on a biomolecule simultaneously using ALEX schemes produces multiple time traces for the same spot that are attributed to the presence of donor and acceptor molecules. Accordingly, we need to check every obtained trace for individual dye behavior and features. Another aspect of data analysis arises from the fact that we normally use more than one detection channel to record the emission intensities of the co-existing labels. This means we need to use some correction factors to account for the differences in the efficiency or sensitivity of the cameras towards different wavelength regions together with the differences in the quantum efficiency of the applied fluorophores. A detailed description of the mentioned steps for TIRF data analysis is included in the attached publications in the appendix, but they are briefly described here for completeness.

- [Channels Mapping](#)

The first step of data analysis after measuring smTIRF with multiple detection channels is to determine the mapping of the corresponding images captured from each EMCCD camera. Although the cameras are well aligned to cover the same field of view, the perfect alignment down to the image pixel size is not mechanically possible and we need to make the final image correspondence via aligning algorithms like transformation matrices that we apply to the corresponding coordinates between the channels. Proper images for mapping purposes should consist of bright spots that are almost evenly distributed over the whole field of view. For this purpose, we use images from zero-mode waveguides. This gives us enough points to cover the entire field of view and make sure we are considering the same molecule or structure in the different channel during the analysis.

- [Trace Extraction and Categorization](#)

After mapping the used detection channels, we need the single molecules to be detected and localized in every channel and co-localized molecules being recognized based on the transformation matrices. It is also important to have the local background level around each detected emitter, to have a more exact estimation of the measured signal. For this purpose, we consider a ring-shaped mask around every detected particle and calculate the local background over the whole number of frames. With this, the real frame-wise intensities from each fluorophore can be extracted according to the defined excitation scheme used to acquire the data. Then final time traces for the whole measurement time duration are extracted along frames.

We need to sort the time traces based on their quality and the information they contain. The categorization depends on the number of involved fluorophores and the information we are interested in. Generally, we make a trash category for the low-quality and noisy trajectories. The traces showing aggregations or artifacts are also sent to this category. The remaining useful traces will be further sorted based on the photobleaching steps to be used for the FRET efficiency histograms and determining the correction factors.

- [Calculating Correction Factors and Distances](#)

A very common motivation for applying FRET methods is to calculate intermolecular distances of a macromolecule or intramolecular separations among small molecules, which are precisely monitored with FRET when not exceeding 10 nanometers in either case. After measuring FRET efficiencies between dye pairs, one can determine the desired distances only after a precise correction of the FRET data. The first parameter to account for is, of course, the background level and then we need to determine the specific FRET correction factors, which include the direct excitation of the acceptor by the donor excitation laser (de), a small crosstalk from the donor emission into the acceptor channel (ct) and the detection efficiency (gamma factor, γ) to correct for the differences in dyes properties and cameras sensitivity towards detecting different wavelengths. When the accurate FRET efficiency is known using equation 2.13, the distance between the centers of any two fluorophores could be determined.

Special regions on traces are used to determine the correction factors, which contain a photobleaching step of either the donor or the acceptor. Depending on the order in which the fluorophores photobleach, different correction factors can be calculated. To calculate the direct excitation from smTIRF data, we need to find traces in which the donor photobleaches before the acceptor. After the donor photobleaching step, the remaining intensity in the acceptor channel after the donor excitation illumination is used to determine the direct excitation. From the ratio of the acceptor intensity upon donor's excitation laser to the acceptor intensity upon its individual laser excitation, the direct excitation is obtained for each trace over the whole category. Traces where the donor photobleaches first, can be corrected individually. Otherwise, the mean or the median value of direct excitation from the whole data can be applied to correct the remaining traces. On the other hand, spectral crosstalk is calculated from the traces in which the acceptor photobleaches before the donor. After the acceptor photobleaching, the emission left in the acceptor channel would be the portion of the donor emission that leaks into the acceptor channel. From the ratio of the donor emission before and after the acceptor's photobleaching step, the spectral crosstalk is calculated for the entire category and is applied to the donor's intensity for accurate FRET calculation. The same category is used to determine the gamma factor by using the ratio between the intensity differences of the acceptor and donor

emissions before and after the acceptor's photobleaching step. As before, the corrections can be done individually when possible and otherwise, the mean or median value can be applied.

Applying the three correction factors to a two-color experiment yields the accurate or corrected FRET efficiency formula as given in equation 3.10.

$$E = \frac{I_{DA} - ct * I_{DD} - de * I_{AA}}{\gamma * I_{DD} + I_{DA} - ct * I_{DD} - de * I_{AA}} \quad (3.10)$$

- [Dynamic Analysis Using Deep Neural Networks](#)

In a conventional way to analyze dynamic FRET data, the dynamic traces are typically fed into Hidden Markov models. The applied model goes through the traces to calculate the state values and transitions probabilities between any of the present states. Afterwards, it will digitalize the traces by the state values and the transition rates using the Viterbi paths. The kinetic rates and the corresponding dwell times are then obtained by fitting the dwell time distributions in the transition plots created by the model. To be able to perform such an analysis, we still need to have all the traces sorted manually to feed the useful and dynamic category into the HMM package.

Recently, the application of deep neural networks (DNN) that could be trained with a wide range of simulated data types, has changed the way we approach data analysis in various research fields. DNN has many advantages in making the whole process of data analysis fast and automatic. Data analysis is usually the limiting step and can be prone to human bias. In the case of single molecule FRET analysis, DNNs became well established and popular because they can analyze huge one-, two- and three-color data sets within a couple minutes. After having the single molecules detected and localized, and the time traces extracted, DNNs perform sorting, determining FRET correction factors and classifying the state transitions in dynamic trajectories. Comparing the kinetic results between DNNs and HMM gave satisfactory results [112-114].

4. Publications Overview

4.1. Deep-LASI: deep-learning assisted, single-molecule imaging analysis of multi-color DNA origami structures

This collaborative work is related to analyzing one-, two- and three-color single-molecule data especially obtained from surface-based smFRET techniques. Using machine learning tools, a deep neural network (DNN) has been created that can analyze smFRET data fast and free of human bias. The software was first trained with simulated time traces and then tested successfully on simulated and experimental data from various L-shaped DNA origami structures obtained by TIRF microscopy. The software helps with handling the high number of collected time traces together with performing an advanced dynamic analysis that has been conventionally done by HMM (hidden Markov model) algorithms.

Deep-LASI: deep-learning assisted, single-molecule imaging analysis of multi-color DNA origami structures by Simon Wanninger, Pooyeh Asadiatouei, Johann Bohlen, Clemens-Bässem Salem, Philip Tinnefeld, Evelyn Ploetz and Don C. Lamb. *Nat Commun* 14, 6564 (2023). <https://doi.org/10.1038/s41467-023-42272-9Motivation>.

4.1.1. Motivation

Single-molecule techniques empowered us to address biological questions on the nanometer scale without ensemble averaging and with the highest sensitivity. This was not possible before such techniques emerged. In particular, single-molecule FRET assays give us information regarding to the inter- and intra-molecular distances in the range of 2.5-10 nm and the underlying dynamic states. Hence, we can visualize the conformational states within complicated structures like proteins and investigate their functions by following the conformational changes [10]. To probe the interactions and functions of biomolecules that were measured with surface-based FRET techniques, one usually needs to acquire a huge amount of data collected as single-molecule time traces. To get the desired information from the traces, each single time trace has to be evaluated to extract the potential information from the measurement. The information involves the time span the molecule is photoactive, the intensity of the emission after direct excitation, the FRET intensity and even the leftovers of intensities after the molecules are photobleached, which is very important for calculating the correction factors. Analyzing this wealth of obtained data, however, has been the limiting factor for researchers because going through thousands of traces is very time-consuming. Moreover, the expected results we have for the system under study might affect the analysis steps we take leading to user bias [114].

The above-mentioned challenges are true for the most conventional approaches of two-color single-molecule FRET assays and it gets even more time-consuming for a three-color experiment regarding to the extraction of information from the collected data. Three-color FRET measurements are attractive because adding a third fluorophore to the system gives more dimensions for visualizing the structure and the underlying dynamics simultaneously. Among various research fields, deep-learning techniques have been developed for analyzing smFRET data as well to overcome the challenges of quantitative multi-color data analysis.

4.1.2. Contribution

After the deep neural networks were inserted into the Deep-LASI software by a colleague in our group (Dr. Simon Wanninger), the neural networks were trained by 200000 simulated time traces. However, to establish the use of DNNs in our software package and prove that it works properly on experimental data as well, we needed to analyze real data with the automatic approach and compare the analysis to that done by the conventional manual approach. To this end, I measured various designs of L-shaped DNA origami structures immobilized on the surface using pTIRF microscopy. After collecting single-, two- and three-color smFRET data with various structures, I extracted the single-molecule traces to be analyzed.

The next step was to analyze the different measurements manually meaning that, after extracting the intensity data, I went through the time traces one by one to sort the traces into categories of suitable or inappropriate (and putting the bad traces to a trash category). Also, within the good traces, I selected the useful regions (i.e. the regions before the photobleaching steps) to have the region of the traces added to the final FRET histograms. Correction factors were calculated based on the intensities before and after the corresponding photobleaching steps. Then, I summarized the apparent FRET, correction factors and accurate FRET after fitting their distributions. The next step was analyzing the kinetics of the system using the HMM packages also available in the Deep-LASI software. In the end, by having the FRET traces digitalized and the transition density plots formed, the dwell times for each discovered state and the transition rates among them were obtained.

After analyzing the two-color data as described, I analyzed three-color time traces, which needed the same approach with additional steps to include each dye pair present in the system. Since we need more categories to correctly analyze such data and address the correction factors for all dye pairs, a larger number of traces are needed for satisfactory statistics in the end. Consequently three-color data analysis takes even more time than a two-color case. After finishing the kinetic analysis, we compared the results from manual and automatic approaches that showed an acceptable overlap for the categorization together with the final values from the two approaches being very close. The time needed for a complete analysis between my manual and DNNs automatic one was, of course, not even comparable!

I have been coordinating a lot with Dr. Simon Wanninger to give feedback about the performance of the inserted neural networks in terms of missing possible functionalities needed for a complete analysis of smFRET data, and comparing the results from manual and automatic analyses.

4.1.3. Key Results

Deep-LASI was designed to take the intensity of FRET time traces as input gathered from different ALEX or non-ALEX experiments. It classifies every frame into a category. For two-color FRET data for instance, Deep-LASI makes in total six categories of traces: dynamic, static, noisy, artifact, aggregate and photobleached categories. The total number of categories varies depending on the number of input channels, i.e. the number of dyes (and alternating light sources) used in the experiment. Traces showing random artifacts, aggregates, or high noise levels are excluded from further analyses. Depending on user-defined thresholds, the output of the trace classifier will be sent for further analysis (on the selected traces and the corresponding regions on each trace). The dynamic sections of each trace will also be sent to a further state classifier network for transition detection.

The trace-classifier network was first trained using the corresponding one-, two- and three-color traces over a wide range of experimental conditions. For the state-classifiers, two states can be distinguished from multi-state trajectories with 98 % or higher precision while four states are predicted with the lowest precision of 86 % for the single-channel model. In the case of the state transition classifiers, the states can be identified with accuracies of $\geq 98\%$, $\geq 90\%$ and $\geq 78\%$ for two-, three- and four-state models respectively. We also compared the performance of the state classifiers with HMM and observed that, at low noise levels, the results are quite similar but at higher noise levels (up to 30 %), DNNs outperform HMM. After various benchmarking approaches, which are described in detail in the paper, the capabilities of Deep-LASI were shown using experimental smTIRF data on L-shaped DNA origami structures. The dynamic structures have a labeled single-stranded DNA pointing out of the origami surface, and together with two or three docking strands with different complementary nucleotides different number of states and kinetic rates were obtained. Then, based on our different designs, one-, two- and three-color experiments were conducted and analyzed. The one-color experiments were done on the two-color DNA origami structures but only the donor was excited using non-ALEX excitation scheme. In this case, the acceptor was serving as a quencher.

The photobleaching steps and the dynamic sections of the corresponding traces were found by the neural networks. Also, the correction factors were automatically determined depending on the photobleaching steps that were found. The analysis results were compared with the same data that were analyzed manually. Usually in a manual analysis, the separated regions where the molecules are active (for instance before and after a blinking event) are taken into account. Deep-LASI determines the correction factors at the first photobleaching step as well. In case of the non-ideal behaviors after the photobleaching steps, Deep-LASI would trash the whole trace as artifacts.

The correction factors from the manual and automatic analysis agree with each other with only about a 3% discrepancy whereas the apparent FRET histograms show an excellent correspondence ([figure 4.1](#)). The little differences in the accurate FRET efficiencies in the end are also rising from the correction factors values but overall the discrepancy between manual and automatic analysis is negligible. The kinetic analyzes from state classifiers vs HMM gave the same TDPs. A comparison of the dwell time distributions showed that Deep-LASI identifies fast and close to the frame time transitions more frequently than HMM.

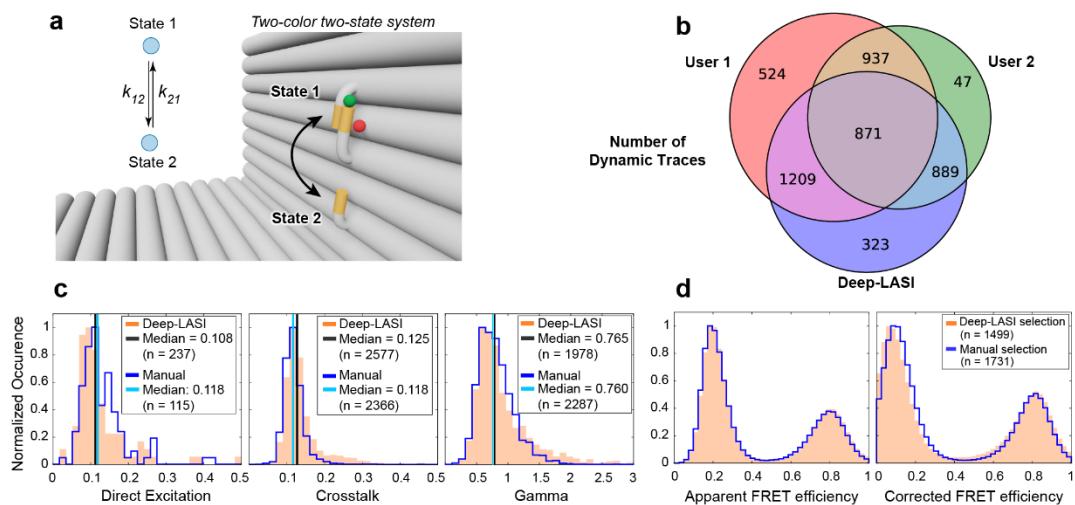


Figure 4.1. Benchmarking Deep-LASI software using L-shaped DNA origami structures. (a) Zoomed-in structure of L-shaped DNA origami structure labeled with Atto647N on the origami at the 12 o'clock position and Cy3B on the tether and corresponding kinetic scheme shown in the inset. The flexible tether contains a 7.5 nt overhang between the pointer and two single-stranded binding sites. Two FRET states with a high (12 o'clock) and a low

FRET (6 o'clock) efficiencies interconvert with rates k_{12} and k_{21} . (b) A Venn diagram comparing the number of categorized dynamic traces determined manually by two users and automatically by the DNN. User 1 tried to gather as much statistics as possible while user 2 was stricter about the quality of traces and added less traces to the dynamic category, while Deep-LASI was somewhat in between. (c) Trace-wise histograms of correction factors including direct excitation, crosstalk and detection efficiency derived automatically by Deep-LASI (orange histograms, median in black) and determined manually (blue lines, median in cyan). (d) Histograms of apparent (left panel) and corrected (right panel) frame-wise smFRET efficiency for 1499 dynamic traces from a total of 6100. The two states have corrected FRET efficiencies of 0.07 and 0.81. The histograms after Deep-LASI sorting are shown in orange and after manual selection in blue. Figure adapted from [113].

The software was then tested on advanced three-color FRET data. The same origami structure shown on [figure 4.1a](#) with a third dye (Atto488) attached to the 12 o'clock position was used to conduct three-color ALEX experiments. After sorting the traces and determining the correction factors, the same values for the YR corresponding FRET efficiencies were obtained (7 % and 81 %) together with two populations for BY (36 % and 81 %) and one broad population for the BR FRET pair (45 %) while the kinetics of the system showed the same transition rate values. Different binding overhangs from 6.5 to 8 nt were also measured and analyzed with a nice trend from fast to slow dynamics as designed by the origami structures (with dwell times of 0.39 s, 0.65 s, 1.90 s and 6.40 s for state 1, and 0.46 s, 0.99 s, 2.70 s and 9.5 s for state 2 respectively). Analyzing a more complex system with the three-color, three-state DNA origami is also described in the paper, which shows the same FRET efficiencies and kinetic rates for similar positions and overhang lengths but different values corresponding to the design of the third binding strand at the 9 o'clock position.

Finally, Deep-LASI was applied to other data types obtained from various biophysical assays beyond TIRF microscopy and from protein samples, which had been published before. For example, the conformational states of a mitochondrial heat-shock protein HSP 70 in yeast known as Ssc1 was examined for the influence of ADP. The nucleotide- and substrate-binding domains of the protein were fluorescently labeled, and the influence of ADP on the inter-domain separation was characterized [115]. Although the signal-to-noise ratio of this experiment was lower than that of photostabilized DNA origami structures data, the two underlying FRET states with the corresponding transition rates were correctly identified by Deep-LASI. The obtained results were in agreement with the docked and undocked conformations that had been previously reported. In addition, we used Deep-LASI to analyze data from confocal measurements on immobilized L-shaped DNA origami structures and was successful in determining the FRET efficiencies and the underlying kinetics. For details about the data and the analysis please refer to the paper in the appendix.

4.1.4. Discussion

Deep-LASI uses a deep-learning algorithm to automatically analyze time traces from different single-molecule assays. Developing Deep-LASI made it possible to extend the automatic two-color FRET analysis to one- and three-color data in a very fast manner that is independent from user assumptions and potential bias. In conclusion, Deep-LASI makes the complicated analysis of FRET data automatic and fast, which opens new possibilities for single-molecule assays in terms of gathering enough data to have proper statistics and higher throughput. Moreover, upon a proper training of the neural networks, the opportunity to address the rare events could be achieved with more measurements without being concerned about the time and effort of the analysis. We believe that the deep-learning approaches combined with single-molecule sensitivity will revolutionize data acquisition and analysis in the single-molecule field.

4.2. Deep-LASI, single-molecule data analysis software

In this work, we introduce Deep-LASI, a MATLAB-based software suite for analyzing single- and multi-color data from single-molecule experiments especially using FRET techniques. Deep-LASI gathers all the modalities needed for handling such data beginning from the raw data files until the final steps of visualizing and plotting the results in various possible ways. The related article has been published in the biophysical journal with an open access. The final edition of the article with a detailed description of the software interfaces and functionalities are provided in the corresponding appendix. This section summarizes the key features of the Deep-LASI software together with some results from multi-color FRET data obtained from smTIRF microscopy.

Deep-LASI, single-molecule data analysis software by Pooyeh Asadiatouei, Clemens-Bässem Salem, Simon Wanninger, Evelyn Ploetz and Don C. Lamb. *Biophys. J.* 2024, ISSN 0006-3495, <https://doi.org/10.1016/j.bpj.2024.02.013>.

4.2.1. Motivation

Single-molecule measurements revolutionized the way we examine the biological world by avoiding ensemble averaging and following the dynamics of individual molecules without the need for synchronization. Hence, the functions, interactions and underlying dynamic behaviors of single biomolecules can be obtained and precisely observed. FRET assays have established a wide range of applications to address crucial biological questions with the spatial resolution down to the Angstrom scale [116]. Since information is gathered from single molecules, a high number of molecules (usually a few thousand) have to be measured and examined to ensure proper statistics. Each measured molecule usually appears as a time trajectory with the desired parameter on the y axis like the fluorescence emission intensity or the FRET efficiency. Thus, the data analysis, in most cases, means going through the single traces one by one and extracting the useful information manually to build up the final results and graphs at a later step.

Having a trustworthy and user-friendly platform to perform a detailed data analysis for various one-, two- and three-color FRET assays was the main motivation of the present work. Deep-LASI (deep learning-assisted single-molecule imaging analysis) is an open-source software suite that helps with performing all the required steps from data extraction to final graphical results and outputs both manually and with an automatic machine learning approach. Another novelty provided by the Deep-LASI analysis package besides the automatic analysis features is related to the three-color data analysis, which is especially complicated when determining accurate FRET values for each dye pair. For such analyses, the amount of time needed for manual analysis and the potential errors arising from human bias can be considerable. The software provides separate interfaces and tabs to register multiple channels, read-in the raw data, extract the single traces, sort them into multiple useful categories, calculate the FRET efficiencies and available correction factors, plot the statistical information about the measurement features, plot various parameters and fit them, extract the kinetic information when dynamic states are present and determine the connectivity among them. Due to the versatility of single-molecule experiments, Deep-LASI also offers environments for the more expert users to train the incorporated neural networks for other types of time traces besides the FRET data and have the software established for their specific experimental assays.

4.2.2. Contribution

The early and first script of the software was developed by colleagues within our research group (the group of Prof. Don C. Lamb at the LMU Munich). After measuring two- and three-color FRET on dynamic L-shaped DNA origami platforms using smTIRF microscopy, under the supervision of Prof. Lamb and Dr. Ploetz, I analyzed the data with the both manual and automatic approaches on Deep-LASI to have a thorough comparison and vision about the system under study. It was necessary to actively use the analysis features of the software suite and give feedback about their accuracy and performance, so I have been coordinating a lot with my colleague (Dr. Simon Wanninger) to make the software more complete especially in case of encountering any missing functionality or feature. After having enough results to establish the software features and functions, I prepared a first draft of the article and the figures to make the use of the software more straightforward for a wide range of users. I was also involved in preparing an online user manual with detailed description and screenshots from all the encountering tabs and interfaces that the users could potentially handle more easily by following the provided steps. The documentation is available online at <https://deep-lasi-tutorial.readthedocs.io/en/latest/documentation.html>.

4.2.3. Key Results

The Deep-LASI software suite offers manual and automatic analysis platforms for one-, two- and three-color single-molecule experiments including: mapping the used detection channels, detecting and localizing the particles on each channel, trace extraction based on the performed excitation scheme, background calculation within a local mask around each molecule, trace categorization and selection of the individual informative time regions, determination of the correction factors based on the photobleaching steps, plotting the histograms of various related parameters and kinetic analyses. There are also environments for data simulations and training of the similar time traces. All the mentioned analysis steps are schematically summarized in the overview figure of the software in [figure 4.2](#).

When measuring with more than one detection channel, which is typical in advanced FRET studies, alignment between the cameras regarding to any discrepancy in shift, magnification and rotation needs to be performed taking one channel as the reference. Deep-LASI uses a phase-correlation algorithm to create a transformation matrix based on the coordinates of well-distributed spots and performs the geometric alignment. The created mapping matrix is then applied to map the respective coordinates between channels and make sure that detected and localized particles are within the same labeled structure. The mapping function is only used to find the corresponding pixels in the various detection channels corresponding to the locations in the reference channel. The actual single-molecule analysis is performed separately on the raw data and images, which no mapping is applied on them.

After mapping, the recorded raw data are loaded to have the time traces extracted from each frame based on the excitation scheme and ALEX order that was used for the measurement. To localize the detected particles, a couple of useful techniques are used: Wavelet, intensity thresholding and regional maxima. They all offer the user a sensitivity threshold to determine the center of mass for each particle. After localizing the particles between all channels, particle positions and intensities are

linked together to make the final time trajectory. The extraction happens on the same order of the selected and loaded files from all involved detection channels based on the ALEX scheme.

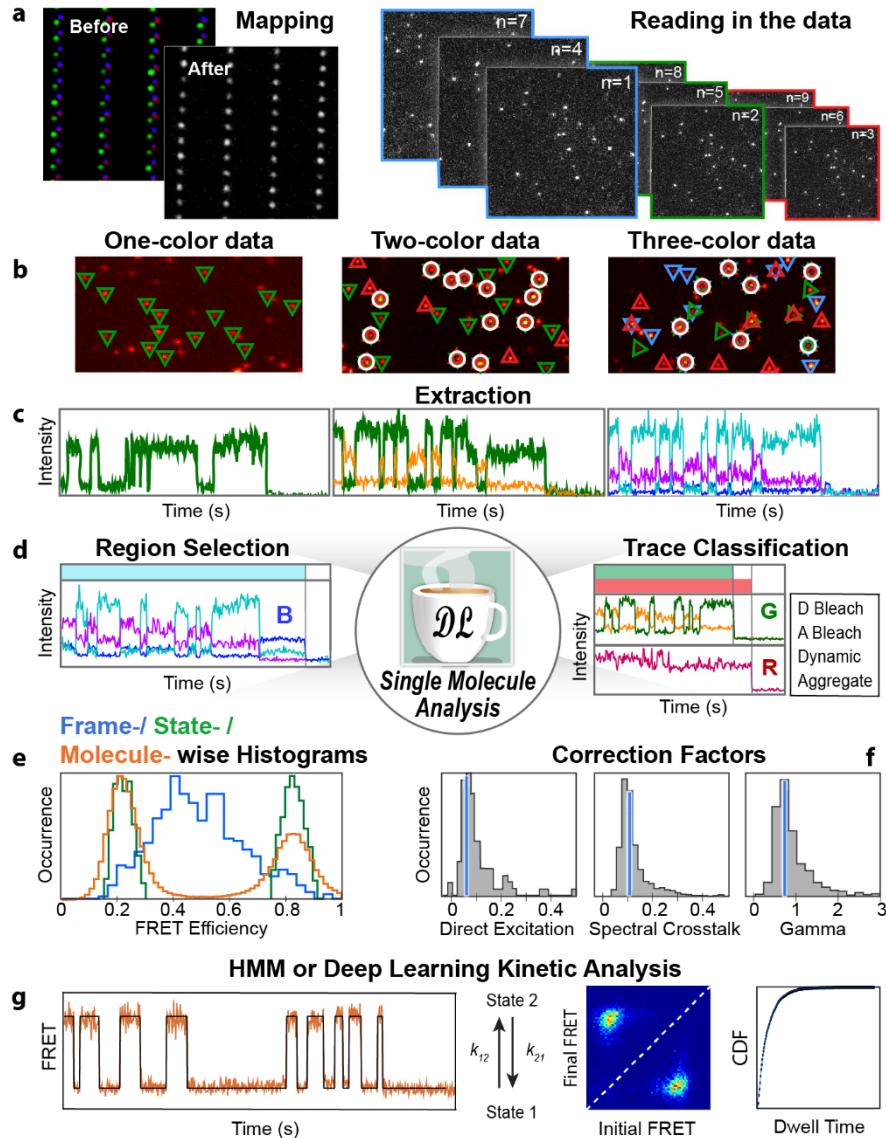


Figure 4.2. Deep-LASI software overview. (a) The mapping process is needed when multiple detectors are used, which is done with a distribution of spots like the zero-mode waveguide image shown here. Reading in the image files is shown with the frame order specified by the ALEX scheme. (b) Single molecules are detected and (co-)localized in the mapped images. (c) The intensity traces are extracted for each individual fluorophore using a particle and background mask. (d) Sorting the traces and selecting the useful regions before photobleaching steps can be done manually or automatically. (e) The results can be plotted with different type of histograms. (f) The correction factors necessary for corrected FRET efficiencies are calculated and plotted for the whole measurement. (g) Dynamic analyses by HMM or DNNs digitalize the FRET traces and build TDP plots for obtaining the exact kinetic rates. Figure taken from [112].

The background calculation and correction is done locally within a mask surrounding the point-spread function of the detected emitter. It is important to have a suitable distribution of the single molecules over the field of view as otherwise the particles emission could interfere with the background

calculation. After the time traces are extracted, the analysis can be performed manually or automatically for the detailed steps mentioned earlier. The dynamic traces can be further analyzed via HMM packages or deep neural networks; both approaches give similar results after estimating the underlying conformational states in time traces. Both approaches result in transition density plots based on the transition probabilities between any two initial and final FRET states. The transition rates can be calculated for each of the transition clusters on the plot and the kinetic rates determined by fitting the dwell time distributions.

The application of the software was demonstrated with three-color FRET experiments on L-shaped DNA origami structures with three FRET states as shown on [figure 4.3](#). The flexible tether was labeled with Cy3B and can bind to either of the protruding strands placed at positions referred to as 6, 9 and 12 o'clock. The binding sites consist of ssDNA strands containing a complementary sequence of 7 nt length at 6 and 12 o'clock and 7.5-nt length (meaning a strand with 8 nt containing a 1 bp mismatch) at 9 o'clock. Binding of the tether occurs by spontaneous base-pairing with the single-stranded protruding strands. The three-color FRET system was introduced with two additional fluorophores, Atto488 and Atto647N on the structure close to the docking strands at the 6 and 12 o'clock positions, respectively.

The ALEX scheme of BYR (Blue-Yellow-Red) excitation was used to excite the sample and three EMCCD cameras were used to record the data. The 8000 obtained traces from co-localized molecules were categorized using the automatic analysis resulting in 740 dynamic traces. After recognizing the photobleaching steps and useful trace regions, the three correction factors for accurate FRET calculations were obtained (spectral crosstalk, direct excitation and the detection correction factor). The results of apparent and corrected FRET efficiencies are shown. The three apparent FRET efficiency were best resolved for the BY FRET pair with values of 0.18, 0.73 and 0.48 corresponding to states 1 to 3. The distribution of dwell times were obtained from the TDP plots as 0.65, 0.69 and 1.40 s for states 1, 2 and 3 respectively.

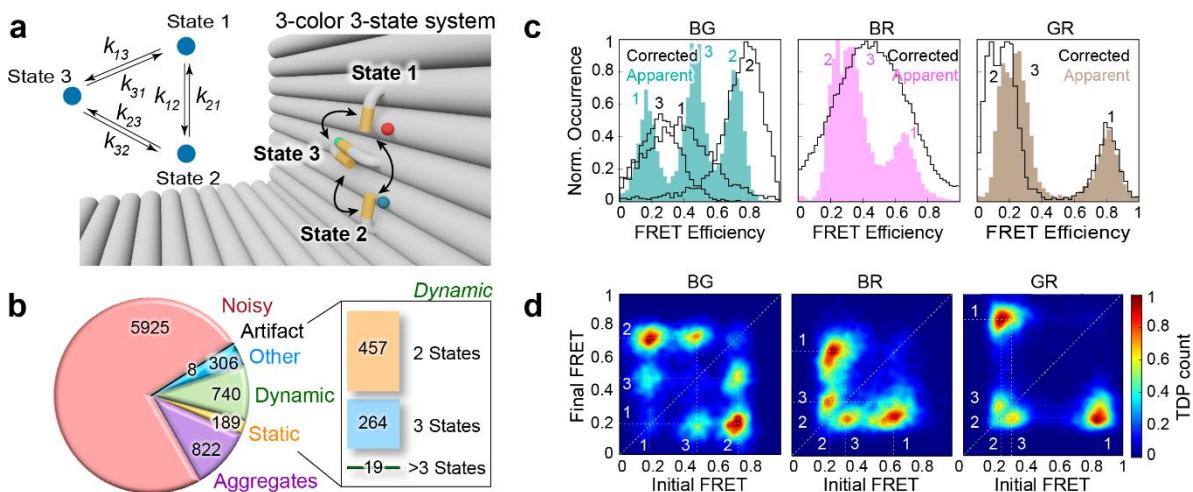


Figure 4.3. Using Deep-LASI software to analyze three-color three-state smFRET data after TIRF microscopy. (a) Schematic structure of L-shaped DNA origami with three binding positions at 6, 9 and 12 o'clock positions for the flexible tether. The three labels are Atto488 at the 6, Atto647N at the 12 o'clock position and Cy3B on the protruding tether. (b) The pie chart shows the distribution of single-molecule traces into various categories performed by the automatic trace sorting. (c) Apparent frame-wise and corrected state-wise FRET efficiency histograms for each dye pair are shown. (d) TDPs are shown for each dye pair with the detected transitions as the clusters of populations. Each cluster can be selected for dwell time fitting and calculating the kinetic rates. Figure adapted from [112].

4.2.4. Discussion

After upgrading our home-built TIRF microscopy setup to support three-color FRET experiments, our analysis software needed the corresponding expansion to perform the required data analysis. The open-source Deep-LASI software allows for straightforward data analysis of single-molecule time traces obtained from various assays. Each analysis step can be performed manually or automatically by trained neural networks. The software also offers environments for simulating data and training neural networks for a special data format based on the assay and the questions to be addressed.

The software suite allows for detailed analysis steps of multi-channel registration, single-molecule trace sorting and categorization, determination of photobleaching steps for each fluorophore, calculation of the correction factors and kinetic analyses. It is also possible to visualize the results by plotting FRET efficiency histograms, correction factors and dwell times. Various fitting models are also provided to fit the resulting distributions and obtaining more exact values using the fitting model parameters.

4.3. Distance and Kinetic Tunability of Dynamic DNA Origami Structures Examined at the Single-Molecule Level

In this work, we focus on to what extent DNA origami nanostructures are programmable compared to the designed structures, especially when they are labeled with fluorophores and/or are used as dynamic systems. We describe some crucial aspects one needs to consider when designing and functionalizing DNA origami structures, selecting the fluorophores and labeling strategies of origami structures to build proper tools and follow the desired interactions and molecular mechanisms accurately. With a friendly collaboration with the group of Prof. Philip Tinnefeld (at the LMU Munich), we show many interesting and useful details related to the DNA nanotechnology that are mostly overlooked due to the high complexity of related data acquisition and analysis. The related article has been recently submitted at the archive and will be submitted for publication shortly. The current edition of the manuscript is provided in the corresponding appendix at the end of this document. This section summarizes the key results from studying various versions of two- and three-color L-shaped DNA origami structures.

Distance and Kinetic Tunability of Dynamic DNA Origami Structures Examined at the Single-Molecule Level by Pooyeh Asadiatouei, Johann Bohlen, Fiona Cole, Fabio Morella, Philip Tinnefeld, Evelyn Ploetz and Don C. Lamb. The manuscript is under preparation.

4.3.1. Motivation

DNA origami platforms with two- or three-dimensional structures offer user-friendly features such as programmability, flexibility, biocompatibility, stability and easy tailoring compared to other biomolecules. These characteristics have made DNA structures as well-studied and precise building blocks with a wide range of applications in various fields of natural sciences. Dynamic nanostructures programmed to undergo controlled conformational changes found important applications in therapeutics, drug delivery, DNA-based logic gates and robotics. Designing the desired motions mainly include transient DNA strand displacement or application of external factors like surrounding ions and biomolecules, effects of light and pH changes.

Since labeling DNA origamis with fluorophores is performed precisely, studying their real-time performance and kinetics by Förster resonance energy transfer (FRET) incorporating alternating laser excitation schemes (ALEX) became very common and desired. Additionally, FRET methods in the single-molecule level reveal the structural heterogeneity among the nanostructures, rare events in dynamic steps and faulty or non-functional structures. These characterizations are crucial for the optimization of the nanostructures and the devices constructed based on them. Hence, the two fields of DNA origami nanotechnology and single-molecule multi-color FRET have been mutually advancing.

Incorporating small fluorescent molecules to DNA origami structures might introduce potential side effects and impacts on the energy landscape of the origami and their conformational activities that need to be carefully considered when designing the assay. For example, some fluorophores might induce dye-DNA interactions and dye-dependent conformational changes. Usually, terminally-labeled fluorophores stabilize hybridization by both stacking and electrostatic interactions when attached to DNA, and positively-charged dyes show a bigger effect than negatively-charged ones. These possible fluorophore-DNA interactions however, do not have a big effect on the reliability and robustness of

single-molecule fluorescence-based methods, and techniques using labeled DNA structures are still extremely powerful tools to address our biophysical questions.

Considering the possible interactions of DNA strands and origami structures with surrounding molecules or the effects of external factors such as temperature, salt concentration and surface effects, one needs a broad background and knowledge about DNA origami structures and incorporating fluorophores, in order to design reliable nanodevices based on them. With this motivation, we performed single-molecule FRET experiments both in solution and on the surface to obtain a thorough understanding of the labeled structures and discuss some critical aspects of dynamic DNA origami from designing to labeling. Various L-shaped DNA origami structures having two or three fluorophores and conformational states were designed and studied in terms of FRET efficiency between multiple dye pairs, kinetics of hybridization and dissociation with different number of overhangs, lifetime and anisotropy of each single fluorophore and distances between labeling positions.

4.3.2. Contribution

After designing the desired DNA origami platforms for two- and three-color FRET experiments, the assembly of over 25 origami structures were performed in the group of Prof. Philip Tinnefeld (LMU Munich) by Dr. Johann Bohlen and Fiona Cole. I measured the static and dynamic L-shaped DNA origami platforms using both wide-field TIRF and confocal microscopy i.e. smTIRF and MFD-PIE microscopy under the supervision of Prof. Don C. Lamb and Dr. Evelyn Ploetz. I analyzed the single-molecule data from both the surface- and solution-based measurements to have a thorough vision and comparison about the various structures under study. I was also involved in extracting the relevant data from the measurements on designed new experiments. After having enough information for publication, I wrote a first draft of the manuscript and prepared the figures.

4.3.3. Key Results

DNA origami structures with an L-shaped design were used throughout this work. The flexible pointer, is a single-strand DNA labeled with Cy3B and, in a couple samples, with Atto542. Two or three docking strands with different complementary sequence lengths were designed to vary the number of states and kinetic rates. The other two fluorophores (Atto488 and Atto647N) were bound to the origami surface close to the two binding strands at the 6 and 12 o'clock positions for the V1 design or the switched positions for the V2 structures ([Figure 4.4](#)). The kinetics of the tether were examined using four different complementary sequence lengths of 7 nt with one nucleotide mismatch (6.5 nt), 7 nt, 7.5 nt and 8 nt for the both V1 and V2 designs.

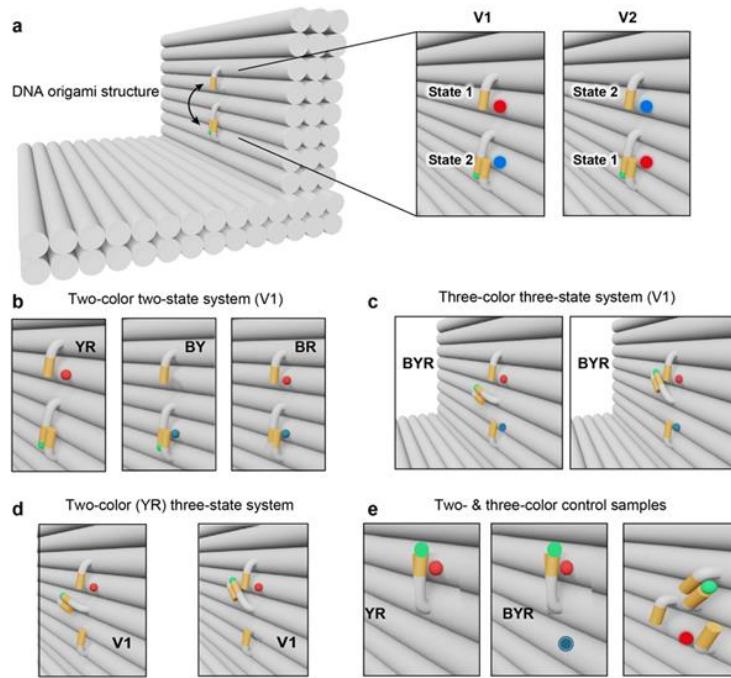


Figure 4.4. Schematic overview of the L-shaped DNA origami structures used in this work. (a) A flexible tether labeled with Cy3B extrudes from the structure surface and can transiently bind to either of the docking strands at the 6 or 12 o'clock positions. The two versions of the structures were designed to include the Atto488 and Atto647N fluorophores attached at the 6 and 12 o'clock positions (V1) or vice versa (V2). (b) Schematics of the two-color variants of the origami structure in panel a with two transient binding states for the V1 samples showing the YR (left panel), BY (middle panel) and BR (right panel) samples. (c) Schematics of three-color, three-state DNA origami structures for the V1 labeling design with binding strands at the 6-9-12 (left panel) and 6-11-12 positions (right panel). (d) Two-color versions of the three-state origami structures shown in panel c. (e) Additional DNA origami structures analyzed including two- (left panel) and three-color (middle panel) control samples with no binding strands around the tether shown for the V1 samples. The right panel shows a three-state system where the binding sites are symmetrically distributed about the tether.

The two-color, two-state samples were all made to have 7.5 nucleotides complementary base pairs. However, the three-state origami structures with the additional docking strand like the ones shown in Figure 4.4c including the 6-9-12 and 6-11-12 binding positions contain 7.5 nt on the third binding strand at both positions of 9 and 11 o'clock positions. For these structures, we reduced the binding sequence lengths of the 6 and 12 o'clock positions from 7.5 nt in the original two-color samples to 7 nt to examine the kinetics tenability of our design. The symmetrically placed binding strands at the 4-8-12 o'clock positions (Figure 4.4e) contain an overhang length of 7 nt. All samples were measured using the both TIRF and MFD-PIE setups except for the TIRF-PAINT measurement and the two-state samples in which the Cy3B on the tether was replaced with Atto542. These constructs were only measured on a passivated surface using TIRF.

The immobilized two- and three-color, two-state DNA origami structures were measured with prism-TIRF microscopy. The FRET efficiencies and the kinetics of the system were measured for samples with 6.5, 7, 7.5 and 8 nt complementary bases. The obtained FRET efficiencies were corrected with FRET correction factors to calculate the distance between labeling positions and compare them with the expected values from the designs. Dynamic switching between the different docking strands is visible in the single-molecule intensity and FRET traces. When the BY FRET efficiency is high the YR efficiency is low and vice versa. The kinetic analyses were carried out using Deep-LASI and the rates determined

from the resulting TDPs by fitting the distribution of dwell times for the different states to exponential-models (Figure 4.5).

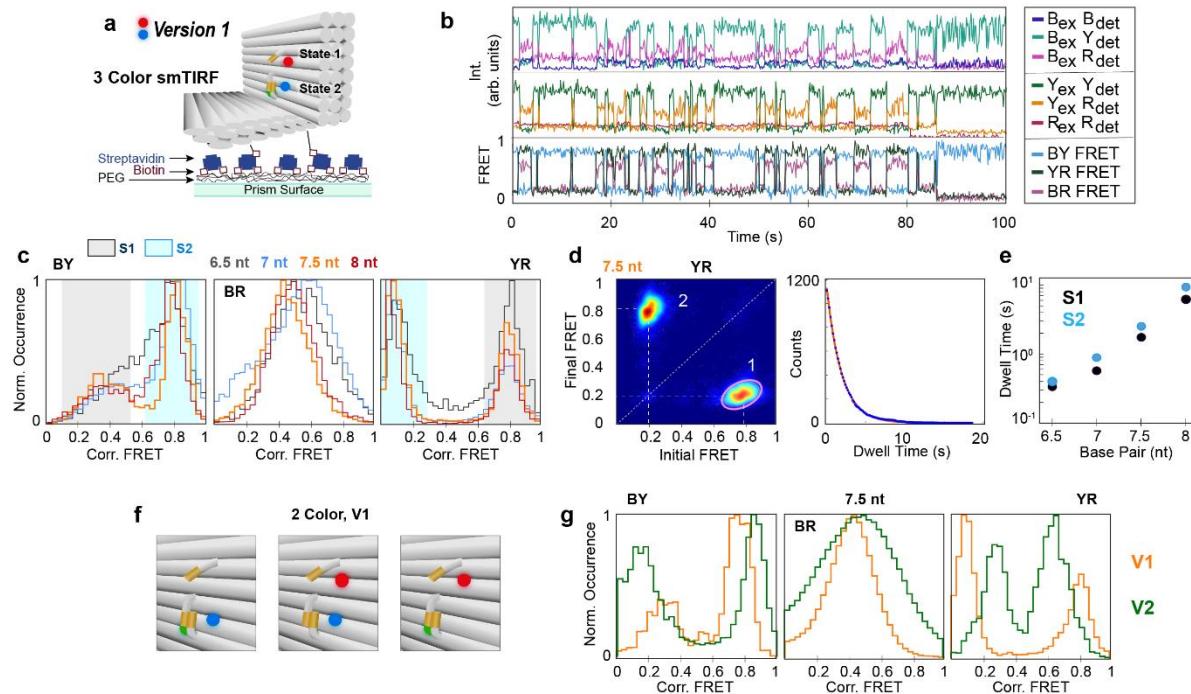


Figure 4.5. Single-molecule TIRF results for the two- and three-color, two-state V1 L-shaped DNA origami structures. (a) A schematic of the smTIRF assay for measuring immobilized DNA origami structures on the passivated prism surface. (b) Single-molecule intensity and FRET traces showing the dynamic switching between different binding strands. Top panel: The intensity traces of blue, yellow and red dyes after blue excitation. Middle panel: The emission of yellow and red dyes after yellow excitation together with the red emission after red excitation. Third panel: FRET traces for all three dye pairs. (c) Frame-wise-weighted state-wise averaged corrected FRET efficiency histograms with three-color ALEX excitation for each dye pair of BY (left panel), BR (middle panel) and YR (right panel) for complementary strands of 6.5 to 8 nucleotides. (d) An overview of the kinetic analyses. Left panel: A transition density plot showing the interconversion events for the YR dye pair. Right panel: A distribution of dwell times for the individual transitions calculated (by selecting the individual clusters). (e) Dwell-times for S1 and S2 as a function of complementary nucleotide lengths on the docking strands for the V1 structures. (f) Zoomed-in schematics of two-color, two-state origami structures containing only one dye pair at a time (Only V1 samples are shown here). (g) Frame-wise-weighted state-wise averaged corrected FRET efficiency histograms for each dye pair for the two-color structures containing 7.5 nt complementary base pairs of V1 and V2 samples. The dynamic BY and YR pairs show two populations.

The same samples were measured in solution to get information about the fast kinetics of the samples and detailed vision about the individual fluorophores' lifetime and their rotational freedom along the labeling axes. The results from FRET efficiency and histograms fitting after corrections were very close to the results we obtained from TIRF microscopy. The information about the dyes lifetime and anisotropy decays helped us understand about the possible interactions among the binding strands and the applied fluorophores. For example, we saw that Atto488 has the lowest residual anisotropy among the dyes implying that it has the most rotational freedom. Atto647N in contrary shows a very high anisotropy that proves this dye is sticking to the surface without proper rotations along its linker axis.

Single-molecule TIRF and MFD-PIE results from various sample designs showed a bias between the two docking strands from their kinetics point of view. The dwell time of state 2 (6 o'clock position) was observed to always be longer than state 1 (12 o'clock position). This observation prompted us to

design two- and three-color control origami structures without any docking strands around the tether. The histograms showed an average FRET efficiency value between the two original values corresponding to states 1 and 2 with almost no bias regarding the favored orientation of the pointer. We concluded that the observed difference in the kinetic rates of S1 and S2 are not as a result of inherent positioning of the tether but related to the uneven accessibility of the docking strands. We tested another approach to check the two binding positions accessibility using PAINT assay. We used the same two-state DNA origami structure (6-12 o'clock position, V1) without any tether, and instead let the 8nt-long tether labeled with Cy3B be free in the chamber and act as an imager strand. The corresponding results proved that our assumption about the harder accessibility of the top binding strand was actually true.

Another interesting factor that we examined was the effect of incorporating different fluorophores either on the flexible pointer or changing the red dye from Atto647N to Cy5 attached to the origami structure. We realized that the overall interactions and the kinetic rates are indeed affected by different fluorophores with different (photo)physical properties. All these contributing factors should be considered carefully when designing origami platforms and addressing desired questions.

4.3.4. Discussion

With this work we show DNA origami nanostructures are tunable and programmable from various designing aspects related to kinetics and distances, allowing researchers to assemble various types of nano-devices based on specific questions at hand. We showed the kinetics of L-shaped DNA origami structures changed in a controlled manner by tuning the number of base pairing; each mismatching nucleotide changed the kinetics by almost a factor of two. However, our data suggests that various interactions between each dye pair and between a fluorophore and DNA region can influence the kinetics of the tether. Hence, we recommend to choose proper fluorophores that does not effectively interact with the environment especially when designing dynamic platforms.

5. Conclusion and Outlook

The present dissertation provides a comprehensive exploration of multi-color single-molecule FRET (Förster Resonance Energy Transfer) principles and techniques both on the surface and in solution applied to two- and three-color DNA origami nanostructures. The high tunability of DNA origami structures coupled with the high sensitivity of single-molecule assays resulted in two important directions that shaped the entire path of this thesis.

The first major research direction, which involved extensive collaboration among various researchers, focused on integrating existing methods and developing a novel approach to create a robust and user-friendly software package. Analyzing one-, two- and three-color single-molecule data has been very time-consuming and increasingly complicated particularly when dealing with systems exhibiting underlying dynamic states. These challenges have discouraged researchers from pursuing three-color approaches especially for the surface experiments. However, the development of Deep-LASI (Deep-Learning Assisted Single-molecule Imaging analysis), a software suit empowered by deep neural networks provides both manual and automatic analysis environments, and drastically reduces the effort and time required for such analyses. Accordingly, we anticipate that, in the near future, three-color experiments will become increasingly desirable, enabling researchers to extract more spatial and dynamic information from complex biological systems.

In order to benchmark and validate the Deep-LASI software to be more universal, a huge amount of data from smFRET experiments were gathered ranging from one- to three-color experiments with various formats of both static and dynamic conformations. The high assembly efficiency of DNA origami structures facilitated obtaining a high amount of statistics from single-molecule time traces. By measuring various assemblies of origami structures and analyzing their features in detail, we observed intriguing patterns in the data that guided us to the second direction of this thesis; characterizing origami structures and the influence of the individual fluorophores attached to them.

Leveraging the high sensitivity of single-molecule FRET techniques to unravel the heterogeneities within a large population of single detected particles, we gained interesting insights into the behavior of origami assemblies and the protruding strands, particularly in their interactions during complementary hybridization and dissociation. Additionally, we observed interactions of the attached fluorophores with the DNA tether and docking strands. Various fluorophores, depending on their net electrical charge and hydrophobicity exhibited different interaction patterns with other fluorophores and the DNA structure backbone. These factors affected the kinetics of natural transient association and dissociation of the DNA strands. Consequently, our findings offer valuable insights for designing DNA origami structures and their labeling strategies.

As shown earlier, we also demonstrated that the kinetics of DNA strands transiently binding and unbinding are highly tunable by manipulating the number of complementary base pairs [117]. Depending on the immediate environment, even a single mismatching nucleotide can alter the system's dynamics by a factor of two or three. This knowledge has broad applications in the design of dynamic systems across various fields such as sensing, drug delivery, and therapeutics, which are of significant interest to the DNA origami community. Moving further, we characterized the distance between the fluorophores with sub-nanometer precision using corrected FRET efficiencies. We report on how closely the determined distances aligned with the theoretical values that one would expect based on the design, assembly and labeling of the DNA origami structures.

The advances made in this research will have profound implications for the broader fields of nanotechnology, molecular biology and biotechnology. The Deep-LASI software suit holds the potential to revolutionize the analysis of complex multi-color single-molecule experiments, facilitating the approaches of three-color and beyond. This will enable researchers to explore intricate molecular interactions with unprecedented resolution. The insights gained from the tunability of DNA origami structures and fluorophore interactions will provide a foundation for future applications in medicine, molecular diagnostics, and advanced materials science. As DNA origami technologies continue to evolve, the principles and methodologies developed in this dissertation will drive innovation and expand the possibilities for dynamic molecular systems and nanodevices in a variety of cutting-edge fields.

The integration of smFRET techniques with DNA origami nanostructures provide a powerful platform for the future of molecular-scale biophysics, engineering and computing. Since the field of DNA origami continues to provide unique precision in the spatial arrangement and motions of molecules, its combination with smFRET will enable even more sophisticated analysis of molecular interactions, dynamics and conformational changes with high temporal and spatial resolution. Moreover, ongoing improvements in deep learning algorithms and analysis tools like Deep-LASI, makes it obvious that multi-color smFRET experiments will become more accessible and adaptable, allowing researchers to monitor complex, multi-state biological processes in real time.

Appendix 1: Single-molecule FRET Measurements on DNA Origami Structures Labeled with an IR dye

Since, in the published data using the DNA origami structures, the measured intensity of the blue dye (Atto488) was always the lowest compared to other present fluorophores like Cy3b and Atto647N, we decided to shift the spectral regions of the three labeling dyes from blue-red to yellow-infrared. The blue dye at the 6 o'clock position corresponding to the V1 design was replaced with an IR dye, namely Cy7. To be able to measure in the IR region, an IR laser was added to the excitation path as shown in [Figure 3.7](#), and a fourth camera was added to the detection path. The IR light was guided to the camera using the proper dichroic mirror and emission filter.

To have a proper comparison between the new samples and the original DNA origami designs (the published structures), three, three-color samples labeled with Cy3B on the tether, Cy7 at 6 o'clock and Atto647N at 12 o'clock positions were assembled and measured ([Figure A.1a](#)). The difference between these three samples was the number of complementary nucleotides between the tether and the binding strands: 7, 8 nt with 1 mismatch and 8 nt to cover a range of kinetics. Moreover, two, two-color DNA origami structures labeled with Cy3B-Cy7 and Atto647N-Cy7 using the V1 design were measured ([Figure A.1g](#)). The two-color samples were prepared with 7 complementary nucleotides between the tether and the binding strands. On the one hand, we observed that the intensity and data quality of Cy7 was even lower than those of Atto488. On the other hand, upgrading the TIRF setup to the IR region and the corresponding experiments were a part of this PhD work. Since the mentioned development is the first step towards four-color assays, we decided to include the measurements in this section of the thesis.

The sample preparation and experimental steps were all similar to other measurements on the smTIRF setup, which are described in detail in the appended publications. Due to the very quick photobleaching of Cy7, it was necessary to use an extra ingredient (Cyclooctatetraene) to the imaging buffer trying to make Cy7 more photostable. To photostabilize the attached fluorophores on the DNA origami structure, an oxygen scavenging system based on PCA/PCD was used [118]. The photostabilization buffer was prepared as follows: 1 μ L of 100 mM Trolox/Ethanol solution was added to 95 μ L storage buffer (1x TAE, 12.5 mM MgCl₂ and 1 mM EDTA, pH=8.2). The mixture was then aged by 254 and 366 nm UV radiation using a UV lamp (M&S Laborgeräte GmbH, UVAC-6U, 2 \times 6 W) until an equal ratio of Trolox and Trolox-quinone was formed (typically 6 minutes) [119]. The ratio was determined using the two absorbance peaks of Trolox and Trolox-quinone measured on a nanodrop spectrophotometer (PeqLab Biotechnologie GmbH).

Immediately before starting the TIRF experiments, 2 μ L COT (Cyclooctatetraene, 98 % solution, Sigma-Aldrich), 1 μ L of 100 mM PCA in methanol and 1 μ L of 100 mM PCD solution (50% glycerol, 50 mM KCl, 100 mM Tris HCl, 1 mM EDTA-Na₂ \times 2H₂O, pH = 8) were added to the 97 μ L mixture of aged Trolox and storage buffer to obtain the photostabilization or imaging buffer. After flushing the sample chamber with the photostabilization buffer, it was sealed using removable sticking rubbers to allow the enzymatic oxygen reaction happen without the entry of additional oxygen into the system. A minimum of 5-minute waiting time preceded the data recordings. The photostabilization buffer was refreshed every 45 minutes until the end of the experiment.

The two- and three-color FRET experiments were carried out using msALEX [15], i.e. two- or three-excitation lasers were alternated framewise. The lasers of different wavelengths were synchronized using an acousto-optical filter (OPTO-ELECTRONIC, France) with the camera frame rate using an FPGA

that synchronizes the excitation and simultaneous detection on the EMCCD cameras at a 50 ms exposure time for both 2000 (two-color) and 2400 (three-color) frames. The IR laser was modulated directly as the wavelength was outside of the specified range of the AOTF. The control software for data collection was updated for controlling the IR laser by Frank Mieskes and Bässem Salem. The laser powers were set to 15 mW for the yellow and red laser lines (561 nm and 640 nm) and 10 mW for the IR laser (730 nm, Toptica Photonics AG, Germany). Both two-color (Y-IR or R-IR) and three-color (Y-R-IR) excitation schemes were used.

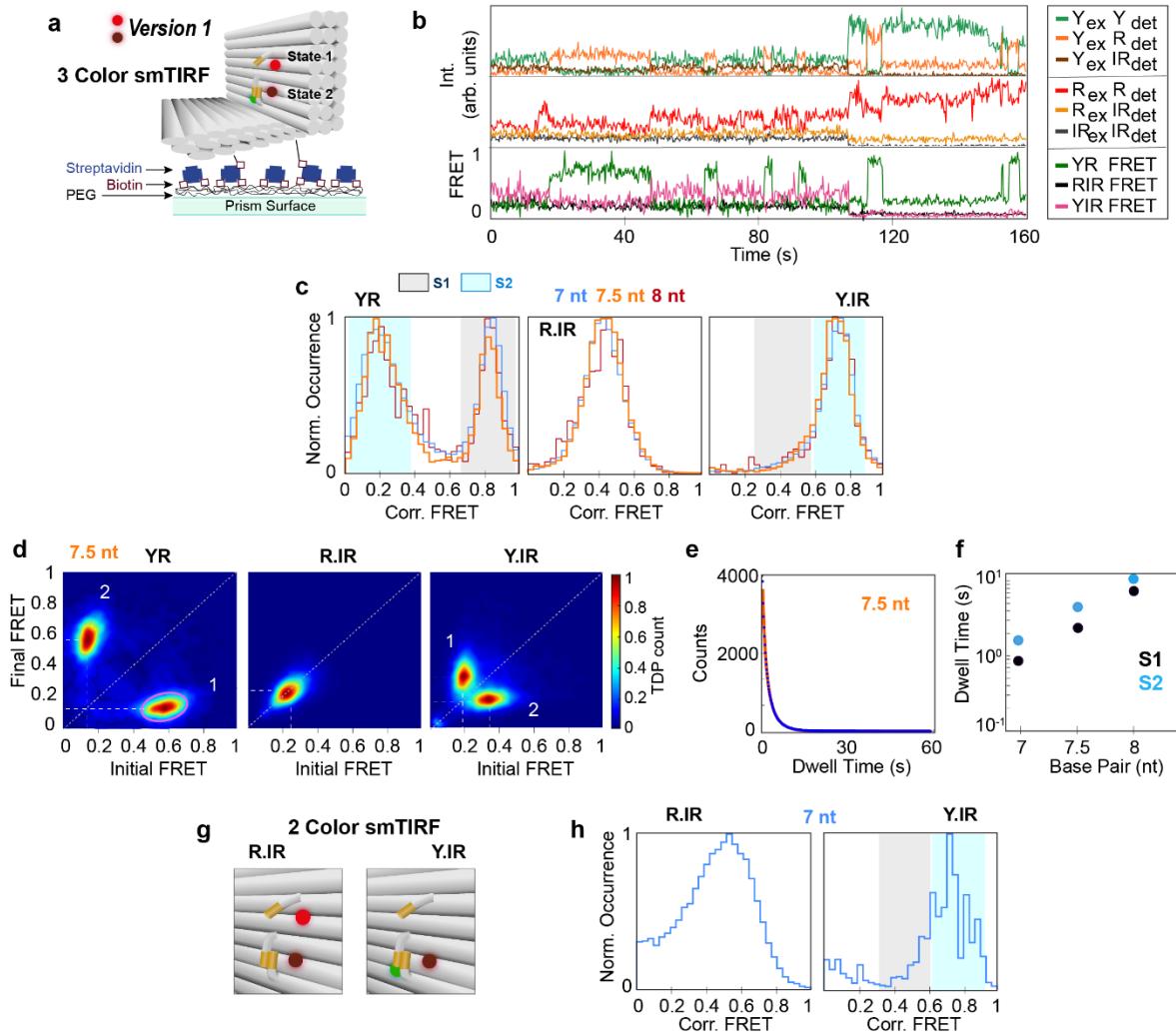


Figure A.1. Single-molecule TIRF results for the three- and two-color, two-state V1 L-shaped DNA origami structures containing an IR dye. (a) A schematic of the smTIRF assay for measuring immobilized DNA origami structures on the passivated prism surface. The PEGylated surface contains biotin to allow a biotin-streptavidin-biotin linkage between the prism surface and the bottom of origamis (the dimensions are changed for clarity). The origami is labeled with three fluorophores: Cy3B on the tether, Atto647N at the 12 o'clock and Cy7 at the 6 o'clock positions. (b) Single-molecule intensity and FRET traces showing the dynamic switching between the different docking strands. The top panel shows the intensity traces of yellow, red and IR dyes after yellow excitation. The middle panel shows the emission of red and IR dyes after red excitation together with the IR emission after IR excitation. The lower panel shows FRET traces for all three dye pairs. (c) Frame-wise-weighted state-wise averaged corrected FRET efficiency histograms from smTIRF experiments with three-color ALEX excitation for each dye pair of YR (left panel), R.IR (middle panel) and Y.IR (right panel) for the three measured structures with a different number of complementary nucleotides: 7 (blue), 7.5 (orange) and 8 (red). The dynamic YR pair gives two populations with rather different FRET efficiency values of 83-84 % for state 1 and a low FRET value around 20-23 % corresponding to state 2. The static R.IR pair gives one broad histogram centered at a fixed

value of 43-44 %. The dynamic Y.IR pair shows two populations with FRET efficiencies of around 50 % and 74-75 % for states 1 and 2 respectively. (d) Transition density plots showing the transitions between any two detected states for each dye pair. For each transition, the initial and final apparent FRET values are plotted. (e) By selecting the individual clusters in panel d, a distribution of dwell times for the individual transitions can be calculated. The dwell-time distribution for the YR FRET pair for S1->S2 (highlighted in panel d in pink) is shown. (f) Dwell-times for S1 and S2 as a function of complementary nucleotides length on the docking strands. The dwell-times were determined from a mono-exponential fit to the dwell-time distribution. (g) Zoomed-in schematics of two-color, two-state origami structures containing one dye pair of R.IR (left panel) and Y.IR (right panel). (h) Frame-wise-weighted state-wise averaged corrected FRET efficiency histograms from smTIRF experiments with two-color ALEX excitation for the R.IR (left panel) and Y.IR (right panel) two-color structures containing 7 nt complementary base pairs. The static R.IR pair gives one population at a FRET efficiency of 52 %. The dynamic Y.IR pair shows two populations with FRET efficiencies of 54 % and 74 % for states 1 and 2 respectively.

Analyzing the data obtained from the smTIRF assay was performed with the MATLAB-based software package, Deep-LASI, which has been developed in our group [112, 113]. Deep-LASI is also designed for four-color data (including the IR channel), and mapping up to four channels is possible. However, four-color analyses are yet to be implemented. Each detected molecule in the smTIRF assay appears as a time trace with the intensity or FRET signal on the y axis. Measuring with the ALEX schemes produces multiple time traces for the same molecule attributing to all fluorophores. Every obtained trace needs to be inspected for individual dye behavior and features. As discussed in section 3.2.5, using more than one detection channel requires some correction factors to account for the differences in the efficiency or sensitivity of the detectors together with the differences in the quantum efficiency of the applied fluorophores. Such data analysis consists of several steps from mapping the channels, trace extraction and categorization, determining the correction factors (including the spectral crosstalk of donor emission into acceptor channel, direct excitation of the acceptor by the donor excitation laser and the gamma factor), plotting and fitting the FRET efficiency and kinetic analysis. The details of these steps are described in our publications, which are attached to the end of this document.

Figure A.1a shows a schematic of an immobilized three-color DNA origami structure measured using smTIRF microscopy on a passivated prism surface. The PEGylated surface of the prism contains biotin groups to make a strong biotin-streptavidin-biotin coupling between the glass surface and the bottom side of L-shaped origami structures. The origami is labeled with three fluorophores: Cy3B on the flexible tether, Atto647N near the upper binding strand (12 o'clock position, state 1) and Cy7 near the lower binding strand (6 o'clock position, state 2). Dynamic switching between the different docking strands can be directly observed in the single-molecule FRET traces (Figure A.1b). When the YR FRET efficiency is high, the Y.IR efficiency is low and vice versa.

The histograms of the FRET efficiencies observed in the dynamic traces are shown in Figure A.1c for the constructs with the 7 nt, 7.5 nt and 8 nt binding sites. The corrected FRET values are averaged over each transition and weighted by the number of frames per transition, which we refer to as the frame-wise-weighted state-wise averaged FRET histograms. They are plotted for each dye pair (YR, R.IR and Y.IR). The results and correction factors are summarized in Table A1 at the end of this section. As expected for the two YR and Y.IR dye pairs, two distinct FRET populations are observed, which are related to the transient binding of the tether to either of the two binding states. The FRET values for the low and high FRET states are similar for all three constructs (S1 = 83-84 % for YR and 40-60 % for Y.IR; S2 = 20-23 % and 74-75 % for YR and Y.IR respectively). The big range of corrected FRET efficiency for the first state using the Y.IR pair is due the not resolvable FRET peaks obtained from the two states, which make the fitting of the populations difficult. This is usually the problem when having a fluorophore with very low emission intensity and the fact that three-color correction factors have

extreme effects on the shape of FRET histograms from such intensity values. For the R.IR dye pair, we observe a single peak at 43-44 % FRET efficiency upon correction ([Figure A.1c, middle panel](#)).

The kinetics of the three different DNA origami structures were also analyzed using the Deep-LASI software. Every state transition is plotted on the transition density plot (TDP), which forms clusters around the values for states 1 and 2. By choosing a cluster, the dwell time distributions for the selected transitions are plotted and fit to a mono-exponential function ([Figure A.1d and e](#)). As designed and expected, the kinetic rates increase as the number of complementary bases decreases ([Figure A.1f, Table A1](#)). The data also consistently shows longer dwell times for state 2 compared to state 1. The dwell-time ratio of S2 to S1 is 1.75 and 1.90 for the two- and three-color versions of the structures, with 7 nt long complementary double strands, 1.8 for the 7.5 nt samples and 1.4 for 8 nt samples. Hence, as observed previously, the tether spends more time bound to the lower position (6 o'clock, state 2) than the top binding strand (12 o'clock, state 1).

We also measured two, two-color origami structures on the surface both having a complementary double strand length of 7 nt ([Figure A.1g](#)). The results for the two dye-pair combinations of R.IR and Y.IR are shown in [Figure A.1h](#) (R.IR on the left panel and Y.IR on the right panel). For the R.IR dye pair, one FRET population is observed with a FRET efficiency of 52 % ([Figure A.1h, left panel](#)) and the Y.IR FRET pair shows two FRET states with values of 54 % and 74 % ([Figure A.1h, right panel](#)). The FRET efficiency values from the two- and three-color samples are in good agreement. Only the low FRET value for the Y.IR pair is different, which is again due to the fact that the peaks for states 1 and 2 are not well resolved. The dwell times obtained from the TDP for the dynamic two-color sample (Y.IR) were 0.63 ± 0.06 s and 1.10 ± 0.02 s for states 1 and 2 respectively, which match very well with the values from the original 7nt DNA origami structures (0.650 ± 0.003 s and 0.99 ± 0.01 s).

The dwell times extracted from the three-color data with the 7 nt structure (Cy3B-Atto647N-Cy7) is a bit longer than those from the corresponding states in the two-color sample (Cy3B-Cy7). We observed the same effect in the original two- and three-color DNA origami structures where that the presence of the red dye (Atto647N) influences the dwell times of the same complementary binding length. This effect is described in the corresponding publication about the origami structures and is attributed to the interactions that Atto647N induces to the system. Comparing the dwell times of three-color samples containing Cy7 with the corresponding structures without it also shows longer dwell times for the 7 and 7.5 nt structures (by a factor of 1.1-1.5). However, for the 8 nt sample, the ones with the IR dye show rather shorter values (by a factor of 0.92-0.95). The only assumption for this observation could be the induced interactions from Cy 7 to the DNA binding and unbinding however, we need more control measurements to have a more correct estimation. The different dwell times values are summarized in [Table A.2](#).

Table A.1. Experimental smTIRF parameters and results for three- and two-color L-shaped DNA origami structures having two states and a different length of complementary base pairs on the docking strands. Correction factors including spectral crosstalk (CT), direct excitation (DE) and the gamma factor (γ) allow for the determination of corrected FRET efficiencies. The dwell time for each state is extracted from the representative transition density plot (TDP) using mono-exponential fitting.

| sm TIRF | FRET Pair | CT | DE | Gamma | Apparent FRET [%] (S1 – S2) | Corrected FRET [%] (S1 – S2) | Population Ratio (%) | Dwell Time t_1 (s) | Dwell Time t_2 (s) |
|----------------|-----------|------|------|-------|-----------------------------|------------------------------|----------------------|----------------------|----------------------|
| 2-color | | | | | | | | | |
| 7 nt | Y-IR | 0.05 | 0.07 | 0.19 | 26 - 42 | 54 - 74 | 15 - 85 | 0.63 ± 0.06 | 1.10 ± 0.02 |
| | R-IR | 0.06 | 0.22 | 0.18 | 21 | 52 | NA | NA | NA |
| 3-color | | | | | | | | | |
| 7 nt | Y-R | 0.11 | 0.04 | 0.73 | 58 - 12 | 84 - 20 | 37 - 63 | 0.79 ± 0.07 | 1.50 ± 0.00 |
| | R-IR | 0.06 | 0.24 | 0.18 | 23 | 43 | NA | | |
| | Y-IR | 0.01 | 0.03 | 0.13 | 18 - 34 | 47 - 74 | 12 - 88 | | |
| 7.5 nt | Y-R | 0.11 | 0.04 | 0.83 | 58 - 11 | 83 - 21 | 35 - 65 | 2.10 ± 0.01 | 3.80 ± 0.00 |
| | R-IR | 0.06 | 0.24 | 0.17 | 22 | 44 | NA | | |
| | Y-IR | 0.01 | 0.02 | 0.14 | 17 - 33 | 60 - 74 | 24 - 76 | | |
| 8 nt | Y-R | 0.11 | 0.03 | 0.77 | 58 - 11 | 84 - 23 | 35 - 65 | 6.10 ± 0.00 | 8.70 ± 0.00 |
| | R-IR | 0.06 | 0.23 | 0.19 | 26 | 44 | NA | | |
| | Y-IR | 0.01 | 0.02 | 0.15 | 18 - 33 | 40 - 75 | 15 - 85 | | |

Table A.2. A comparison between the dwell time values from smTIRF experiments for state 1 and state 2 of three-color, two-state DNA origami structures (V1) having different dye combinations (B-Y-R vs Y-R-IR) but the same number of complementary nucleotides between the tether and the docking strands.

| 3-color Samples | | Dwell Time t_1 (s) | Dwell Time t_2 (s) |
|-----------------|--------|----------------------|----------------------|
| 7 nt | B-Y-R | 0.65 ± 0.003 | 0.99 ± 0.01 |
| | Y-R-IR | 0.79 ± 0.07 | 1.50 ± 0.00 |
| 7.5 nt | B-Y-R | 1.90 ± 0.02 | 2.70 ± 0.03 |
| | Y-R-IR | 2.10 ± 0.01 | 3.80 ± 0.00 |
| 8 nt | B-Y-R | 6.40 ± 0.04 | 9.50 ± 0.10 |
| | Y-R-IR | 6.10 ± 0.00 | 8.70 ± 0.00 |

Appendix 2: Appended Publications

Paper 1.

Deep-LASI: deep-learning assisted, single-molecule imaging analysis of multi-color DNA origami structures¹¹³

Wanninger S., Asadiatouei P., Bohlen J., Salem C.B., Tinnefeld P., Ploetz E., Lamb D.C.

Nature Communications, 14, 6564 (2023).

DOI: <https://doi.org/10.1038/s41467-023-42272-9>



Article

<https://doi.org/10.1038/s41467-023-42272-9>

Deep-LASI: deep-learning assisted, single-molecule imaging analysis of multi-color DNA origami structures

Received: 9 February 2023

Accepted: 5 October 2023

Published online: 17 October 2023

Check for updates

Simon Wanninger ¹, Pooyeh Asadiatouei ¹, Johann Bohlen ¹, Clemens-Bässem Salem ¹, Philip Tinnefeld ¹, Evelyn Ploetz ¹ & Don C. Lamb ¹

Single-molecule experiments have changed the way we explore the physical world, yet data analysis remains time-consuming and prone to human bias. Here, we introduce Deep-LASI (Deep-Learning Assisted Single-molecule Imaging analysis), a software suite powered by deep neural networks to rapidly analyze single-, two- and three-color single-molecule data, especially from single-molecule Förster Resonance Energy Transfer (smFRET) experiments. Deep-LASI automatically sorts recorded traces, determines FRET correction factors and classifies the state transitions of dynamic traces all in ~20–100 ms per trajectory. We benchmarked Deep-LASI using ground truth simulations as well as experimental data analyzed manually by an expert user and compared the results with a conventional Hidden Markov Model analysis. We illustrate the capabilities of the technique using a highly tunable L-shaped DNA origami structure and use Deep-LASI to perform titrations, analyze protein conformational dynamics and demonstrate its versatility for analyzing both total internal reflection fluorescence microscopy and confocal smFRET data.

Single-molecule spectroscopy has revolutionized how we investigate the mechanism of processes on the nanometer scale. In particular, optical fluorescence imaging allows contact-free investigations of single, dynamic biomolecules, one at a time, in cells, membranes and in solutions. Single-molecule Förster Resonance Energy Transfer (smFRET) in combination with confocal microscopy or Total Internal Reflection Fluorescence (TIRF) microscopy probe distances on the nanometer scale (2.5–10 nm). While solution measurements can provide information on sub-millisecond dynamics, measurements with immobilized molecules give access to the temporal evolution of single molecules on the timescale of microseconds to minutes¹. By removing ensemble averaging, it is possible to directly measure the underlying conformational states and molecular dynamics of biomolecules. Its ability to measure accurate distances and kinetics turned smFRET into a powerful tool for deciphering molecular interaction mechanisms and structures of biomolecules^{1–3}. Typically, FRET experiments are

performed using two colors and used to probe conformational distributions and distance changes. However, also other single-molecule approaches can be used to investigate small distance changes or interactions (e.g., Metal-Induced Energy Transfer (MIET)⁴, Graphene Energy Transfer (GET)⁵, or Protein-Induced Fluorescence Enhancement (PIFE)^{6–8}).

When combining three- or more labels, multi-color FRET can probe molecular interactions between different binding partners and also measure multiple distances simultaneously, i.e. correlated motion within the same molecule^{9–11}. However, multi-color analyses remain challenging. Quantitative smFRET data analysis is strongly hampered by experimental restrictions due to, for example, a low number of usable single molecule traces, data with a low signal-to-noise ratio (SNR), or short traces due to photochemistry. Overcoming these limitations requires large data volumes as very few molecules contain the desired information with suitable quality, which significantly increases

¹Department of Chemistry and Center for NanoScience (CeNS) Ludwig-Maximilians-Universität München Butenandtstr. 5-13, 81377 Munich, Germany.
 e-mail: evelyn.ploetz@lmu.de; d.lamb@lmu.de

the efforts involved in sorting through the data when performed manually. Low statistics result from various reasons including molecular events exhibiting slow kinetics or rare transition probability, insufficient labeling efficiency, low SNR, quick photobleaching or spurious background. In addition, arbitrary fluctuations due to unwanted interactions and/or aggregations between binding partners hamper a concise analysis of the underlying state and kinetics.

Various approaches have been developed to overcome these time-consuming burdens, employing user-defined thresholds on the channel count rate, signal-to-noise ratio, FRET values, FRET lifetime, and donor/acceptor correlation¹²⁻¹⁹. However, setting appropriate thresholds requires a substantial amount of expertise. Depending on the user, the data evaluation is prone to cognitive biases and poses a challenge to reproducible analysis results. Recently, software packages have been published that use deep-learning techniques to rapidly automate trace classification and keep user bias to a minimum²⁰⁻²². In particular, Thomsen et al. comprehensively demonstrated that artificial neural networks could match manual classifications and even outperform conventional methods of commonly used programs to extract valid single-molecule FRET traces²². So far, deep learning has been solely applied to single-channel and two-color FRET data to categorize the time trajectories for downstream analysis. To study structural dynamics, reflected by changes in intensity and FRET efficiencies, the kinetics are then analyzed separately typically using Hidden Markov Models (HMMs)^{23,24} approaches. Training an HMM requires knowledge of the number of states and modeling of the emission probabilities. Moreover, it assumes that the probability of a transition to the next state only depends on the current state. While the initial HMM settings are straightforward for simple systems, obtaining the optimal parameters for multi-color FRET becomes a challenging task. To date, only one software package, SMACKS¹³, allows an ensemble HMM for three-color FRET data. As the complexity of the datasets grows, the effort and the required knowledge about the system also grow.

To alleviate the shortcomings of HMM analyses, the hybridization of HMMs with Deep Neural Networks (DNN) has gained popularity²⁵⁻²⁹. In contrast to HMMs, DNNs are capable of learning higher-order dependencies without prior assumptions about the number and properties of the states. A long-short-term memory (LSTM) neural network was developed to automate stoichiometry determination via photobleaching steps in fluorescence intensity traces³⁰. However, the use of DNNs for extracting quantitative kinetic information from single-molecule data has not yet been explored.

Here, we present the Deep-Learning Assisted, Single-molecule Imaging (Deep-LASI) approach, an ensemble of DNNs with architectures specifically designed to perform a fully automated analysis of single-color traces as well as two-color and three-color single-molecule FRET data. Deep-LASI begins with raw intensity traces and provides corrected FRET efficiencies, state determination, and dwell times without any prior knowledge or assumptions about the system. It classifies each time trace into different categories, identifies which fluorophores are active in each frame, which is then used for determining FRET correction factors for spectral crosstalk, direct acceptor excitation and detection efficiency, and performs a state transition analysis of the different states in dynamic traces. Deep-LASI also includes optional number-of-state classifiers to estimate the actual number of observed states within one trace. Since the pre-trained neural networks operate locally on each trace, they do not neglect rare events, which would be missed in global analysis approaches. We benchmark the performance of Deep-LASI using ground truth simulations and experimental one-, two- and three-color data using an L-shaped DNA origami structure with tunable dynamic behavior^{5,31}. The results are further compared to the manual evaluation of the data and the extracted dwell times obtained with HMM. Finally, we demonstrate the power of Deep-LASI with multiple applications: (1) titration

experiments, which would be unfeasible without Deep-LASI; (2) smFRET on a mitochondrial Hsp70 to extract substrate-specific dwell times and conformational states; and (3) the applicability of Deep-LASI to another experimental setup.

Results

The Deep-LASI approach

Deep-LASI utilizes an ensemble of pre-trained deep neural networks designed for the fully automated analysis of one-, two- and three-color single-molecule data including multi-color FRET correction and kinetic analyses (Fig. 1; Supplementary Note 1). The designed input for Deep-LASI is a single-molecule fluorescence intensity trace or traces measured directly using confocal microscopy or extracted from movies using wide-field or TIRF microscopy. In the case of two-color fluorescence data, continuous wave excitation or Alternating Laser EXcitation (ALEX) modalities can be analyzed. For three-color smFRET measurements, ALEX data is required. All available channels are fed into a combination of a Convolutional Neural Network (CNN) using the omni-scale feature learning approach and a Long Short-Term Memory (LSTM) model (Supplementary Fig. 1.1).

Deep-LASI extracts spatial and temporal sequence features simultaneously and classifies every frame into a specific category (Fig. 1a). Building upon Deep-FRET for two-color FRET analysis²², we separate the traces into six categories: dynamic, static, noisy, artifact, aggregate as well as photobleached (see Supplementary Note 2 for details). The total number of categories depends on the number of input channels, i.e. the number of dyes (and alternating light sources) used in the experiment. Traces containing random artifacts, aggregates, or high noise are excluded from further analyses. The final output of the state classifier provides an estimation of the probability for each category. The summed probabilities over all non-photobleached frames serve as confidence levels for each trace. Here, user-defined thresholds can be set to increase or decrease the tolerance towards non-ideal traces to be included in further analyses. In contrast to previous networks, Deep-LASI detects photobleaching events of individual dyes and, therefore, allows the calculation of correction factors obtainable for that molecule. Traces showing no apparent state transition are classified as static and can be included, e.g. in the final corrected FRET histograms.

All sections in each trajectory identified as dynamic are transferred to the state classifier network (Fig. 1b), which is designed to detect transitions based only on the intensity data and not via the FRET efficiency. The state classifier assigns every frame to one of the multiple states present in a dynamic trace section and again provides a confidence value of state occupancy that can be used for additional thresholding. Given the state transition classifications, a Transition Density Plot (TDP) is calculated and the kinetic rates of all identified states can be extracted by fitting the corresponding dwell-time distributions (Fig. 1c). Starting from trace extraction, the TDP marks the first necessary point of human intervention, i.e., the manual selection of state transitions and the fitting procedure. Thus, user bias is kept to a minimum. No assumptions are needed regarding the number of states, state-specific emission probabilities, or other settings required for conventional methods such as Hidden Markov Models (HMM). Of course, as for any deep-learning algorithm, the output of the analysis is dependent on the quality and appropriateness of the training data used. Depending on (1) the total number of frames, (2) the yield of valid frames, (3) the computer performance, and (4) the desired confidence threshold, a given dataset can be fully categorized on a time scale of 20–100 ms per trace.

Training of Deep-LASI

To use Deep-LASI for analyzing single molecule data, we first trained the trace-classifier network on datasets appropriate for the corresponding network (i.e., one-color data, two-color data without ALEX,

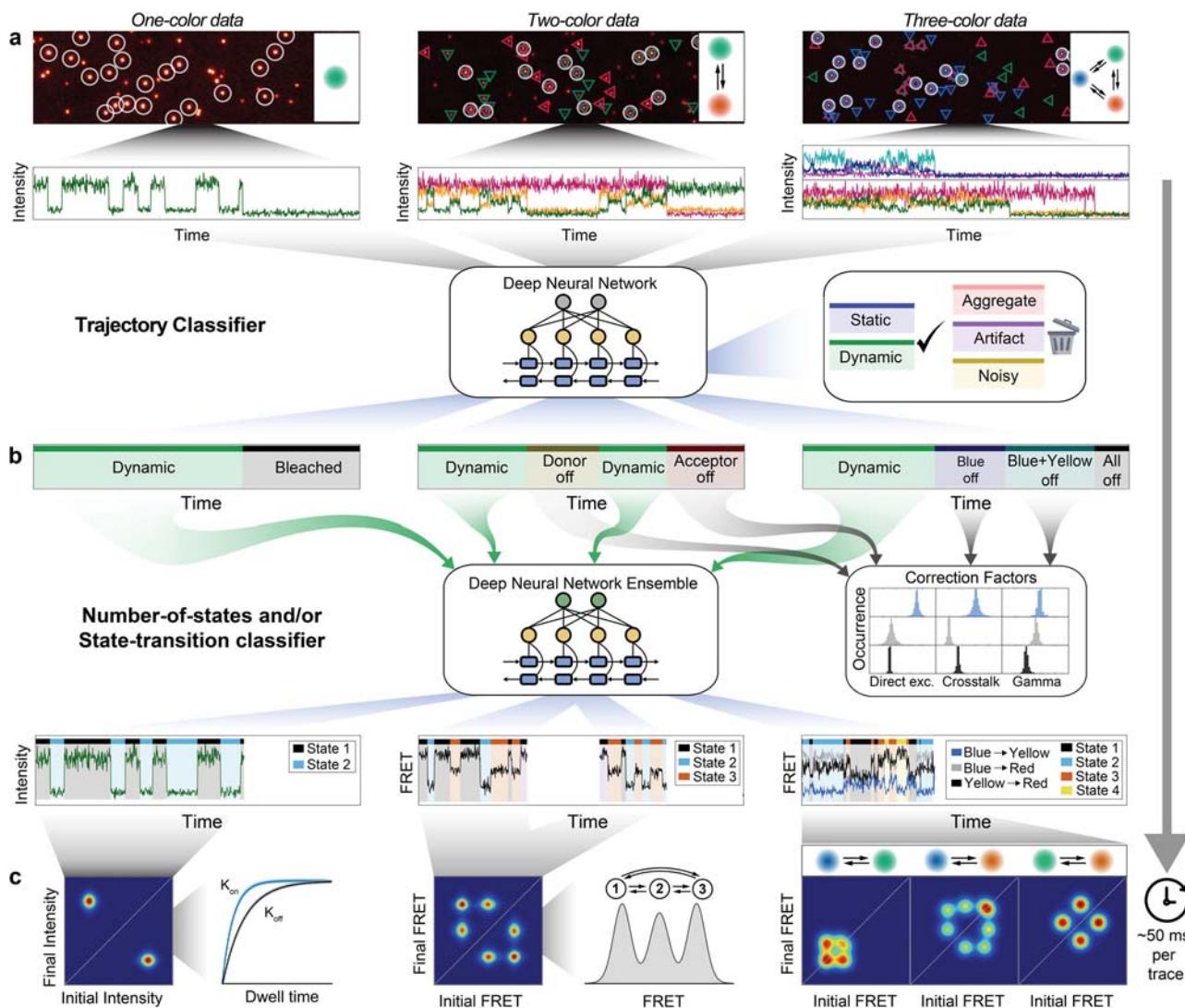


Fig. 1 | Overview of data extraction, evaluation, and analysis using Deep-LASI. **a** Single-molecule data of up to three separate channels after direct and alternating laser excitation are identified, extracted, and presorted for further analyses. Each frame within the time traces is classified into categories using a hybrid CNN-LSTM. **b** A second hybrid CNN-LSTM evaluates the kinetics and state information in the

presorted data. The photobleaching information can be used for determining the correction factors to obtain accurate FRET values between two and three fluorophores. **c** Next, the interconversion rates between underlying states and absolute, distance-related FRET values are extracted from multi-color datasets.

two-color data with ALEX or three-color data with ALEX). As the noise sources in single-molecule fluorescence intensity data are well understood, simulated traces are well suited for training the neural network. In addition, it has the advantages of being able to minimize biases and quickly retrain neural network models to adjust for specific circumstances. The training datasets were designed to cover a wide range of experimental conditions and FRET efficiencies. Hence, no initial estimation of the number of states and expected FRET efficiencies are needed. A detailed description of the program architecture, simulations, training datasets and benchmarking can be found in the Methods section as well as in Supplementary Notes 1–4.

Deep-LASI contains a total of 16 pre-trained deep neural networks for state classification. Four models account for the classification and segmentation of time trajectories obtained from measurements using single-channel data acquisition, two-color FRET with continuous-wave excitation, two-color FRET with ALEX, and three-color FRET with ALEX. For each type of experiment, we provide three state-transition-classifiers trained on either two, three or four observed states, which take the output category dynamic as the input. Note that the acceptor intensity after direct excitation does not contain relevant kinetic

information and is not used in the state classifier networks. In addition, a deep neural network is provided that has been optimized for detecting the actual number of observed states and can be utilized for model selection. One network has been trained for each type of dataset (one-, two- and three-color data). The number-of-states neural networks are not essential in the automated analysis process but can serve as a safeguard against trajectories that may be out of the scope of the state transition classifiers.

Performance of Deep-LASI

A common approach to benchmark classifier models is using ground truth labeled data and calculating confusion matrices, which summarize the correct and incorrect predictions. For every trained model (using ~200,000 traces), we generated approximately 20,000 new traces for testing, which were not part of the training dataset. Each of the validation datasets was then fed into the corresponding model. The output predictions were compared to the ground truth labels for every frame to obtain the percentage values of true positive, false positive and false negative classifications. All trace classifier models achieve a minimum combined precision of 97% in predicting smFRET

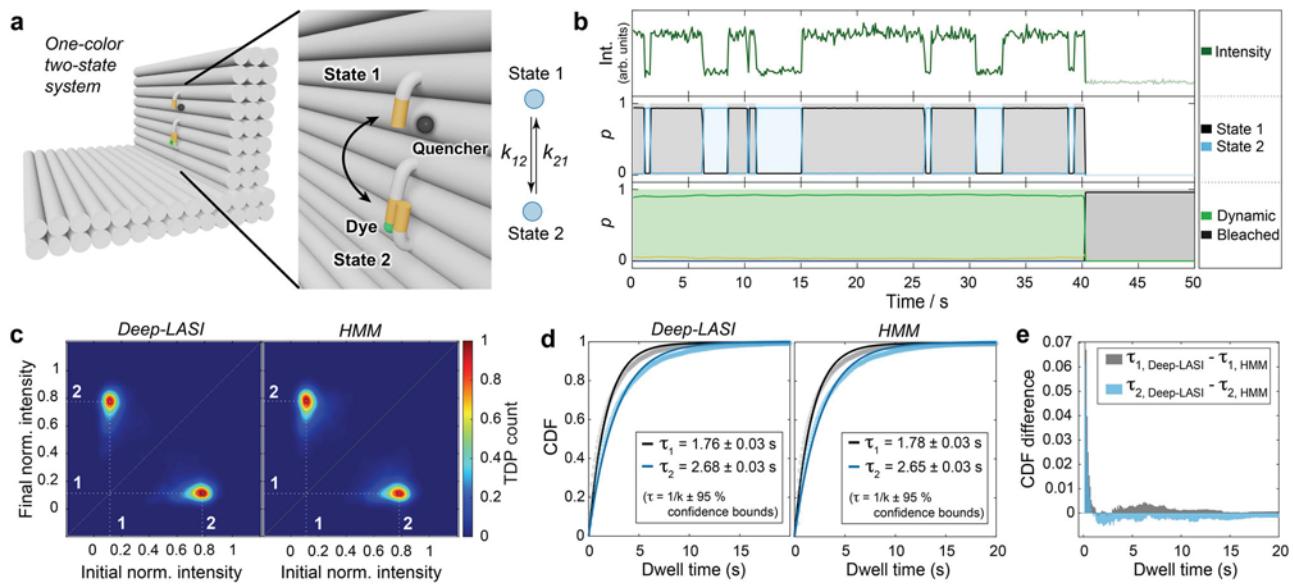


Fig. 2 | State analysis of single-color single-molecule data. **a** Sketch of the used L-shaped DNA origami structure with a single fluorophore (Cy3B) attached to a flexible tether, which changes position from state 1 \rightarrow 2 at the rate k_{12} and from state 2 \rightarrow 1 at the rate k_{21} . The zoom-in shows the two single-stranded binding sites (orange) in close and distant proximity to a quencher dye (Atto647N) bound to the DNA origami structure. **b** Representative time transient for a DNA origami structure with 7.5 nt binding strands after classification and kinetic evaluation by Deep-LASI. **c** Transition-density plots depicting the interconversion events between the two detected states 1 and 2 after trace kinetics evaluation by Deep-LASI (left, number of transitions $n = 25,948$) and by Hidden-Markov Modeling (HMM) analysis (right, number of transitions $n = 19,390$). Both approaches identify identical states.

d Cumulative probability Distribution functions (CDFs) of the dwell times: The mono-exponential fits obtained by both methods reveal equivalent dwell times of approximately 1.75 and 2.65 s for the upper (State 1) and lower (State 2) binding sites, respectively. The errors in the dwell times are the 95% confidence intervals returned by the fitting procedure (estimated from the Jacobian matrix). **e** A comparison of the CDFs was determined using Deep-LASI and HMM. Deep-LASI is already sensitive at time scales on the order of the acquisition time. The average difference is less than 1% between both methods. Source data are provided as a Source Data file.

categories, i.e. static or dynamic, and 96% in predicting non-smFRET categories (Supplementary Figs. 3.1 and 3.2).

Our number-of-states and state-transition classifiers were benchmarked analogously. For the number-of-state classifiers, two states can be distinguished from multi-state trajectories with at least 98% precision whereas four states are predicted with the lowest precision of 86% for the single-channel model (Supplementary Fig. 3.3). For the state-transition classifiers, the states can be identified with accuracies of $\geq 98\%$, $\geq 90\%$ and $\geq 78\%$ for two-state, three-state and four-state models respectively (Supplementary Fig. 3.4). The comparison between all state-transition classifiers reveals a clear trend of decreasing accuracies with an increasing number of states and increasing accuracy with an increasing number of available channels. This is expected since a higher number of states have a larger probability of lower contrast, and a higher number of channels improves the robustness towards uncorrelated noise. Since confusion matrices do not reveal any underlying dependencies, we additionally benchmarked the state transition classifiers with HMM by calculating the precision of the state label prediction for a broad range of noise levels, FRET state differences and dynamic time scales (Supplementary Fig. 3.5). Overall, the performance of state classifiers is at least on par with HMM at low noise levels and outperforms HMM at high noise levels by up to 30%. To investigate the advantage of using the information in the entire dataset for the HMM analysis, we also compared the performance of Deep-LASI with a local and a global HMM on idealized synthetic data (Supplementary Fig. 3.6). Global HMM performs significantly better than local HMM in this case and is on par with Deep-LASI.

As a last test, we compared the performance of Deep-LASI with other kinetic analysis routines that have been recently published in a multi-laboratory study³². We chose to analyze the two-state datasets as these require no user input and the analysis can be performed without bias. Deep-LASI returned values corresponding to the ground truth for

the simulated dataset and close to the average values obtained for the experimental dataset (Supplementary Fig. 3.8).

Deep-LASI analyses of DNA origami structures

Next, we benchmarked the potential of Deep-LASI to automatically analyze experimental data obtained from DNA origami structures. DNA origami is extensively used in bio-nanotechnology and has the advantage of being programmable with high precision and controllability. In particular, we choose an L-shaped DNA nanostructure with a dynamic, fluorescently labeled 19 nucleotide (nt) single-stranded DNA pointer. The geometry of the DNA structure was originally designed for measuring energy transfer to a graphene surface^{5,31}. The single-stranded DNA pointer, along with two or three exchangeable docking strands of different complementary sequences, allows the number of states, position of the dyes, and kinetic rate to be programmed as desired. Hence, it is an ideal test system for measuring and extracting kinetic information from smFRET traces. FRET efficiencies and kinetic rates could be tuned by varying the position and complementary sequence length of binding strands on the DNA origami platform. We designed various DNA origami structures with one-, two-, and three-color labels and measured them on the single-molecule level.

In the first assay, we assessed Deep-LASI's capability to evaluate single-color data. For this, we probed one-color single-molecule kinetics where the flexible pointer was labeled with Cy3B at the 3'-end. Two complementary binding sites with 8 nt complementary nucleotides containing a 1 nt mismatch at the 5'-end (referred to as 7.5 nt) were placed about 6 nm below and above the pointer position (Fig. 2a). Binding occurred by spontaneous base-pairing to single-stranded protruding strands. A single red dye, Atto647N, acting as a quencher, was attached about 3 nm aside from the upper binding site (state 1). Figure 2b shows an exemplary intensity trajectory of Cy3B classified as dynamic until photobleaching was detected by the trace classifier with

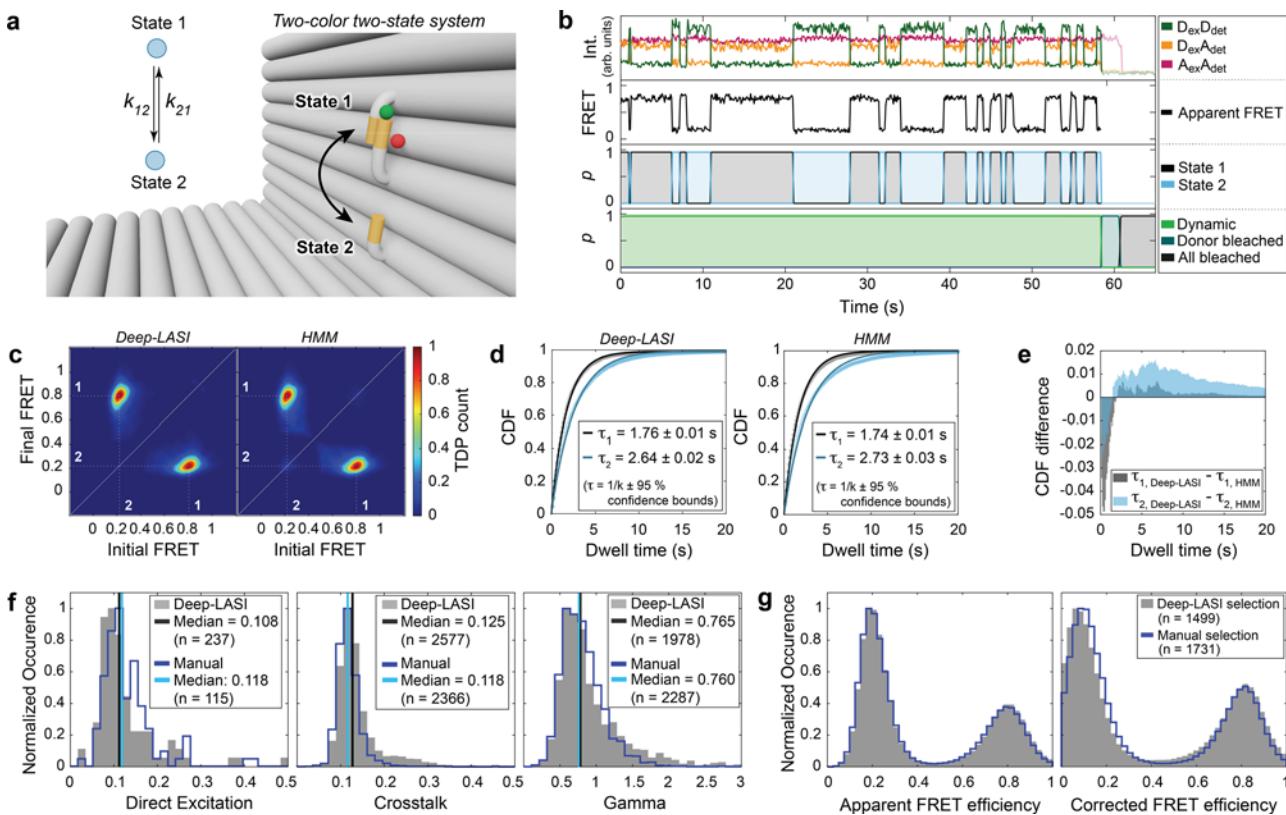


Fig. 3 | Single-molecule analysis of two-color FRET data. Experiments were performed with DNA origami structures exhibiting two FRET states. **a** Zoom-in of an L-shaped DNA origami structure labeled with Atto647N and Cy3B and corresponding kinetic scheme. The donor is attached to the flexible tether with a 7.5 nt overhang between the pointer and two single-stranded binding sites. FRET is expected between a high FRET state 1 (12 o'clock) and a low FRET state 2 (6 o'clock) interconverting at rates k_{12} and k_{21} . **b** Representative single-molecule and apparent FRET trace after alternating red-yellow (RY) laser excitation. Deep-LASI classifies the trace and determines the underlying state for each frame. D: donor; A: acceptor; ex: excitation; det: detection. **c** TDPs determined using Deep-LASI (left) and HMM (right) are shown revealing two interconverting states with apparent FRET values of 0.8 and 0.2. The two states are labeled in white. Total number of transitions: $n_{\text{Deep-LASI}} = 15,958$, $n_{\text{HMM}} = 21,243$. **d** CDFs extracted from the TDPs

shown in (c) and mono-exponential fits yield dwell times of 1.76 s and 2.64 s, respectively. The errors in the dwell times are the 95% confidence intervals returned by the fitting procedure (estimated from the Jacobian matrix). **e** A comparison of the cumulative dwell-time distribution determined using Deep-LASI - HMM for τ_1 (gray) and τ_2 (cyan). **f** Histograms of trace-wise determined correction factors for direct excitation, crosstalk and detection efficiency, either derived automatically by Deep-LASI (gray histograms, median in black) or determined manually (blue lines, median in cyan) (see Supplementary Note 5). **g** Apparent (left) and corrected (right) frame-wise smFRET efficiency histograms for 1499 dynamic traces from a total of 6100 traces. The states have corrected peak FRET efficiencies of 0.07 and 0.81. The histograms from traces selected by Deep-LASI are shown in gray and by manual selection in blue. Source data are provided as a Source Data file.

two corresponding states determined by the state classifier as the linker moves up and down.

We compared the results from Deep-LASI with a Hidden-Markov-Model analysis (HMM) trained on the same dataset. Since the state classifier does not directly predict a pre-trained intensity value for each state, the TDP was generated by averaging the normalized intensity between transitions. Both methods yield identical TDPs (Fig. 2c). The residence time of the DNA tether in both states was determined by fitting the cumulative dwell-time distribution functions (CDFs) derived from the state-classifier of Deep-LASI with a mono-exponential fit and compared to the results from HMM. The dwell times of 1.76 s versus 1.78 s (State 1) and 2.68 s versus 2.65 s (State 2) for Deep-LASI and HMM, respectively, are in excellent agreement (Fig. 2d). The differences between the CDFs obtained by Deep-LASI and HMM (Fig. 2e) indicate that Deep-LASI identifies fast transitions close to the frame time more frequently than HMM. The overall difference at longer dwell times remains well below 1%, which proves that Deep-LASI obtains identical results to HMM with negligible differences in the extracted rates. Interestingly, although the DNA binding strands are identical in sequence and length, there are clear differences in the dwell times. We attribute this to an inherent bias in the equilibrium position of the DNA pointer and non-symmetric, non-specific dye-origami interactions. In

addition, it is unlikely that the distance to each docking strand and potentially induced stress upon binding are identical for the two binding sites, even though the binding sequence is the same. We note that the kinetics we measure here are not directly comparable to other DNA-hybridization experiments due to both interacting DNA strands being tethered to the DNA origami platform. This leads to a high local concentration of the binder strand, and multiple dissociation and rebinding events can occur before the tether switches binding sites.

In the next step, we studied Deep-LASI's ability to deal with two-color data. We investigated two-color FRET assays with two states and compared the results with a pure manual evaluation of the same data. Here, both donor and acceptor signals from the same DNA origami sample system as shown in Fig. 2a were analyzed (Fig. 3a). TIRF measurements were performed using msALEX³³ yielding donor signal (Cy3B, Channel $D_{\text{ex}}D_{\text{em}}$), sensitized emission (Channel $D_{\text{ex}}A_{\text{em}}$) and acceptor signal (Atto647N, Channel $A_{\text{ex}}A_{\text{em}}$) to obtain information about acceptor photobleaching and direct excitation. Figure 3b shows a fully classified example trace with the signals on top and the derived FRET trace below. From the trace classifier, Deep-LASI identified dynamic sections and individual photobleaching events (Fig. 3b; bottom). The dynamic section was further classified in the state transition classifier according to their state occupancy using only the two

channels of the donor and acceptor intensity after donor excitation (Fig. 3b; middle). The channel of acceptor excitation and detection does not serve as input for the state transition classifier since it does not contain valuable kinetic information. From a total of 6100 recorded traces in the dataset, 1499 traces were classified as dynamic smFRET trajectories with at least one transition.

The same traces were also sorted manually and the 1731 selected dynamic traces were analyzed using HMM³⁴ (see Supplementary Note 5 for details). TDPs from the state transition classifier and from the HMM analysis are nearly identical (Fig. 3c). Also, the corresponding dwell times, determined via mono-exponential fits to the CDFs, are similar (Fig. 3d) and correspond to the expected dwell times of the one-color sample shown in Fig. 2 (~1.75 s for state 1 and 2.68 s for State 2). A comparison of the CDFs from Deep-LASI and HMM indicates that manually selected traces contained more fast transitions than the traces selected by Deep-LASI in this case (Fig. 3e). We looked into the differences between manually selected traces and traces selected by Deep-LASI. The most common classification discrepancies between the two are discussed in Supplementary Note 4.1. Based on individual example traces, we observed a stronger influence of the non-ideal behavior of the traces outside the regions of interest (e.g. a noisy ALEX signal or nonconstant signal intensities in photobleached regions of the trace) on Deep-LASI's classification compared to that of manual selection. We also compared the output of Deep-LASI with that of a global HMM analysis executed on the same dataset (Supplementary Fig. 4.2). As expected, the global HMM was prone to miss transitions due to slight heterogeneities in the dataset.

Next, we investigated how sensitive Deep-LASI is to the training dataset. Hence, we trained two additional classifier networks using newly simulated datasets. Details are given in Supplementary Note 4.2. The consistency between the differently trained neural networks is ~90%, similar to what would be expected from analysis run on the validation datasets (Supplementary Fig. 3.1c). Interestingly, the consistency between the different neural networks is higher than between two independent users (Supplementary Fig. 4.3b).

To determine the distance between both dyes in the two FRET states, the smFRET data needs to be corrected. Deep-LASI uses the frames classified as photobleached to automatically derive the correction factors necessary for an accurate FRET calculation^{1,35,36}. In the manual analysis, the relevant regions are selected by hand (Fig. 3f, Supplementary Note 5). The correction factors agree within ~3%. Using the derived correction factors, the correct FRET efficiency is determined. The apparent (left) and corrected FRET histograms (right) of the Deep-LASI (gray histograms) and manually (blue lines) selected traces are shown in Fig. 3g. There is excellent agreement between the Deep-LASI and manually analyzed apparent FRET histograms. The difference between the corrected histograms is due to the difference in the correction factors determined and applied from the two analyses. In this case, as Deep-LASI classifies photobleaching on a per-frame basis, more frames can be used for determining the correction factors and are, thus, most likely, more accurate here. The corrected peak FRET efficiencies are 0.81 and 0.82 (State 1) and 0.08 and 0.14 (State 2) for Deep-LASI and manual evaluation, respectively, and correspond to distances of 53 and 53 Å, and 103 and 92 Å (assuming an R_0 of 68 Å).

In the last step, we then tested the performance of Deep-LASI for analyzing three-color data by labeling the DNA origami structure with an additional blue dye, Atto488, at ~3 Å distance to the binding site for State 2 (Fig. 4a). The labeling sites of the yellow (Cy3b) and red (Atto647N) dyes were left unchanged to provide consistency with the previous two-color experiments. The use of three FRET pairs provides three distances simultaneously and allows the resolution of states that are degenerate for two-color FRET.

Using the six available intensity traces, each frame is categorized by the fluorophores that are active and whether the trace is static,

dynamic or should be discarded. As the acceptor intensity after acceptor excitation ($R_{ex}R_{em}$) does not contain valuable kinetic information, the other 5 intensity channels for dynamic traces (before photobleaching) are given as input for the state transition classifier (Fig. 4b). Movement of the flexible tether results in an anti-correlated change in the FRET efficiency of blue to yellow (BY) and yellow to red (YR), visible in the apparent FRET panel of the example trace in Fig. 4b. For each FRET pair, a TDP can be calculated, which allows the assignment of the state number to the actual FRET populations (Fig. 4c). Note, the apparent FRET efficiency of blue to red (BR) varies with the YR FRET efficiency due to the different energy transfer pathways taken upon blue excitation. Deep-LASI classifies a state regardless of which dye is undergoing a transition, i.e. the extracted dwell time distribution of a given state is the same for all FRET pairs when there is no overlap of multiple states in the TDP. The dwell times for states 1 and 2 match with those for the one-color and two-color samples, which indicates that the transition rates are not influenced by the acceptor dyes close by (Figs. 2d, 3d, Supplementary Fig. 6.1). From a total of 2545 recorded molecules, 581 were classified as valid, dynamic three-color FRET traces. The uncorrected, framewise smFRET histograms of BY, BR and YR FRET pairs are very similar to those from the 694 manually selected traces (Supplementary Fig. 4.4a). A detailed comparison between the manual analysis of the results from Deep-LASI is given in Supplementary Notes 4.3 and 4.4.

As for two-color FRET, Deep-LASI automatically determines all correction factors obtainable per trace depending on which dyes are photoactive. The results of the automated extraction of correction factors are summarized and compared to manually derived correction factors in Supplementary Fig. 4.4b. The corresponding apparent and state-wise, corrected FRET efficiency histograms for each FRET pair are shown in Fig. 4d. While the YR FRET efficiency can be directly calculated, the corrected BY and BR FRET efficiencies are subjected to higher uncertainties due to the large number of correction factors involved (see Supplementary Note 5). In particular, their dependency on the YR FRET efficiency leads to the broadening of the distributions. To minimize this influence, we perform the correction using the state-averaged FRET efficiencies. After correction, the FRET efficiencies of State 1 (0.81) and State 2 (0.08) for the YR FRET pair are virtually identical as for the two-color system. For the BY FRET pair, State 1 and State 2 correspond to peak FRET efficiencies of 0.36 and 0.81, respectively. As expected, the two populations of the apparent BR FRET efficiency merge into one static population in the corrected histogram with a peak FRET efficiency of 0.36.

To probe the performance of the kinetic analysis from Deep-LASI, we used the tunability of the L-shaped DNA origami structure to vary the timescale of the dynamics. In addition to the 7.5 nt binding sites (Fig. 4a–d), we measured three samples using binding sites of length 7 nt with a 1 nt mismatch (referred to as 6.5 nt), 7 nt, and 8 nt (Fig. 4e). The summary of all extracted dwell times (Fig. 4f, Supplementary Figure 6.1) shows an exponential increase in the dwell times of both states with increasing binding site lengths ranging from 0.33 s to 9.5 s. Considering the camera exposure times of 32 ms (6.5 nt), 50 ms (7 nt and 7.5 nt datasets) and 150 ms (8 nt dataset) and frame shift time of 2.2 ms, a dwell-time to frame-time ratio ranges from 9 (6.5 nt State 1) to 62 (8 nt, State 2).

To test Deep-LASI with more complex dynamics with multiple states, we constructed a three-state system with three-color labels using 7 nt binding strands at positions 6 and 12 o'clock and an additional 7.5 nt complementary binding strand at 9 o'clock (Fig. 5a). An example trace containing all possible transitions identified by Deep-LASI is shown in Fig. 5b. The TDP of the BY FRET pair (Fig. 5c, left panel) yields clearly distinguishable populations, while the TDP of the YR FRET pair (Fig. 5c, right panel) shows a degeneracy of state 3 transitions. Using the BY TDP, we determined the dwell time distributions with residence times between 0.65 s and 1.43 s (Supplementary

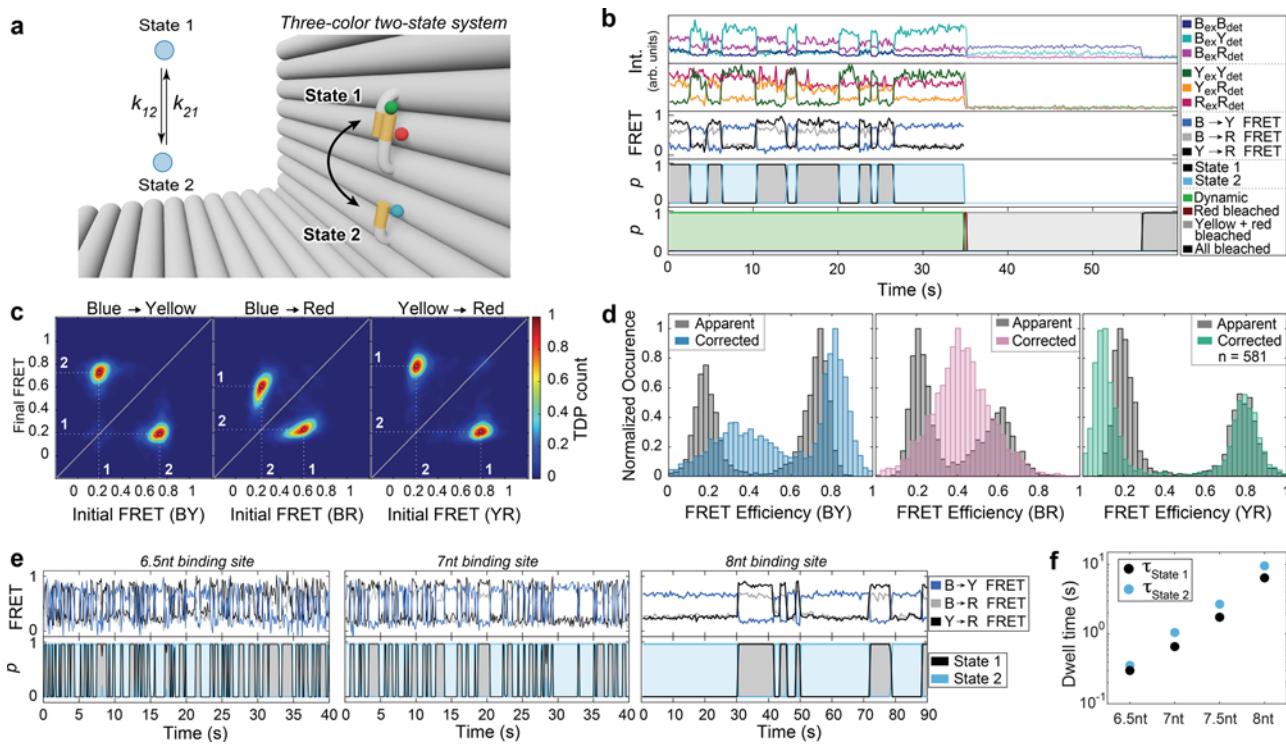


Fig. 4 | Single-molecule analysis of three-color FRET data. Experiments were performed on two-state DNA origami structures that were labeled with Atto647N, Cy3B and Atto488. While Cy3B is attached to a flexible tether, Atto647N and Atto488 are positioned close to the top (12 o'clock; state 1) and bottom (6 o'clock; state 2) binding sites, respectively. **a** Zoom-in of the L-shaped DNA origami structure and corresponding kinetic scheme. **b** Representative single-molecule intensity and FRET trajectories for binding sites with 7.5 nt overhang after alternating red-yellow-blue laser excitation. First panel: Intensities after blue excitation. Second panel: intensities after yellow and red excitation. Third panel: corresponding three-color FRET efficiencies. Fourth and fifth panels: Deep-LASI output for state transition and trace classification. B: blue; Y: yellow; R: red; ex: excitation; det: detection. **c** TDPs of the apparent FRET efficiency states reveal an apparent distance change for all three FRET pairs (BY (left), BR (middle), and YR channel (right)) with dwell times of 1.75 s and 2.69 s for the upper and lower binding site, respectively, nearly

identical to the two-color DNA origami structures (Fig. 3c). Total number of transitions: 5,013. **d** Frame-wise weighted state-wise apparent (gray) and corrected (color) smFRET efficiency histograms of the BY (left), BR (middle), and YR (right) FRET pairs. As expected, the accurate FRET efficiency of the BR pair is static ($E = 0.36$). As the position of Cy3B changes from state 1 to state 2, the accurate FRET efficiency changes from 0.36 to 0.81 (BY pair) and from 0.81 to 0.08 (YR pair). **e** Upper panel: Representative three-color smFRET traces for binding sites with 6.5 nt (7 nt with 1 nt mismatch), 7 nt and 8 nt overhangs after alternating RYB laser excitation. Bottom Panel: The corresponding state determined by Deep-LASI. **f** Extracted dwell times from mono-exponential fits for the lower (blue) and upper positions (black) for 6.5 nt ($\tau_1: 0.31$ s, $\tau_2: 0.4$ s), 7 nt ($\tau_1: 0.66$ s, $\tau_2: 1.05$ s), 7.5 nt ($\tau_1: 1.75$ s, $\tau_2: 2.69$ s) and 8 nt overhangs ($\tau_1: 6.41$ s, $\tau_2: 9.54$ s) (see Supplementary Figure 6.1 for more details). nt: nucleotides. Source data are provided as a Source Data file.

Figure 6.2). The three states are well-resolved in the framewise apparent BY FRET histogram, while state 2 and state 3 are degenerate for the BR and YR FRET pairs (Fig. 5c). Applying all correction factors yields peak YR FRET efficiencies of 0.81 (state 1), 0.08 (state 2) and 0.19 (state 3). Upon correction, States 1 and 3 in the BY FRET histogram merge into a broad degenerate FRET population. However, using the state information for all three fluorophores allows us to separate out the BY FRET histograms of the individual states.

For three-color FRET, the corrected BY and BR FRET efficiencies depend on the YR FRET efficiency, and the additional corrections broaden the population. However, even though the data may be noisier, three-color experiments contain additional information, which typically allows one to resolve states that are degenerate in two-color experiments. This is exemplified in two-color FRET experiments on the same construct missing the blue fluorophore near the 6 o'clock binding site (Supplementary Note 6.3). For distinguishable states, the determined corrected FRET efficiencies and kinetic rates from two- and three-color experiments are the same. However, three-color FRET experiments enable the lifting of this degeneracy between states 2 and 3. To minimize the influence of the increased noise in three-color experiments, it is advantageous to analyze the data in proximity ratio and only convert it to corrected FRET efficiencies when necessary¹⁰. Deep-LASI can rapidly classify a large number of molecules and quickly

provide an overview of multi-state dynamics with easy access to the kinetic information.

Further applications of Deep-LASI

After extensive benchmarking, we applied Deep-LASI to various single-molecule datasets originating from biophysical assays, protein samples and experimental systems beyond TIRF microscopy. With the speedup in analysis time from days to minutes, experiments become possible that would have been unthinkable when performing the analysis manually. One example is a titration experiment where the biochemical conditions are varied. Here, we measured the influence of glycerol concentration on the dynamics of the 3-colored L-shaped DNA origami introduced in Fig. 4a with 7.5 nt overhangs. Interestingly, we observed a decrease in residence time in both states with increasing glycerol concentrations (Fig. 6a, b). Dwell times start at 1.75 s (state 1) and 2.69 s (state 2) for pure imaging buffer and decrease down to 0.62 s and 0.85 s in buffer containing 30% (v/v) glycerol. The multi-fold increase in binding kinetics can be explained by a reported destabilization of base-pairing due to changes in the ssDNA hydration shell³⁷ and concomitantly disturbed hydrogen bonding due to the osmolyte-DNA interaction. The melting enthalpy and melting temperature decreases linearly with glycerol concentration at about 0.2 °C per % (v/v)^{38,39} in line with our observations (Fig. 5b). With Deep-LASI at hand,

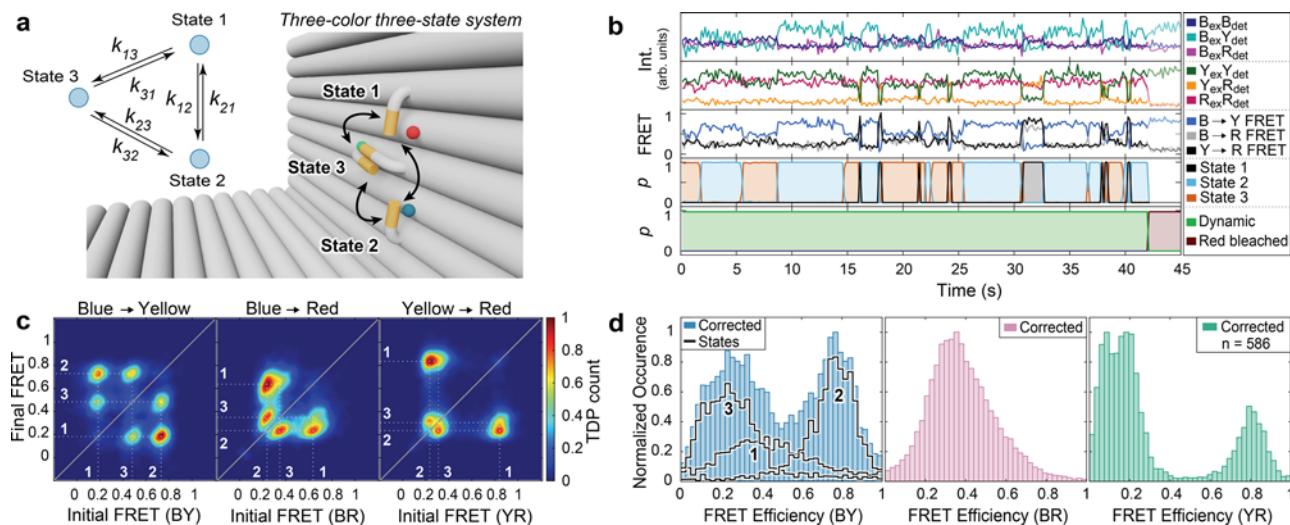


Fig. 5 | Single-molecule analysis of three-state, three-color FRET data. **a** Zoom-in of the L-shaped DNA origami structure with an additional binding site for the tether (state 3 at 9 o'clock) and corresponding kinetic scheme. **b** A representative single-molecule three-color FRET trace and apparent FRET for the 3-state system. The upper panel shows the intensity in the blue, yellow and red channels after blue excitation. The second panel shows the intensity in the yellow and red channels after yellow excitation and the red intensity after red excitation. The middle panel shows the corresponding FRET efficiencies for the three dye pairs. The fourth and fifth panels show the output of the Deep-LASI analysis for state transition and trace classification, respectively. B: blue; Y: yellow; R: red; ex: excitation; det: detection.

c Transition density plots of the apparent FRET efficiency states are shown for each FRET pair revealing an interconversion between 3 binding sites. Total number of transitions: $n = 17,136$. **d** Frame-wise weighted, state-wise corrected smFRET efficiency histograms. Corrected, distance-related FRET values are best resolved for the YR pair, showing three populations at 0.81, 0.19 and 0.09. The BY FRET shows one population at 0.8, corresponding to state 2, and a broad population at 0.3 for states 1 and 3. Individually-corrected states are indicated with the highlighted lines, showing the actual BY FRET efficiencies of state 1 (0.4) and state 3 (0.21). The apparent FRET states for the BR channel merge into one broad, static state with a value of 0.35. Source data are provided as a Source Data file.

local screening and time-consuming parameterization of imaging conditions become feasible.

Next, we applied Deep-LASI to smFRET measurements on proteins. We previously used dual-color FRET studies to probe the nucleotide-dependent conformational states⁴⁰ of Ssc1, a mitochondrial heat-shock protein Hsp70 in yeast. By fluorescently labeling both the nucleotide-binding domain and the substrate-binding domain, we investigated the influence of ADP on the inter-domain separation via smFRET. As the proteins were immobilized by incorporation in vesicles, a photostabilization buffer could not be used for the experiments. Hence, the signal-to-noise ratio in these experiments is lower than those exhibited by the photostabilized DNA origami structures. A comparison of traces selected manually and/or by the Deep-LASI analysis is discussed in Supplementary Note 6.5. For the different ADP concentrations, Deep-LASI identifies the underlying FRET states in line with the manually evaluated data⁴⁰ (Fig. 6c). It correctly identifies transitions between two distinct states, a loosely docked conformation with high FRET efficiency ($E = 81\%$) and a separated undocked state ($E = 50\%$), as shown in Fig. 6d). The automated data analysis of Deep-LASI confirmed the ADP-dependent kinetics of the domain sensor in good agreement with previous, manually evaluated results⁴⁰ (Fig. 6e, f). This demonstrates the proficiency of Deep-LASI for unsupervised data evaluation of smFRET data on proteins.

Finally, we tested the automated analysis of Deep-LASI applied to a different microscopy approach for smFRET, i.e. confocal single-molecule data on immobilized molecules that can be collected with microsecond time resolution. We chose the same double-labeled DNA origami structure introduced in Fig. 2a but with different combinations of docking strands. For these constructs, the measured FRET efficiencies will be the same but with different dynamics. By changing the hybridization length or by adding mismatched bases in the docking strand, the interaction time of each binding site can be tuned individually from ~1 ms to 10 s of seconds by adjusting the stabilization energy of DNA hybridization. As expected, shorter hybridization sequences lead to fast dynamics. Figure 6g shows a representative

intensity trajectory of a DNA origami structure (containing 6 nt complementary overhangs) that was classified as dynamic and the corresponding predictions of the state classifier. Although the unquenched state (state 2) shows a high variance in intensity, the state classifier predicts transitions with high accuracy and confidence. In the case of the 5 nt complementary overhangs, the dwell times approach 1 ms (Fig. 6h), and the output probability, p , of the state classifier decreases significantly due to the low signal-to-noise ratio of the trace. Thus, the probability value is an important parameter indicating the confidence the state classifier has in the assignment of the state and can be used as a threshold. Figure 6i (colored symbols) compares the mean dwell times extracted by Deep-LASI for all the confocal datasets with the results obtained by a newly developed shrinking-gate fluorescence correlation spectroscopy (sg-FCS) approach⁴¹. In sg-FCS, a pulsed light source is used such that the fluorescence lifetime information can be incorporated into the analysis. By shrinking the analysis window of photons based on their detected arrival time after excitation, we vary the relative brightness of two species with different fluorescence lifetimes (e.g. the low FRET and high FRET states). This leads to a robust extraction of the kinetic rates between the two states from the auto-correlation analysis of the FCS data. For all binding site combinations with 6 nt to 7 nt complementary overhangs, dwell times obtained by both methods are in excellent agreement. The largest deviation was found for the 6 nt binding sites in the asymmetric 6 nt/7 nt sample (Fig. 6i, purple) (a factor of 2) where there is large heterogeneity and limited statistics⁴¹. The dwell times for the sample with 5 nt complementary overhangs follow the exponential trend observed for longer binding sites but the binning of 0.6 ms, together with the resulting low signal-to-noise ratio, reach the current limit of Deep-LASI's state classifier. For completeness, we have included the results from Fig. 4e, f in Fig. 6i (gray triangles). There is a shift in dwell times between TIRF and confocal data due to the different temperatures of the two laboratories (~19 °C confocal, -22 °C TIRF, see Supplementary Note 6.4). Lower temperatures lead to a higher standard free energy and concomitantly longer binding time^{42,43}. In the case of the 6.5 nt

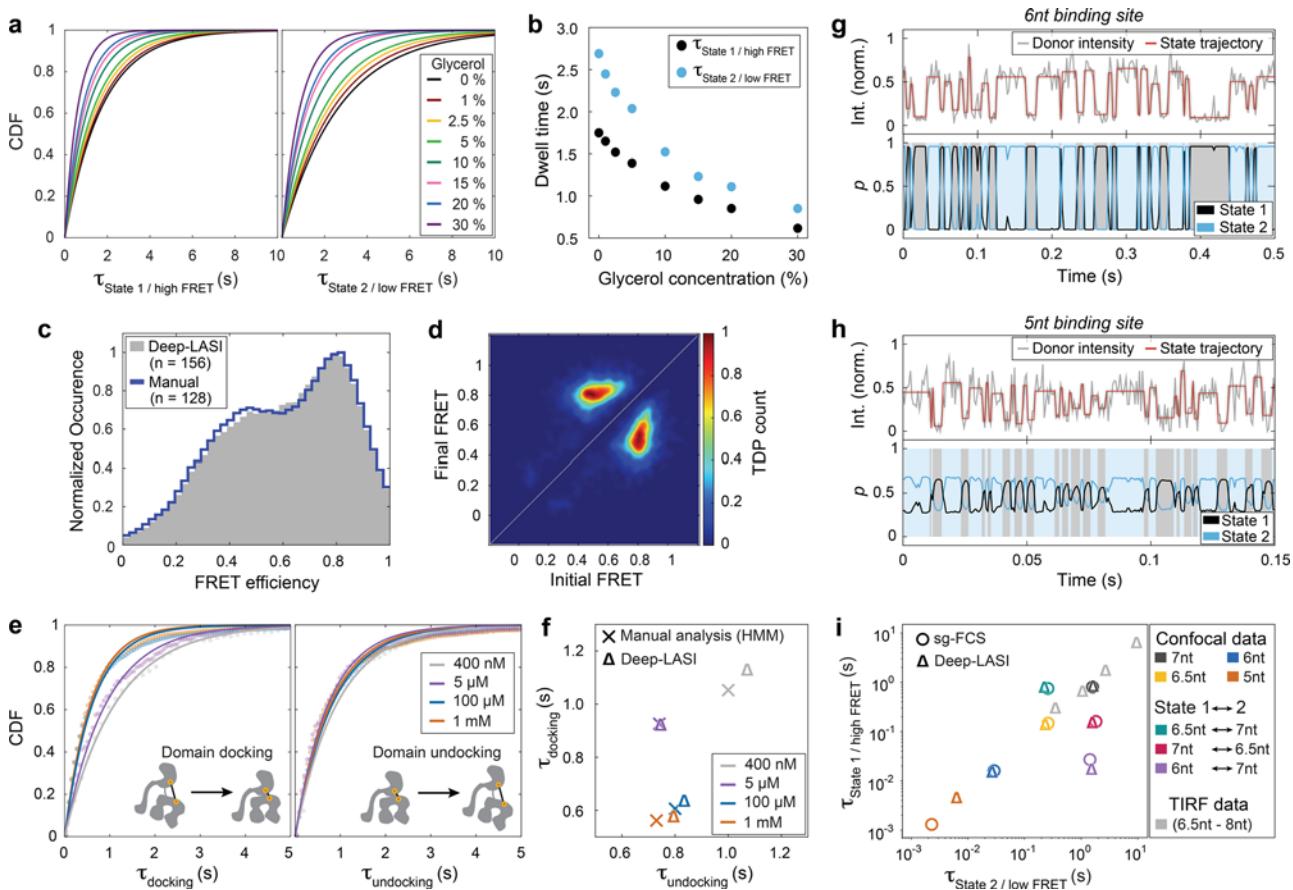


Fig. 6 | Use of Deep-LASI on titration experiments, protein data, and confocal data. **a, b** 3cFRET: Tuning the dissociation thermodynamics between protruding ssDNA strands by osmolytes. **a** CDFs of the dwell times, assessed through mono-exponential fits, for state 1 (left) and state 2 (right) of the L-shaped DNA origami structure from Fig. 4a decrease with increasing glycerol concentration. **b** Dependence of dwell times for both states versus glycerol concentration. **c–f** 2cFRET: Probing domain-domain interactions in Ssc1, a mitochondrial Hsp70. **c** Frame-wise smFRET distributions of Hsp70 molecules in the presence of 1 mM ADP classified as dynamic by Deep-LASI (gray) and evaluated manually (blue) from a total of 3534 traces. **d** The TDP generated by Deep-LASI aligns with the data plotted in (c), illustrating the interconversion between the undocked (-0.5) and docked (-0.8) conformations. Total number of transitions, $n = 3914$. **e** CDFs of the dwell times and mono-exponential fits to the dwell time distributions derived by Deep-LASI for domain docking (left panel) and domain undocking (right panel) depending on the ADP concentration. **f** Comparison between average dwell times extracted by Deep-

LASI (triangles) and by manual evaluation (crosses) using HMM. Deep-LASI matches the published trend with similar dwell times⁴⁰. **g–i** 1c-FRET: Deep-LASI analysis of ssDNA binding kinetics observed via confocal microscopy. **g** Confocal trace (with 2 ms binning) of the DNA origami structure from Fig. 2a with 6 nt binding sites and corresponding states predicted with high confidence. **h** Confocal trace (with 0.6 ms binning) of a DNA origami structure with 5 nt binding sites and predicted states. Due to the low SNR of the data, the confidence output of Deep-LASI reaches its lower limit. **i** Mean dwell times obtained from confocal data for various binding site lengths analyzed by sg-FCS⁴¹ (circles) and Deep-LASI (triangles). The results align well, except for dwell times extracted from the 5 nt sample, which was predicted with a low confidence distribution due to low SNR and a limited amount of information in the one-channel input. Dwell times obtained from TIRF data are displayed in light gray for comparison. nt: nucleotides. Source data are provided as a Source Data file.

binding sites sample (Fig. 6i, yellow), lower dwell times are consistently observed for the TIRF data. This discrepancy is due to the difference in temporal resolution of the two measurements (2 ms for confocal vs 30 ms for TIRF). The lower temporal resolution of the TIRF measurements led to a higher probability of fast transitions being averaged out and an underestimation of the actual transition time. This is a limitation of the real experimental data and is not attributable to Deep-LASI. On the contrary, Deep-LASI can back-trace shortcomings of either technique, identify rare events and monitor conformational changes over several time scales in an unsupervised manner.

Discussion

Deep-LASI is a deep-learning algorithm for the rapid and fully automated analysis of one-, two- and three-color single-molecule assays. Employing state-of-the-art neural network architectures optimized for time series data, we extended the classification of two-color FRET trajectories to include single- and three-color data, analyzed the

photobleaching information and incorporated a full state transition classification.

The utilization of deep-learning approaches for single-molecule analysis comes with both advantages and potential pitfalls. One major advantage is the ability of neural networks to capture intricate temporal dependencies and complex patterns in time-series data. This allows for improved classification accuracy and the identification of subtle transitions or states that may be challenging to discern using traditional analysis methods. Additionally, deep-learning models can learn from large amounts of data, reducing the dependence on prior assumptions that may introduce user bias.

It is essential to consider potential pitfalls when using deep neural networks for single molecule analysis. One challenge is the interpretability of the neural networks' decisions. While mathematical models and simpler thresholding techniques introduce user bias, they provide explicit confidence levels or probabilities derived from the user's modeling choices. In contrast, the output generated by neural

networks can be viewed as an artificial confidence level, minimizing user bias and increasing the consistency of the analysis. However, potential unknown biases inherent in the network itself may be introduced. Although neural networks can demonstrate high accuracy on validation datasets, understanding the underlying features and mechanisms influencing their predictions can be more challenging compared to conventional methods with explicit assumptions.

Ideal single-molecule traces are straightforward to model, and Deep-LASI is trained to be stringent when encountering non-ideal traces at a high confidence threshold. Hence, the default output of Deep-LASI when analyzing data with poor SNR is to discard the majority of the traces. By adjusting this threshold, users can instantaneously modify the accepted traces and monitor changes in the final results. This approach ensures a very low false positive classification rate when using a high confidence threshold and allows for gradual threshold reduction to increase statistical coverage. However, the outcome should be continuously monitored by the user. This iterative process effectively balances stringent classification and the need for increased statistical robustness in the analysis of new datasets.

Furthermore, Deep-LASI offers an advantage in terms of interpretability as it is trained solely on editable and extendable simulations. This characteristic provides users with greater control and knowledge over potential biases and enables them to tailor Deep-LASI to a wide range of experimental conditions. It is important to note that neural networks are data-driven models and heavily rely on the quality and representativeness of the training data. Therefore, careful consideration must be given to curating the training dataset to avoid biases and ensure the generalizability of the model to diverse experimental conditions. Regular validation and testing using independent datasets are crucial steps to assess the robustness and reliability of the model's performance. In addition, when measuring an unknown experimental system for the first time, it is helpful to visually inspect the traces that are being discarded to verify that the classification is still reasonable. By following these practices, researchers can enhance the trustworthiness and applicability of Deep-LASI in real-world scenarios.

In conclusion, Deep-LASI addresses the need for rapid, high-throughput screening of fluorescence intensity trajectories. This opens new possibilities for single-molecule assays and enables a timely analysis of complex experimental approaches thanks to the efficient and retrainable neural network architecture of Deep-LASI. It has a high potential for applications in a myriad of fields including biotheranostics, sensing, DNA barcoding, proteomics and single-molecule protein sequencing. We envision that deep-learning approaches, along with single-molecule sensitivity, will dramatically assist and accelerate analytics and be indispensable in the future.

Methods

Chemicals

Chemicals were purchased from Sigma-Aldrich and used without further purification, if not stated otherwise. Chemicals include acetic acid, agarose, ammonium persulfate, (3-aminopropyl)- triethoxysilane (APTES), biotin-poly(ethylene glycol)-silane (*biotin-PEG*, MW3000, PG2-BNSL-3k, Nanocs, NY; USA), bovine serum albumin (BSA; New England Biolabs, Ipswich, MA, USA), Blue Juice gel loading buffer (ThermoFisher Scientific), ethylene-diamine-tetraacetic acid sodium salt dehydrate (EDTA-Na₂ × 2H₂O), glycerol, magnesium chloride (MgCl₂ × 6H₂O), 2-[methoxy(polyethyleneoxy)propyl]trimethoxysilane (*mPEG*, #AB111226, abcr; Germany), phosphate-buffered saline (PBS), protocatechuate 3,4-dioxygenase from *Pseudomonas* sp. (PCD), protocatechuic acid (PCA), streptavidin, sodium chloride, Tris base, Tris HCl, and 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox) and beta-mercaptoethanol (βME).

All unmodified staple strands (Supplementary Note 7, Supplementary Table 7.2) used for DNA origami structure folding are

commercially available and were purchased from Integrated DNA Technologies®. Staple strands with modifications (Supplementary Tables 7.3 and 7.4) were obtained from Biomers (Supplementary Table 7.3: Biotin; Supplementary Table 7.4: Atto488) and Eurofins Genomics (Supplementary Table 7.4: binding sites, Cy3b and Atto647N).

DNA origami structures: assembly, purification and characterization

Preparation of the L-shaped DNA origami structures follows the procedures described previously by Tinnefeld et al.^{5,31}. In brief, the L-shaped DNA origami structures were folded with a 10-fold excess of 252 different, unmodified and labeled oligonucleotides to the complementary 8064 bp scaffold strand in folding buffer, which contained 40 mM Tris base, 20 mM acetic acid, 20 mM MgCl₂ × 6 H₂O, and 1 mM EDTA-Na₂ × 2 H₂O. A complete list with sequences of oligonucleotides used for producing the DNA origami structure is given in Supplementary Note 7. For folding, a nonlinear thermal annealing ramp over 16 hours was used⁴⁴.

After folding, the DNA origami solution was cleaned via gel electrophoresis in 50 mL 1.5% agarose-gel containing 1× gel buffer (40 mM Tris base, 20 mM acetic acid, 12 mM MgCl₂ × 6 H₂O, and 1 mM EDTA-Na₂ × 2 H₂O). The gel pockets were filled with a solution of 1× Blue Juice gel loading buffer and the DNA origami solution. The ice-cooled gel was run for 2 h at 60 V. When samples were to be recovered from the gel, the staining step was omitted and the Cy3b fluorescence was used instead to identify the correct DNA origami structures. Gel extraction was performed via cutting with a scalpel and squeezing the gel with a Parafilm® (Bernis®) wrapped glass slide. The concentration was determined by absorption spectroscopy on a NanoDrop 2000 device (ThermoFisher Scientific). Purified DNA origami structures were kept in storage buffer, i.e. in 1× TAE buffer (40 mM Tris base, 20 mM acetic acid and 1 mM EDTA-Na₂ × 2 H₂O) with 12.5 mM MgCl₂ × 6 H₂O (pH = 8.4).

The correct folding of the DNA origami structures was confirmed using atomic force microscopy and transition electron microscopy (see Supplementary Figure 7.2).

Sample preparation for multicolor prism-type TIRF experiments

Labeled DNA origami molecules were immobilized in flow channels formed between a coverslip and a surface-functionalized quartz prism. The surfaces were sandwiched on top of each other and sealed by a molten, pre-cut Nesco film (Nesco) channel. The employed prism surface was functionalized before with a biotin-PEG/mPEG coating to achieve surface passivation and prevent unspecific binding. Before the TIRF experiments, the prisms were first flushed with PBS and then incubated with a streptavidin solution (0.2 mg/mL) for 15 min. Afterwards, the sample holder was washed 3× with PBS to remove free streptavidin and then with storage buffer (1× TAE, 12.5 mM MgCl₂, pH = 8.4). Next, the DNA origami sample was diluted to 40 pM in storage buffer, added to the flow chamber and immobilized to the prism surface via the biotin-streptavidin linkage. After 5 min, untethered DNA origami structures were removed by rinsing the chamber 3× with storage buffer. Next, the attached fluorophores on the DNA origami structure were photostabilized by a combination of ROXS and an oxygen scavenging system based on PCA/PCD⁴⁵. The photostabilization buffer was mixed as follows: 1 μL of 100 mM Trolox/Ethanol solution was added to 97 μL storage buffer. The sample was then aged using a UV Lamp (M&S Laborgeräte GmbH, UVAC-6U, 2 × 6 W; 254 and 366 nm) until an equal ratio of Trolox and Trolox-quinone was formed (typically 6 minutes)⁴⁶. Immediately before starting the TIRF experiments, 1 μL of 100 mM PCA in methanol and 1 μL of 100 mM PCD solution (50% glycerol, 50 mM KCl, 100 mM Tris HCl, 1 mM EDTA-Na₂ × 2 H₂O, pH 8) was added to the total volume. In the case of samples containing Atto488, 1 μL of 14.3 M βME was added to the

photostabilization buffer. The sample chamber was flushed with photostabilization buffer and sealed to allow for enzymatic oxygen depletion. A minimum of 5 minutes waiting time preceded the experimental recordings. The photostabilization buffer was refreshed every 45 minutes until the end of the experiment.

All two- and three-color FRET experiments were carried out using msALEX³³, i.e. two- or three excitation lasers were alternated frame-wise. The lasers of different excitation wavelengths were synchronized using an acousto-optical filter (OPTO-ELECTRONIC, France) with the camera frame rate using an FPGA that synchronizes the excitation and simultaneous detection on the EMCCD cameras at 32 ms, 50 ms or 150 ms exposure times (depending on the sample) for 2000 (two-color) and 2400 (three-color) frames. The laser powers were set to 28 mW (0.022 mm², 491 nm), 16 mW (0.040 mm², 561 nm) and 10 mW (0.022 mm², 640 nm) for B-Y-R excitation.

Multi-color TIRF setup

Single-pair FRET experiments on surface-immobilized DNA origami structures were carried out on a home-built TIRF microscope with prism-type excitation as previously published⁴⁷. Three laser sources (Cobolt, Solna; Sweden) at 491 nm, 561 nm and 640 nm are available, and used for triple-color TIRF experiments with an alternation rate of 27 Hz (including a 2.2 ms frame transfer rate) between the B-Y-R laser excitation. The resulting emission was collected by a 60×water immersion objective (60×/1.27 WI Plan Apo IR, Nikon), cleaned up with a notch filter (Stopline® Notch 488/647, AHF), and the red emission was separated from the blue/yellow emission by a dichroic mirror (630DCXR AHF; Germany) followed by separation of the blue and yellow emission (560DCXR AHF). The emission was spectrally filtered (AHF Analysentechnik, Tübingen, Germany) for the blue (ET525/50), yellow (HQ595/50) and red (ET685/40) collection channels and afterwards detected on three EMCCD cameras (Andor iXon (1×)/iXon Ultra (2×), Andor Technologies, Belfast; UK) via the supplier's software Andor Solis (Version 4.29.30005.0; Oxford Instruments). Synchronization and alternation of the exciting laser sources, as well as the frame-wise data acquisition on three separate cameras, were achieved using a LabView-written program that controls a field programmable gate array (FPGA; NI cRIO-9073). While the program starts the measurement, the FPGA synchronizes the execution of the hardware via TTL pulses, i.e. it controls switching on/off the excitation sources by direct modulation of the AOTF (491, 561, 640 nm), while simultaneously starting the data acquisition by the three cameras. The videos were analyzed afterward by a custom-written MATLAB program (Version 9.13.0.2166757; Mathworks, Massachusetts, USA).

Single-molecule data analysis, data evaluation and representation

Time traces of individual, fluorescently labeled DNA origami structures were extracted from measurements using one, two or three cameras for one-, two- and three-color experiments, respectively, using Deep-LASI. Deep-LASI is written in MATLAB (Version 9.13.0.2166757; Mathworks, Massachusetts, USA) and uses neural networks trained with the Python library TensorFlow (Version 2.8.0). All raw data were recorded by EMCCD cameras (iXon 897, i.e. frames with 512 × 512 pixels containing fluorescence intensity information) and stored as TIFF stacks using the supplier's software Andor Solis. The resulting traces are then analyzed either using the pre-trained neural networks (Supplementary Notes 2, 3) or manually (Supplementary Note 5). The regions of single-molecule traces that were classified as dynamic with photoactive fluorophores were selected for downstream analysis. In the automated analysis procedure, the state transitions and state dwell times were predicted by a neural network model. All manually selected traces were analyzed using Hidden Markov Models, locally fit to each intensity trace (1-color data) or FRET trace (2-color data) assuming two states with Gaussian emission distribution functions and using the

Baum-Welch algorithm. The Gaussian emission distribution functions serve as the prior for the HMM, which are iteratively updated during the analysis. The convergence threshold was set to 10⁻⁹ and the maximum number of iterations was set to 10⁸. All predicted transitions were extracted from a transition density plot and the corresponding dwell times were fit to an exponential function. All correction factors for calculating the corrected FRET efficiency were determined using the manual or automated classification of photobleaching steps. All employed methods, automated and manual, were performed using the Deep-LASI user interface. Final panels were all presented using MATLAB 2022b (Version 9.13.0.2166757; Mathworks, Massachusetts, USA), exported as vector graphics, and assembled into figures using Adobe Illustrator CS2022 (Adobe Inc.; USA). 3D representations of the DNA origami structures were rendered in Blender (Version 2.93.6), and further assembled and labeled in Illustrator. The AFM images were plotted using JPK Desktop Software (Version 6.1.200 A).

Statistics & reproducibility

For training the neural network, we used a simulated dataset with ~200,000 traces as it is sufficient to cover an extensive range of realistic experimental parameters and thereby avoid any bias in the analysis. This includes FRET efficiencies between 0.01 and 0.99, dwell times of 1 to 100 frames and SNR of ~0.3 to 50. Experimentally, we typically measured 100 movies for each condition, as this usually generates several thousand acceptable traces. The full datasets were analyzed. The program, as part of its function, determines which intensity traces are suitable for further analysis. The computer selection was tested against simulated traces as well as compared with human analyses.

The Deep-LASI software was trained on three independently generated datasets. Deep-LASI was also compared with two users who manually analyzed the same datasets. The number of states, FRET efficiency histograms and kinetic rates extracted from the different analyses are consistent and, when available, are within the confidence intervals from the fits. Experiments were not randomized. The researchers were not blinded as knowledge regarding the sample did not influence the manual selection or analysis of the data. For the neural network, the advantage is that it operates only based on the data that it has been trained with. Hence, blinding is not applicable.

For analysis of the dwell-time distributions, a mono-exponential function was fit to the cumulative distribution function in MATLAB 2022b. The optimal fit values, along with the 95% confidence intervals, are given in the text and figures.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data for all figures and all supplementary figures have been deposited in the Zenodo database⁴⁸ [<https://zenodo.org/record/7561162>], with the exception of previously published data (HSP70 SSC1 and DNA origami confocal data in Fig. 6). Source data are provided with this paper.

Code availability

The program is available on GitLab [<https://gitlab.com/simon71/deeplasi>]. Extensive documentation for the Deep-LASI software package can be found at <https://deep-lasi-tutorial.readthedocs.io/en/latest/index.html>.

References

1. Lerner, E. et al. FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *eLife* **10**, e60416 (2021).

2. Hellenkamp, B. et al. Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nat. Methods* **15**, 669–676 (2018).
3. Agam, G. et al. Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins. *Nat. Methods* **20**, 523–535 (2023).
4. Chizhik, A. I., Rother, J., Gregor, I., Janshoff, A. & Enderlein, J. Metal-induced energy transfer for live cell nanoscopy. *Nat. Photon.* **8**, 124–127 (2014).
5. Kamińska, I. et al. Graphene energy transfer for single-molecule biophysics, biosensing, and super-resolution microscopy. *Adv. Mater.* **33**, 2101099 (2021).
6. Ploetz, E. et al. Förster resonance energy transfer and protein-induced fluorescence enhancement as synergetic multi-scale molecular rulers. *Sci. Rep.* **6**, 33257 (2016).
7. Lerner, E., Ploetz, E., Hohlbein, J., Cordes, T. & Weiss, S. A Quantitative Theoretical Framework For Protein-Induced Fluorescence Enhancement–Förster-Type Resonance Energy Transfer (PIFE-FRET). *J. Phys. Chem. B* **120**, 6401–6410 (2016).
8. Ploetz, E. et al. A new twist on PIFE: photoisomerisation-related fluorescence enhancement. *Methods and Applications in Fluorescence* (in press), <https://doi.org/10.1088/2050-6120/acfb58> (2023).
9. Hohng, S., Joo, C. & Ha, T. Single-molecule three-color FRET. *Biophys. J.* **87**, 1328–1337 (2004).
10. Barth, A., Voith von Voithenberg, L. & Lamb, D. C. Quantitative single-molecule three-color Förster resonance energy transfer by photon distribution analysis. *J. Phys. Chem. B* **123**, 6901–6916 (2019).
11. Voith von Voithenberg, L. et al. Comparative analysis of the coordinated motion of Hsp70s from different organelles observed by single-molecule three-color FRET. *Proc. Natl. Acad. Sci. USA* **118**, e2025578118 (2021).
12. Holden, S. J. et al. Defining the Limits of Single-Molecule FRET Resolution in TIRF Microscopy. *Biophys. J.* **99**, 3102–3111 (2010).
13. Schmid, S., Götz, M. & Hugel, T. Single-Molecule Analysis beyond Dwell Times: Demonstration and Assessment in and out of Equilibrium. *Biophys. J.* **111**, 1375–1384 (2016).
14. Van de Meent, J.-W., Bronson, J. E., Wiggins, C. H. & Gonzales, J. R. Empirical Bayes Methods Enable Advanced Population-Level Analyses of Single-Molecule FRET Experiments. *Biophys. J.* **106**, 1327–1337 (2014).
15. McKinney, S. A., Joo, C. & Ha, T. Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling. *Biophys. J.* **91**, 1941–1951 (2006).
16. Kaur, A., Ellison, M. & Dhakal, S. MASH-FRET: A Simplified Approach for Single-Molecule Multiplexing Using FRET. *Anal. Chem.* **93**, 8856–8863 (2021).
17. Preus, S., Noer, S. L., Hildebrandt, L. L., Gudnason, D. & Birkedal, V. iSMS: single-molecule FRET microscopy software. *Nat. Methods* **12**, 593–594 (2015).
18. Greenfeld, M., Pavlichin, D. S., Mabuchi, H. & Herschlag, D. Single Molecule Analysis Research Tool (SMART): An Integrated Approach for Analyzing Single Molecule Data. *PLoS ONE* **7**, e30024 (2012).
19. Juette, M. F. et al. Single-molecule imaging of non-equilibrium molecular ensembles on the millisecond timescale. *Nat. Methods* **13**, 341–344 (2016).
20. Ge, L., Liu, F. & Luo, J. Highly-efficient quantitative fluorescence resonance energy transfer measurements based on deep learning. *J. Innov. Opt. Health Sci.* **13**, 2050021 (2020).
21. Li, J., Zhang, L., Johnson-Buck, A. & Walter, N. G. Automatic classification and segmentation of single-molecule fluorescence time traces with deep learning. *Nat. Commun.* **11**, 5833 (2020).
22. Thomsen, J. et al. DeepFRET, a software for rapid and automated single-molecule FRET data classification using deep learning. *eLife* **9**, e60404 (2020).
23. Zarrabi, N., Schluesche, P., Meisterernst, M., Börsch, M. & Lamb, D. C. Analyzing the Dynamics of Single TBP-DNA-NC2 Complexes Using Hidden Markov Models. *Biophys. J.* **115**, 2310–2326 (2018).
24. Salem, C.-B., Ploetz, E. & Lamb, D. C. in *Spectroscopy and Dynamics of Single Molecules* (ed Carey K. Johnson) 71–115 (Elsevier, 2019).
25. Liu, D., Honoré, A., Chatterjee, S. & Rasmussen, L. K. Powering Hidden Markov Model by Neural Network based Generative Models. *24th European Conference on Artificial Intelligence (ECAI 2020)*, Santiago de Compostela, Spain, <https://doi.org/10.48550/arxiv.1910.05744> (2020).
26. Buys, J., Bisk, Y. & Choi, Y. Bridging HMMs and RNNs through Architectural Transformations. *32nd Conference on Neural Information Processing Systems (NIPS 2018)*, Montréal, Canada (2018).
27. Liu, L., Lin, Y.-C. & Reid, J. Improving the Performance of the LSTM and HMM Model via Hybridization. *ArXiv*, <https://doi.org/10.48550/arxiv.1907.04670> (2021).
28. Rallabandi, P. K. & Patidar, K. C. A Hybrid System of Hidden Markov Models and Recurrent Neural Networks for Learning Deterministic Finite State Automata. *Int. J. Comp. Sci. Inf. Eng.*, <https://doi.org/10.5281/zenodo.1109543> (2015).
29. Kravonna, V. & Doshi-Velez, F. Increasing the Interpretability of Recurrent Neural Networks Using Hidden Markov Models. *2016 ICML Workshop on Human Interpretability in Machine Learning (WHI 2016)*, New York, NY, <https://doi.org/10.48550/arxiv.1606.05320> (2016).
30. Xu, J. et al. Automated stoichiometry analysis of single-molecule fluorescence imaging traces via deep learning. *J. Am. Chem. Soc.* **141**, 6976–6985 (2019).
31. Krause, S. et al. Graphene-on-glass preparation and cleaning methods characterized by single-molecule DNA origami fluorescent probes and Raman spectroscopy. *ACS Nano* **15**, 6430–6438 (2021).
32. Gotz, M. et al. A blind benchmark of analysis tools to infer kinetic rate constants from single-molecule FRET trajectories. *Nat. Commun.* **13**, 5402 (2022).
33. Kapanidis, A. N. et al. Alternating-Laser Excitation of Single Molecules. *Acc. Chem. Res.* **38**, 523–533 (2005).
34. Schreiber, J. Pomegranate: fast and flexible probabilistic modeling in python. *J. Mach. Learn. Res.* **18**, 1–6 (2018).
35. Kudryavtsev, V. et al. Combining MFD and PIE for accurate single-pair Förster resonance energy transfer measurements. *Chem. Phys. Chem.* **13**, 1060–1078 (2012).
36. Lee, N. K. et al. Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys. J.* **88**, 2939–2953 (2005).
37. Diaz, A., Jothiraman, H. B. & Ramakrishnan, V. Effect of glycerol on free DNA: A molecular dynamics simulation study. *J. Mol. Graph. Model.* **114**, 108169 (2022).
38. Wang, C., Altieri, F., Ferraro, A., Giartosio, A. & Turano, C. The effect of polyols on the stability of duplex DNA. *Physiol. Chem. Phys. Med. NMR* **25**, 273–280 (1993).
39. Spink, C. H., Garbett, N. & Chaires, J. B. Enthalpies of DNA melting in the presence of osmolytes. *Biophys. Chem.* **126**, 176–185 (2007).
40. Sikor, M., Mapa, K., von Voithenberg, L. V., Mokranjac, D. & Lamb, D. C. Real-time observation of the conformational dynamics of mitochondrial Hsp70 by spFRET. *EMBO J.* **32**, 1639–1649 (2013).
41. Schroder, T. et al. Shrinking gate fluorescence correlation spectroscopy yields equilibrium constants and separates photophysics from structural dynamics. *Proc. Natl. Acad. Sci. USA* **120**, e2211896120 (2023).
42. Fornace, M. E., Porubsky, N. J. & Pierce, N. A. A Unified Dynamic Programming Framework for the Analysis of Interacting Nucleic Acid Strands: Enhanced Models, Scalability, and Speed. *ACS Synth. Biol.* **9**, 2665–2678 (2020).

43. Dirks, R. M., Bois, J. S., Schaeffer, J. M., Winfree, E. & Pierce, N. A. Thermodynamic Analysis of Interacting Nucleic Acid Strands. *SIAM Rev.* **49**, 65–88 (2007).
44. Nickels, P. C. et al. Molecular force spectroscopy with a DNA origami-based nanoscopic force clamp. *Science* **354**, 305–307 (2016).
45. Aitken, C. E., Marshall, R. A. & Puglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys. J.* **94**, 1826–1835 (2008).
46. Cordes, T., Vogelsang, J. & Tinnefeld, P. On the mechanism of Trolox as antiblinking and antibleaching reagent. *J. Am. Chem. Soc.* **131**, 5018–5019 (2009).
47. Kopperger, E. et al. A self-assembled nanoscale robotic arm controlled by electric fields. *Science* **359**, 296 (2018).
48. Wanninger, S. et al. Deep-Learning assisted, Single-molecule Imaging analysis of multi-color DNA Origami structures. Zenodo, URL: <https://zenodo.org/record/7561162> (2023).

Acknowledgements

We thank Julian Heeg for helping with data collection. We thankfully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 201269156 – SFB 1032 Project B03 (to D.C.L.), Project A13 (to P.T.), individual grants to PL696/4-1 (to E.P.), TI 329/9-2, project number 267681426, TI 329/14-1, and TI 329/15-1 (to P.T.) and Germany's Excellence Strategy – EXC 089/1 – 390776260. D.C.L. and P.T. gratefully acknowledge funding from the Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the Federal Government and the Länder through the ONE MUNICH Project Munich Multiscale Biofabrication. P.T. acknowledges the support of BMBF (SIBOF, 03VP03891). D.C.L., P.T., and E.P. gratefully acknowledge the financial support of the Ludwig-Maximilians-Universität München via the Department of Chemistry, the Center for NanoScience (CeNS) and the LMUinovativ program Biolmaging Network (BIN).

Author contributions

S.W. developed and implemented the deep-learning algorithm Deep-LASI and performed the Deep-learning assisted analyses, J.B. prepared the DNA origami samples under the supervision of P.T., P.A. collected the single-molecule TIRF data, P.A. and E.P. performed the manual analysis of the smFRET data, C.B.S. wrote the manual analysis software in which Deep-LASI was incorporated, S.W. and E.P. wrote the first draft of the manuscript, S.W. and E.P. designed the figures, all authors

contributed to revising the manuscript. E.P. and D.C.L. supervised the project.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

All authors (S.W., P.A., J.B., C.B.S., P.T., E.P., and D.C.L.) declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-42272-9>.

Correspondence and requests for materials should be addressed to Evelyn Ploetz or Don C. Lamb.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

Supplementary Information

Deep-LASI: Deep-Learning Assisted, Single-molecule Imaging analysis of multi-color DNA Origami structures

Simon Wanninger¹, Pooyeh Asadiatouei¹, Johann Bohlen¹, Clemens-Bässem Salem¹,
Philip Tinnefeld¹, Evelyn Ploetz^{1,*} and Don C. Lamb^{1,*}

¹ Department of Chemistry and Center for NanoScience (CeNS)
Ludwig-Maximilians-Universität München
Butenandtstr. 5-13 81377 Munich,
Germany

* Corresponding authors: evelyn.ploetz@lmu.de; d.lamb@lmu.de

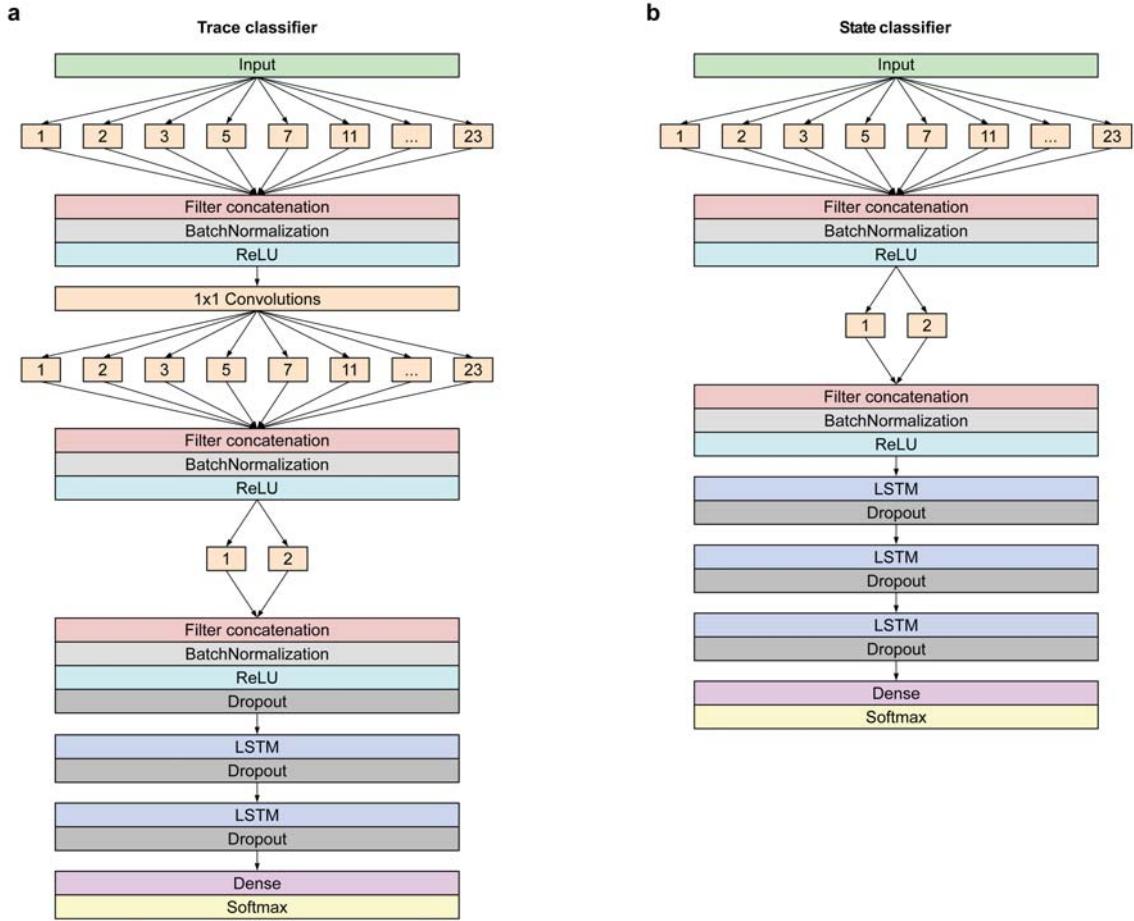
TABLE OF CONTENTS

| | |
|---|-----------|
| Supplementary Note 1: Neural network | 3 |
| 1.1. Architecture | 3 |
| 1.2. Trace classifier architecture | 4 |
| 1.3. State transition classifier and number of states classifier architecture..... | 5 |
| Supplementary Note 2: Training | 6 |
| 2.1. Training procedure..... | 6 |
| 2.2. Training dataset preparation | 6 |
| 2.3. Simulation of single molecule traces..... | 8 |
| 2.4. Simulation settings for training the state classifier network | 12 |
| Supplementary Note 3: Training validation..... | 14 |
| 3.1. Trace classifiers..... | 14 |
| 3.2. Number-of-states classifiers..... | 16 |
| 3.3. State-transition classifiers | 16 |
| 3.4. Limitations of the state classifiers and a comparison with HMM..... | 17 |
| 3.5. Comparison of the state classifiers with local and global HMM | 19 |
| 3.6. Training and validation loss..... | 20 |
| 3.7. Analysis of kinetic data from the kinsoft challenge | 22 |
| Supplementary Note 4: Deep-LASI versus manual analyses | 23 |
| 4.1 Comparison of Deep-LASI and manual analysis of 2-color DNA origami traces .. | 23 |
| 4.2. Influence of different training datasets and comparison to user classification | 24 |
| 4.3 Deep-LASI versus manual analyses for 3-color DNA origami samples..... | 25 |
| 4.4 Comparison of 3-color DNA origami traces selected manually and/or via DNN classification..... | 27 |
| Supplementary Note 5: Manual Analysis of Single-molecule TIRF Data | 28 |
| 5.1. Work-flow | 28 |
| 5.2. Camera mapping for FRET traces..... | 29 |
| 5.3. Trace extraction and background subtraction | 29 |
| 5.4. Manual trace selection and analysis..... | 30 |
| 5.5. Accurate FRET determination..... | 30 |
| 5.6. Hidden-Markov modeling | 32 |
| 5.7. Parameters for Hidden Markov Modeling | 34 |
| 5.8. Evaluation of involved FRET states and interconversion rates | 34 |
| Supplementary Note 6: Details of Deep-LASI Analyses..... | 35 |
| 6.1 Results for the three-color, two-state DNA origami structure with different binding site lengths | 35 |
| 6.2 Kinetics of the three-color, three-state DNA origami. | 36 |
| 6.3 Results for the two-color, three-state DNA origami structure..... | 37 |
| 6.4 Kinetics as a function of Temperature..... | 39 |
| 6.5 Analysis of previous published 2-color Hsp70 Ssc1 using Deep-LASI | 39 |
| Supplementary Note 7: DNA sequences..... | 41 |
| Supplementary Note 8: Statistics Single-Molecule Data | 50 |
| Supplementary References..... | 51 |

SUPPLEMENTARY NOTE 1: NEURAL NETWORK

1.1. Architecture

For the Deep-LASI software package, two different neural-network architectures are used. One architecture is for trace classification and another for the number of states and state transition classification (Supplementary Figure 1.1). Both architectures are hybrids of a convolutional neural network (CNN) and a long short-term memory (LSTM) model, which were designed using TensorFlow with Keras API.¹ The CNN framework was inspired by an omni-scale 1D-CNN, which elegantly solves the problem of finding the optimal kernel sizes by making it part of the training process.² Unlike traditional CNNs that operate on 2D spatial grids, the omni-scale CNN processes time series data directly by utilizing one-dimensional convolutions. These convolutional operations consider the temporal dependencies among data points, enabling the network to effectively model the sequential nature of time series data. Single-molecule Förster Resonance Energy Transfer (smFRET) data often exhibit complex patterns on different timescales, which can be crucial for accurate classification. The omni-scale CNN architecture incorporates multiple convolutional layers with different kernel sizes at the same level of depth, enabling the extraction of features at various temporal resolutions. However, the pure CNN architecture can only produce a single classification for the entire input sequence, i.e. one category per single molecule trace. The inclusion of LSTM after feature extraction of the CNN enables the sequential classification of each trace, producing a classification output at each time step. This characteristic of LSTM allows for the fine-grained detection of valid smFRET frames, photobleaching steps and state transitions. For all of our presented tasks, the omni-scale CNN LSTM hybrid architecture outperformed pure LSTM or ResNet³ models. We did not employ the full range of prime numbers suggested for the kernel sizes as we found the accuracy did not increase above 23. Hence, the number of trainable parameters was greatly reduced. In the trace classifier model, we added a 1x1 convolution layer for dimensionality reduction to further increase efficiency without a trade-off in validation accuracy. Any down-sampling of the time dimension was avoided since the loss of this information significantly decreased the validation accuracy. This was achieved by omitting any kind of pooling or averaging layers, by zero-padding all inputs for the convolutional layers and setting the stride of all convolutional filters to 1. For kernel initialization, we used the He Normal distribution⁴ as it showed the fastest convergence rate. Each convolution layer is followed by a batch normalization layer and an activation layer using the rectified linear unit (ReLU) activation function.⁵



Supplementary Figure 1.1: The deep neural network architectures used for the trace classifier (a), and for the state transition classifier and the number of states classifier (b).

1.2. Trace classifier architecture

The trace classifier consists of four convolution layers followed by two LSTM layers and one fully connected layer as the feature extraction module (Supplementary Figure 1.1a). In the first convolution layer, the input is fed into 10 layers with 32 filters each. The kernel filter size is varied between layers with sizes given by the prime numbers from 1 to 23. All layers are stacked sequentially, i.e. they operate on the same level of depth. The second convolution layer serves as a dimensionality reduction layer with 32 filters and a kernel size of 1. The third convolution layer has the same hyperparameters as the first layer. A fourth convolution layer is added, composed of two branches with 32 filters each and kernel sizes of one and two, which allows the receptive fields of the network to cover all possible integers. The output of the CNN is fed into a LSTM layer with 128 units, followed by a second LSTM layer with 32 units and the final dense layer for classification. For the training procedure, we placed a dropout layer at a rate of 0.22 before the first LSTM layer and two dropout layers at a rate of 0.5 after the two consecutive LSTM layers to maximize the validation accuracy and reduce overfitting.

1.3. State transition classifier and number of states classifier architecture

The main difference in architecture between the transition classifier and the trace classifier are the depths and widths of the CNN and LSTM structure. The state transition classifier is composed of two convolution layers, three LSTM layers, and one final dense layer ([Supplementary Figure 1.1b](#)). The kernel sizes of the first convolution layer are prime numbers in the range of 1 to 23 with 64 filters each. The second convolution layer has kernel sizes of one and two with 32 filters each. The CNN substructure is directly followed by three LSTM layers with 128 units. Dropout layers are placed after each LSTM layer using a rate of 0.5. At the end, a fully connected layer (or dense layer) is used to reduce the output of the network into the number of given categories.

SUPPLEMENTARY NOTE 2: TRAINING

Although an optimized architecture is important and improves the efficiency of a neural network, the functionality of the network rises and falls with the dataset used for training the network. There are a number of important factors to consider when training a neural network. Typical pitfalls include using a dataset that is too small in size or contains an intrinsic bias, or overfitting the training data. A neural network is biased towards features it has seen before. Hence, the training dataset should include the various possibilities (e.g. number of FRET states, kinetic rates, signal-to-noise ratios). If the training dataset includes any bias, this will also be reflected in the output of the algorithm. One way that bias can be introduced into the training dataset is from unbalanced sampling of categories. For example, for the trace classifier models, it is important that the training dataset includes the same number of traces from each category. It is also important to know when to stop the training process. Neural networks can be overtrained, meaning that they memorize the training data but do not learn the general principles behind it. Below we discuss the details of the training procedure and how we optimized the training process.

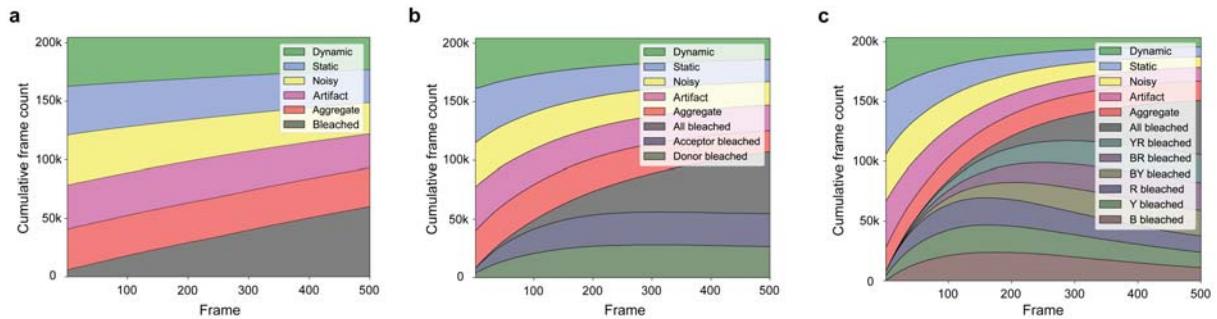
2.1. Training procedure

Both the trace classifier and transition classifier models were trained using the Adam optimizer with the default settings⁶. We used a hybrid method of increasing the batch size and lowering the learning rate during training. The entire training set of ca 200,000 traces is feed into the neural network in batch sizes of 32 traces until the network has seen all traces (referred to as an epoch). After the network has seen all traces once, some input units are randomly set to 0 using dropout layers and the dataset is fed again in batches to the neural network in the next epoch. The dropout layers reduce overfitting and allow generalization of the learned information. When the validation loss is not significantly lowered within 4 epochs, the batch size is doubled. An initial learning rate of 0.001 was decreased analogously by factors of 10 after a batch size of 512 was reached.

2.2. Training dataset preparation

To generate training datasets, we found the approach of using simulations, originally described in Thomsen et al.⁷, to be the most promising. This is especially true for three-color models capable of detecting state transitions or photobleaching events of each dye individually. A manual collection of labeled traces on a scale large enough for adequate training would be prone to biases and/or errors due to incorrect trace identification. In addition, the datasets would not be optimized for microscope setups with different characteristics. The signal and noise characteristics of smFRET data is well enough understood that simulated data can accurately reproduce the characteristics of real data. Beside the architecture itself, the main differences between our trace classifiers and the DeepFRET model⁷ is the ability to classify one-color and three-color data and to predict the photobleached frames of each fluorophore separately. We adopted the categories ‘dynamic’, ‘static’, ‘noisy’, and ‘aggregate’ while implementing

additional categories for all possible photobleaching events. The ‘artifact’ category includes false localizations, overestimated background and random perturbations of the intensity traces. For the simulated data, idealized intensity traces for each primary category (i.e. non-photobleaching category, ‘dynamic’, ‘static’, ‘noisy’, ‘aggregate’ and ‘artifact’) are generated and then photobleaching steps are added for the different dyes by randomly determining the survival time of the dye from a given exponential distribution. The addition of photobleaching as well as other processes in the simulation may lead to alterations in the label of the given trace. For example, a simulated ‘dynamic’ trace that does not undergo a transition before photobleaching or by the end of the trace would be recategorized as ‘static’. To ensure that the network sees the same number of frames for the different categories including all the photobleached categories, the number of traces selected for the training set needs to be balanced. Hence, we begin by simulating ~250,000 traces of 500 frames for each primary category. After including photobleaching, the number of labeled frames for each category is determined. The category with the minimum number of frames determines a threshold at which additional traces, depending on their present classification, are added or excluded from the final training dataset. The typical number of traces included in the final training dataset is approximately 200,000. This balancing procedure ensures that no category is over- or underrepresented across all frames and minimizes biases of the trained deep neural networks. **Supplementary Figure 2.1** shows the cumulative distribution of category labels in each training set used for trace classification.



Supplementary Figure 2.1: Cumulative distribution of labeled categories in training datasets for one-color (a), two-color (b) and three-color data (c). The same dataset was used for training the continuous wave two-color network as for the two-color ALEX network with the exception that the ALEX channel was not included.

For training of the state classifiers, only frames where all dyes are photoactive are included and hence photobleaching can be ignored for training these classifiers. The number of categories in the training sets then equals the number of states in the model. The visible states are first counted and sorted according to their chronological order before the state of each frame is assigned. This results in the first observed state always receiving the first label regardless of the FRET efficiency value and hence a state label only corresponds to a particular FRET value when a given dataset is analyzed globally.

2.3. Simulation of single molecule traces

Single-molecule intensities traces are simulated by first initializing the number of traces to be simulated (250,000 in our case) and the probability of the trace being a single-molecule trajectory or an 'aggregate'. All parameters used for simulating single-molecule traces are given in [Supplementary Table 2.1](#). An idealized FRET efficiency trajectory is then generated using an HMM routine⁸ in one, two or three colors where the FRET efficiency (or efficiencies) is randomly selected between 0.01 and 0.99. For single molecule trajectories, the number of states in the trajectory is selected with a probability of containing only a single state (45 % in our case for 'static' traces) and the remaining probability (55 % here) is equally distributed between two, three and four states ('dynamic'). In the case of 'dynamic' traces, the FRET efficiency or efficiencies of each state are randomly selected from a uniform distribution between 0.01 and 0.99. The difference in FRET efficiency for the different states has to be above a given threshold (0.1 in our cases). If this is not the case, new FRET efficiencies are randomly selected until this criterion is fulfilled. The transition rates are generated by taking the inverse of the dwell time to exposure ratio drawn from a uniform distribution between 1 and 100. The generated transition matrices are then, in general, non-symmetric. Hence, we use the transition rate matrix to calculate the probability of which state is observed first. We do this by using the least-squares solution to the matrix equation $Ax = b$, where A is the transition matrix and x is the probabilities for observing the different states. While the calculation of the state equilibrium is not mandatory for the classification accuracy, it ensures that the output matches the ground truth input of a defined transition matrix, which was used for benchmarking the transition classifier. Once the initial state has been selected, the parameters are fed into an HMM routine⁸ and a state trajectory is generated. 'Aggregate' traces are always assumed to be static (uncorrelated dynamics are categorized as 'artifact'), but the number of dye-pairs is generated from a Poisson distribution with randomly selected FRET efficiencies between each pair. Next, the idealized FRET efficiency trace is converted into normalized fluorescence intensity traces for the donor and acceptor molecules based on the FRET efficiency (discussed in more detail below). Next, photobleaching of the fluorophores are included into the trajectories. The frame at which each fluorophore photobleaches is randomly drawn from an exponential distribution. Upon photobleaching, the affected channel intensities are either set to 0 for both channels for donor photobleaching, 1 for the donor intensity upon acceptor photobleaching or recalculated using the two-color FRET equations (for three-color simulations). Blinking is then added to a fraction of the traces where each dye has a probability of being in a short-lived dark state ([Supplementary Table 2.1](#)). At this point, 'artifact' traces are generated from 'static' or 'dynamic' traces with a given probability by subtracting a constant from the trace (to simulate overestimation of the background correction), adding random fluctuations to the total intensity (to simulate among other things new molecules or aggregates flowing through the observation volume or simulating molecules in the background mask), flipping the traces (to simulate molecules that turn on during the experiment) and/or adding non-correlated signal in the different channels. To account for non-uniform brightness of the individual molecules, all excitation channels are multiplied by a scaling factor that is randomly selected from uniform distributions. In particular, the red channel after red excitation, I_{RR} , can reach scaling factors

up to three times higher than the other two channels. This allows the trace classifier to correctly analyze datasets in which high red laser powers were used to increase statistics for the calculation of correction factors and for making sure only a single red fluorophore is present. Without intensity scaling, the trace classifier strongly favors the aggregate category for traces with imperfect stoichiometry even when no second bleach step is present. With a given probability, additional small fluctuations in the total intensity are also added to the traces to simulate experimentally observed system instabilities (assuming sinusoidal oscillations of randomly determined frequency and amplitude). Next, we incorporate spectral crosstalk, direct excitation and differences in detection efficiency into the data by randomly selecting the respective parameters from a uniform distribution (see [Supplementary Table 2.1](#)). In the last step, we add two or three types of random noise to the traces. The first component is intensity-independent background noise drawn from a Poisson distribution. The second component considers intensity-dependent noise contributions (i.e. shot-noise) by drawing values from Gaussian distributions. There are different descriptions of how to treat noise from EMCCD cameras. According to Basden et al., the variance in shot noise due to the EM gain is increased by a factor of two.⁹ This corresponds to a rescaling of the Gaussian distribution mentioned above. Hirsch et al describe the additional noise from EM-CCD cameras using a gamma distribution.¹⁰ Hence, with a given probability, we also add a third component to the noise modeled using a gamma distribution with random amplitude. After adding noise to the trajectories, we then recategorize traces with high noise as 'noisy'. This is done by recalculating the FRET efficiency trace or traces from the intensity data. When the standard deviation for static traces or individual states of a dynamic trace are above the given threshold ([Supplementary Table 2.1](#), we used 0.25), the trace is categorized as 'noisy'. Finally, the classification of the individual traces is checked and, if necessary (for example a dynamic trace that photobleaches before a transition is observed), recategorized. The dataset is then balanced, as discussed above, each trace normalized to its maximum value and then used for training. For one-color traces, we simulated the intensity of the donor molecule although, for a single channel, it does not make a difference. The donor intensity (we refer to it as YY here) is given by:

$$I_{YY} = 1 - E_{YR} \quad \text{Eq. 2.1}$$

Since only one dye and one channel is observed, there is only one photobleaching category and no correction factors are included. The photoactive state of the acceptor molecule is still calculated and its influence on the donor intensity incorporated into the trace. For calculation of 'aggregates', fluorescent dye-pairs are added to the trace but only the donor signal is considered. Furthermore, the amount of noise is not quantified by the standard deviation of the FRET efficiency but by the signal-to-noise ratio of the channel intensity, which is defined as:

$$\text{SNR} = \frac{\mu}{\sigma} \quad \text{Eq. 2.2}$$

where μ is the mean signal intensity of the observed state and σ is its standard deviation. When the signal-to-noise ratio falls below the given threshold, the trace is classified as 'noisy'.

For two-color FRET simulations, we use the same approach as described in the following section for 3-color FRET but we only consider the equations necessary for 2-color FRET, i.e. all equations including the yellow/red FRET pair.

For generating three-color FRET data, the distances and Forster radii between all three fluorophores need to be considered as they are interrelated. Assuming a minimum FRET efficiency of 0.01, the generated FRET states of the first two randomly drawn FRET pairs put constraints on the maximum possible distance for the third FRET pair. To guarantee a uniform distribution of possible FRET combinations, we randomly select two of the three FRET pairs and their corresponding FRET efficiencies. Using the two selected FRET efficiencies, a lower limit is calculated for the third FRET pair, which depends on the Förster radii of the first two FRET pairs. For example, when the yellow-red dye-pair is generated last, the dye-dye separation for r_{BY} and r_{BR} are calculated and then used to determine the minimum FRET efficiency (i.e. maximum separation for the third dye-pair) as given below:

$$r_{BY} = R_{0,BY} \left(\frac{1}{E_{BY}} - 1 \right)^{\frac{1}{6}} \quad \text{Eq. 2.3}$$

$$r_{BR} = R_{0,BR} \left(\frac{1}{E_{BR}} - 1 \right)^{\frac{1}{6}} \quad \text{Eq. 2.4}$$

$$E_{YR,min} = \frac{1 - 0.01}{1 + \left(\frac{r_{BY} + r_{BR}}{R_{0,YR}} \right)^6} + 0.01 \quad \text{Eq. 2.5}$$

where the Forster radii $R_{0,BY}$, $R_{0,BR}$ and $R_{0,YR}$ are sampled over values that are typically available using commercially available dyes pairs. The FRET efficiency $E_{YR,min}$ represents the lower boundary used to randomly scale the FRET trace of the yellow-red FRET pair in a correlated or anti-correlated manner. When a different dye-pair is generated last, the same equations are used where the indices are changed accordingly. For dynamic traces, this procedure is performed for all states. Once we have selected the FRET efficiencies for the different dye-pairs and states, we then convert them what would be observed for a two-color experiment. The YR dye-pair is already a two-color FRET efficiency and does not need to be corrected. When all three fluorophores are photoactive, the blue dye may be quenched by two acceptors. In this case, the distance-related FRET efficiencies E_{BY} and E_{BR} need to be converted into the apparent FRET efficiencies $E_{BY,app}$ and $E_{BR,app}$ via:

$$E_{BY,app} = \frac{E_{BY}(1 - E_{BR})}{1 - E_{BY}E_{BR}} \quad \text{Eq. 2.6}$$

$$E_{BR,app} = \frac{E_{BR}(1 - E_{BY})}{1 - E_{BY}E_{BR}} \quad \text{Eq. 2.7}$$

Since the input data for the neural networks are normalized, the channel intensities are initialized as follows:

$$I_{BB} = 1 - E_{BY,app} - E_{BR,app} \quad \text{Eq. 2.8}$$

$$I_{BY} = E_{BY,app}(1 - E_{YR}) \quad \text{Eq. 2.9}$$

$$I_{BR} = E_{BR,app} + E_{BY,app}E_{YR} \quad \text{Eq. 2.10}$$

$$I_{YY} = 1 - E_{YR} \quad \text{Eq. 2.11}$$

$$I_{YR} = E_{YR} \quad \text{Eq. 2.12}$$

$$I_{RR} = 1 \quad \text{Eq. 2.13}$$

Upon photobleaching of one of the dyes in the three-color experiments, the system then reverts into the two-color case:

$$I_{BY,2c} = -\frac{I_{BB}E_{BY}}{E_{BY} - 1} \quad \text{Eq. 2.14}$$

$$I_{BR,2c} = -\frac{I_{BB}E_{BR}}{E_{BR} - 1} \quad \text{Eq. 2.15}$$

I_{BB} is still determined by using

$$I_{BB,2c} = 1 - E_{BY} \quad \text{Eq. 2.16}$$

$$I_{BB,2c} = 1 - E_{BR} \quad \text{Eq. 2.17}$$

where FRET to the blinking fluorophore is set equal to zero in two-color sections of the trace. Blinking events are treated the same way as photobleaching during the frames where the one dye is off and the channel intensities are either set to 0 or recalculated using Eq. 2.16/2.17.

In three-color experiments, each channel has its own set of correction factors for differences in detection efficiency and quantum yield, γ , direct excitation, de , and spectral crosstalk, ct . The values are randomly drawn from a wide uniform range and implemented in the following order. First, the FRET channels are multiplied by the corresponding γ -factor:

$$I_{BY} = \gamma_{BY}I_{BY} \quad \text{Eq. 2.18}$$

$$I_{BR} = \gamma_{BR}I_{BR} \quad \text{Eq. 2.19}$$

$$I_{YR} = \gamma_{YR}I_{YR} \quad \text{Eq. 2.20}$$

The crosstalk of the blue fluorophore leaking into the yellow and red channel after blue excitation are given by:

$$I_{BY}^{ct} = \frac{ct_{BY}}{1 + (ct_{BY} + ct_{BR})} I_{BB} \quad \text{Eq. 2.21}$$

$$I_{BR}^{ct} = \frac{ct_{BR}}{1 + (ct_{BY} + ct_{BR})} I_{BB} \quad \text{Eq. 2.22}$$

where ct_{BY} and ct_{BR} denote the randomly drawn crosstalk factors (Supplementary Table 2.1). Spectral crosstalk of the yellow fluorophore into the red channel is calculated using:

$$I_{(BY)R}^{ct} = \frac{ct_{YR} I_{BY}}{(1 + ct_{YR})} \quad \text{Eq. 2.23}$$

$$I_{YR}^{ct} = \frac{ct_{YR} I_{YY}}{(1 + ct_{YR})} \quad \text{Eq. 2.24}$$

The observed intensities including all correction factors are determined by:

$$I_{BB,obs} = I_{BB} \quad \text{Eq. 2.25}$$

$$I_{BY,obs} = \gamma_{BY} I_{BY} + de_{BY} I_{YY} - I_{(BY)R}^{ct} \quad \text{Eq. 2.26}$$

$$I_{BR,obs} = \gamma_{BR} I_{BR} + de_{BR} I_{RR} + E_{YR} \frac{de_{YR} I_{YY}}{(1 - E_{YR})} \quad \text{Eq. 2.27}$$

$$I_{YY,obs} = I_{YY} - I_{YR}^{ct} \quad \text{Eq. 2.28}$$

$$I_{YR,obs} = \gamma_{BY} I_{YR} + de_{YR} I_{RR} \quad \text{Eq. 2.29}$$

where de_{BY} , de_{BR} and de_{YR} are the randomly drawn direct excitation factors from a uniform distribution (Supplementary Table 2.1).

While the non-smFRET categories ‘noisy’ and ‘aggregate’ mimic experimental data, the category ‘artifact’ is primarily designed to increase the robustness of the trace classifier. It is important to note that the accuracy of a trained neural network to distinguish between an ‘artifact’ and any other category depends on the number of traces which are labeled as ‘artifact’ but maintain a strong resemblance to the original trace. For the goal of increasing robustness, it is therefore not desirable to achieve 100% prediction accuracy as it would be caused by too easily identifiable perturbations in the training dataset.

2.4. Simulation settings for training the state classifier network

Sixteen pre-trained deep neural networks are provided for state classification. Four models account for the classification and segmentation of time trajectories obtained from measurements using single-channel data acquisition, two-color FRET with continuous wave excitation, two-color FRET with ALEX, and three-color FRET with ALEX. For each type of experiment (one, two and three-color), we provide three state-transition-classifiers trained on either two, three or four observed states. The state classifier networks only use traces as input

that are categorized as dynamic. Hence, the training datasets only contain valid FRET traces with at least one transition, a minimum state difference of 0.1 in FRET efficiency and no photobleaching. The transition rates are generated by drawing random dwell time to exposure ratios between 1 and 100 from a uniform distribution. Traces with a state-wise FRET distribution width above 0.25 on average are excluded from the training dataset. After a dynamic trace is simulated, it is labeled according to state occupancy. Here, the first observed state always receives the first label regardless of the FRET efficiency, followed by the next observed states until the maximum number of states is reached. For three-color FRET data, every transition regardless of the dye is treated as a new state. For a two-state model, transitions of one dye can be described whereas the multi-state model also considers transitions of two dyes. Thus, we have trained the network to recognize four different states. For a system with three independently moving dyes, a minimum of 9 states would be possible in one trace. Expert users can generate a corresponding training dataset by setting the algorithm parameter ‘static dyes’ to ‘None’.

Supplementary Table 2.1. Simulation parameters for the training datasets.

| Number of Frames | Maximum number of states | Minimum state difference | Dwell time / exposure ratio | Artifact probability | Stoichiometry tolerance |
|-----------------------|--------------------------|--------------------------|-------------------------------|----------------------------|------------------------------|
| 500 | 4 | 0.1 | 1-100 | 0.25 | 0.1-0.9 |
| Aggregate probability | Blue intensity scaling | Yellow intensity scaling | Red intensity scaling | Mean bleaching frame | Blinking probability |
| 0.1 | 0.7-1.3 | 0.7-1.3 | 0.5-2 | 400 | 0.2 |
| Gamma blue/yellow | Gamma blue/red | Gamma yellow/red | Direct excitation blue/yellow | Direct excitation blue/red | Direct excitation yellow/red |
| 0.7-2 | 0.49-2.6 | 0.7-1.3 | 0-0.4 | 0-0.2 | 0-0.3 |
| Crosstalk blue/yellow | Crosstalk blue/red | Crosstalk yellow/red | Noise scaling factor | Gamma noise prob. | Noisy threshold |
| 0-0.6 | 0-0.2 | 0-0.3 | 0-0.9 | 0.8 | 0.25 (SNR: 1.5) |

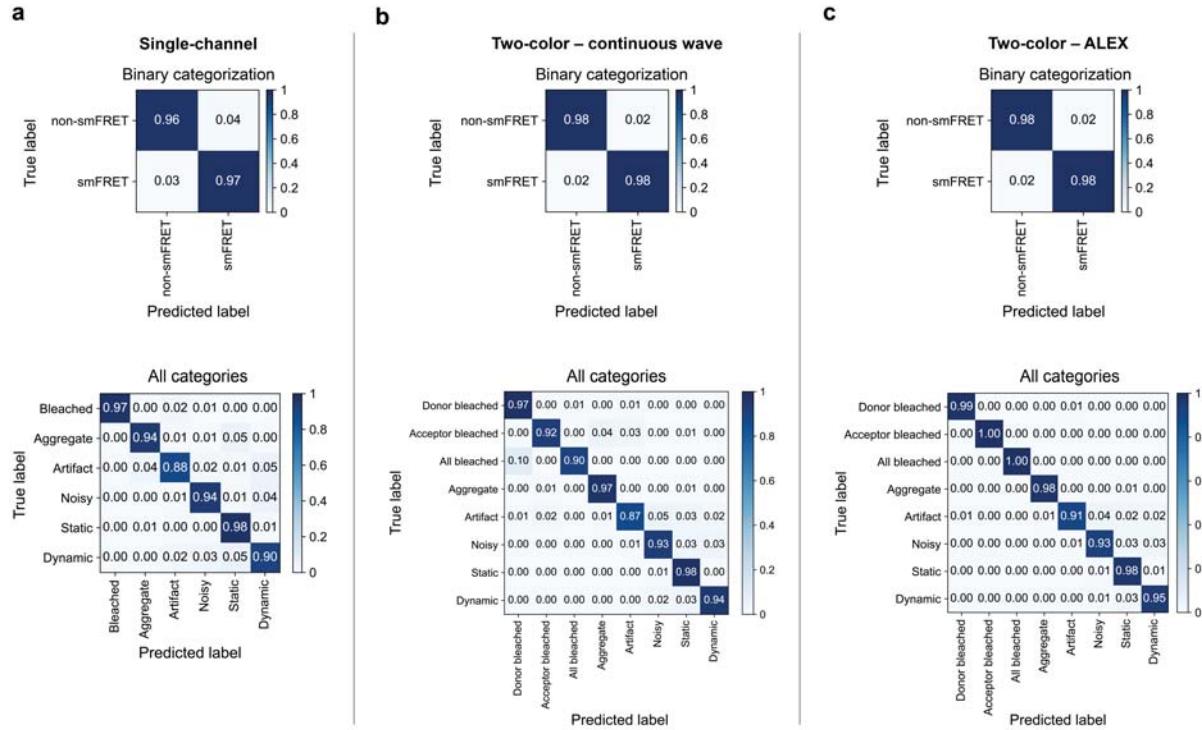
SUPPLEMENTARY NOTE 3: TRAINING VALIDATION

In the following sections (3.1–3.3), the final validation of every deep neural network is shown via confusion matrices. Approximately 20,000 new simulated traces were generated and fed into each trained model. Each row of the confusion matrices represents the instances in a ground truth category while each column represents the instances in a predicted category. The diagonal values report the percentage of true positives and true negatives whereas the off-diagonal values are the false negatives and false positives.

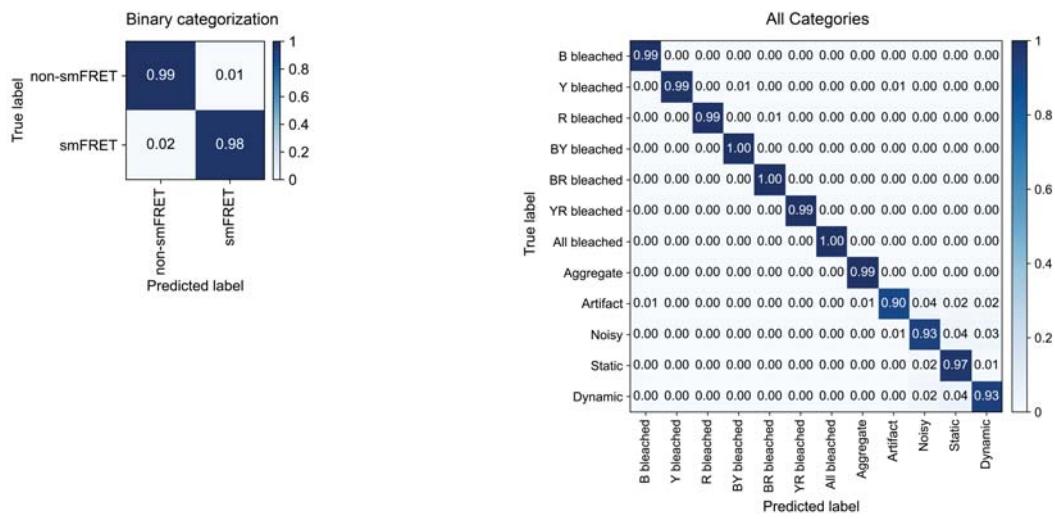
3.1. Trace classifiers

Confusion matrices for the trace-classifier networks are shown in [Supplementary Figures 3.1 and 3.2](#). The single-channel classifier has the lowest overall performance, in particular, due to a higher rate of falsely classifying random perturbations in ‘artifact’ frames (88 % precision) and misinterpreting ‘dynamic’ traces as ‘static’ (5 % false negative rate). The two-color and three-color models achieve similar accuracies for recovering smFRET frames with at least 93 % precision in correctly predicting ‘dynamic’ frames and 96 % precision for ‘static’ frames. In general, most of the false predictions concerning smFRET categories come from the high resemblance of ‘static’ frames, ‘dynamic’ frames with low contrast between states and ‘noisy’ frames close to the defined threshold. Here, the tolerance towards noise, defined as the mean standard deviation of the observed FRET efficiencies for all states, was set to 0.25. The highest sensitivity for detecting photobleached dyes (>98 %) is achieved by ALEX-enabled models for two- and three-color data. The continuous wave models depend on the contrast in intensity between the quenched and photobleached dyes, causing a significant decrease in sensitivity down to 91 % for detecting a photobleached acceptor. However, falsely predicted ‘acceptor bleached’ frames were mostly misclassified as either ‘aggregate’ or ‘artifact’ and would still be excluded from further analysis.

In addition to the confusion matrix for all available categories, we also calculated a binary trace classifier confusion matrix where we separated the frames into those that were accepted for further analysis (i.e. from ‘static’ and ‘dynamic’ traces without photobleaching) and those that were rejected (‘photobleached’, ‘aggregate’, ‘artifact’ and ‘noisy’ traces and/or frames). All trace classifier models achieve a minimum combined precision of 97 % in predicting smFRET categories, i.e. ‘static’ or ‘dynamic’, and 96 % in predicting non-smFRET categories ([Supplementary Figures 3.1 and 3.2](#)).



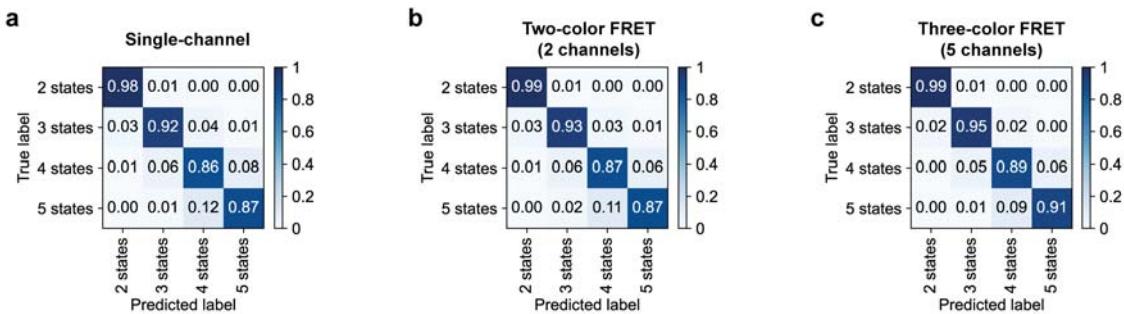
Supplementary Figure 3.1: Confusion matrices for 1-color and 2-color trace classification. Prediction accuracies depicted as confusion matrices for the (a) single-channel, (b) two-color continuous wave and (c) two-color ALEX models. The upper panels show the binary assignments into valid smFRET and non-smFRET categories. The detailed categorization is shown in the lower panels.



Supplementary Figure 3.2: Confusion matrices for 3-color trace classification. The left panel shows the binary assignments into valid smFRET and non-smFRET categories. The detailed categorization is shown in the right panel.

3.2. Number-of-states classifiers

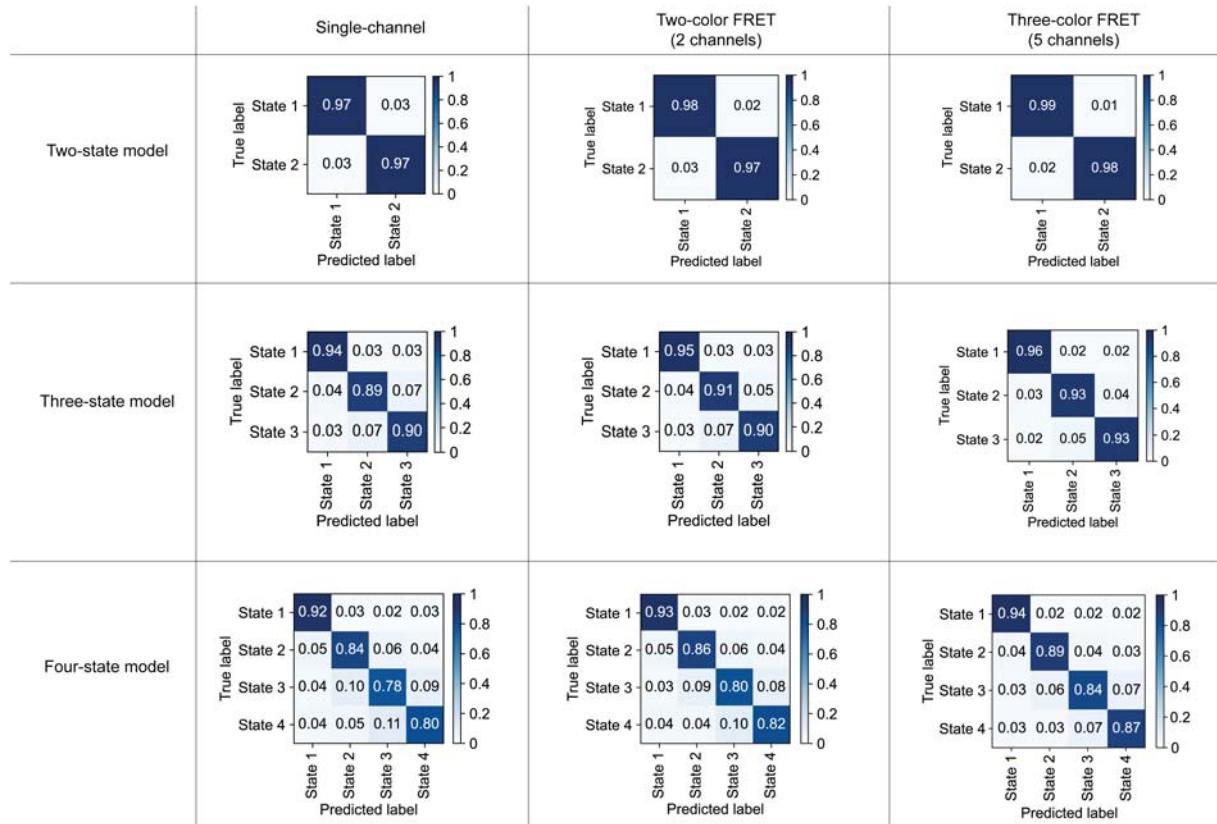
After classifying the individual traces, the dynamics are analyzed. One option is to classify the number of states in a particular trace, i.e. to run the number of states classifier for the type of data measured. [Supplementary Figure 3.3](#) shows validation of the deep neural networks trained on traces containing the given number of observed states. Only traces classified as ‘dynamic’ by the trace classifiers serve as input, hence the first category is for two observed states. The category of five observed states serves as a safeguard against traces that may be out of the scope of the pretrained state transition classifiers. All models achieve a high accuracy of at least 98 % in distinguishing two-state from multi state traces. The lowest accuracies are achieved in separating four-state from five-state traces, ranging from 86 % (single-channel) to 89 % (5-channels). The overall performance increases with increased number of available channels. As only dynamic information is considered in the state classifiers, the presence of the ALEX channel, though very useful for the trace classification, is no longer relevant.



Supplementary Figure 3.3: Confusion matrices for number of states classification. Confusion matrices for the (a) single-channel, (b) two-color FRET and (c) three-color number of states classifiers.

3.3. State-transition classifiers

After estimating the number of states in a dataset, the state trajectories of the individual dynamic traces are determined. This section summarizes the validation of the deep neural networks trained on the state occupancy and therefore also on the state transitions ([Supplementary Figure 3.4](#)). The performance does not differ significantly for the two-state models with a minimum of 97 % precision for predicting the correct state and a minimum of 84 % precision for four-state systems. The performance for multi-state models increases when more channels are available

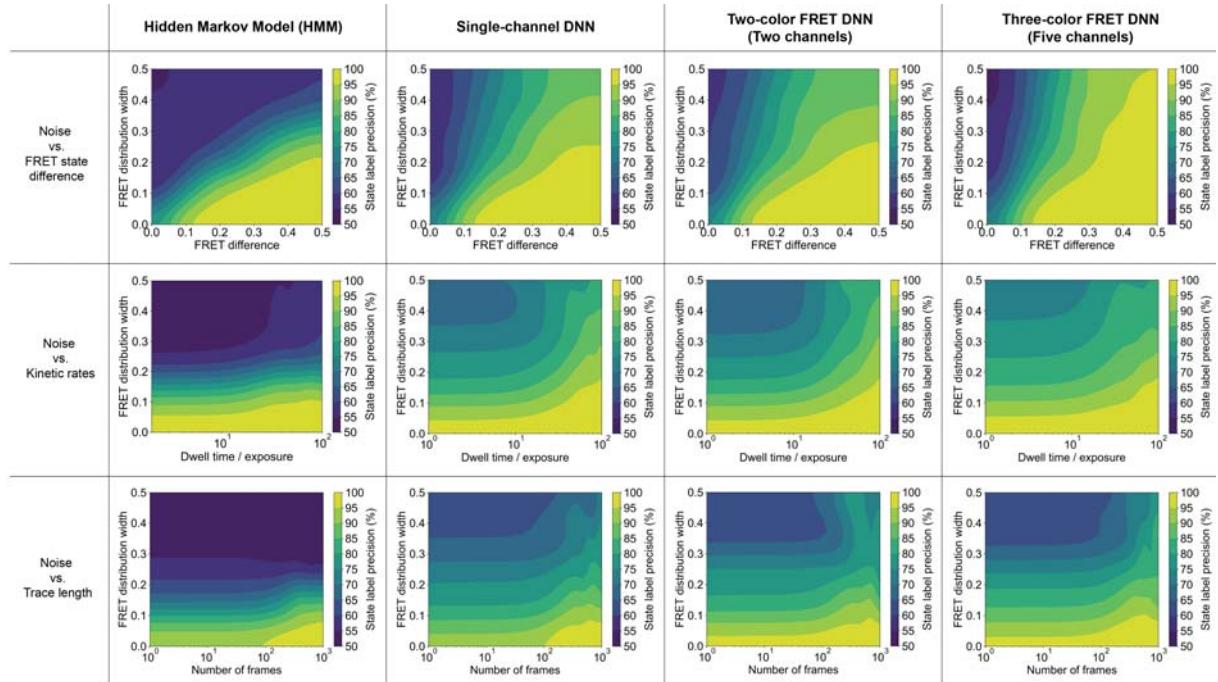


Supplementary Figure 3.4: Confusion matrices for state classification. Confusion matrices for the single-channel (first column), two-color FRET (second column) and three-color state classifiers (third column) and their corresponding two-state (first row), three-state (second row) and four-state models (third row).

3.4. Limitations of the state classifiers and a comparison with HMM

This section provides additional benchmarks and a comparison of the results from the state classifiers with HMM. First, we investigated how the performance of HMM and our state classifiers depends on noise (i.e. the width of the FRET distribution), difference between FRET states, the kinetic rates (dwell time to exposure ratio) and the length of observation time (number of frames) for dynamic transitions between two FRET states. For the three-color simulations, only the FRET distribution width of the yellow-red dye pair was used as the ground truth parameter to keep the continuity with two-color FRET traces and avoid averaging inconsistencies. [Supplementary Figure 3.5](#) shows interpolated maps of the precision of state label recovery for all models and were generated using approximately 300,000 simulated traces for each condition. The precision is the fraction of true positives divided by the sum of true positives and false positives for the state label predictions. Each map shows the precision dependency on the amount of signal noise with two of three parameters being fixed, namely the FRET state contrast (0.2), the transition rate (0.05/frame) and the number of frames (500). In general, at a fixed transition probability and number of frames (top row), the precision decreases with broader FRET distributions and smaller differences between the FRET states. All models are able to achieve a precision of at least 90% for FRET differences above 0.2 and FRET distribution widths below 0.10 with the state classifiers outperforming HMM only at high noise

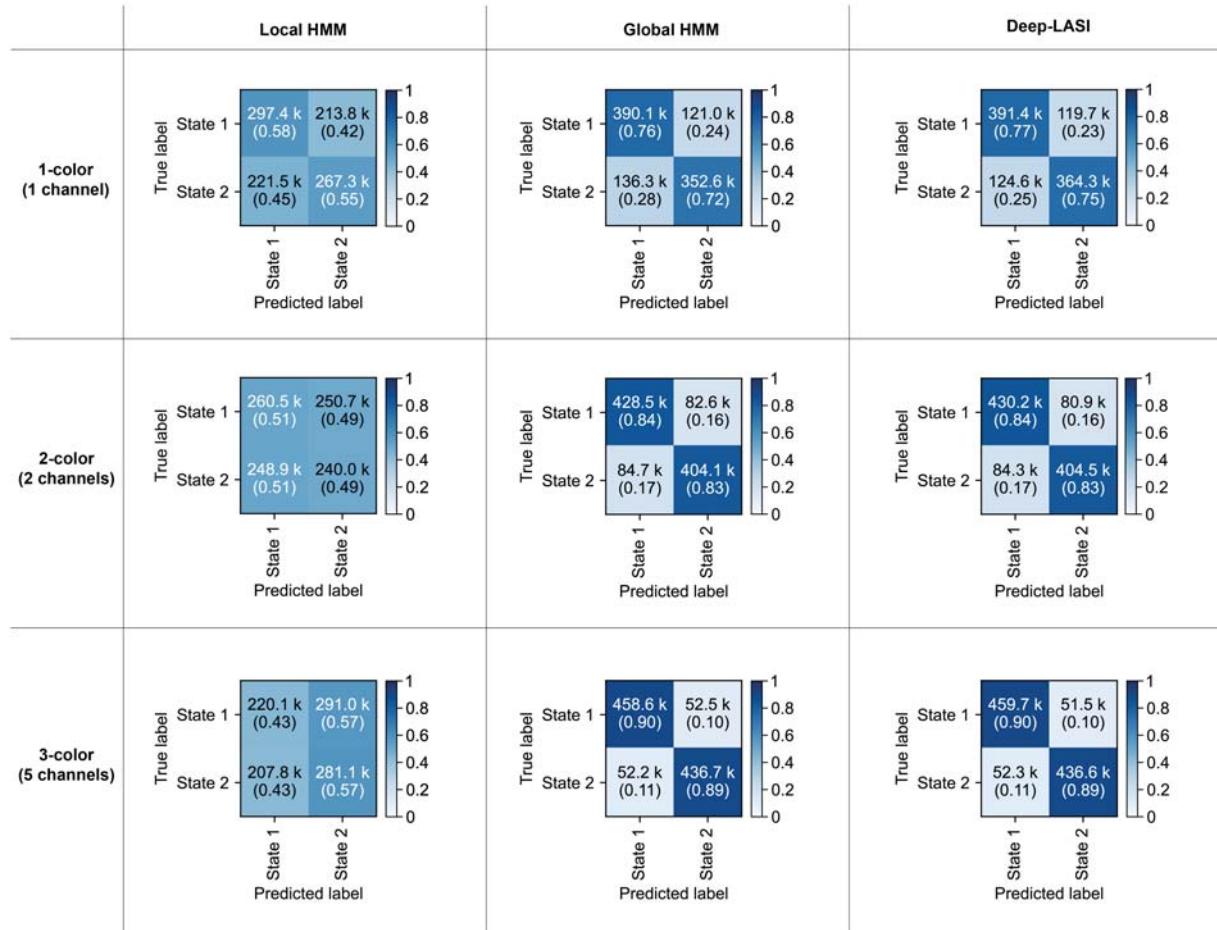
levels above 0.25. For a fixed contrast between FRET states (0.2) and total number of frames (500) (Supplementary Figure 3.5, middle row), the precision of HMM remains largely independent of the dwell time to exposure ratio at a constant noise level. All DNN state classifiers show a similar overall performance but achieves a higher precision than HMM at higher noise levels for larger dwell time to exposure ratios. For fixed FRET states and kinetic rates (Supplementary Figure 3.5, bottom row), trace length has little influence on the precision of all models below \sim 100 frames and the precision slightly increases for all models/classifiers above 100 frames. Again, the DNN outperform HMM at higher noise levels. In summary, while the precision does not differ significantly between the single-channel and two-channel state classifiers, the five-channel model used for three-color FRET shows an increased performance of up to \sim 10 % at high noise levels. Due to the five available channels, the signal-to-noise ratio is effectively increased which leads to higher precision and accuracy as soon as the signal noise becomes a limiting factor for the other models. In addition, DNN models still predict transitions in high noise trajectories, however with decreased confidence, whereas HMM no longer finds transitions at high noise.



Supplementary Figure 3.5: Deep-LASI state prediction compared to HMM. Precision of the state-label recovery for HMM and for the state transition classifiers as a function of noise (i.e. width of the FRET distribution), contrast between FRET states and the kinetic rates (dwell time to exposure ratio). Each map shows the precision dependency on the noise and one additional parameter: the contrast between FRET states (top row), the kinetic rate (middle row) and the number of frames (bottom row). One dataset with \sim 300,000 traces was generated for each row with the corresponding two of the three parameters fixed (FRET efficiency for yellow/red: 0.4 and 0.6, transition probability: 0.05/frame, and number of frames: 500). The noise is defined as the mean standard deviation of the FRET signal from both states. The lower limit of the precision is set to 50 % since it represents the highest amount of uncertainty for two states.

3.5. Comparison of the state classifiers with local and global HMM

We also compared the performance of Deep-LASI with a local and global HMM analysis ([Supplementary Figure 3.6](#)). In a global HMM analysis, the algorithm can learn from a large dataset rather than being limited to the number of frames in a single trace. For the comparison, we simulated 2000 traces of 500 frames (1 million data points) of three-color data having two states. To avoid difficulties for the global HMM analysis, traces were simulated with time-independent, normally-distributed noise. From the three-color data, we analyzed the yellow channel alone (1-color), yellow and red together (2-color) as well as the full three-color data (3-color). As the ground truth is known, we were able to initialize a three-color HMM analysis such that it could eventually converge. Whereas local HMM struggles to analyze the traces yielding an almost random guestimate of the state, both global HMM and Deep-LASI perform similarly well in all cases. As more channels become available, the state classification becomes more accurate.



Supplementary Figure 3.6: Confusion matrices of frame-wise state predictions for local HMM, global HMM and Deep-LASI, performed on a global data set containing 2000 traces with 500 frames each. Three-color traces were simulated and the yellow channel with yellow excitation (1-color), yellow and red channels with yellow excitation (2-color) and all five channels for the three-color data with blue and yellow excitation were analyzed (3-color). In all cases, local HMM is unable to learn states and transitions due to the limited amount of training data. Global HMM and Deep-LASI are both able to effectively remove background noise, showing similar accuracies. All models show increased performance with higher number of channels due to effectively decreasing the signal-to-noise ratio. The data set was simulated based on 3-color FRET using the following parameters: Two states with only the yellow dye transitioning between $E(YR) = 60\%$ and $E(BY) = 40\%$ to $E(YR) = 40\%$ and $E(BY) = 60\%$, a static blue to red FRET efficiency with $E(BR) = 20\%$, a symmetric transition probability of 0.05 per frame, and

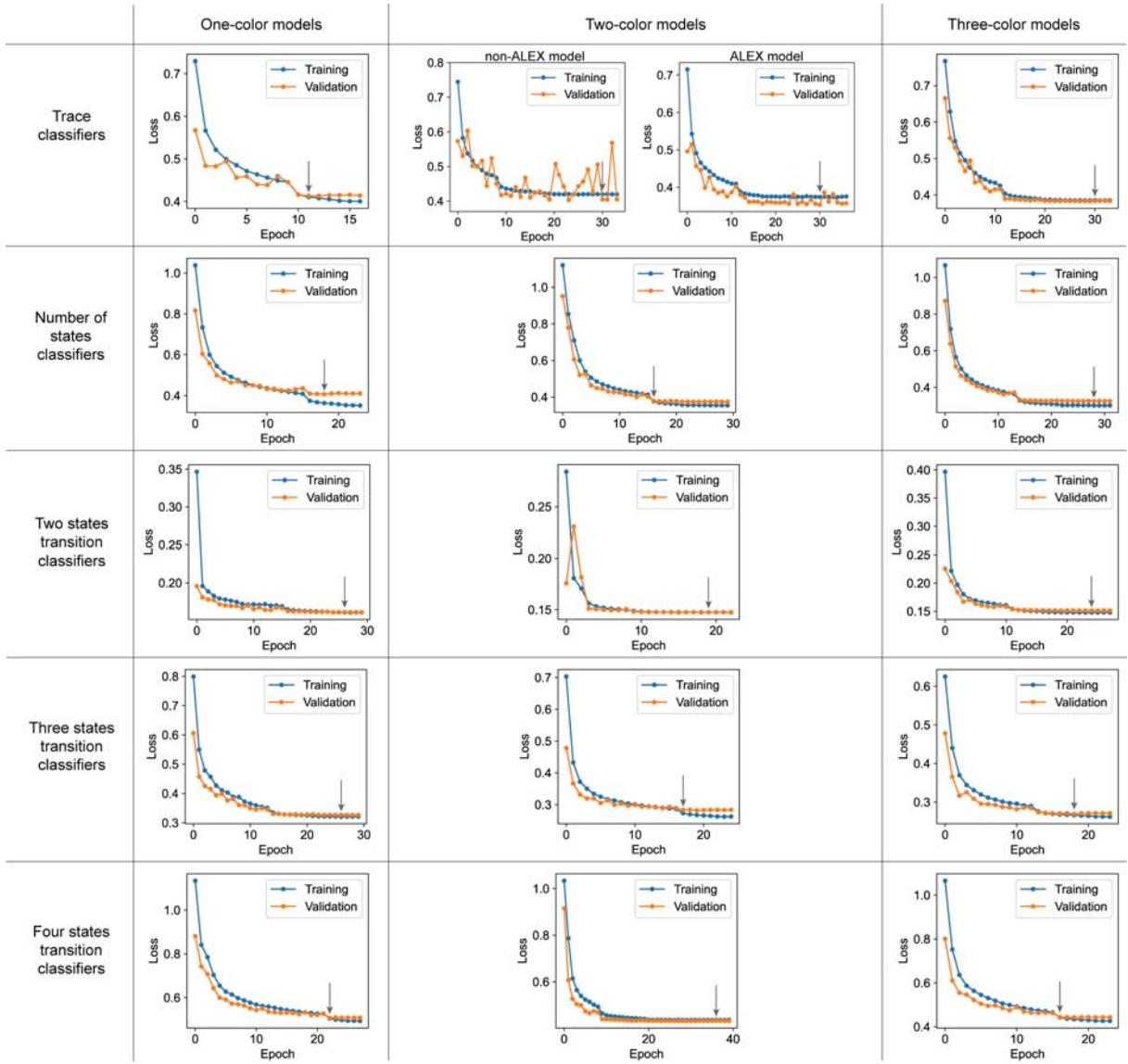
the addition of normally distributed noise resulting in a mean FRET distribution width of 0.23 (averaged over the yellow-red FRET efficiency states).

3.6. Training and validation loss

The training and validation loss for all pre-trained deep-neural networks are shown in [Supplementary Figure 3.7](#). Of the ~200,000 traces generated for training, ~160,000 were used in each epoch for training and then the capacity of the network to generalize what it learned was tested with the remaining ~40,000 traces. The error was calculated using the categorical cross-entropy, i.e. the loss function:

$$Loss = - \sum_i^C t_i \log(f(s)_i) \quad \text{Eq. 3.1}$$

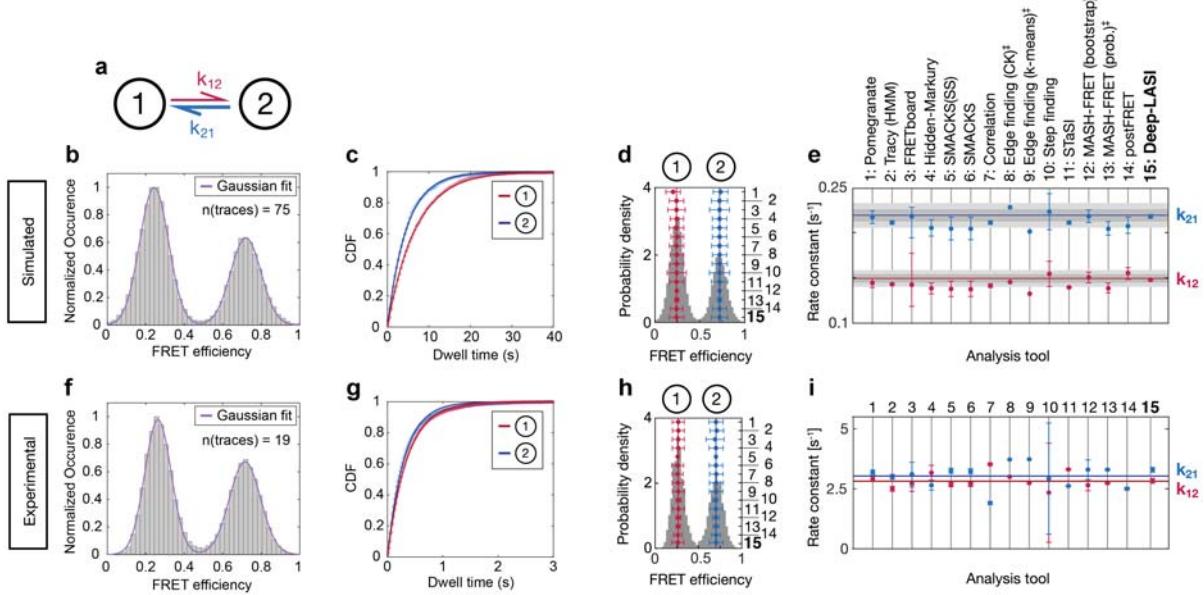
where C is the total number of classes, t is the target vector and $f(s)$ is the one-hot encoded vector of scores.¹¹ The categorical cross-entropy is specifically designed for multi-class classification problems, where each input belongs to exactly one class out of multiple mutually exclusive classes. It calculates the dissimilarity between the predicted class probabilities and the true class labels, providing a measure of how well the model captures the correct class assignments. Therefore, the model is able to produce probabilistic outputs in the form of class probabilities. By optimizing the model to minimize the cross-entropy loss, it learns to assign higher probabilities to the correct class and lower probabilities to the incorrect classes. During training of the model, the loss should decrease but maintain similar values for the training dataset as for the validation set. If a lower loss is observed for the training dataset, then for the validation dataset, the network is overfitting (i.e. it is memorizing the traces rather than learning the features of the categories). All models show no or a minimal amount of overfitting. While the full number of epochs are displayed in each plot, the model with the lowest validation loss and lowest amount of overfitting was saved and implemented (indicated with an arrow).



Supplementary Figure 3.7: Training and validation loss of all Deep-LASI models. Each row refers to the type of classifier and each column refers to the corresponding data type. For two-color data, there are two trace classifier models, one for ALEX and a second for non-ALEX measurements. Black arrows mark the saved model used when following epochs did not decrease the validation loss and indicated overfitting.

3.7. Analysis of kinetic data from the kinsoft challenge

We tested the performance of Deep-LASI on datasets provided by a recently published multi-laboratory software comparison study for extracting kinetics from smFRET data. As we contributed to this study using conventional HMM, we chose to analyze the datasets that did not require additional human input for interpretation of the data. The results are shown in **Supplementary Figure 3.8**. Deep-LASI returned values corresponding to the ground truth for the simulated dataset and close to the average values obtained for the experimental dataset.

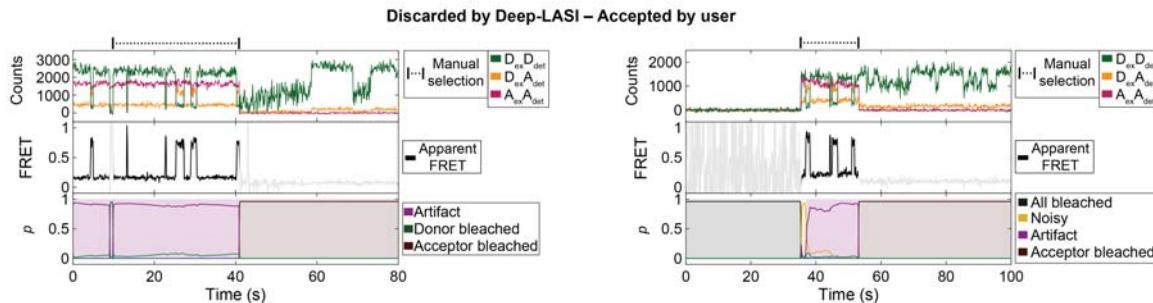


SUPPLEMENTARY NOTE 4: DEEP-LASI VERSUS MANUAL ANALYSES

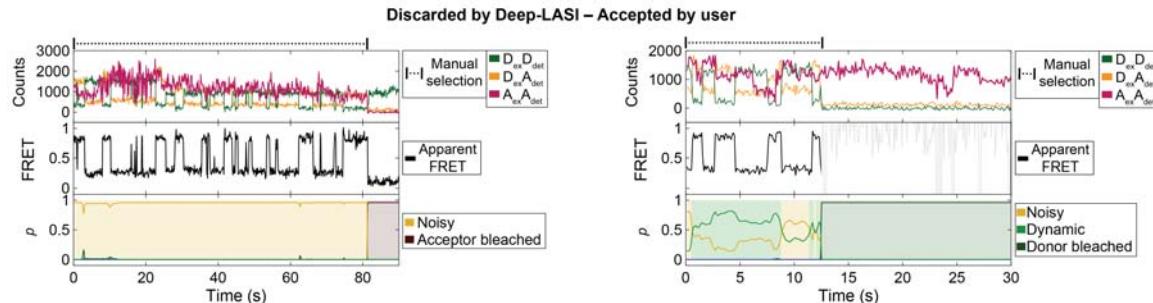
4.1 Comparison of Deep-LASI and manual analysis of 2-color DNA origami traces |

We investigated the disagreements between Deep-LASI and manual classification and summarize the primary causes, illustrated through specific examples in [Supplementary Figure 4.1](#). First, traces that exhibit non-ideal behavior after photobleaching are often thrown out by Deep-LASI as the entire traces is than categorized as an artifact whereas users may ignore characteristics of the traces in the non-accepted regions. Secondly, in cases where the leading frames are photobleached and the molecules reactivate, Deep-LASI tends to classify the entire trace as an artifact instead of extracting potentially useful information from the middle section of the trace. This is due to the fact that the training dataset does not yet include valid single molecule FRET traces starting with inactive dyes. Thirdly, Deep-LASI categorizes high noise or intensity fluctuations in the acceptor channel as noisy or classifies them as dynamic, but with insufficient confidence ($> 70\%$) to include in further analyses. In such cases, the user has the flexibility to adjust the confidence threshold or consider the unfiltered dynamic category based on the maximum confidence output. Lastly, Deep-LASI may select traces that were discarded by manual evaluation, particularly short traces with fast dynamics that may be overlooked or considered noisy by the user.

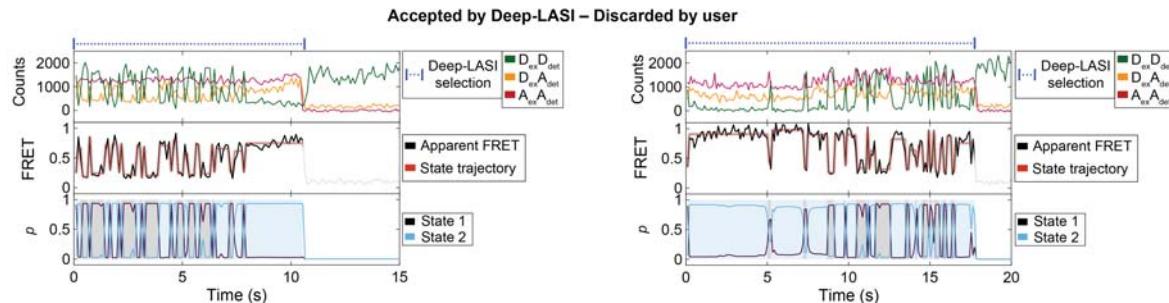
a



b

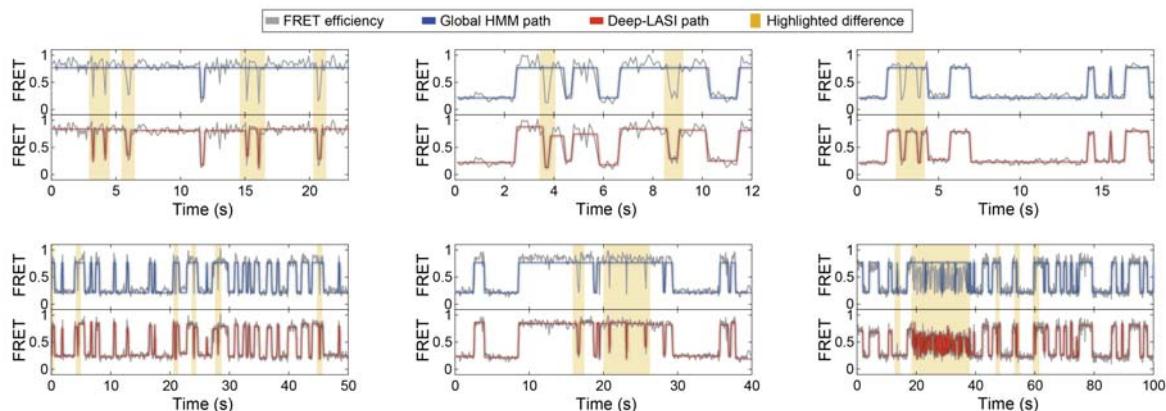


c



Supplementary Figure 4.1: Representative 2-color DNA origami traces with disagreements between manual and Deep-LASI selection. (a) Non-ideal donor signals. In the case of non-ideal donor intensities after photobleaching of the acceptor (left panel) or leading photobleached frames (right panel), Deep-LASI tends to classify the whole trace as an artifact while the manual selection still includes the middle region between the photobleached frames. (b) Non-ideal acceptor signals. Even though the FRET efficiency trace looks ideal, due to the high noise or intensity fluctuations in the acceptor channel, Deep-LASI classifies the trace as noisy (left panel) or is not confident enough ($> 70\%$) to classify the trace as dynamic (right panel). In the latter case, the user could influence the selection of this trace by lowering the confidence threshold. (c) Traces selected by Deep-LASI that were discarded by manual evaluation. Short traces (pay attention to the timescale on the x-axes) with fast dynamics can be either overlooked or deemed noisy by the user.

We further compare representative FRET traces analyzed using global HMM and Deep-LASI (Supplementary Figure 4.2). Global HMM tends to struggle in accurately capturing fast transitions, which can be attributed to the difficulties of HMM in distinguishing between fast transitions and noise or due to the inherent (non-Markovian) heterogeneities commonly encountered in single molecule experiments. Local HMM analyses can more easily deal with these heterogeneities. In comparison, Deep-LASI exhibits enhanced performance in detecting and characterizing these fast kinetics, suggesting its potential advantages over global analysis approaches.

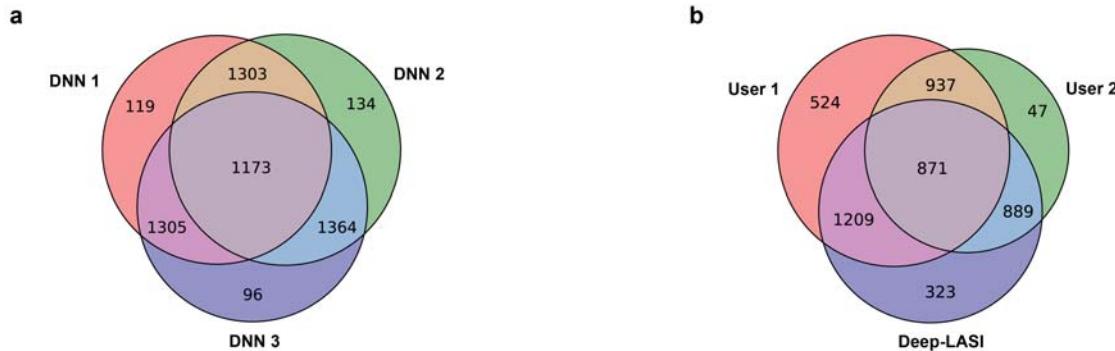


Supplementary Figure 4.2: Comparison of global HMM and Deep-LASI analyses. Representative 2-color DNA origami smFRET traces highlighting the differences between a global HMM and Deep-LASI analyses. In contrast to Deep-LASI, global HMM frequently misses fast transitions due to heterogeneities in the single molecule data.

4.2. Influence of different training datasets and comparison to user classification

The quality of a neural network rises and falls with the data by which it has been trained. To see the influence of different training datasets, we simulated three different datasets using the same parameters and used them to train Deep-LASI. The three different networks were then used to classify the two-color, two-state data shown in Figure 3. The results are shown in Supplementary Figure 4.3a. From the confusion matrix shown in Supplementary Figure 3.1, one would expect a consistency on the order of 95 %, provided the experimental data are similar to the training datasets. Each pair of networks agree within ca 93 %. Interestingly, the consistency between the neural networks is higher than that from two individual users (Supplementary Figure 4.3b). Here, user 1 tried to maximize the statistics and selected subsections of traces whereas user 2 was very conservative, only classifying the best traces as

dynamic. This suggests that neural networks may be more consistent in the analysis than different users.



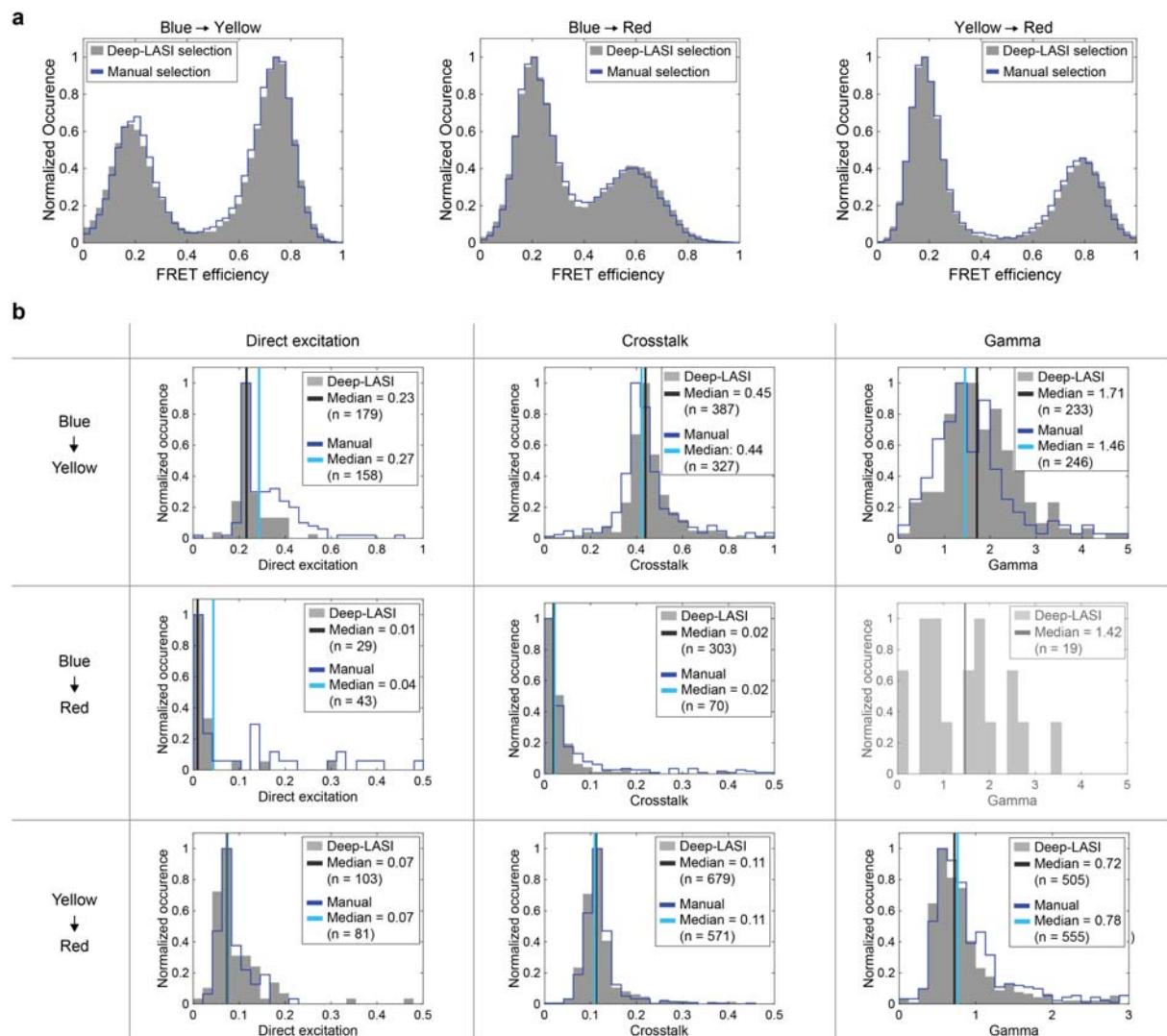
Supplementary Figure 4.3: Venn diagrams for differently trained neural networks and users. a) A Venn diagram showing the consistency between traces classified as dynamic (with a 70% confidence threshold) for three networks trained using different datasets (DNN1, DNN1 and DNN3) and applied to the experimental data shown in Figure 3. b) A Venn diagram comparing the number of traces classified as dynamic by two different users as well as DNN1 (with a 70% confidence threshold) for the same dataset used in panel a.

4.3 Deep-LASI versus manual analyses for 3-color DNA origami samples

We emphasize the importance of also using experimental data for testing deep learning methods trained on synthetic data since the simulations used for validation are usually generated by the same algorithm as the training dataset. Deep neural networks can easily learn biases of any kind in the training data, which may have no relationship to the respective category under new conditions. Hence, the prediction of categories with respect to ground truth simulations can produce high accuracies, which may not be directly translatable to real-world examples. Therefore, we compared the performance of our network models on real data with that of experts who manually analyzed the same dataset.

We benchmarked the three-color performance of Deep-LASI by comparing the automated analysis with traces manually selected by an expert user (Supplementary Note 5). We used the three-color L-shaped DNA origami structure with two binding locations spaced at 6 and 12 o'clock with complementary binding regions of 7.5 nt (Figure 4). Deep-LASI yielded 581 usable smFRET traces versus 694 for manual selection out of a total of 2545 extracted traces (Supplementary Figure 4.4a). The two uncorrected, framewise smFRET histograms are almost identical. The automatically extracted FRET correction factors, which are based on the predictions of the three-color trace classifier, were compared to those determined manually. The expert user selected the relevant regions of the traces for determining various FRET correction factors by hand. Very similar distributions and median values were obtained for the YR correction factors (Supplementary Figure 4.4b). For BR, both direction excitation and spectral crosstalk terms are small and the differences are not significant here. Due to the high stability of the yellow fluorophore, it is challenging to collect enough statistics to directly derive the detection correction factor. Hence, it is calculated from the product of the BY and YR γ factors. For BY, the distributions for spectral crosstalk from Deep-LASI and manual selection are consistent. However, for both direct excitation and the detection correction factor, there are

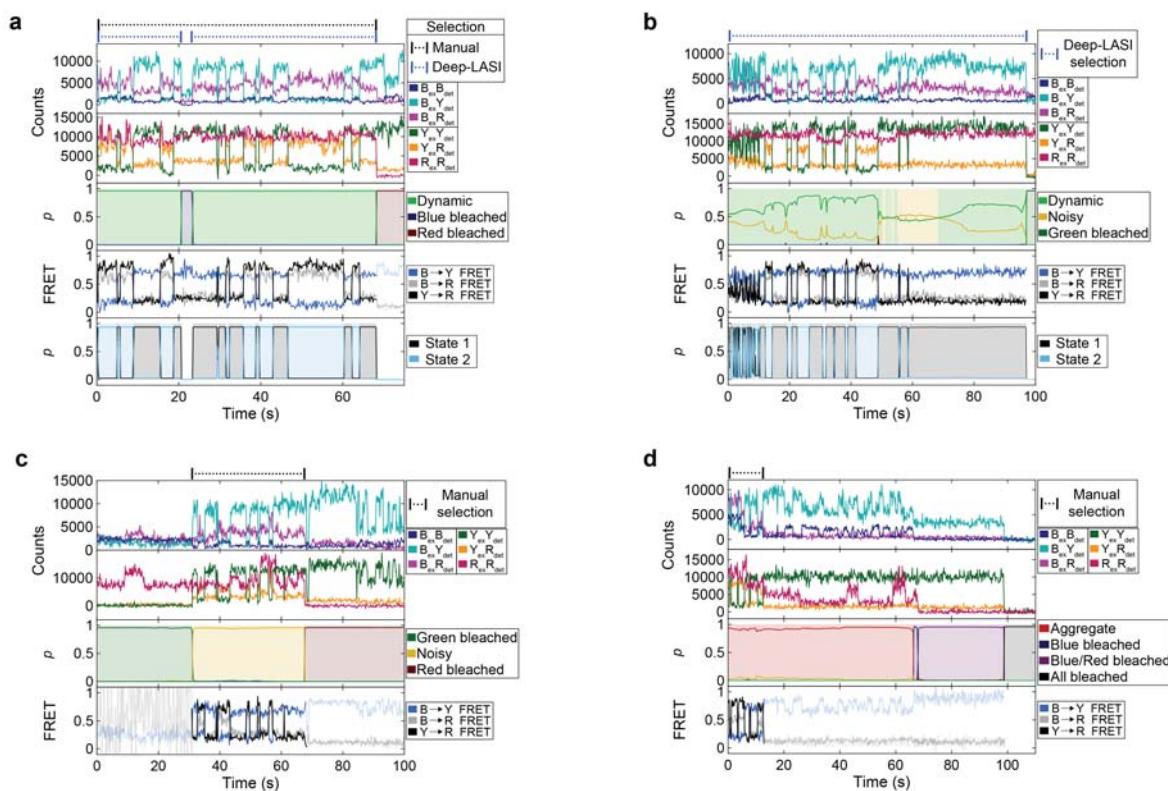
differences of $\sim 15\%$. Manual selection with the blue fluorophore is difficult because of the low fluorescence intensity of the blue dye. In the manually selected regions for direct excitation, a second population is visible due to difficulties of distinguishing between a Y only fluorophore and a dim B fluorophore undergoing high FRET. Similarly, there are differences in the detection correction factor distribution. As Deep-LASI has more flexibility in choosing relevant regions of the traces for determining the correction factors, it is most likely that Deep-LASI is more accurate in these cases. FRET correction-factor determination is a potential source of human bias in the analysis of smFRET data and we demonstrate here an advantage of using a well-trained neural network for automated analysis.



Supplementary Figure 4.4: Uncorrected smFRET histograms and correction factors extracted by Deep-LASI for the 3-color 2-state DNA origami. (a) Uncorrected framewise smFRET histograms for BY, BR and YR calculated from traces selected manually ($n=694$, blue line) versus the histograms determined by Deep-LASI ($n=581$, gray histograms). There is excellent correspondence between the histograms. (b) Each panel displays the normalized distribution of available correction factors from all traces categorized as ‘dynamic’ by Deep-LASI (gray filled histograms) or manually labeled as dynamic (blue histogram line). Due to the high stability of the yellow dye compared to the blue and red dyes, the number of usable traces to calculate blue/red detection correction factor was too low to be determinable. Therefore, we used the theoretical value of 1.23 (for Deep-LASI compared to 1.15 for manual selection) for the blue/red gamma factor determined from the product of the gamma factors for blue/yellow and yellow/red.

4.4 Comparison of 3-color DNA origami traces selected manually and/or via DNN classification.

To gain insights into the selection criteria of traces performed manually and via Deep-LASI, we examined the traces in detail that were selected differently. Examples are shown in [Supplementary Figure 4.5](#). In general, similar differences arise as observed in the two-color FRET classifications ([Supplementary Figure 4.1](#)). In case of three-color FRET, blinking and dark states of the blue dye are inherently difficult to spot during manual inspection of the trace, whereas Deep-LASI predicts dark frames with high accuracy. Single-molecule traces often exhibit ambiguity. Deep-LASI tends to disregard traces displaying non-ideal intensities due to the way it was trained. Conversely, when manually selecting traces, users may incline towards including more non-ideal traces to improve the statistics and their selection can be subjectively influenced by a myriad of parameters.



Supplementary Figure 4.5: Representative 3-color DNA origami traces with disagreements between manual and Deep-LASI selection. (a) Short blinking events. In this trace, Deep-LASI and manual selection agree in general but Deep-LASI excludes frames with an inactive blue dye. Blinking events of the strongly quenched blue dye in 3-color experiments can be easily missed during manual selection. (b) Regions of traces with high-noise or fast dynamics. The trace shown here was excluded manually due to seemingly high noise at the beginning of the trace. Deep-LASI predictions show the competing categories of ‘dynamic’ and ‘noisy’. The summed confidence for the ‘dynamic’ classification exceeds the user defined threshold of 70 % and the state classifier predicts state transitions with high confidence. (c) Initial dark frames and non-ideal intensities. In the training datasets, we currently do not start with photobleached molecules the begin to fluoresce during the traces. Here is an example of Deep-LASI’s tendency to classify valid sections as noisy due to leading bleached frames or erratic intensities during bleached frames. (d) Short traces. Short section in the beginning of the trace is manually selected whereas Deep-LASI classifies the whole trace as an aggregate due to intensity spikes in the acceptor channel after the valid section.

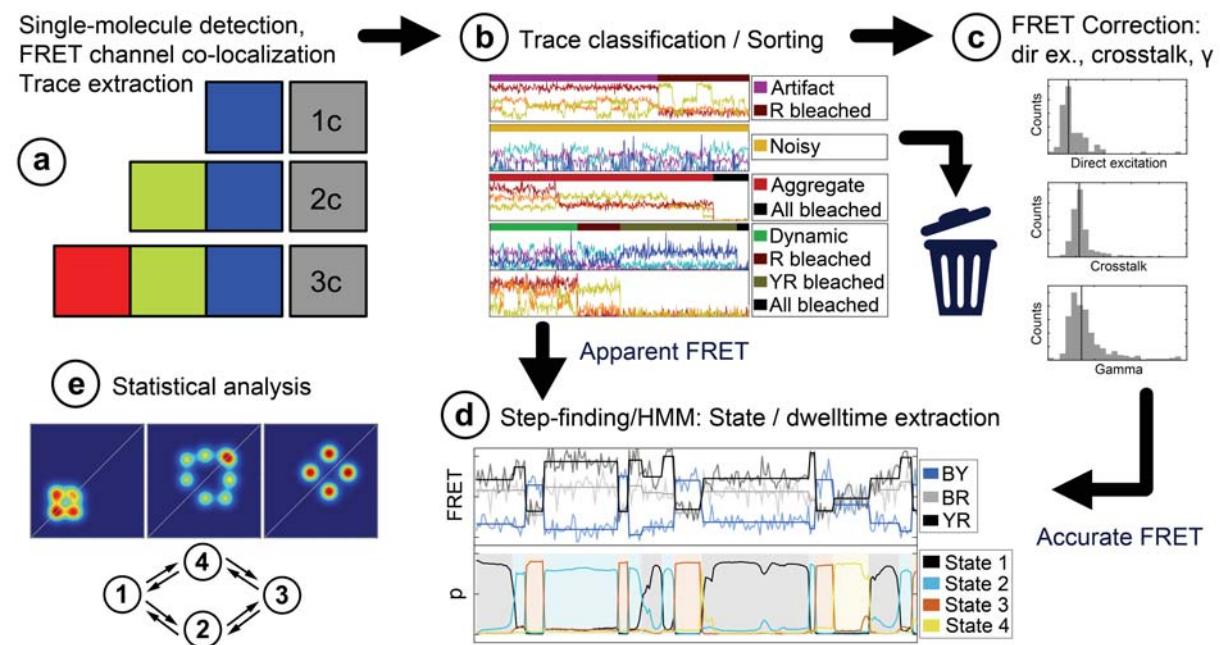
SUPPLEMENTARY NOTE 5: MANUAL ANALYSIS OF SINGLE-MOLECULE TIRF DATA

5.1. Work-flow

We benchmarked the performance of Deep-LASI by comparing it to manually analyzed single-molecule data from an expert user. Starting from individual movies, the procedure for extracting the intensity information over time is highlighted in [Supplementary Figure 5.1](#).

The procedure begins with:

- (1) a pixel-wise mapping of the position between two or three cameras for two- and three-color experiments,
- (2) camera-wise localization and excitation-cycle dependent assignment of intensities, and
- (3) extraction of intensities and background correction for each detection channel.



Supplementary Figure 5.1: Work-flow of data extraction, sorting, analysis and evaluation. a) One-, two-, or three-color data is collected with various excitation schemes and the time-dependent intensity traces extracted and corrected for background. b) The traces are then visually inspected and sorted either for further analysis or marked as junk. Regions of the trace can be selected for smFRET evaluation or for correction factor determination. c) After manual selection, trace-wise correction factors are extracted. d) For the dynamic traces selected for further analysis, the dwell-time distributions are determined using a Hidden Markov Model approach. e) From the HMM analysis, transition density plots are extracted from the smFRET data.

Next, the recorded traces were analyzed ([Supplementary Figure 5.1b-e](#)) either manually ([Section 5.4-5.9](#)) or assisted by neural networks (cf. [Supplementary Note 1](#)). Manual evaluation of single-molecule or multi-color FRET traces involves:

- (4) the pre-sorting of traces suitable for 1c, 2c-, and 3c- smFRET analyses
- (5) determining consecutive regions of the trace for evaluation and for determination of local and global correction factors

- (6) Hidden Markov Modeling of smFRET traces to identify underlying states and dwell-times
- (7) kinetic evaluation of transition rates and states using transition density plots (TDP), state-wise histograms and dwell-time analyses.

5.2. Camera mapping for FRET traces

In order to extract the fluorescent intensity traces of individual, fluorescently labeled DNA origami structures detected in various channels, an accurate localization and mapping of the detected emission channels across the three cameras or detection areas needs to be achieved. To compensate for potential chromatic aberrations and non-ideal alignment, an image transformation was used to map the corresponding pixels between different cameras/regions onto each other. The associated transformation matrix describing the potential shifts, tilts, etc. was obtained by imaging a calibration pattern on all three detection channels ([Supplementary Figure 5.1a](#)). As a calibration pattern, we typically use a zero-mode waveguide array.

5.3. Trace extraction and background subtraction

After mapping the different detection channels, the location of the individual single emitters needs to be determined and the intensity extracted. When msALEX excitation is used, the alternating laser excitation scheme needs to be taken into account and the intensity traces separated based on both the detection and excitation channels. The most blue-shifted detection channel serves as the reference channel (Channel 1). This refers to the blue excitation, blue channel (BB) for BY-, BR- and BYR-labeled samples or the yellow detection channel with yellow excitation (YY) for YR-labeled samples ([Supplementary Figure 5.1a](#)).

Individual molecules in the reference channel are identified by searching for the brightest spot in the summed projection of the movie. After calculating the central position of the molecule using a wavelet approach¹², the corresponding position in all other channels is calculated using the transformation matrix. Molecules in the projection images exhibiting detectable intensity in all desired channels are then selected. The intensities and background are extracted using different masks. For the signal, the pixels within an approximate circle of roughly 3 pixels radius around the central coordinates of the molecule are summed together. With a pixel size of 124 nm, the fluorescence signal of a single molecule is accumulated within an area of $614 \times 614 \text{ nm}^2$. For the background, a mask representing roughly a circle with radius of 7.5 pixels (850 nm) and width of 2 pixel centered on the molecule is used. The background is calculated as the median value of all pixels inside the ring-shaped mask and averaged over a five-frame sliding window depending on the excitation cycle and the detection channel. Afterwards, the determined background is scaled to the signal mask and subtracted from the framewise intensity per each channel for each molecule. When analyzing single molecule traces from hand, trajectories which contain molecules within the background mask are discarded. In DeepLASI, these traces are typically discarded in the 'artifact' category during the first characterization step.

5.4. Manual trace selection and analysis

The background-corrected fluorescence intensity traces of the individual molecules are then inspected and sorted ([Supplementary Figure 5.1b](#)). The properties of the extracted traces are generally very heterogeneous. This stems from different sources including photochemistry, dye blinking, aggregates and impurities within the sample of different brightness. In all cases, molecules were rejected automatically if they exhibited (1) a low SNR or (2) a brightness that is significantly higher than expected for a single fluorophore (aggregates or impurities). We further classified traces according to

- (1) their static and dynamic behavior
- (2) the existence of photobleaching steps in the different intensity channels
- (3) the order of bleaching steps between the different intensity channels
- (4) the degree of labeling efficiency

With the presorted trajectories at hand, we next prepared the data either for (1) correction factor determination to obtain accurate FRET efficiencies ([Supplementary Note 5.5](#)) or (2) directly to kinetic and state evaluation based on background-corrected trajectories (1-color data) or apparent FRET efficiencies (2/3-color data). In the first case, we first derived the correction factors per trace and marked regions for trace evaluation by HMM afterwards ([Supplementary Note 5.6](#)). In the second case, we manually marked the regions in traces to be analyzed and added them to the ‘HMM’ category.

5.5. Accurate FRET determination

In real smFRET experiments, the intensity of the acceptor signal needs to be additionally corrected for direct excitation of the acceptor fluorophore and spectral crosstalk from donor into the acceptor channel. In addition, one needs to correct for the difference in the detection sensitivity between the donor and acceptor fluorophores. The correction factors are denoted as:

d_{XY} for direct excitation of the acceptor fluorophore Y during excitation with X ,

ct_{XY} for spectral crosstalk from the fluorophore X in the detector channel Y ,

and γ_{XY} compensates for differences in detection sensitivities between channels.

We denote the background-corrected intensities as I_{XY} and the corrected Intensity as $I_{XY,corr}$, where x stands for the excitation source and y for the emission channel, i.e. $I_{BR,corr}$ denotes the background corrected emission of the acceptor within the red channel (R) after donor excitation in the blue channel (B).

Trace-wise and global correction factors

¹³Depending on when individual fluorophores photobleach, some of the correction factors can be extracted from the trace itself. However, in the vast majority of the traces, one cannot extract all correction factors individually. When a trace-wise correction factor is unavailable or unreasonable, the *median* value of the corresponding distribution of trace-wise correction factors for the particular correction factor is used to calculate the accurate FRET values, i.e. a global correction factor. Using traces that were presorted and categorized as ‘Blue dye

bleached' (or 'yellow / red dye bleached', respectively), we first determined the trace-wise correction factors for direction excitation de_{XY} and spectral crosstalk ct_{XY} . Having corrected the background-corrected intensities against both contributions, we next determined the trace-wise correction factor γ_{XY} .

To derive the contribution of spectral crosstalk from the donor channel X in the acceptor channel Y , we determine the trace-wise correction factor ct_{XY} using the intensity information after photobleaching of the acceptor:

$$ct_{XY} = \frac{\langle I_{XY} \rangle}{\langle I_{XX} \rangle} \Big|_{\text{no acceptor}} \quad \text{Eq. 5.1}$$

Here, $\langle I_{XX} \rangle$ refers the mean donor intensity and $\langle I_{XY} \rangle$ to the mean acceptor intensity after donor excitation in the region of the trace where there is no acceptor fluorescence.

Similarly, we determined the correction factors for direct excitation of the acceptor during donor excitation using traces in which the donor fluorophore bleached first:

$$de_{XY} = \frac{\langle I_{XY} \rangle}{\langle I_{YY} \rangle} \Big|_{\text{no donor}} \quad \text{Eq. 5.2}$$

where $\langle I_{XY} \rangle$ and $\langle I_{YY} \rangle$ describes the mean acceptor emission after donor excitation or acceptor excitation, respectively.

Lastly, we determined the detection correction factors γ_{XY} compensating for differences in detection sensitivities between different channels. For this, we used traces where the acceptor photobleaches before the donor. The acceptor intensity is first corrected for direct excitation de_{XY} and spectral crosstalk ct_{XY} . We then derive the detection correction factor γ_{XY} per trace from the ratio of changes in donor and acceptor emission before and after photobleaching of the acceptor. The correction factors are denoted as:

$$\gamma_{XY} = \frac{\langle \Delta I_{XY,cor} \rangle}{\langle \Delta I_{XX,cor} \rangle} \Big|_{A \text{ bleaches}} \quad \text{Eq. 5.3}$$

where $\langle \Delta I_{XX,cor} \rangle$ and $\langle \Delta I_{XY,cor} \rangle$ refer to the intensity difference for the mean donor and acceptor emission after donor excitation before and after acceptor photobleaching.

Data Correction

Once all correction factors are determined, every trace is corrected using the local, trace-wise correction factors, when available and suitable. Otherwise, the global correction factor is used. In three-color experiments, the corrected FRET efficiency for E_{YR} is calculated first since it is required for subsequent corrections. Upon yellow excitation, the same approach is used as for two-color FRET experiments:

$$I_{YY,corr} = I_{YY} \quad \text{Eq. 5.4}$$

$$I_{YR,corr} = I_{YR} - ct_{YR}I_{YY} - de_{YR}I_{RR} \quad \text{Eq. 5.5}$$

The corrected FRET efficiency is then given by the ratio of both corrected intensities

$$E_{YR} = \frac{I_{YR,corr}}{\gamma_{YR}I_{YY,corr} + I_{YR,corr}} \quad \text{Eq. 5.6}$$

For the BY FRET pair, the fully corrected intensities after blue excitation read as:

$$I_{BB,corr} = I_{BB} \quad \text{Eq. 5.7}$$

$$I_{BY,corr} = I_{BY} - ct_{BY}I_{BB} - de_{BY}I_{YY} \quad \text{Eq. 5.8}$$

The accurate BY FRET efficiency follows equation 5.5 with an additional term which takes into account the reduction in brightness of the yellow dye due to the FRET process between the YR pair:

$$E_{BY} = \frac{I_{BY,corr}}{\gamma_{BY}I_{BB,corr}(1 - E_{YR}) + I_{BY,corr}} \quad \text{Eq. 5.9}$$

The intensity of the red fluorophore after blue excitation needs to be corrected against direct excitation, contributions of both the blue and yellow dye due to crosstalk into the red channel and due to cascading of FRET from the blue dye over the yellow dye into the red channel:

$$I_{BR,corr} = I_{BR} - de_{BR}I_{RR} - ct_{BR}I_{BB} - ct_{YR}(I_{BY} - ct_{BY}I_{BB}) - de_{BY}E_{YR}(1 - E_{YR})^{-1}I_{YY} \quad \text{Eq. 5.10}$$

The accurate FRET efficiency of the BR FRET pair is then given by:

$$E_{BR} = \frac{I_{BR,corr} - E_{YR}(\gamma_{YR}I_{BY,corr} + I_{BR,corr})}{\gamma_{BR}I_{BB,corr} + I_{BR,corr} - E_{YR}(\gamma_{BR}I_{BB,corr} + \gamma_{YR}I_{BY,corr} + I_{BR,corr})} \quad \text{Eq. 5.11}$$

5.6. Hidden-Markov modeling

The kinetics and underlying states within the selected trajectories, i.e. either smFRET or intensity traces, were evaluated using Hidden Markov Modeling. The input data of both assays vary between 0 and 1. We anticipate that every molecule undergoes transitions between a fixed numbers of conformations described by a discrete number of states q_i ($i = 1, \dots, Q$). The behavior of the system can be captured by the joint distribution of the observed data $\mathbf{x} = (x_1, x_2, \dots, x_T)$ and the corresponding hidden state sequence $\mathbf{q} = (q_1, q_2, \dots, q_T)$. The joint distribution can be factorized as follows:

$$p(\mathbf{x}, \mathbf{q}) = p(x_1|q_1) \cdot p(q_1) \cdot \prod_{t=2}^T p(x_t|q_t) \cdot p(q_t|q_{t-1}) \quad \text{Eq. 5.12}$$

Here, $p(x_t|q_t)$ represents the conditional probability of observing x_t given the system is in the hidden state q_t , $p(q_1)$ represents the probability of being in the initial state, q_1 , and $p(q_t|q_{t-1})$ represents the conditional probability of transitioning from state q_{t-1} to state q_t .

For a system with Q states in total, the transition probability matrix $\bar{\mathbf{K}}$ comprises $Q \times (Q - 1)$ independent transition probabilities k_{ij} describing the likelihood for going from state i to state j . Here, it is a prerequisite for the Markovian process, that the row-wise sum of transition probabilities is normalized to 1. For a Hidden-Markovian process, the state sequence is not directly observable but buried in random noise of the system. It can only be inferred from measured observables \mathbf{x} , i.e. the single-molecule trajectory, with a length of T data points. Here, the emission probabilities $f_{q_i}(x_t | \theta_q)$ serve as parameter to represent the relative likelihood for observing a specific FRET value (or intensity value) for a given set of model parameters θ_q and the molecule being in state q_i . For intensity measurements and single molecule FRET traces, it is appropriate to model the emission probability of a state q_i as a Gaussian distribution:^{14,15}

$$f_{q_i}(x_t | \theta_q = \{\mu_{q_i}, \sigma_{q_i}\}) = \frac{1}{\sqrt{2\pi}\sigma_{q_i}} \cdot e^{-\frac{(x_t - \mu_{q_i})^2}{2\sigma_{q_i}^2}} \quad \text{Eq. 5.13}$$

The parameters estimators are: the mean value μ_{q_i} and covariance σ_{q_i}

$$\langle \mu_{q_i} \rangle = \frac{\sum_{t=0}^T w_{q_i,t} x_t}{\sum_{t=0}^T w_{q_i,t}} \quad \text{Eq. 5.14}$$

$$\sigma_{q_i}^2 = \frac{\sum_{t=0}^T w_{q_i,t} x_t^2}{\sum_{t=0}^T w_{q_i,t}} - \langle \mu_{q_i} \rangle^2 \quad \text{Eq. 5.15}$$

For this, we introduce the relative occurrence probability $w_{q_i,t}$, i.e. the conditional probability $w_{q_i,t}$ of being in state q_i given the data x_t at a time t , which is linked to the fraction of time spent in state q , W_q

$$W_q = \frac{1}{T} \sum_{t=1}^T w_{q_i,t} \quad \text{Eq. 5.16}$$

and emission probability $f_{q_i}(\mathbf{x} | \theta_{q_i})$.

$$w_{q_i,t} = \frac{W_q f_{q_i}(x_t | \mu_{q_i}, \sigma_{q_i})}{\sum_{i=1}^Q W_i f_{q_i}(x_t | \mu_{q_i}, \sigma_{q_i})} \quad \text{Eq. 5.17}$$

As equations 5.16 and 5.17 are recursive, we have to optimize them iteratively. To do this, we indirectly maximize the likelihood function by optimizing an expectation-maximization (EM) criterion function. For the EM criterion, we use the log likelihood function as it allows for an efficient estimation of the model parameters, even in cases where the likelihood function is intractable or difficult to optimize directly. The log likelihood function for determining a sequence of states \mathbf{q} given the observed FRET trajectory \mathbf{x} , is computed as the product of the emission probabilities weighted by the relative occurrence probabilities summed over all trajectories and is given by¹⁵:

$$\log L = \sum_{i=1}^Q \sum_{t=1}^T w_{q_i,t} \log(f_{q_i}(x_t | \mu_{q_i}, \sigma_{q_i})) \quad \text{Eq. 5.18}$$

To optimize the log likelihood function, we employ the Baum-Welch algorithm, also known as the forward-backward algorithm. During each iteration of the Baum-Welch algorithm, an expectation step (E-step) and a maximization step (M-step) are performed. In the E-step, the algorithm calculates the expected values of the hidden states given the observed data and the current parameter estimates. These expected values are then used in the M-step to update the model parameters, aiming to improve the fit between the model and the observed data. The training of the HMM continues until the relative improvement in the log likelihood between consecutive iterations falls below a predefined threshold. In all of our HMM training procedures, we set the convergence threshold to 10^{-9} . To derive the transition density matrix, the HMM needs to be trained on the dataset to be analyzed. This working step can be carried out in two different approaches: the analysis can be carried out either (1) trace-wise or (2) globally. In the first case, the transition probability and emission probability are optimized for each individual trace while, in the second case, one uses a shared single transition probability matrix and parameters for the emission probability for all trajectories together. When analyzing 1-color and apparent FRET traces, we use a local HMM as the exact values of the states can be shifted due to the above-mentioned background contributions.

5.7. Parameters for Hidden Markov Modeling

FRET efficiencies of dynamic DNA Origami structures were conventionally analyzed by HMM using the HMM Pomegranate toolbox written by Jacob Schreiber (2016). Molecules were manually classified as dynamic or static, and time windows were selected for the data analysis. For molecules showing dynamic transitions, we choose an HMM model with two or three states, depending on the designed Origami structures. Start parameters were chosen assuming a self-adapting width μ , and a standard deviation σ of 0.05 with random uniform distribution. Convergence between the experimental data and the fit was assumed, when the change in likelihood during consecutive iterations was less than 10^{-9} . For visualization via TDP, each transition as superimposed as a 2D Gaussian function with a fixed width.

5.8. Evaluation of involved FRET states and interconversion rates

The last step involves the visualization of determined rates, i.e. dwell times, and states determined from the FRET and / or normalized intensity traces. We employed so called transition density plots (TDPs), which depict each transition that was identified by the HMM algorithm or the Deep-LASI state classifier in the recorded time traces as a single event in a 2D diagram. The diagram, hence, depicts and links the FRET value before and after an identified transition visually. In the case of 1-color data, we normalized the traces between the minimal and maximal value of observed counts of all measured single traces. The TDPs were generated as described by McKinney et al., i.e. all transitions are depicted as summed up two-dimensional Gaussian functions with an amplitude equal to the total number of transitions and a fixed variance of 0.0005.¹⁴

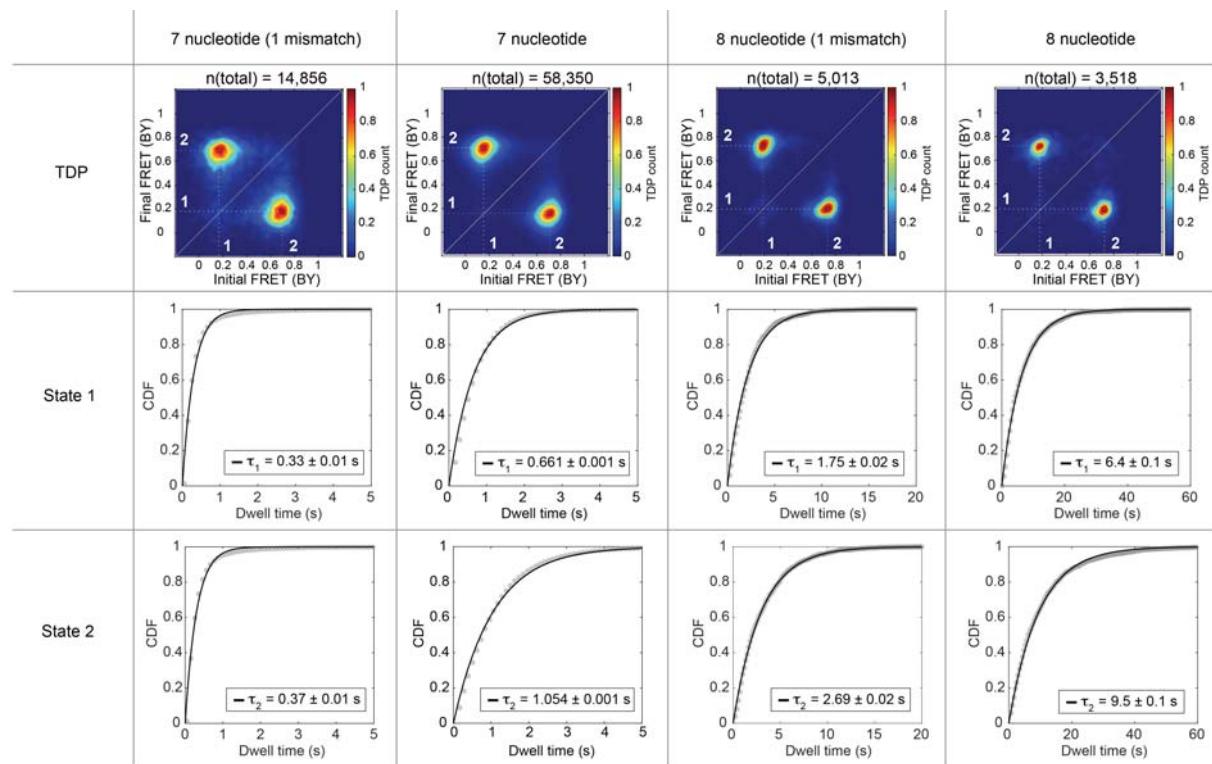
SUPPLEMENTARY NOTE 6: DETAILS OF DEEP-LASI ANALYSES

6.1 Results for the three-color, two-state DNA origami structure with different binding site lengths

The three-color DNA origami structures were measured with four different lengths of complementary DNA for the two binding sites. The two binding sites contained the identical DNA sequence and lengths. The dwell time distributions determined from the state classifiers of Deep-LASI for the different three-color DNA origami structures are shown in [Supplementary Figure 6.1](#). The same analysis workflow was followed for each sample: a fully automated categorization and prediction of state occupancy in traces labeled as ‘dynamic’ were performed with Deep-LASI followed by a manual selection of the different states and fit to a mono-exponential function:

$$f(x) = 1 - e^{-b \cdot x} \quad \text{Eq. 6.1}$$

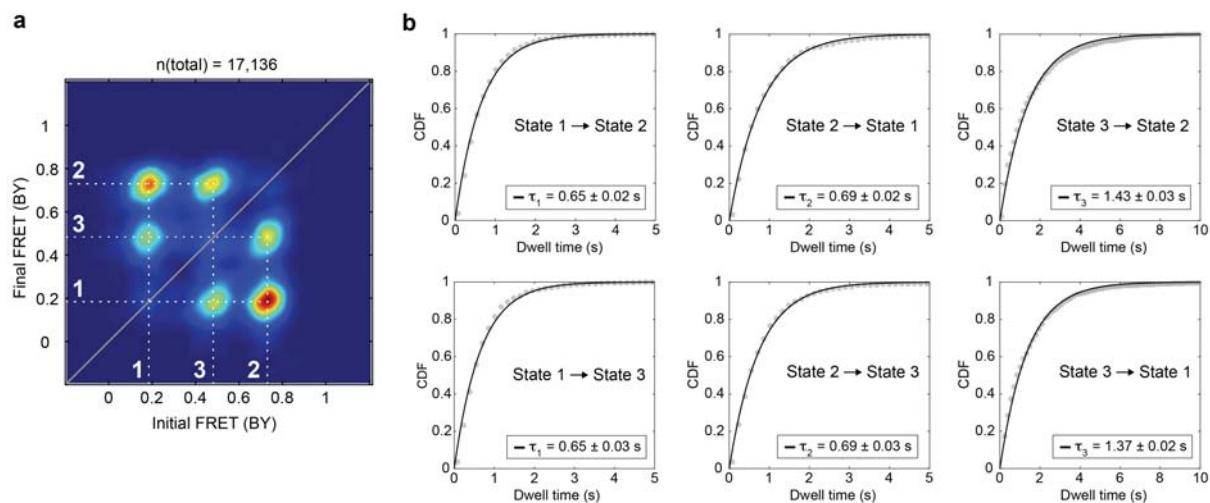
These experiments confirm that Deep-LASI is capable of extracting mono-exponentially distributed dwell times over a large range of kinetic rates.



Supplementary Figure 6.1: Dwell-time distributions of the three-color, two-state DNA origamis with different binding site lengths. Each row corresponds to a specific state and each column depicts the TDPs (top) and dwell-time distributions (middle, bottom) extracted from the uncorrected blue-yellow transition density plots and fitted with a mono-exponential for each binding site length. The errors on the dwell times are the 95% confidence intervals returned by the fitting procedure (estimated from the Jacobian matrix).

6.2 Kinetics of the three-color, three-state DNA origami.

From the three-color, three-state DNA origami with 7 nt binding strands at positions 6 and 12 o'clock and a 7.5 nt complementary binding strand at 9 o'clock. Three populations were extracted automatically from the traces identified by Deep-LASI as dynamic. The dwell-time distributions of all 6 populations observed in the blue/yellow TDP plot (Figure 5c) were extracted manually and fit with an exponential function (Supplementary Figure 6.2). The dwell times of each state are in excellent agreement with the two-color, three-state DNA origami sample (Supplementary Figure 6.3), indicating that the additional blue dye in close proximity of state 2 does not influence the kinetic rates.



Supplementary Figure 6.2: Dwell-time distributions of the 3-color 3-state DNA origami. (a) The blue/yellow transition density plot and (b) the dwell-time distributions extracted from the BY-TDP and fit using a mono-exponential. The errors on the dwell times are the 95% confidence intervals returned by the fitting procedure (estimated from the Jacobian matrix).

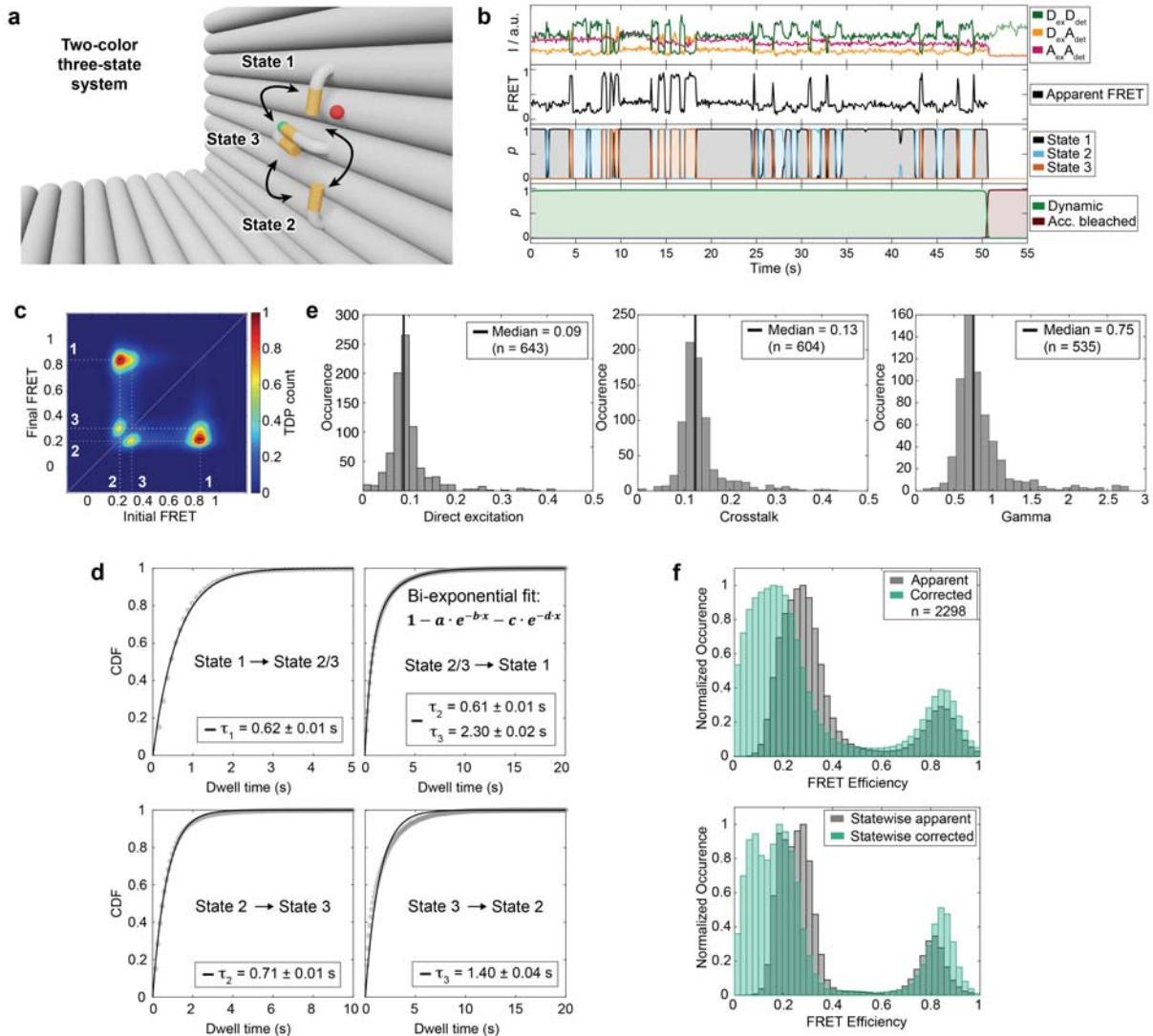
6.3 Results for the two-color, three-state DNA origami structure

Next, we tested the performance of Deep-LASI on a more complex, two-color, multi-state system by introducing a third binding site on the DNA origami (Supplementary Figure 6.3a) and increasing the average transition rates. In contrast to the two-state system described above, State 1 and State 2 at the 6 o'clock and 12 o'clock positions are now characterized by 7 nt binding sites in the three-state DNA origami. The added State 3 at 9 o'clock has a 7.5 nt overhang. In the example trace shown in Supplementary Figure 6.3b, Deep-LASI extracts the dynamic section and identifies all transitions between the three states summarized in the TDP of apparent FRET efficiencies (Supplementary Figure 6.3c). As expected, the FRET efficiency of state 1 (0.83) and state 2 (0.21) do not change significantly compared to the two-state system. In addition, a third state with an apparent FRET efficiency of 0.31 is observed. However, as states 2 and 3 show a similar distance to the acceptor, the states and thereby the transitions are not easily separable. When looking at the dwell-time distributions, the transition out of state 1 is not affected by the degeneracy of states 2 and 3. However, the transition rates from state 2 or state 3 to state 1 differ significantly due to the different binding site lengths and can only be extracted using a bi-exponential fit (Supplementary Figure 6.3d):

$$f(x) = 1 - a \cdot e^{-b \cdot x} - c \cdot e^{-d \cdot x} \quad \text{Eq. 6.2}$$

From the TDP, we can also extract the transitions between states 2 and 3. The transition from state 2 to state 3 can be well described by a mono-exponential distribution whereas the reverse transition from state 3 to state 2 has a second component due to the difficulties of clearly separating the different states.

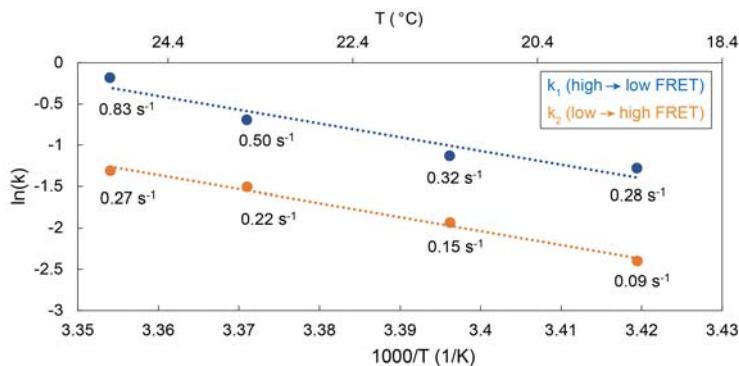
From the single molecule trajectories, Deep-LASI also extracts the regions of the trace that can be used for determining the different correction factors. The FRET correction factor distributions determined by Deep-LASI are shown in Supplementary Figure 6.3e and are consistent with the correction factors of the two-state DNA origami dataset shown in Figure 3f. The framewise apparent smFRET histogram is shown in Supplementary Figure 6.3f (top, gray). In this histogram, states 2 and 3 merge into one degenerate state (0.27) due to heterogeneous broadening of the two populations. After correction (Supplementary Figure 6.3f, top, orange), the degeneracy is decreased and the low-FRET peak broadens. However, they are still not clearly separable. It is only after using the state-label information, which allows us to average the state FRET efficiencies that the two low-FRET populations become distinguishable and the individual FRET populations observed (Supplementary Figure 6.3f).



Supplementary Figure 6.3: Analysis of 2-color, 3-state DNA origami measurements. **(a)** Zoom-in of the L-shaped DNA origami structure with three binding sites. FRET is expected between a high FRET state 1 (12 o'clock), a low FRET state 2 (6 o'clock), and an intermediate FRET state 3 (9 o'clock). **(b)** A representative single molecule intensity trace and FRET trajectory. The upper panel shows the intensity in the yellow and red channels after yellow excitation and the red intensity after red excitation. The middle panel shows the corresponding FRET efficiencies for the dye pair. The third and fourth panels show the output of the Deep-LASI analysis for state-transition and trace classification respectively. **(c)** The TDP of the apparent FRET efficiency states are shown. Interconversion between three conformations with apparent FRET efficiencies of 0.21, 0.31 and 0.83 are observed. The three states are labeled in white. Total number of transitions: 174,697. **(d)** Exponential fits of the dwell time distributions for all states are plotted. The transitions from state 2 and 3 to state 1 were pooled together due to the high overlap and fit with a bi-exponential function. While the dwell time of state 2 in the bi-exponential fit is close to the dwell time extracted from the single population (state 2 to state 3), the dwell time of state 3 is significantly overestimated compared to the single population of transitions from state 3 to state 2. The errors on the dwell times are the 95% confidence intervals returned by the fitting procedure (estimated from the Jacobian matrix). **(e)** Correction factors for direct excitation, crosstalk and gamma extracted by Deep-LASI. **(f)** *top* Frame-wise weighted state-wise smFRET histograms of apparent and accurate smFRET efficiencies. A broadening of the low-FRET population is observed as the correction of the FRET efficiency begins to lift the degeneracy. *bottom* Plotting the framewise-weighted statewise smFRET histograms of apparent and accurate FRET efficiencies improves the contrast. Three peaks are now observable with corrected FRET efficiencies of 0.09 and 0.84 (in line with the two-state system), and a new third state at 0.19.

6.4 Kinetics as a function of Temperature

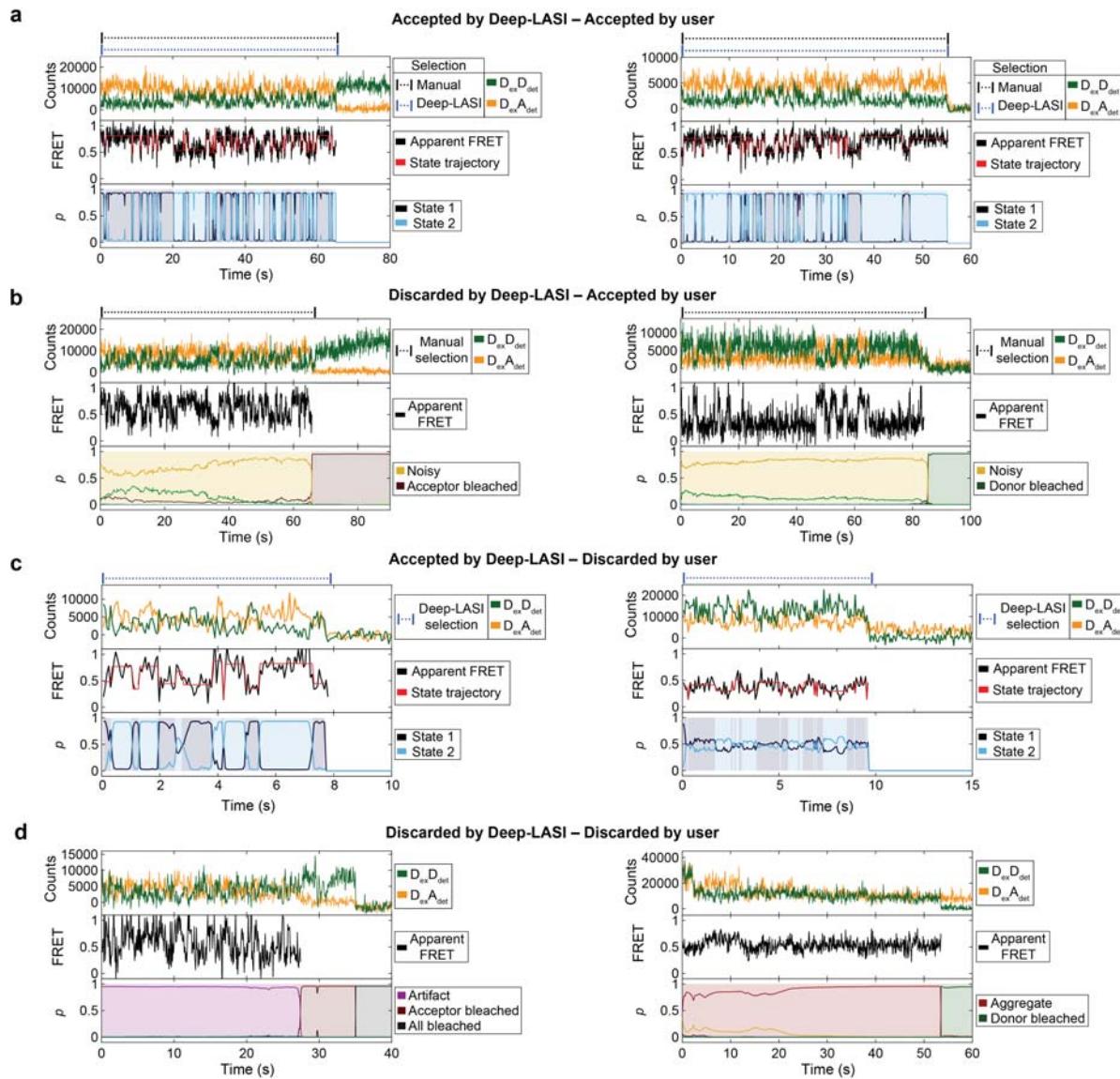
To investigate the influence of temperature on the binding kinetics, we used the two-color two-state L-shaped origami structure with minor changes (exchanged staple strands are noted with asterisks in [Supplementary Table 7.2](#)). Single molecule dynamics were measured between 19.3 °C and 25.0 °C. The transitions rates between state 1 and state 2 are given in [Supplementary Figure 6.4](#). The rates change roughly by a factor of 2 per 2 °C.



Supplementary Figure 6.4: Binding kinetics as a function of temperature. An Arrhenius plot of the transition rates between state 1 and state 2 of the two-color, two-state L-shaped DNA-origami structure. Linear fits (dotted lines) are shown to guide the eye.

6.5 Analysis of previous published 2-color Hsp70 Ssc1 using Deep-LASI

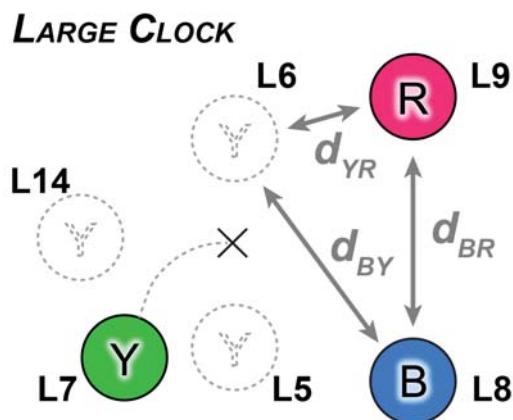
To test Deep-LASI on single-molecule FRET data from proteins, we reanalyzed data that we published previously¹⁶. These data were collected of proteins that were encapsulated in ~200 nm liposomes. Due to vesicle encapsulation, a photostabilization buffer could not be used. Hence, the protein data on this system had a lower signal-to-noise ratio than we typically had with the photostabilized DNA origami structures. In [Supplementary Figure 6.5](#), we show examples of individual traces that were evaluated similarly or differently by the user and Deep-LASI. From these comparisons, we see that: 1) the user and Deep-LASI agree on traces that have sufficient SNR and otherwise show no anomalies, 2) manually selected dynamic traces with erratic intensities after photobleaching or broad FRET distributions tend to be categorized by Deep-LASI as noisy rather than dynamic, 3) Deep-LASI tends to include short traces that are trashed in the manual analysis and 4) traces that show clear features of overestimated background, multiple bleaching steps or other anomalies are discarded by both Deep-LASI and the user.



Supplementary Figure 6.5: Representative 2-color SSC1 traces with disagreements between manual and Deep-LASI selection. (a) Two true positive examples of manual and Deep-LASI selection agreeing on the validity of dynamic sections. The state trajectory (red) and confidence levels of the state transition classifier are shown in the two lower panels. (b) Two example traces that were included in the manual analysis but classified by Deep-LASI as noisy. The user could include these traces by lowering the confidence threshold of the ‘dynamic’ category. (c) Two examples of short traces classified as ‘dynamic’ by Deep-LASI but not selected manually. The left panel shows a valid dynamic trace, which was likely missed during evaluation. The right panel shows an apparent false positive classification by Deep-LASI. The low confidence of the state transition classifier (lower panel) allows this trace to be easily excluded the user. (d) True negative examples excluded from further analysis by both manual evaluation and Deep-LASI. The left panel shows a trace with overestimated background correction and artifact prediction. The right panel shows a trace two bleaching steps of the donor and aggregate prediction.

SUPPLEMENTARY NOTE 7: DNA SEQUENCES

Here, we describe the details of the L-shaped DNA origami structures. The structures were previously published by Timnefeld et al.^{17,18}. As a scaffold, we used the p8064 scaffold derived from M13mp18 bacteriophages. An overview of all designed DNA origami structures including name, the strand IDs of the introduced modified staple strands as well as the binding sites is given in [Supplementary Table 7.1](#). The unlabeled staple strands are specified in [Supplementary Table 7.2](#), staple strands with biotin modifications for surface immobilization are listed in [Supplementary Table 7.3](#) and staple strands with fluorescent modifications for single-molecule FRET are summarized in [Supplementary Table 7.3](#).

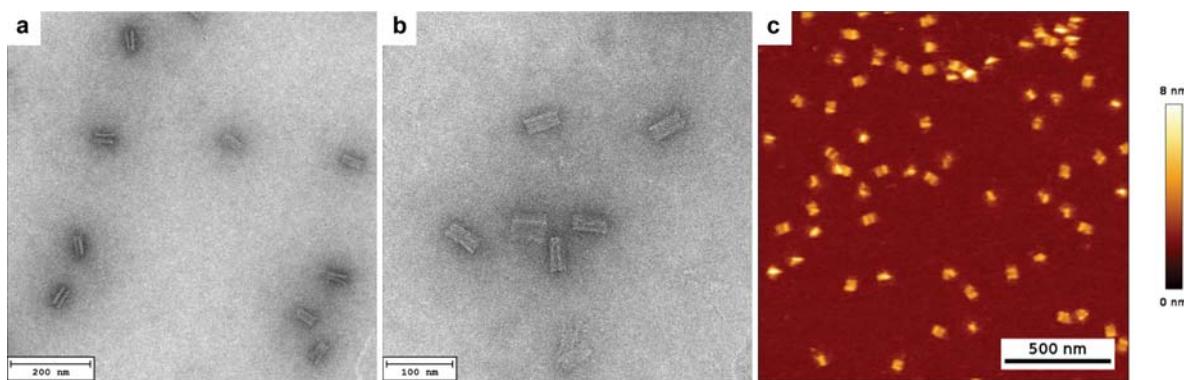


Supplementary Figure 7.1: Schematic of the replacement staple strands forming the 3 state, 3color FRET clock on the L-origami. The strands either carry one of the three fluorophores (L7, L8, and L9) or represent a binding site at 6, 9 and 12 o'clock (L5, L14 and L6 respectively).

additional binding site was introduced on staple strands L14 at 9 o'clock. In addition, for the implementation of the 9 o'clock binding site, the staple strands L12 and L13 are replaced by L12-13-I, L12-13-II and L12-13-III ([Supplementary Table 7.4](#)). All samples share biotinylated attachment sites at positions L1-L4 ([Supplementary Table 7.3](#)).

After folding and purification of the origami samples, the correct folding was confirmed via transmission electron microscopy (TEM) and atomic force microscopy (AFM) as shown in [Supplementary Figure 7.2](#). They form compact structures of roughly $60 \times 20 \times 20 \text{ nm}^3$.

The L-shaped DNA origami structures are made of 252 ssDNA staple strands annealed to a circular complementary ssDNA scaffold strand of 8064 nucleotides. The three fluorophores ATTO488, Cy3b and ATTO647N are introduced into the structures, by replacing the unlabeled ssDNA strands L7, L8 and L9 ([Supplementary Figure 7.1](#)) with strands containing the appropriate label ([Supplementary Table 7.4](#)). Binding sites for the L7-attached tether strands consisting of different lengths are introduced at position L5 and L6 for the 2 state systems with low and high FRET values and different binding rates, even with identical sequences. We refer to the binding site for staple strand L5 as 6 o'clock and staple strand L6 as 12 o'clock. For generating a 3 state FRET system, an



Supplementary Figure 7.2: Structural characterization of L-shaped DNA origami structures. (a-b) Transmission electron microscopy and (c) atomic force microscopy images indicated the efficient folding of the L-shaped DNA origami structures. Color table: Dark Gold.

AFM imaging was performed on a NanoWizard® 3 ultra AFM (JPK BioAFM AG/Bruker; Germany) in solution using 1x folding buffer. The DNA origami structures were immobilized on a freshly cleaved Mica surface (Quality V1, Plano GmbH; Germany) by Ni^{2+} ions, which were incubated on the Mica plate for 5 minutes with a 10 mM NiCl_2 solution. Afterwards, the mica was washed three times with ultra-pure water and dried by pressurized air. 10 μL of 1 nM DNA origami solution were then added and incubated for 5 minutes. Measurements were performed with a USC-F0.3-k0.3-10 cantilever (Nano World; Switzerland).

TEM imaging of the L-Origami structures was carried out on Ar-plasma cleaned TEM grids (Formvar/carbon, 400 mesh, Cu, TedPella, Inc.; USA). The DNA origami structures were stained with a 2 % uranyl formate solution. The imaging was performed on a JEM-1100 microscope (JEOL GmbH; Japan) with an acceleration voltage of 80 kV.

Supplementary Table 7.1. The applied nomenclature used for the designed L-shaped DNA origami structures with the corresponding staple strand IDs that carry the fluorescent dyes or the attachment of the pointer. The laser excitation scheme for the 3cFRET B-Y-R samples involves excitation at 488, 561 and 640 nm.

| # | Name | Blue | Yellow | Red | Binding sites | | | Replaced |
|---|-------------------|--------|--------|--------|---------------|----------|-----------|--------------|
| 1 | BYR-Pos6/12-6.5nt | 488-L9 | 561-L7 | 640-L8 | 6.5nt-L5 | 6.5nt-L6 | --- | --- |
| 2 | BYR-Pos6/12-7nt | 488-L9 | 561-L7 | 640-L8 | 7nt-L5 | 7nt-L6 | --- | --- |
| 3 | BYR-Pos6/12-7.5nt | 488-L9 | 561-L7 | 640-L8 | 7.5nt-L5 | 7.5nt-L6 | --- | --- |
| 4 | BYR-Pos6/12-8nt | 488-L9 | 561-L7 | 640-L8 | 8nt-L5 | 8nt-L6 | --- | --- |
| 5 | BY-Pos6/12-7.5nt | 488-L9 | 561-L7 | --- | 7.5nt-L5 | 7.5nt-L6 | --- | --- |
| 6 | BR-Pos6/12-7.5nt | 488-L9 | --- | 640-L8 | 7.5nt-L5 | 7.5nt-L6 | --- | --- |
| 7 | YR-Pos6/12-7.5nt | --- | 561-L7 | 640-L8 | 7.5nt-L5 | 7.5nt-L6 | --- | --- |
| 8 | YR-Pos6/9/12 | --- | 561-L7 | 640-L8 | 7nt-L5 | 7nt-L6 | 7.5nt-L14 | L12-13-I-III |
| 9 | BYR-Pos6/9/12 | 488-L9 | 561-L7 | 640-L8 | 7nt-L5 | 7nt-L6 | 7.5nt-L14 | L12-13-I-III |

Supplementary Table 7.2. Unmodified staple strands used for the L-shaped DNA origami structure given from the 5' to 3' end. All oligonucleotides were purchased from Integrated DNA Technologies. *In the origami used for the measurements shown in [Supplementary Figure 6.4](#), the staple strands L141, L153 and L165 were replaced with the sequences highlighted with the asterisks at the end of the table.

| Staple ID | Sequence (5' to 3') |
|-----------|---|
| L1 | ATCCAGAACAAATTAGTCATCAGGAACGGT |
| L2 | CGTGCCTGTTCTCGCATCCAGCGCCGGGTTA |
| L3 | ATAATCAGAAAAGCCAACATCCACTGTAATA |
| L4 | CATAGGTCTGAGAGACAAATCGTCGAATTACC |
| L5 | AGAAACAGCTTAAAGGAAGAAAAATCTACGATTTAAGCATATAAC |
| L6 | GCACCCCTCCGTCAAGGTACGTTAGTAATGAATAGTTAGCGTCAATCAT |
| L7 | ACGATAAACCTAAAACAAAGAATACACTAAAACATTACCAACAAAGC |
| L8 | TGCTCATTCTTATGCGTTAATAAAACGAACATATTCAATTGGCTTTG |
| L9 | AAGGGAACCGGATATTCACTCATCTTGACCGTAATGCCATCGGAAC |
| L10 | CGGAATCTCAGGTCTGTTAAATATGCATGCGAACGAATCATTG |
| L11 | TGAATTACCACTGAAATGGAATTACGAGGCATATAGCAGAGAATCCCC |
| L12 | CATTATACGGTTACCCATAACCCTCGAAATACAATGTTAACAGGG |
| L13 | ATTCATATCAGTGATTGGCATCAGGACGTTGTAACATAAACCGACG |
| L14 | TAATAAGAAGAGCCACCCATTAGCGTTGCCATTCAACAATAGAAA |
| L15 | GGCACCAAAACCAAAAGTAAGAGCAACACTATAGCAACGTAATGCC |
| L16 | ATAAAAATATCGCCTCTCCTTTGATAAGAGCTATAT |
| L17 | TACCACTACGCTAACAGTTGCTATTTGCACCCATCCT |
| L18 | GAGGGTAGTTGCAGGGTGCTAAACAACCTTCACGCCGGAAAGAG |
| L19 | AGAGCCGAAACAAATGAGACTCCTCAAGAGATTAGCGGGCAGTAGCA |
| L20 | AGTTGATTAGCTGAAAAGAGTACCTTAATTGTTAATTCGGACCATAA |
| L21 | TCGATAGCAGCACCGTAAATCACGTTTGCT |
| L22 | AAAGACAAATTAGCAAGTCACCAATGAAACCA |
| L23 | ATATTCAACGCCAGCATTGACAGGCAAAATCA |
| L24 | TTTCCCTTACACTGGTTG |
| L25 | CTCCAATCGTCTGAAATT |
| L26 | TTTTGCCTGAGTAGAAGAA |
| L27 | TTTCCGACTTACAAATAAACAGTTT |
| L28 | ATACGAAAGAAATTATTCAATTAAAGGTGAATT |
| L29 | TTTCTTACAAACAATTG |
| L30 | TTTTAAGTTACCAGGGTAATTGAGCTTT |
| L31 | TTTTAACGATGCTGATGG |
| L32 | TTTCAGGGTGGTTTCTT |
| L33 | ACAAAGTATGAGGAAGCTTGAGGACTAAAGATT |
| L34 | CCGAATCTAACGATCTTT |
| L35 | AGATGAAGGGTAAAGTTT |
| L36 | TTTCGAAATGGTCAATAACCATTAGATGC |
| L37 | TCGAAGATGATGAAACTTT |
| L38 | AGAGCAAATCCTGTCAGATACCGACAAAAGGTAAATT |
| L39 | TTCCGGAATCATAATT |

| | |
|-----|--------------------------------------|
| L40 | TTTTGGATTATTCACAGAA |
| L41 | TGCGGCCAGAATGCGGTTT |
| L42 | TCAGCAGCAACCGCAATTT |
| L43 | TTTAGAGCGGGAGCTAGAT |
| L44 | TTTGCTAATATCAGAGAGATAACCCCGCCACCGCG |
| L45 | ACCTCGTCATAAACATTTT |
| L46 | TTTTTGAGGGGACGACGAC |
| L47 | TTTAACAGTACCTTTACA |
| L48 | TTTTGGCCTCCTGTATAA |
| L49 | TTTGGCGCATAGGCTGGCTAACGGTGTAAATTGT |
| L50 | TCACCGGAAGCATAAATTT |
| L51 | TTTTATCATCGCCTGAACAGACCATT |
| L52 | TTCATAGGGTTGAGTGT |
| L53 | TTTTAGCCCGGAATAGCCTATTCTTT |
| L54 | TTTCCCTCAGAGCCACCACCCCTCAGAAAGCGCTTA |
| L55 | TAGTAATAACATCACTTTT |
| L56 | TTTTTGTCCAGTTGGAACAAGA |
| L57 | TTTCGGGCCGTTTCACGG |
| L58 | TTTATTGCTGAATATAATACATT |
| L59 | TTTAGTAATTCAATCGCAAGACAATT |
| L60 | TTTGAATGCCAACGGCAGC |
| L61 | CAGATGAATATAACAGTT |
| L62 | TTTCATATTATTATCCAAATCCAAAGTCAGAGA |
| L63 | TTTTTATCACCGTCACAGCGTCAGTT |
| L64 | TTTCTTTTCACAACGGAGATTGTT |
| L65 | TTTGTGTAGGTAAAGATT |
| L66 | TTTTTTTTAAACTAG |
| L67 | TTTGATTAAGACGCTGAGA |
| L68 | TTTGCATTGGCGCTT |
| L69 | ATTATAGCGTCATAAGTAAAATGTTT |
| L70 | TAGTCAGAACGAAAGCGGATT |
| L71 | TTTTAGACTGGCATCAGTTGAGATT |
| L72 | CATAATAATTGCGT |
| L73 | ATATATATAAGCGACATCGGCTGCTTCCTTATCATT |
| L74 | AAAACGGAATCGTTT |
| L75 | ACAAATTATCATCATATT |
| L76 | TTTTTCCTGATTATCACGT |
| L77 | TTTCATATAAAAGAAAGCCGAACATT |
| L78 | TTTGTGTAAAGCCTGGCG |
| L79 | TTTAAACATCAAGAAAAAA |
| L80 | AATGCAATAGATTAAGGGCTAGAGCTATT |
| L81 | ACATAGCGATAGCTATT |
| L82 | TTTTGCATAAAAGCCTGAGTAATT |

| | |
|------|---|
| L83 | CTGATAGCCCTAAACTTT |
| L84 | GAAAGGAGCGGGCGCTAGGTTT |
| L85 | TTTGCCTCAGAGCATAAAGAAAATTAAGCAATAAATT |
| L86 | TTTACTGTAGCCTCAGAACGCCATT |
| L87 | TTTACCGTCCAGTAAGCGTCATACATGGCTCAGTTAAT |
| L88 | AGTGTGCTGCAAGGC |
| L89 | TTAATTAAACCATAACATACATAAAGGTGGCAATT |
| L90 | CCGTGCATCTGCCAGTTT |
| L91 | TTTTTAGGAATACCAACAGTAGTAATT |
| L92 | TTTATTGGGCTTGAGATGCCAGAACGATT |
| L93 | TTTGAAACAACAAAGAACACTGATT |
| L94 | TTTACTAGAAAAAGCCTGTT |
| L95 | TTTCGACTTGATCGAGAGGGTTGATATAAGTATT |
| L96 | ATTTAGAAGTATTAGATT |
| L97 | TTTACCTGCTGAACCAGG |
| L98 | TTTTCCAAGAACGGGTGCGAACCTT |
| L99 | TTTACGCATAATGAGAATAGAAAGTT |
| L100 | TTTAGAACGCGAGAAAATT |
| L101 | TTTGTTCGTCACCAGTACTGTACCGTAAT |
| L102 | CATGTTTACCAAGTCCCTT |
| L103 | TTTGGAAATTGTGAGAGAT |
| L104 | TTTATTAAAGTGGGTACGC |
| L105 | TTTGGAACCTAAGTCTCTGAATT |
| L106 | TTTATGCCATTAAAATA |
| L107 | GAGCGATATAACACAAACATGCCCTT |
| L108 | CGGCCTCGTTAGAATCTT |
| L109 | TTTGCCTGGCAAGTGTAG |
| L110 | TAGTTGCCAGTTGGGAGGTTGAAGATCAATAA |
| L111 | ATGGCTACAATCAACTGAGAGCCAGCAGCAAATGAAAACGAACCTAATGCGCTGGCAGA |
| L112 | TCATCAACAAGGAAATATGTACCCGGTTG |
| L113 | TTCAAATTAGAAAAACAGGAGCAAACAAGAGAATCGATGAAGGGTGAGATATT |
| L114 | CAACTAATGCAGACAGAGGGCAACTG |
| L115 | GTACTATGGTTGCTTTAGACACGCAAATT |
| L116 | TGTAGCTAACATTACCGTCGAAAGAC |
| L117 | ATCAAAAGTCATAAAACGGAACAACATTATCAACTTAGTAGAT |
| L118 | AACGTCAATAGACGGGAATACCCAAAAGAACAGACTCCGTTTAT |
| L119 | GGAGGGAAAGAGCCAGCAATCAGTAGCGACAGACCAGAACCGCCTC |
| L120 | AGCGAACAGAACGCTGGAGAATCACAAAGGCTATCAGGT |
| L121 | GCCCCCTGGTGTATCACCGTACTC |
| L122 | TACAGGCATTAAATTAAACCAATAGGAACGCCATCAAAGTCATCAGAATTAGCCTAAATCG |
| L123 | TATTTTGAGAGATCTGCCATTTCCTCTACTCAATTGA |
| L124 | CATTGCCTGAGAGTCTTATGACCATAATCATTCAATT |
| L125 | CCAGCCAGCTTCCGGTAATGGGTAACAC |

| | |
|------|--|
| L126 | ATCGGCAAAATCCCTTACGTGGACTCCAACGT |
| L127 | CCTGCAGCCATAACGGGGTGTCCAGCATCAGC |
| L128 | GGGCCTTCGCTATTACGTTGACCTCACCG |
| L129 | ACCCTCATGCCCTCATTTCTGTATGGGATTAGTTAAAGCAGCTTGA |
| L130 | GTTGTACCACCCCTCATAAAGGCCGGAGACAG |
| L131 | TCTTTAGGCTGAATAATGCTCATTAGTAACAT |
| L132 | CTCTCACGGAAAAAGAACGGATAAAAACGACG |
| L133 | ACGCCAGATGACGGGGCGCCGCTAGCCCCAGC |
| L134 | TTAATTCATGTTCTATAACTATATGTAATGCTGATGTCAATAGAATCCTGACAAAATT |
| L135 | TTTCATCGAATAATATCCAGCTACAATACTCCAGCAATTCTTACAG |
| L136 | AATAAGTTAGCAAAAACGCAATAATAACGAGAATTAAAAGCCAA |
| L137 | GACCGTGTGATAAAATACAAATTCT |
| L138 | ACAAGAACCGAACTGATGTTACTTAGCCGGAAAAGACAGCACTACGAA |
| L139 | ATCAAACCTAAATTCTGGAAGGGCCATATCA |
| L140 | CGCTGGCACCAACAGGGTACCGCATTGCAA |
| L141 | GAGAAACATTAAATTTCAGGGTAGAAAG |
| L142 | CTGCGCGGCTAACTCACAATTCCACACAACATACGAGTACCGGGGCTGTGGGTGTTCAG |
| L143 | CCGAGTAAGCCAACAGGGTACCGCATTGCAA |
| L144 | AAACGGCGCAAGCTTGAAGGGCGATCGGTGC |
| L145 | CAAAGAAATAAAATACCCAGCGATTATACCAAGCGCGAA |
| L146 | CTTAATTGAGACCGGAAACAGGTAGGATTAGAGGTGGCA |
| L147 | GCCAGTGCAGGTTGACCCACCCTCTGGTGC |
| L148 | CCAGAATGGAGCCGCCAATCAAGTTGCC |
| L149 | CCCCCTGCGCCGCTTAGCTGTTCTGTGT |
| L150 | GGAAACCAGGCAAAGCGTACATAAGTGA |
| L151 | AAATCAACACGTGGCATCAGTATTCTCAATCC |
| L152 | AGGAGGTGGCGATAAGTATTAGAGGCTAAATCCTCTACAGGAG |
| L153 | GACAGATGGACCTTCATCAAGAGCCCTGAC |
| L154 | CTGAGGCCAACGGCTACAGAGGTTCCATT |
| L155 | ATAACCTTATCAACAAAAATTGTATAACCTCC |
| L156 | AAATCAGCTATTGAGGATCTCAAGAGGCTTT |
| L157 | CACAGACATTCAGGGATCTCAAAAAAAAGGTTCTAAAGCCGCTT |
| L158 | TAATAGTATTCTCCGTGCATTAATTTGTT |
| L159 | CGTTGGTAGTCACGACGCCAGCTGGCGAAAGGGGATATCGGCCTGCGCATGGCCAGCTT |
| L160 | CTTCTGACCTAAATTGAGGCCAGAACGCAATTACG |
| L161 | GCTGCGCAACTGTTGGCAGACCTATTAGAAGG |
| L162 | AGAACGTTAACGGCGTAATGGGAAAGGTTCTTGCCTGGTGTGGCTGGCTTGCGTT |
| L163 | TTAGTTGCCTGTTAGGTCAATTGCGGATAGGAAGCCGACTATTA |
| L164 | AATTACATAGATTTCAATAACGGATTGCC |
| L165 | AAACGGGGTTTGCTACATAACGCCAAAAAAGGCTGTAATCTT |
| L166 | TGCGAATAATAATCGACAATGTCGGTCG |
| L167 | TTATACTTAGCACTAAAAAGTTGTGCCGCCA |
| L168 | GCCGTCACAATATAAAAGAAACCACCAGAAGGAGCGGACTCGTATTACATTGTCAAATAT |

| | |
|------|---|
| L169 | ATTGCGTTAACACATTCAATTACCTGAGCAAAAGGGAGAACAGGTTAACGATGATGG |
| L170 | GGAGCCTTCACCCCTCAGAGCCACC |
| L171 | CCAGCTTACGGCTGGAAACGTGCCGTCTCGT |
| L172 | TTCGTAATCATGGTCATCCATCAGTTATAAGT |
| L173 | AACAGAGGTGAGGCCGGCAGACAATTAAAAGGG |
| L174 | TTGAGTAAGCCACCCCTCAGAACCG |
| L175 | CAGTATGTTATTTGCGAAGGCCCTTTAACGTTCTGAACA |
| L176 | GCCTGTTGCTTCTGTTACCTTTAACGTTAA |
| L177 | ATAAACAAATCCCTAGTGAATTATCAAAAT |
| L178 | CAAAGGGCTGTCGTGGCCCTGAGAGAGTT |
| L179 | CTCAAATGTTAGAAATGGAAGTTCACCGCATTACTCAACTGGCT |
| L180 | CCCGCCCGCTTAATGAAAGCCGGCGAACGTG |
| L181 | TTCACCAGGTAGCAATGCCCTGCTGGTAAT |
| L182 | GTCGAAATCCCGCACCTGCTCCACCAACTTTAGCATT |
| L183 | TGATTGCTTGAATACAAACAGAATGTTGGA |
| L184 | TTCTGAAACATGAAAGTGCCGCCATTG |
| L185 | AACCGTTCACACGGAAATACCTACATTGACGCTAAACTATCACTCTTTAACAGGAG |
| L186 | CGTTGAAAATAGCAAGCCCAATA |
| L187 | CTTTGCGTTATTCAATGATATTCAACCGTT |
| L188 | AAATCCCGTAAAAAAACGTTTTGGACTTGT |
| L189 | TATCATTGCGAACATCCTGATATAAGAA |
| L190 | AAATTATTGGAAACAGCCATTGAAAATCGC |
| L191 | GCAGCAAGCGGTCCACAAGTGTGAGGCCA |
| L192 | CCAACATGACGCTCAATGCCGGAGGAATACC |
| L193 | TATTTGTTAAAATTGGTATATCAAAAC |
| L194 | TGTTGCCCTGCGCTGATCAGATGCAGTGTCA |
| L195 | TGCGGGATAGCAGCGACGAGGCGCAGAGAACGGCCCGGTAACGATC |
| L196 | TACCGATAGTTGCGTTTCA |
| L197 | TCAAATCACCATCAATACGCAAGG |
| L198 | GTAAGAATAGTTGAAACTTCGCAAACACCGC |
| L199 | ATTGCCCTCACCGCCCCAGCTGCTGCGTTG |
| L200 | AAGCGATAATGAAACAGATATAGAAGGCTTAGCAAGCCTTATTACG |
| L201 | GGAATTAGGTAAATTTGGTCATGCCAACCGGAACCACCACC |
| L202 | GTTCGCGCTGAGATGGCAGGAAGATCGCACT |
| L203 | GCGAGAAAAGGGATGACGAGCACGTATAACGTGCTTTCACGCTGAAGAAAGC |
| L204 | GGGGCGGCCAATTCACTAAAGTACGGTGTACGAGAACAGCTCAA |
| L205 | GAAATTGTTATCCGCTCACATTAATTAATGA |
| L206 | TTTTTAATGCACGTACAAGTTACCCATTCA |
| L207 | CAATTCAATAGATAATAATCCTTGCCCG |
| L208 | CCTCAGAGCACAAGAAGAAAAGTAAGCAG |
| L209 | CGCTCACTATCAGACGGTCCGTGAGCCTCTC |
| L210 | GCAGAGGCGAATTATTTCAATTGCTATTAA |
| L211 | TTAGAGCTATCCTGAGGCTGGTTCAGGGCGC |

| | |
|------|--|
| L212 | GCCAGTACGTTATAAGCGTTAATAAGAATAAACACAAAT |
| L213 | AACGTTATTAATTTACAACTAATCAGTTGGC |
| L214 | GCCGGGCGCGGTTGCGCCGCTGACCCCTGTG |
| L215 | CTGCAACAGTGCCACGTATCTGGTAGATTAGA |
| L216 | TAAAGTTAGAACCGCTAATTGTATCGGGGTTAAGTTGGCCTTG |
| L217 | GAAACAACGCGGTGCCGCACAGGCGGCTTAGTGACTTCTCACGTACAGACGCCAGG |
| L218 | GTCCACTAAACGCGCGGACGGCAACAGCTG |
| L219 | GGAACCCAAAACATACAAACAGTTCAGCG |
| L220 | ATCGGCCTTAAAGAATAATCAAAGAATAGCCCAGACCAGTGAGGGAGAGGGGTGCCTA |
| L221 | ACAGTTGAGGATCCCCAGATAGAACTGAAAGC |
| L222 | CCGGAACCGCAAGAAAGCAATAGCTATCTTACTCACAATCCGATTGAG |
| L223 | GCAGTTGGCGGTTGTCCAGTTATGGAAGGAG |
| L224 | GCCGATTAAGGAAGGGCGCGTAACCACCACA |
| L225 | TGTACTGGTAATAAGTTCACTG |
| L226 | CAAATCGTCAGCGTGGTGCCATCCCACGCAA |
| L227 | TCTTACCATAAAGCCATAATTAGAATGGTTAGGGTAGC |
| L228 | AGGCAGAAAATCCTGTTGTCTATCACCCCCGAT |
| L229 | GCCTAATTATCATATGATAAGAGATTTAGTTAATTTCAT |
| L230 | TTTCATCGGCATATTGACGGCACCACGG |
| L231 | CTAGCTGATAAATTAAACAGTAGGG |
| L232 | CCCTGAACAAATAAGAAACCGCGAGGCAGT |
| L233 | CACATCCTCAGCGGTGGTATGAGCCGGGTAC |
| L234 | CAGGAAAAACGCTCATACCAGTAAATTTTGA |
| L235 | CCACCCCTCTGTTAGGAAGGATCGTCTTCAGCAGACGATTATCAGCT |
| L236 | CAAACCCTTAGTCTTACCAAGCAGAAGATAA |
| L237 | GGCTTAGGTTGGGTTAAGCTAATGATTTCGA |
| L238 | CCGTCGGAGTAGCATTCAAAACAGGAAGATT |
| L239 | ATGAGTGACCTGTGCAGTTCTGCCAGCACG |
| L240 | CCGGCAAATCGCGAAGTGGTAAGGGATAG |
| L241 | ACAAGAAATAGGAATCCCAATAGCAAGCAAATATAGCAGCATCCTGAA |
| L242 | CCATTACCAAGGGCGACATCTTCATAGGCAGAAAGAATAGGTTGAG |
| L243 | TGGAGCCGGCCTCCGGTACATCGACATAAAA |
| L244 | CACTCATGAAACCACCTTAAATCAAGATTGAGCGTCTTTGTT |
| L245 | GTATAAGCAAATTTTAGATAAGTAACACG |
| L246 | AGGAAACCGAGGACGTAGAAAAAGTACCG |
| L247 | CGGGAAACGAAAAACCTGATGGTGGTCCGAA |
| L248 | AGCATGTACGAGAACAAATCCGGTATTCTAAGAACGATTTCCAGA |
| L249 | ACATTCTGAAGAGTCTCCGCCAGCAGCTCGAA |
| L250 | GGGTCAATTGCAAGGGGGATTGACTAAATA |
| L251 | TGCTTCGAGGTGAATCTCAAAA |
| L252 | CAGTACCATAGTACCCAGTGCCGTATAAATTGATGAATTAAAG |

| | |
|-------|-----------------------------------|
| *L141 | TGCCCTGACGAGAACATTAAACAGGTAGAAAG |
| *L153 | GACAGATGGACCTTCATCAAGAGTAATCTG |
| *L165 | AAACGGGGTTTGCTACATAACGCCAAAAAGGCT |

Supplementary Table 7.3. Modified staple strands given from the 5' to 3' end for the L-shaped DNA origami structures used. The biotin was used for surface-immobilization via a biotin/avidin interaction. All oligonucleotides were purchased from Biomers.

| ID | Sequence (5' to 3') | Function | Replace |
|--------|---|-------------------------|---------|
| Bio-L1 | Biotin -ATCCAGAACAAATTAGTCCATCAGGAACGGT | Attachment Biotin at 5' | L1 |
| Bio-L2 | Biotin -CGTGCCTGTTCTCGCATCCAGCGCCGGGTTA | Attachment Biotin at 5' | L2 |
| Bio-L3 | Biotin -ATAATCAGAAAAGCCCAACATCCACTGTAATA | Attachment Biotin at 5' | L3 |
| Bio-L4 | Biotin -CATAGGTCTGAGAGACAAATCGTCGAATTACC | Attachment Biotin at 5' | L4 |

Supplementary Table 7.4. Modified staple strands given from the 5' to 3' end for the fluorescently-labeled L-shaped DNA origami structures. The complementary docking sequences are highlighted in orange. The docking strands have a three-base extension from the DNA origami structure and the pointer contains nine single-stranded thymine bases plus GC and the docking sequence, highlighted in grey.

| ID | Sequence (5' to 3') | Supplier | Function | Replace |
|------------|--|--------------------|------------------------------------|-------------|
| 488-L8 | TGC TCA TTC TXA TGC GTT AAT AAA ACG AAC TAT ATT CAT TGG CTT TTG; X = dT- Atto488 | biomers | Lower Label – V1 | L8 |
| 640-L9 | AAG GGA ACC GYA TAT TCA CTC ATC TTT GAC CCG TAA TGC CAT CGG AAC; Y = dT- Atto647N | Eurofines Genomics | Upper Label – V1 | L9 |
| 561-L7 | GGCACCAAAACCAAAAGTAAGAGACACTATA GCAACGTAATGCC TTTTTTTCG GGCATTAA - Cy3b | Eurofines Genomics | Pointer - dye at 3' | L7 |
| 6.5nt-L6 | GCACCCCTCCGTCAAGGTACGTTAGTAAATGAATAG TTAGCGTCAATCAT TTTCAAATGC | Eurofines Genomics | Pos 12 catching site 7 nt mismatch | L6 |
| 7nt-L6 | GCACCCCTCCGTCAAGGTACGTTAGTAAATGAATAG TTAGCGTCAATCAT TTTCAAATGC | Eurofines Genomics | Pos 12 catching site 7 nt | L6 |
| 7.5nt-L6 | GCACCCCTCCGTCAAGGTACGTTAGTAAATGAATAG TTAGCGTCAATCAT TTTCAAATGCC | Eurofines Genomics | Pos 12 catching site 8 nt mismatch | L6 |
| 8nt-L6 | GCACCCCTCCGTCAAGGTACGTTAGTAAATGAATAG TTAGCGTCAATCAT TTTCAAATGCC | Eurofines Genomics | Pos 12 catching site 8 nt | L6 |
| 6.5nt-L5 | AGAACAGCTTAGAAGGAAGAAAAATCTACGAT TTAAGCATATAACT TTTCAAATGC | Eurofines Genomics | Pos 6 catching site 7 nt mismatch | L5 |
| 7nt-L5 | AGAACAGCTTAGAAGGAAGAAAAATCTACGAT TTAAGCATATAACT TTTCAAATGC | Eurofines Genomics | Pos 6 catching site 7 nt | L5 |
| 7.5nt-L5 | AGAACAGCTTAGAAGGAAGAAAAATCTACGAT TTAAGCATATAACT TTTCAAATGCC | Eurofines Genomics | Pos 6 catching site 8 nt mismatch | L5 |
| 8nt-L5 | AGAACAGCTTAGAAGGAAGAAAAATCTACGAT TTAAGCATATAACT TTTCAAATGCC | Eurofines Genomics | Pos 6 catching site 8 nt | L5 |
| 7.5nt-L14 | TGCCATTCAACAATAGAAAATTCATATGGTTTCAT AATGCC | Eurofines Genomics | Pos 9 catching site 8 nt mismatch | L14 |
| L12-13-I | CATTATACCAAGTGATTGGCATCAGGACGTTGTA ACATAAACCGACG | Eurofines Genomics | Replacement for Pos 9 | L12 and L13 |
| L12-13-II | TAATAAGAAGAGCCACCCTTATTAGCGTT | Eurofines Genomics | Replacement for Pos 9 | L12 and L13 |
| L12-13-III | TTACCCATAACCCTCGAAATACAATGTTAAACA GGG | Eurofines Genomics | Replacement for Pos 9 | L12 and L13 |

SUPPLEMENTARY NOTE 8: STATISTICS SINGLE-MOLECULE DATA

The following section summarizes the results for all Hsp70 and origami datasets, newly designed, presented, and analyzed in this work. For each dataset, the total number (#) of recorded traces, the confidence interval, the number of dynamic traces, the retrieved number of states, and dwell times are specified. For the origami datasets, we refer to State 1 as the 12 o'clock position and State 2 as the 6 o'clock position for the 2-state systems. In 3-state systems (as presented in [Figure 5](#)), we refer to States 1, 2, and 3 as the 12 o'clock, 9 o'clock, and 6 o'clock positions, respectively. For the Hsp70 datasets, we refer to State 1 as the undocked conformation (low FRET) and State 2 as the docked conformation (high FRET). More details can be found together with the raw data on Zenodo [<https://zenodo.org/record/7561162>].

Supplementary Table 8.1. Statistics summarizing the single-molecule results obtained for the origami and Hsp70 datasets. N.A.: not applicable; max: highest confidence of all possible classes.

| Fig. | Name | Binding Site | # of Traces | Confidence | Frame time (ms) | # of Frames | # of Dynamic Tracers | # of States | Dwell time t_1 (s) | Dwell time t_2 (s) | Dwell time t_3 (s) |
|------|------------------------------------|------------------------------------|-------------|------------|-----------------|-------------|----------------------|-------------|----------------------|----------------------|----------------------|
| 2 | 1c origami 2 states | 8 nt / 1 MM | 7448 | 70 % | 52.2 | 998 | 2510 | 2 | 1.75 | 2.65 | N.A. |
| 3 | 2c origami 2 states | 8 nt / 1 MM | 6100 | 70 % | 52.2 | 1998 | 1499 | 2 | 1.76 | 2.64 | N.A. |
| 4 | 3c origami 2 states | 7 nt / 1 MM | 5731 | 20 % | 34.2 | 2997 | 482 | 2 | 0.33 | 0.40 | N.A. |
| | | 7 nt | 5093 | 70 % | 37.5 | 2997 | 1885 | 2 | 0.66 | 1.05 | N.A. |
| | | 8 nt / 1 MM | 2545 | 70 % | 52.5 | 2397 | 581 | 2 | 1.75 | 2.69 | N.A. |
| | | 8 nt | 8097 | 70 % | 152.2 | 1797 | 1545 | 2 | 6.41 | 9.54 | N.A. |
| 5 | 3c origami 3 states | 7 nt @ 6 / 12h 8 nt / 1MM @ 9h | 7990 | 70 % | 52.5 | 2997 | 586 | 3 | 0.65 | 0.69 | ~1.40 |
| 6i | 2c origami 2 states | 5 nt | 95 | N.A. | N.A. | N.A. | 95 | 2 | 0.0045 | 0.0063 | N.A. |
| | | 6 nt | 104 | N.A. | N.A. | N.A. | 104 | 2 | 0.015 | 0.026 | N.A. |
| | | 7 nt / 1 MM | 99 | N.A. | N.A. | N.A. | 99 | 2 | 0.14 | 0.23 | N.A. |
| | | 7 nt | 97 | N.A. | N.A. | N.A. | 97 | 2 | 0.84 | 1.62 | N.A. |
| | | 6 nt \leftrightarrow 7 nt | 11 | N.A. | N.A. | N.A. | 11 | 2 | 0.017 | 1.524 | N.A. |
| | | 7 nt \leftrightarrow 7 nt / 1 MM | 102 | N.A. | N.A. | N.A. | 102 | 2 | 0.79 | 0.22 | N.A. |
| | | 7 nt / 1 MM \leftrightarrow 7 nt | 21 | N.A. | N.A. | N.A. | 21 | 2 | 0.15 | 1.57 | N.A. |
| | | | | | | | | | | | |
| Fig. | Name | Glycerol concentration (%) | # of Traces | Confidence | Frame time (ms) | # of Frames | # of Dynamic Tracers | # of States | Dwell time t_1 (s) | Dwell time t_2 (s) | Dwell time t_3 (s) |
| 6a-b | 3c origami 2 states 8 nt / 1 MM | 1 % | 1167 | 50 % | 52.2 | 2397 | 176 | 2 | 1.65 | 2.45 | N.A. |
| | | 2.5 % | 1087 | 30 % | 52.2 | 2397 | 150 | 2 | 1.52 | 2.23 | N.A. |
| | | 5 % | 1814 | 50 % | 52.2 | 2397 | 382 | 2 | 1.38 | 2.02 | N.A. |
| | | 10 % | 1040 | 50 % | 52.2 | 2397 | 120 | 2 | 1.11 | 1.52 | N.A. |
| | | 15 % | 1006 | 50 % | 52.2 | 2397 | 204 | 2 | 0.96 | 1.23 | N.A. |
| | | 20 % | 1295 | 50 % | 52.2 | 2397 | 101 | 2 | 0.85 | 1.11 | N.A. |
| | | 30 % | 1207 | 30 % | 52.2 | 2397 | 101 | 2 | 0.62 | 0.85 | N.A. |
| Fig. | Name | ADP concentration (M) | # of Traces | Confidence | Frame time (ms) | # of Frames | # of Dynamic Tracers | # of States | Dwell time t_1 (s) | Dwell time t_2 (s) | Dwell time t_3 (s) |
| 6e-f | Hsp70 SSC1 | 100 nM | 12319 | max | 32.5 | 2000 | 466 | 2 | 1.13 | 0.93 | N.A. |
| | | 5 μ M | 9246 | max | 32.5 | 2000 | 155 | 2 | 0.92 | 0.75 | N.A. |
| | | 100 μ M | 5035 | max | 32.5 | 2000 | 258 | 2 | 0.64 | 0.83 | N.A. |
| | | 1 mM | 3534 | max | 32.5 | 2000 | 156 | 2 | 0.58 | 0.79 | N.A. |

SUPPLEMENTARY REFERENCES

- 1 Abadi, M. *et al.* TensorFlow: Large-scale machine learning on heterogeneous systems. *ArXiv*, doi:10.48550/arXiv.1603.04467 (2015).
- 2 Tang, W. *et al.* Omni-Scale CNNs: a simple and effective kernel size configuration for time series classification. *ArXiv*, doi:10.48550/arXiv.2002.10061 (2022).
- 3 He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. *ArXiv*, doi:10.48550/arXiv.1512.03385 (2015).
- 4 He, K., Zhang, X., Ren, S. & Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile*, 1026-1034, doi:10.1109/ICCV.2015.123 (2015).
- 5 Fukushima, K. Cognitron: a self-organizing multilayered neural network. *Biol. Cybern.* **20**, 121-136, doi:10.1007/BF00342633 (1975).
- 6 Smith, S. L., Kindermans, P.-J., Ying, C. & Le, Q. V. Don't Decay the Learning Rate, Increase the Batch Size. *ArXiv*, doi:10.48550/arXiv.1711.00489 (2018).
- 7 Thomsen, J. *et al.* DeepFRET, a software for rapid and automated single-molecule FRET data classification using deep learning. *eLife* **9**, e60404, doi:10.7554/eLife.60404 (2020).
- 8 Schreiber, J. Pomegranate: fast and flexible probabilistic modeling in python. *J. Mach. Learn. Res.* **18**, 1-6, doi:10.48550/arxiv.1711.00137 (2018).
- 9 Basden, A. G., Haniff, C. A. & Mackay, C. D. Photon counting strategies with low light level CCDs. *Mon. Not. R. Astron. Soc.* **345**, 985-991, doi:10.1046/j.1365-8711.2003.07020.x (2003).
- 10 Hirsch, M., Wareham, R. J., Martin-Fernandez, M. L., Hobson, M. P. & Rolfe, D. J. A Stochastic Model for Electron Multiplication Charge-Coupled Devices – From Theory to Practice. *PLoS ONE* **8**, e53671, doi:10.1371/journal.pone.0053671 (2013).
- 11 Murphy, K. P. *Probabilistic machine learning: an introduction.* (MIT press, 2022).
- 12 Messer, P. K., Henss, A. K., Lamb, D. C. & Winterlin, J. A multiscale wavelet algorithm for atom tracking in STM movies. *New J. Phys.* **24**, 14, doi:10.1088/1367-2630/ac4ad5 (2022).
- 13 Wanninger, S. *et al.* Deep-Learning assisted, Single-molecule Imaging analysis of multi-color DNA Origami structures. *Zenodo*, doi:10.1101/2023.01.31.526220 (2023).
- 14 McKinney, S. A., Joo, C. & Ha, T. Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling. *Biophys. J.* **91**, 1941-1951, doi:10.1529/biophysj.106.082487 (2006).
- 15 Zarabi, N., Schluesche, P., Meisterernst, M., Börsch, M. & Lamb, D. C. Analyzing the Dynamics of Single TBP-DNA-NC2 Complexes Using Hidden Markov Models. *Biophys. J.* **115**, 2310-2326, doi:10.1016/j.bpj.2018.11.015 (2018).
- 16 Sikor, M., Mapa, K., von Voithenberg, L. V., Mokranjac, D. & Lamb, D. C. Real-time observation of the conformational dynamics of mitochondrial Hsp70 by spFRET. *The EMBO Journal* **32**, 1639-1649, doi:10.1038/emboj.2013.89 (2013).
- 17 Krause, S. *et al.* Graphene-on-glass preparation and cleaning methods characterized by single-molecule DNA origami fluorescent probes and Raman spectroscopy. *ACS Nano* **15**, 6430-6438, doi:10.1021/acsnano.0c08383 (2021).
- 18 Kamińska, I. *et al.* Graphene Energy Transfer for Single-Molecule Biophysics, Biosensing, and Super-Resolution Microscopy. *Adv. Mater.* **33**, 2101099, doi:10.1002/adma.202101099 (2021).

Paper 2.

Deep-LASI, single-molecule data analysis software¹¹²

Asadiatouei, P., Salem, C.B., Wanninger, S., Ploetz, E., Lamb, D. C.

Biophysical Journal, 123, 2682–2695, September 3, 2024

DOI: 10.1016/j.bpj.2024.02.013

Deep-LASI, single-molecule data analysis software

Pooyeh Asadiatouei,¹ Clemens-Bässem Salem,¹ Simon Wanninger,¹ Evelyn Ploetz,^{1,*} and Don C. Lamb^{1,*}

¹Department of Chemistry and Center for NanoScience (CeNS), Ludwig-Maximilians-Universität München, Munich, Germany

ABSTRACT By avoiding ensemble averaging, single-molecule methods provide novel means of extracting mechanistic insights into function of material and molecules at the nanoscale. However, one of the big limitations is the vast amount of data required for analyzing and extracting the desired information, which is time-consuming and user dependent. Here, we introduce Deep-LASI, a software suite for the manual and automatic analysis of single-molecule traces, interactions, and the underlying kinetics. The software can handle data from one-, two- and three-color fluorescence data, and was particularly designed for the analysis of two- and three-color single-molecule fluorescence resonance energy transfer experiments. The functionalities of the software include: the registration of multiple-channels, trace sorting and categorization, determination of the photobleaching steps, calculation of fluorescence resonance energy transfer correction factors, and kinetic analyses based on hidden Markov modeling or deep neural networks. After a kinetic analysis, the ensuing transition density plots are generated, which can be used for further quantification of the kinetic parameters of the system. Each step in the workflow can be performed manually or with the support of machine learning algorithms. Upon reading in the initial data set, it is also possible to perform the remaining analysis steps automatically without additional supervision. Hence, the time dedicated to the analysis of single-molecule experiments can be reduced from days/weeks to minutes. After a thorough description of the functionalities of the software, we also demonstrate the capabilities of the software via the analysis of a previously published dynamic three-color DNA origami structure fluctuating between three states. With the drastic time reduction in data analysis, new types of experiments become realistically possible that complement our currently available palette of methodologies for investigating the nanoworld.

SIGNIFICANCE Single-molecule experiments are very powerful but, at the same time, the analysis can be very time intensive. Here, we present a software that eases the analysis of single-molecule time traces. We have incorporated machine learning methods to support the data analysis. The software performs all steps required for such an analysis either manually or automatically starting from data extraction through to the final graphical outputs. Hence, the time investment needed for the analysis of single-molecule data can be reduced from days or even weeks to minutes.

INTRODUCTION

The ability to detect individual molecules has revolutionized the way we investigate the physical world. When measurements are no longer limited by ensemble averaging, sample heterogeneities, subpopulations and dynamic processes are directly observable. With such high sensitivity, a minimal amount of sample is necessary and, as the analysis is done one molecule at a time, high purification of the sample can be performed in the analysis (1). When performing measurements on immobilized molecules with methods such as

atomic force microscopy, optical and magnetic tweezers (2), or total internal reflection fluorescence (TIRF) microscopy, the dynamic processes of a single molecule can be observed as a function of time (3,4). As a result of such measurements, a huge number of trajectories are typically produced that need to be analyzed to extract the desired information from the system of interest.

Among the various valuable single-molecule techniques, Förster resonance energy transfer (FRET) experiments stand out as a noncontact method that can detect distances on the 2–10 nm scale and measure dynamics processes from nanoseconds to kiloseconds. Recent studies have shown that single-molecule FRET (smFRET) experiments are reproducible with an accuracy of 0.6 nm (5,6). With smFRET, it becomes possible to gain insights about the structural features and dynamics of materials, such as the structural fluctuations in biomolecules resulting from complex biological interactions (7). Notably, smFRET promises

Submitted November 30, 2023, and accepted for publication February 19, 2024.

*Correspondence: evelyn.ploetz@lmu.de or d.lamb@lmu.de

Pooyeh Asadiatouei, Clemens-Bässem Salem, and Simon Wanninger contributed equally to this work.

Editor: Robert Best.

<https://doi.org/10.1016/j.bpj.2024.02.013>

© 2024 Biophysical Society.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

to be an important method for the upcoming age of dynamic structural biology (8). For many experiments, it is possible to detect subpopulations and measure dynamics directly from the collected data. For a detailed quantitative analysis, there are additional steps that need to be performed. Here, it is useful to monitor the fluorescence of the acceptor directly, which can be done using alternating laser excitation (ALEX) (9). In ALEX, the donor and acceptor molecules are excited alternately. Hence, the photophysical state of the acceptor can be probed during the smFRET experiment, and correction factors for the determination of accurate FRET efficiencies can be extracted (10–12). When analyzing ALEX data, the excitation scheme needs to be determined and incorporated into the analysis. As experiments grow in complexity, the intricacy of the analysis increases as well, and the availability of more advanced analysis tools becomes increasingly important.

Numerous software packages have been developed to aid in the analysis of single-molecule measurements. The choice of the optimal analysis tool depends upon the specifics of the experimental system and analysis required for the study. For smFRET experiments, Lerner et al. recently summarized and published an extensive list of analysis tools that were released until the year 2021 including tools to analyze time trajectories from surface experiments (8). In a later study, Götz et al. compared the performance of 11 widely used smFRET analysis tools regarding the determination of kinetic models and extraction of the rate constants (13). With respect to smFRET experiments, the vast majority of software has been developed for two-color FRET experiments (14–19). Our group has also developed a MATLAB-based software for data analysis on two-color FRET systems for surface-immobilized molecules called Tracy (13,20).

Upon expanding our single-molecule TIRF setup to accommodate three-color FRET experiments, we needed to expand our analysis software. With this paper, we introduce our new software, Deep-LASI (deep learning-assisted single-molecule imaging analysis), an open-source software package using MATLAB (but also available as a runtime version) that incorporates Python and C++ routines. The Deep-LASI software offers both manual and automatic analysis environments for a wide range of one-, two- and three-color single-molecule experiments (21). The features of the software include mapping of multiple detection channels, extraction and background correction of one-, two- and three-color FRET data, trace classification and selection of relevant time points for the analyses, determination of the correction factors for the calculation of accurate FRET efficiencies, histogram generation of various parameters, and kinetic analyses using hidden Markov models (HMMs) and deep neural networks. Deep learning techniques are emerging in virtually all data-driven fields and are having a big impact in the life sciences, in particular in microscopy (22–27). Inspired by these developments, we incorporated deep learning to help in trace classification,

determination of the relevant regions of relevant traces, automated FRET correction and kinetic analyses.

Deep-LASI supports various data file formats with extendable support to read in new formats into the software. Although originally written for smFRET data, the software is adept at handling any data as long as it results in time traces. Furthermore, the software offers environments for simulating and training single-molecule time traces. We also provide example data sets and tutorials to help users quickly gain proficiency in using the software (28).

RESULTS AND DISCUSSION

Deep-LASI is a user-friendly software package with a high degree of automation and compatibility for the analysis of time-resolved single-molecule intensity traces. It is designed to help with the data analysis of one-, two- and three-color FRET experiments with interactive graphical user interfaces (GUIs) to provide enough freedom so that the user can extract the desired information based on their analysis needs. The source code is available such that the software can be adapted and further developed by expert users and software developers (29). A description of the implemented features is given in the following sections.

The information to be extracted from the intensity traces of single molecules and the necessary steps will vary depending on the measurement assay and question of interest. An overview of the most common procedures in single-molecule data analysis is summarized in Fig. 1. The main analysis steps include reading in the raw data, mapping the detection channels, (co-)localizing the particles and extracting the intensity information over the measurement time. In the next step, the software allows for classifying traces, determining usable regions within each time trace, plotting the distributions of the extracted parameters (such as FRET values, labeling stoichiometry or dwell times), calculating the necessary correction factors and performing a kinetic analysis in the case of a dynamic system. In the case of the kinetics analyses, transition density plots (TDPs) are automatically generated and provide access to the cumulative dwell-time distribution functions (CDFs).

Deep-LASI also offers the opportunity for expert users to simulate multicolor smFRET traces and to train neural networks for new single-molecule assays. To ensure flexibility and accessibility overall, the extracted and analyzed data can be saved and reloaded at any time, and can be additionally imported and exported, from and into standard data files. In the following section, we discuss the individual working steps and underlying mechanisms that define the software's functionality.

Main functionalities of Deep-LASI

Typically, the initial step in analyzing single-molecule experiments involves reading in the raw data. The most

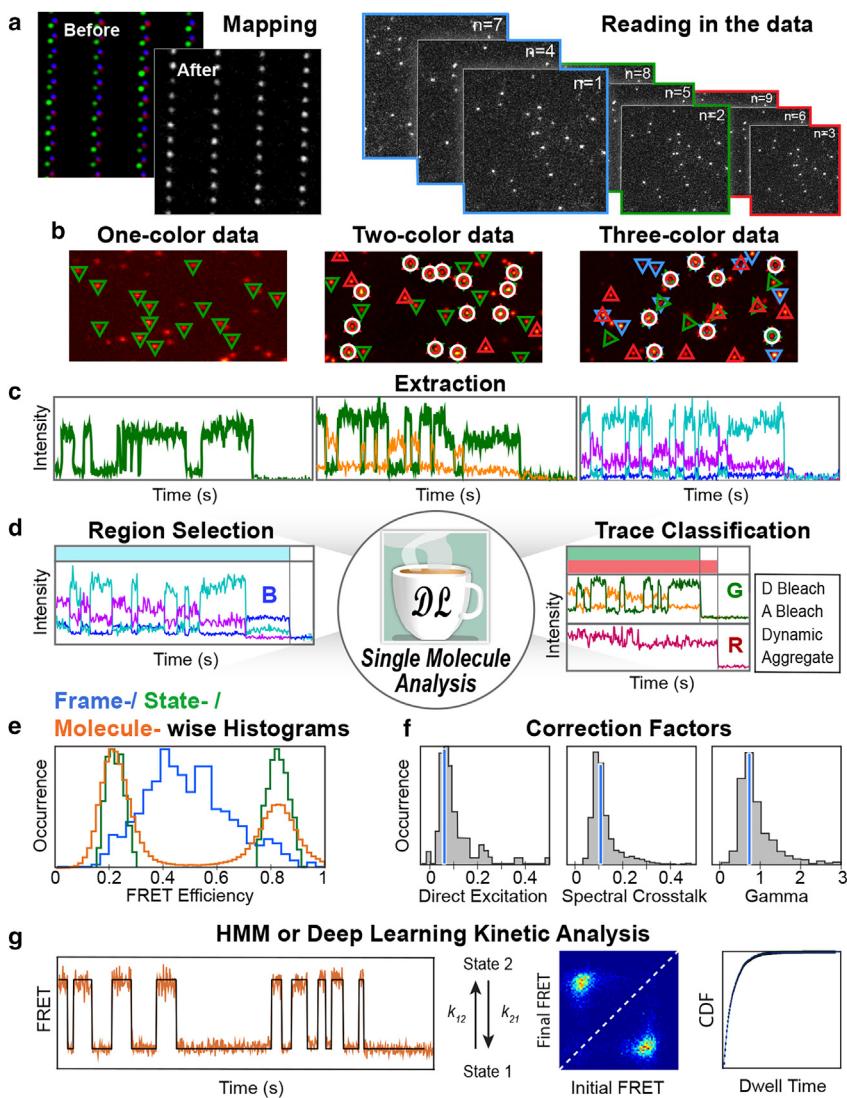


FIGURE 1 Overview of the functionalities of the Deep-LASI software package. The main applications of the software package are to extract, sort and analyze intensity traces from single-molecule data. This process involves a series of key steps: (a) for multicolor experiments, the different channels need to be registered to each other (i.e. mapped). Afterward, the raw data is read in for each channel from a stack of frames based on the excitation scheme. (b) Single molecules are localized and, when desired, colocalized across different channels based on the created map. (c) The intensity traces are extracted from each detected (and colocalized) particle and corrected for background. (d) The analysis of extracted intensity traces starts with trace classification and selection of the useful region of each channel where the corresponding fluorophores are active. (e) The results can then be visualized by the means of various histograms with frame-, state- and molecule-wise approaches. (f) Optionally, the method-specific correction factors are determined. (g) For dynamic traces, a kinetic analysis can be performed by hidden Markov modeling (HMM) or deep learning approaches. The panels show a typical state transition path inferred by HMM (i.e. Viterbi path) and transition density plots with state transition information and the cumulative dwell-time distribution function (CDF) determined by fitting, respectively. To see this figure in color, go online.

elaborate features of the software are designed to work with images or movies from cameras like EMCCD (electron-multiplying charge-coupled device) or sCMOS (scientific complementary metal oxide semiconductor) cameras. However, it also accommodates the direct read in of custom, nonimage data file formats encoding a time series (Fig. 2, blue boxes). For detailed information and the latest list of supported file formats, please refer to the online tutorial available for Deep-LASI on Read the Docs (28).

Given the diversity of fluorescence-based assays and methods resulting in time traces suitable for analysis through Deep-LASI, we focus on the main functionalities of the software. This includes importing data files, detecting and mapping molecules, extracting traces, calculating background, and manually or automatically sorting the collected data (Fig. 2, green boxes). Following these steps, the software provides a set of different tools to spectrally

correct, kinetically analyze and summarize the single-molecule data (Fig. 2; pink boxes). These tools can be used for 1) determining correction factors, 2) plotting representative properties of the results via their distributions (e.g. of apparent or accurate FRET values of single molecules, of states or frames), 3) allowing unsupervised, kinetic analysis of selected regions of the appropriate traces using HMM or deep learning algorithms and finally, 4) visualizing the data using TDPs and CDF plots. For an in-depth understanding of each feature, including the algorithms involved, the philosophy behind the GUIs, and tutorials featuring specific analysis examples, please refer to the comprehensive software documentation and manual (28).

Channel mapping

For multicolor experiments using separate detection channels, a registration of the different detection channels is

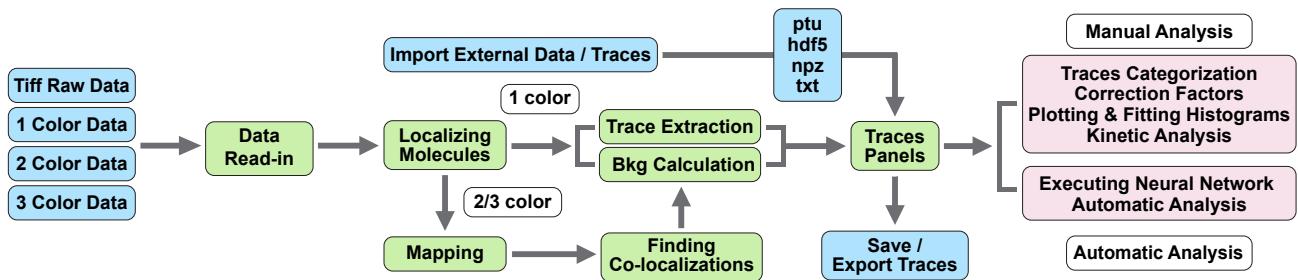


FIGURE 2 Schematic representation of the data-handling workflow using the Deep-LASI software. The blue boxes show the raw movie data and intensity trace files that can be loaded into the software as well as the saving routines for storage and export of the analyzed data. The green boxes show the initial steps for data extraction and preparation with the Deep-LASI software. The pink boxes summarize the tools for data analysis and representation of results available through both manual and automatic modalities. To see this figure in color, go online.

needed. When measuring at different wavelengths, perfect alignment between channels in terms of shift, magnification, and rotation can be challenging and time-consuming. Hence, a mapping process between cameras or regions of interest is required to ensure that the fluorescence signatures visible in the different channels originate from the same immobilized molecule. When performing fluorescence-based single-molecule experiments using imaging, the optimal pixel size is usually in the range of 40–100 nm. Thus, the fluorescence emitted by a single molecule spans multiple pixels, and alignment within a single pixel is sufficient. The mapping is performed using a set of emitters well distributed across the detectors' field of view. We commonly employ a zero-mode waveguide pattern or a surface covered with emitting or scattering particles, such as fluorescent beads. Alternatively, mapping can be accomplished using the actual single-molecule data. One channel is selected as a reference channel. Our software then utilizes a phase-correlation algorithm to estimate the geometric transformation necessary to align the other channels to the reference image (Fig. 3 a)(30). This geometric transformation involves scaling, rotation, and translation of the read-in images. Individual emitters are detected based on a user-selected threshold and their localizations are utilized to further refine the mapping using a 2D polynomial of order up to 3. The prerequisite for this refinement operation is the colocalization of individual particles within 2 pixels after application of the geometric transformation. Their positions are determined using a stationary wavelet algorithm with adjustable sensitivity (31). A transformation matrix is generated, which is then used to map the respective coordinates between channels. The mapping step corrects small misalignments between the cameras originating from tilts and shifts of cameras and different magnifications, as well as aberrations in the detection paths. Notably, the mapping function is only used to find the corresponding pixels in the various detection channels corresponding to the location in the reference channel. The actual single-molecule analysis is performed separately on the raw data. No mapping of the images via the transformation matrix is performed

except for inspection of the quality of the transformation matrix.

Loading imaging data collected using various excitation schemes

Once the detection channels have been mapped onto each other, the actual single-molecule data one wishes to analyze is loaded. Upon loading the data, the frames are segregated based on the excitation scheme used (when necessary). For accurate smFRET experiments using camera-based data acquisition, it is advantageous to use millisecond ALEX schemes (5,32). For two-color experiments, alternating frames are collected using donor and accepter excitation respectively. Acceptor excitation is used to probe the presence and photoactive state of the acceptor molecule enabling the calculation of labeling stoichiometry. However, frames with acceptor excitation have to be excluded when calculating the FRET efficiency. When expanding to three-color experiments, ALEX is essential for analyzing the data and three excitation lasers are alternated, respectively. There are also experiments where one wishes to detect the presence of one color at the beginning of the experiment, but then perform a smFRET experiment with different colors. One example would be measuring the conformation of DNA using smFRET in the absence or binding of a DNA-binding protein, which is labeled with a third color. In this case, the first few frames are used to detect the presence of the third color at the beginning of the measurement and used for selecting the traces that are to be analyzed. The remainder of the frames from the selected traces are then used to extract the smFRET information.

Typically, a series of consecutive measurements is performed using the same measurement parameters (excitation scheme, detection channels, exposure time, etc.) to gather sufficient statistics. This results in a collection of data files originating from each camera. To initiate the analysis of the entire experiment, the Deep-LASI read-in process begins by collectively selecting all files from a single camera at once, usually starting with the most blue-shifted detection channel. Next, the first movie of the chosen file set is loaded. Here, the user is prompted to define the frame range,

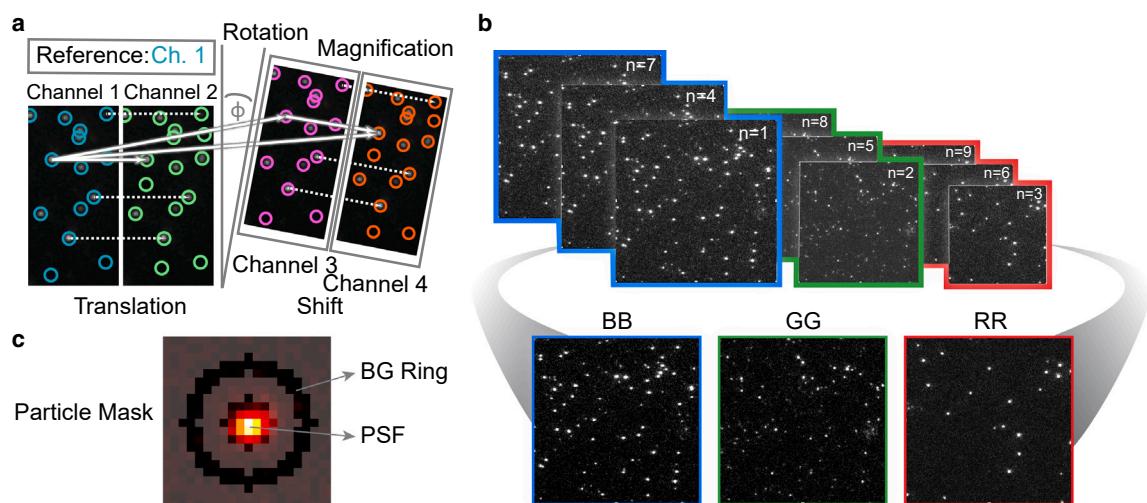


FIGURE 3 Particle detection, localization and mapping. (a) Mapping process with Deep-LASI for multicolor experiments: a reference channel is chosen and all other channels are registered with respect to the reference channel. Here, the chosen reference channel is the blue channel. (b) For trace extraction, Deep-LASI first opens the images registered on a specific camera and segregates the detection channels according to the excitation scheme. Here, we show the BB, GG, and RR channels. To identify the position of single emitters, we use the cumulative image taken from the direct excitation frames for the respective channels. (c) To calculate the measured intensity coming from a detected particle, a mask function is selected. We typically use a particle detection mask having a circular geometry of 7 pixels in diameter. The outer dark ring with a width of 2 pixels is used to determine the background contribution. To see this figure in color, go online.

excitation sequence, and detection channel. Next, a cumulative intensity is displayed over the user-selected range (Fig. 3 b) to improve the signal-to-noise ratio and facilitate a user-friendly, interactive parameterization for the trace extraction. The loading procedure is then repeated for the corresponding movies of the remaining channels. The particle detection method and threshold for each detection channel are then determined. In the last step, the extraction parameters are provided by the user: particle and background mask (Fig. 3 c), molecule selection criteria and the frame range used for extraction. Once the detection and extraction thresholds are established for all channels, the corresponding sets of files are sequentially loaded and single-molecule traces are extracted according to the given selection mode: Deep-LASI extracts the trajectories 1) for all detected molecules, 2) for colocalized molecules only, or 3) for molecules detected in a given detection channel.

Particle detection

To extract single-molecule trajectories, Deep-LASI provides three different techniques for single-molecule localization. For each technique, a sensitivity threshold is applied based on the normalized reconstruction from weighted wavelet coefficients or intensity values. Based on the selected threshold, a binary image is generated that encodes the detected particles. The position of each particle is determined by the center-of-mass of the pixels associated with the particle.

Wavelet. Wavelets are filters that can be applied to images (or time series) to enhance features with particular spatial (or temporal) frequencies. As the fluorescence signal coming from single molecules are diffraction limited, the de-

tected fluorescence should be symmetric with the size given by the point-spread function. By applying different wavelet filters, the original image is decomposed into a finite number of wavelets where particular spatial features are enhanced and others suppressed. More specifically, by mathematically applying low-pass and high-pass filters on the signal and repeating the procedure, a set of wavelet planes are generated at different resolutions (33–35). Based on the median absolute deviation of the wavelet coefficients for each plane, insignificant features are removed automatically.

Intensity thresholding. Another approach is to use intensity thresholding to detect molecules emitting intensities higher than a user-defined level. When enough adjacent pixels are above the threshold, the area would be considered a particle and the central point taken for trace extraction. This rather easy method works well as long as the signal and background are homogenous over the field-of-view and the signal from the molecules is sufficiently stronger than the background.

Regional maxima. An alternative method based on intensity thresholding is the regional maxima approach. With this method, a Gaussian filter of 9 pixels is first convoluted with the image and then the MATLAB function *imregionalmax* is used to find the local maxima. This is done by locating pixels where all eight neighboring pixels are lower in intensity. The routine then returns all the regional pixels which are a local maximum and are considered as the center of single emitting particles. Pixel intensities below the user-defined threshold are set to zero. These selection criteria have an advantage over normal intensity thresholding when

analyzing data with heterogeneous particle intensities and noise within a frame.

Trace extraction

From the binary image generated from the detected particles, the particle positions are extracted using the MATLAB in-built function *regionprops*. This calculates the centers-of-mass for connected pixels. Using the central position of the individual particles, the particle mask (Fig. 3 c) is then used to determine the total number of detected photons for the particle as well as the background contribution. Typically, we use a circular particle mask with a diameter of 7 pixels. The size is chosen to optimize collection of photons within the point-spread function of the molecule while minimizing the inclusion of additional pixels and hence potential overlaps between neighboring particles. The user can also adjust the particle mask settings based on their specific needs.

The particle positions are then linked in consecutive frames to generate time trajectories. To extract the intensity traces from each detected single emitter, frame-wise intensities for each channel are determined, and plotted over the whole measurement or selected frame range.

Background determination

The size and shape of the particle mask surrounding each particle's point-spread function (Fig. 3 c) and the method of background determination have a considerable impact on the signal-to-noise ratio, the quality of traces and finally on the resulting histograms. There are multiple approaches to background correction. Fortunately, the number of pixels that can be used to calculate the background intensity far outnumber the number of pixels within the point-spread function and hence can be subtracted with high accuracy. Deep-LASI extracts frame-wise intensities for each molecule detected in the various channel(s). To avoid any potential heterogeneity from the illumination profile, a nonconstant background level within a frame or differences between cameras, the background signal is calculated and subtracted from the accumulated intensity within each particle mask. At any time during the analysis, the user can view the raw intensity traces without background subtraction.

As the signal is averaged in the background mask, no molecules should be present in the region used for determining the background. For densely populated surfaces, the default mask can be adjusted, e.g., by reducing the radius of the mask. To decrease the uncertainty in the background estimation, the background is measured in approximately twice as many pixels as the signal. In addition, as the background does not typically change strongly with time, an 11-frame sliding window (± 5 frames) is used to average the background value. The average background signal (scaled to the number of pixels in the particle mask) is then subtracted from the total measured intensity. The total measured intensity and the local background are determined for each frame

and the background-corrected intensity traces stored. By visually checking the intensity level of a trace after photo-bleaching of all fluorophores, the quality of the background subtraction routine can be controlled. At this point, all the extracted traces from the experiment are saved into a single datafile with a filename adapted from the name of the first movie file with the extension of .tdat.

Trace read-in options

At this point in the analysis, one has extracted and saved the single-molecule time traces from one or more channels for a given excitation scheme. Here, it is also possible to reload the traces as well as to directly import intensity traces extracted using other software for any type of single-molecule time-series data. Several data importing options are incorporated including ptu, hdf5, npz, and txt files. For example, we have also used Deep-LASI to analyze single-molecule intensity traces collected one at a time on a confocal microscope (21). The txt file format is provided (13,28) such that users can convert their data into a format that can be read into Deep-LASI.

Analysis options

Deep-LASI offers diverse tools for analyzing and presenting information derived from single-molecule time traces, irrespective of the methods employed for data acquisition, ranging from one- to three-color measurements: the software facilitates both manual and automatic processes for 1) trace categorization into, for example, usable static and dynamic traces and 2) selection of specific regions within individual traces for further analysis. In addition, 3) Deep-LASI provides an overview of parameters characterizing selected regions in the intensity traces including brightness, background intensity, signal-to-noise ratio, and photo-bleaching time. Beyond these basic functionalities, Deep-LASI supports manual and automatic analyses of one-, two- and three-color FRET assays. Moreover, the software enables 4) extraction of kinetic information from dynamic traces. Two distinct approaches are available for kinetics analyses: the first involves conventional HMM with selectable algorithms for up to three channels, as detailed below (36,37). The second approach employs neural networks for automated data analysis wherein Deep-LASI outputs a confidence level of the time trajectory being in a specific state for each frame. In addition to kinetic analyses, Deep-LASI allows for 5) the calculation of accurate FRET efficiencies by extracting the necessary FRET correction factors from the data. Finally, the software provides 6) state-of-the-art tools for summarizing the FRET states and kinetics extracted during the analysis. These include histograms illustrating distributions of, e.g., FRET efficiencies (apparent and accurate FRET) of static and dynamic traces, stoichiometry, or FRET correction factors. Furthermore, Deep-LASI provides TDPs and CDFs for summarizing the kinetics information found in the single-molecule data.

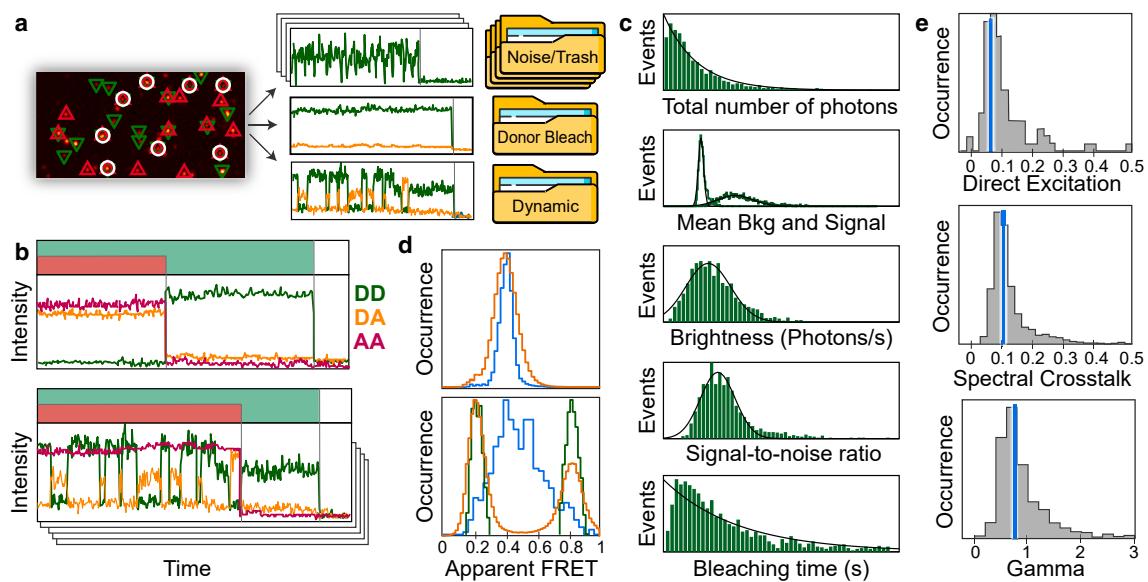


FIGURE 4 Single-molecule trace analysis of 2c FRET data with the Deep-LASI software. (a) The analysis of extracted traces begins with categorization based on the 1) underlying single-molecule assay, 2) numbers of dyes, 3) trace quality, 4) photochemistry involved, and 5) dynamic information available. (b) Two exemplary intensity traces are shown for two-color smFRET measurements using the ALEX scheme. (Top) A molecule exhibiting static FRET and (bottom) a dynamic molecule. In both cases, the acceptor molecule photobleaches first leaving fluorescence only from the donor molecule until it photobleaches at the end of the traces. The donor signal (green), sensitized acceptor emission (orange), and direct acceptor excitation (red) are shown. Shadowed regions in green and red mark the selected regions in each detection channel to be considered in further analyses. (c) Deep-LASI offers an overview of information from selected regions of the single-molecule traces. In addition to FRET, it displays molecule-wise histograms of signal (Sig) and background intensities (Bkg), brightness (Hz), signal-to-noise ratio, and photobleaching time (s) with the corresponding fits. The histograms shown are for green detection after green excitation. (d) SmFRET histograms of (top) a static sample and (bottom) a dynamic sample interconverting between two states characterized by low and high FRET efficiencies. The apparent FRET efficiency of the sample is depicted frame-wise (orange), molecule-wise (blue), and state-wise (bottom panel, dark green). (e) Distributions of the FRET correction factors are shown, which are used to correct for direct excitation, spectral crosstalk, and differences in detection sensitivity (gamma) for a simulated two-color smFRET sample. The median values of each histogram are indicated as a blue line. To see this figure in color, go online.

The subsequent sections provide a brief introduction to manually using Deep-LASI for categorizing single-molecule traces, selecting regions, and analyzing static experimental parameters. Subsequently, we discuss how to obtain accurate FRET measurements and extract kinetic information from single-molecule data.

Trace categorization and static analysis

After extracting or loading single-molecule traces, the next step involves the categorization and sorting of the molecules. In a typical single-molecule experiment, the data set can easily comprise several thousands of traces. Many of the traces may be noninformative due to rapid photobleaching, the presence of aggregates, incomplete labeling, or inadequate signal to noise ratios. Hence, the primary objective in trace categorization is to select the suitable regions of appropriate traces for further analysis. This starts by separating out traces that are unsuitable. For this, Deep-LASI provides dedicated panels and GUIs for systematically reviewing and categorizing traces (see Fig. 4 a). Typical categories include “static”, “dynamic” and “trash” although users have the flexibility to add custom categories as needed for their experiment. Furthermore, Deep-LASI facilitates the sorting of traces based on the number of photoactive flu-

orphores by considering which fluorophores are active in each frame. For two-color FRET assays, for instance, traces can be sorted into categories like “donor bleach” and “acceptor bleach” which proves instrumental in determining FRET correction factors at a later stage. Notably, users have the flexibility to assign multiple categories to individual traces, allowing classifications such as “static” and “acceptor bleach” simultaneously. This functionality becomes particularly advantageous in three-color FRET experiments where additional statistics for FRET correction factors can be obtained from analyzing constructs that contain only two of the three fluorescent dyes.

In the second step of the characterization procedure, it is necessary to mark the regions of the useful traces to be included in further analyses (Fig. 4 b). Selection of the desired regions is possible with an activated cursor on the intensity trace panels. The selection can be general to define the regions in all channels to be included in the final histograms or kinetic analyses, or can be specific to each detection channel (shaded in the corresponding color) for determining individual photobleaching steps and regions to be used for the calculation of FRET correction parameters later on. Once correction factors have been estimated, users can choose to visualize the data at the level of apparent or

accurate FRET. Correction factors are used to account for donor leakage into the acceptor channel, direct excitation of the acceptor, and differences in detection sensitivity of the donor and acceptor molecules. In addition, the individual intensity traces can be displayed with or without background correction.

Statistical overview of selected traces

The Deep-LASI software offers the possibility to evaluate and visualize the characteristics of selected frames, traces and, ultimately, the analytical results. From the first interface, the fluorescence properties of the different fluorophores can be assessed (Fig. 4 c). For EMCCD cameras, the characteristics of the amplification can be included to convert the camera counts into approximate photon numbers (otherwise, the signal in camera counts will be plotted). These distributions showcase the total signal until photobleaching (number of photons), the total signal and mean background per molecule (in kHz), the background-corrected brightness for the corresponding channels (in Hz), the individual signal to noise ratios, and the time until photobleaching of the respective fluorophores (in s). The histograms for each channel are automatically fitted to mono-exponential or Gaussian functions. The fit results are given in the respective panels.

smFRET analysis

In our research group, we specialize in smFRET experiments and evaluation. Hence, parts of the software are optimized for smFRET analysis from experiments carried out on immobilized molecules. With FRET, it is possible to investigate structural properties or dynamics due to FRET's strong dependency on the distance between fluorophores (8). From the selected regions of the corresponding molecules, it is possible to calculate the apparent FRET efficiency histograms, that is the FRET efficiency determined from background-corrected intensities without any further corrections. These can be plotted for each frame and molecule (frame-wise) or averaged value determined individually for each molecule (molecule-wise) (Fig. 4 d). Frame-wise FRET histograms contain all FRET values obtained across different molecules and frames, giving a comprehensive projection of accessible FRET states in the sample from all selected molecules (Fig. 4 d, orange line). Alternatively, the molecule-wise (or trace-wise) histogram reports an average FRET value for each single molecule over the selected frame range (Fig. 4 d, blue line). Notably, for static samples, molecule-wise and frame-wise histograms will coincide whereas, for dynamic molecules, they will not. To overcome this, it is possible to plot histograms state-wise when analyzing dynamic traces (see dynamic analysis below; Fig. 4 d, lower panel, green line).

To capitalize on the ability of FRET to measure distances accurately on the subnanometer regime, it is necessary to correct the apparent FRET efficiency for direct excitation

of the acceptor, spectral crosstalk of the donor fluorophore into the acceptor channel and variations in detection sensitivity to the various fluorophores. Depending on which molecule photobleaches first, it is possible to determine a subset of the correction factors directly from the individual traces. In the case where the donor undergoes photobleaching before the acceptor, the software calculates the direct excitation correction factor using the residual emission of the acceptor directly excited by the donor laser excitation (Fig. 4 e, top panel). Conversely, if the acceptor photobleaches before the donor molecule, the spectral crosstalk correction factor is determined as the residual donor emission detected in the acceptor channel (Fig. 4 e, middle panel). After correcting the trace for direct excitation and spectral crosstalk, the same trace can be used for determination of the detection efficiency correction factor from the ratio of the changes in acceptor and donor intensity after the acceptor's photobleaching step (Fig. 4 e, bottom panel). Once all individual traces are assessed for possible contributions to the correction factors, the distribution is plotted and the software computes the average, median, and mode of the distribution for each correction factor and dye pair. For an accurate estimate of the various correction factors, a minimum number of continuous frames after the photobleaching step should be included (we use a minimum of 20 frames). To kick out spurious values from the distributions, a maximum tolerable value for all correction factors can be entered. Values above the maximum will not be included in the calculation of the average, median and mean. The correction factors that cannot be determined directly from the traces are taken from the distribution (referred to as global correction factors). We typically use the median of the distribution as it was found to be most robust given typical statistics, but the average or mode can also be selected (Fig. 4 e). The user also has the option to use the global correction factors for all traces or to enter the values individually for each trace. Once the correction factors are determined, accurate FRET values as well as distances can also be displayed. These together with additional parameters such as stoichiometry and FRET efficiencies (both accurate and apparent FRET) can be viewed in a second interface. All histograms can be normalized and/or fit to a wide variety of functions.

Machine learning analysis of dynamic trajectories

To analyze dynamic samples, additional functionalities are available in the Deep-LASI software. One can choose from two HMM analyses (Murphy (37) or Schreiber (36)) or automatically via deep neural networks (DNNs) (21) (Fig. 5 a). The results provide an estimation of the underlying states and kinetics within the individual trajectories. Hence, in the end, one generates a "digitalized" version of the state pathway, which allows determination of the transition rates via the CDFs.

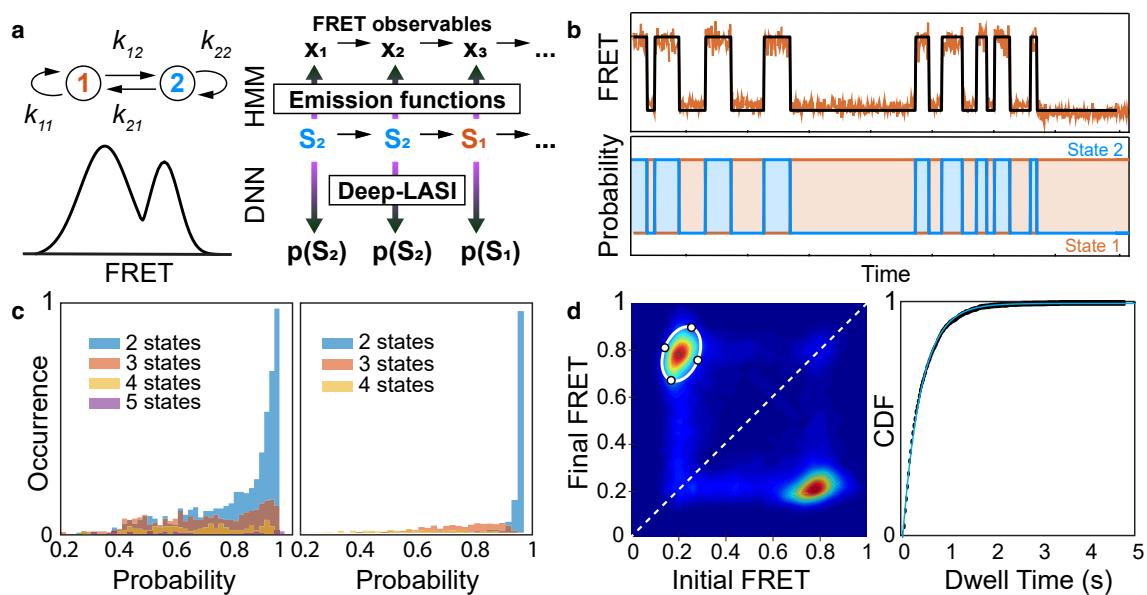


FIGURE 5 Kinetic analysis with the Deep-LASI software. (a) Left: schematic representation of a two-state dynamic FRET system with the transition rates (k_{ij}) between the two states. Right: the software extracts the dynamic interconversion rates between the states through an HMM analysis or via state probabilities derived using a deep learning algorithm based either on the measured FRET values or on the intensities directly. (b) Results of a kinetic analysis for an exemplary two-state smFRET trace. Top: the FRET trace plotted along with the most probable state path (black) generated through an HMM analysis (i.e., the Viterbi path) or via the neural network and (bottom) the corresponding state probabilities. (c) For an automated kinetic analysis, the neural network identifies the most likely number of detected states in each individual trace (left panel). The state path is then determined based on the number of detected states and a histogram of the average confidence of the assigned states in each trace is given (right panel). (d) Left: to extract the underlying dwell times and kinetic rates, TDPs are generated. By selecting individual clusters in the TDP, which summarize transitions of dynamic molecules from an initial to a final FRET state, the dwell-time distribution for the cluster is calculated. Right: the corresponding CDF for the selected population shown on the left. To see this figure in color, go online.

HMM can be performed on one-, two- and three-color data. For each data type, the FRET efficiencies or the intensities can be used as input. The number of states, mean values, standard deviations and the transition matrix can be initialized either using prior knowledge of the user, random uniform distributions or estimations based on k-means clustering. Other adjustable model parameters include the convergence threshold, the maximum number of iterations, and the choice between local or global HMM. Local HMM creates a new model for each trace, whereas a global HMM utilizes one model (rates and states) for all traces of a selected category. Like other analysis tabs, this analysis can also be exploited on any desired category(s). The states and kinetics of one- and two-color data are straightforward to model as they inherently represent distances in only one dimension. In three-color FRET experiments, distances can be extracted in three dimensions by combining the FRET efficiencies of all fluorophore pairs. However, three-color FRET is complicated by the strong interdependence of the FRET efficiencies and the numerous correction terms that are necessary to convert the apparent FRET efficiencies to actual distances. These corrections introduce significant uncertainties, making it difficult to properly model the system using HMM and identifying state transition. Therefore, the software focuses on using apparent FRET efficiencies and uncorrected intensities to accurately

analyze three-color FRET kinetics, treating the states of a given molecule as unique combinations of FRET efficiencies or intensities (38). This is achieved by employing multivariate HMM, where each trace is 2D and each observation is a multivariate vector. In the case of three-color FRET, the multivariate vectors can contain either the three FRET efficiencies or the five intensity channels that are relevant for determining the kinetic information. Direct excitation of the last acceptor is excluded as it provides no kinetic information. Regardless of the number of colors, the software provides the option to use HMM on traces that were manually selected or classified by a DNN.

After running the HMM, multiple corresponding panels with the number of states, state values and transition probability matrices are updated. These will depend upon the executed mode, i.e., local or global. Fig. 5 a shows an example of a dynamic, two-state system with independent transitions between these states. On the trace panel of the HMM tab, the individual traces and their corresponding Viterbi path are shown (Fig. 5 b, top panel). One can click through all traces present in the selected category to check the accuracy of the predicted states and transitions sequences.

DNNs can also be used for the kinetic analysis after the state classification step has been performed (e.g., Fig. 5 b, bottom panel). Here, there is the option to run a "number

of states" classifier to determine the predicted number of states observed in each trace (Fig. 5 *c, left panel*). The user can then run a particular state-classifier model (i.e., for two states, three states, or four states) on the selected data or use the output of the number of states classifier to automatically use the corresponding state-classifier model for state assignment on a trace-by-trace basis. The software plots the average confidence level for the state assignment over the individual traces (Fig. 5 *c, right panel*). For the example shown here, the software is very confident regarding the existence of two states and their corresponding state assignments.

After running the machine learning approach of choice (i.e., HMM or DNN), a digitalized state pathway is generated for each trace. This allows one to calculate a state-wise distribution from the state trajectories (Fig. 4 *d, bottom panel, green line*). The state-wise trajectories can be plotted normalized to the number of transitions or weighted by the number of frames contributing to each state. To analyze the underlying dynamics, the Deep-LASI software utilizes TDPs to visualize the detected transitions between initial and final states within the data (Fig. 5 *d, left panel*). From the TDP, the number of states, their corresponding values (e.g. FRET efficiencies), their connectivity and the number of transitions between different pairs of states can be revealed. Transition rates can be obtained directly from the output of a global HMM analysis. Alternatively, they can be calculated by selecting individual populations in the TDP and then fitting the corresponding dwell-time distribution (Fig. 5 *d, right panel*).

Automatic analysis

The analysis of single-molecule data, especially for molecules immobilized on a surface, usually takes days or weeks, even for a single day of measurement, and is prone to bias from the person analyzing the data. In addition, due to the time necessary for obtaining a reasonable amount of statistics, the parameter space that can be analyzed via such experiments is limited. To overcome such issues, automatic analysis tools using trained DNNs are available in the Deep-LASI software. Hence, each step of the analysis workflow described for manual evaluation above can be performed automatically. DNNs are available for trace classification (with region selection), number of state determination, and state trajectory analyses, and can be applied individually. In addition, the entire workflow from sorting and categorizing time traces, determining the photobleaching steps, calculating method-specific correction factors and state pathway determination can also be done automatically with the click of a single button (21). After running the automated analysis, the workflow continues with evaluation of the TDPs by selecting clusters and fitting their dwell-time distributions. On typical smFRET measurements with 4000 frames, the automated analysis is performed

within 20–100 ms per intensity trace and has been implemented for one-, two- and three-color data.

Currently, we have tested the Deep-LASI software on DNA origami structures as well as protein systems (21). However, the number of possible sorting categories and traces characteristics that the DNN has been trained on is not exhaustive. Therefore, for advanced users, the Deep-LASI software has the option to simulate one-color single-molecule data as well as two- and three-color smFRET traces. In addition, the simulated data can also be used to train new neural networks, if desired.

Additional attributes of Deep-LASI

The motivation for designing and publishing Deep-LASI was manifold. We wanted to develop a software package for the community that 1) is easy to use for everyone independent of their scientific maturity or disciplinary background, 2) contains a high degree of automatization with respect to data extraction and analysis to save time and remove user bias, 3) is compatible with other single-molecule methods and setups with temporal resolution and 4) provides state-of-the-art tools for analyzing single-molecule trajectories. In addition, the software should provide 5) advanced tools for analyzing experimental data from multicolor FRET experiment up to three colors (and potentially up to four in the future). These include the ability to simulate single-molecule data and train new machine learning approaches (HMM and neural networks) that are extendable for future single-molecule assays. Here, we summarize these general aspects of the developed software suite Deep-LASI.

Ease-of use/user friendliness

To make the software easily accessible to a broad range of users and establish a universal analysis environment, Deep-LASI was designed to be easy to learn with a clear workflow. Each step in the workflow is accompanied with its own GUI guiding from reading-in the raw data over data extraction and analysis to visualization of the results. After each step of data processing and analysis, it is possible to save the current status of the project, giving the user complete freedom to stop the analysis and resume at a later point in time. The resulting parameters and plots can be exported to external tools for presentation or publication purposes.

Automation

Gathering sufficient statistics in single-molecule experiments requires collecting data from thousands of single molecules. Hence, the Deep-LASI software can read-in hundreds of consecutive movies and extract the single-molecule traces automatically after the conditions have been determined for the first movie. As data extraction and evaluation can be performed without supervision, new analyses and experiments are now realizable. For example, it is now possible to utilize more of the collected information such

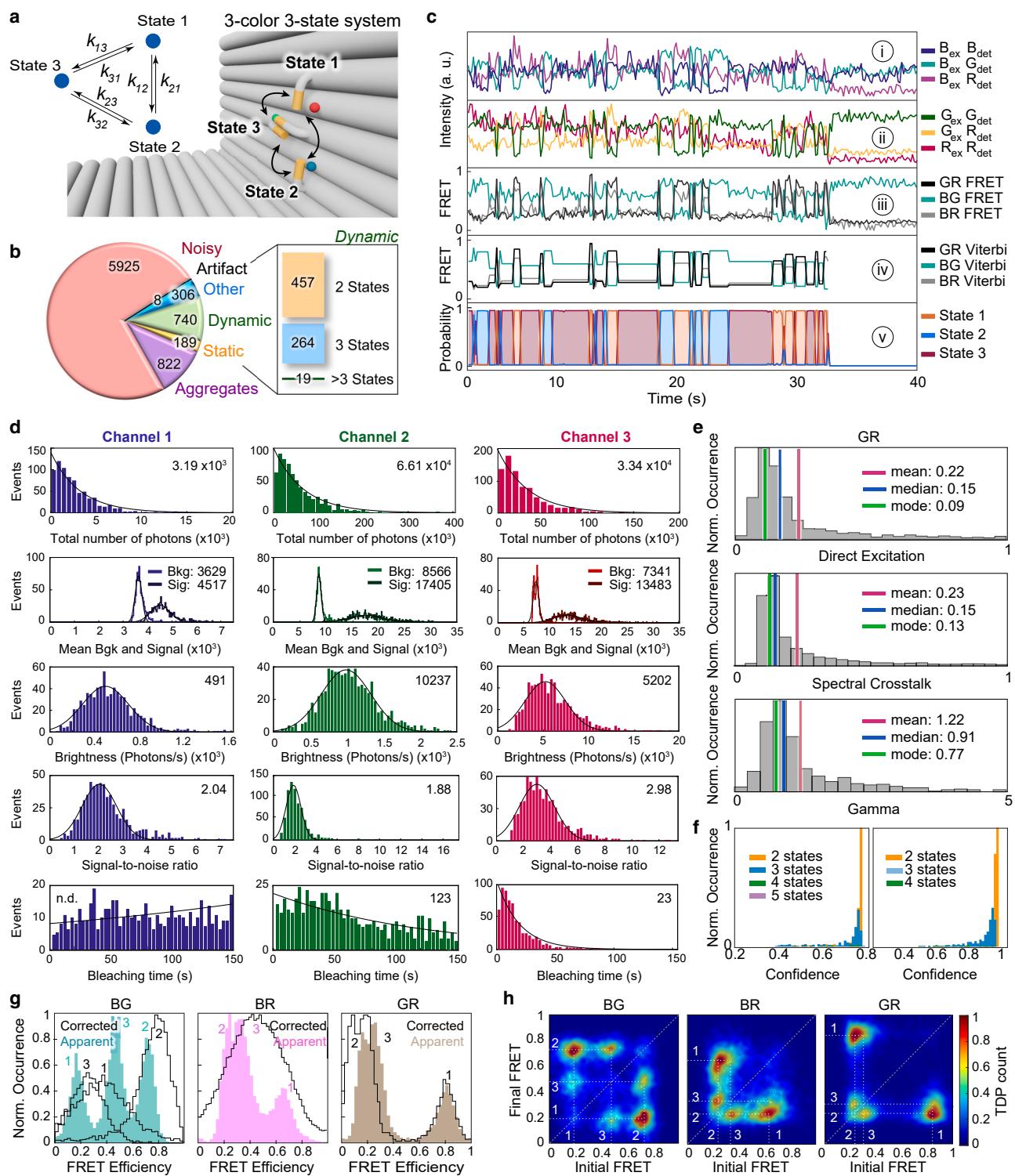


FIGURE 6 Application of the Deep-LASI software for analyzing three-color, three-state smFRET data. (a) Schematic overview of an L-shaped DNA origami structure with three binding positions, which are defined as 6, 9 and 12 o'clock. To probe the movement of the tether (labeled with Cy3B), Atto488 and Atto647N are placed near two of the binding sites (6 and 12 o'clock, respectively). The inset shows the corresponding kinetic scheme. (b) A pie chart showing the distribution of the 7990 traces that have been sorted into the given categories (Static, Dynamic, Noisy, Aggregate, Artifact, and Other). The 740 dynamic traces have been further categorized based on the number of states observed on each dynamic trace. (c) An exemplary single-molecule trace showing the six intensity and three apparent FRET traces for the three-state system. (i) The intensity traces detected in the blue, yellow, and red channels after blue excitation. (ii) The intensity traces after yellow and red excitation. (iii) The corresponding FRET efficiencies for each of the three dye pairs followed by the Viterbi path (iv) and state probabilities (v) derived by three-color HMM and a deep-learning-assisted kinetic analysis respectively. (d) Statistical analysis of photon counts, signal-to-noise ratio, and bleaching time for Channel 1, Channel 2, and Channel 3. (e) Histograms of occurrence for Direct Excitation, Spectral Crosstalk, and Gamma. (f) Confidence distribution for 2, 3, 4, and 5 states. (g) FRET efficiency histograms for BG, BR, and GR. (h) Heatmaps of FRET efficiency distributions.

(legend continued on next page)

as also analyzing partially labeled molecules or performing a series of measurements as a function of experimental conditions that would be otherwise unthinkable when performing a manual analysis.

Compatibility

Although written for the analysis of smFRET data, Deep-LASI is applicable to all single-molecule experiments that end up measuring time traces independent of the setup and raw data file format. The software is compatible with ALEX excitation schemes as well as a non-ALEX excitation for two-color experiments and can read in diverse file formats such as tif, ptu, hdf5, and npz. It is also possible to import other file formats by converting them into a pre-given txt format (13,29). Notably, for measurements with significantly different noise characteristics than those encountered in fluorescence-based methods, the currently incorporated DNNs may need to be retrained.

Adaptability

The development of single-molecule methodologies is a quickly advancing field, and the analysis needs are exceedingly dynamic and often specific for each single project. Accordingly, it should be possible to easily modify and adapt the analysis approaches. Therefore, the source code for Deep-LASI is freely available on the GitLab platform where active feedback and comments can be given (e.g., reporting bugs and suggesting improvements) and users can contribute new functionalities. This is not only possible, but also appreciated.

Unique methods

Deep-LASI combines an easy-to-use manual trace analysis software with state-of-the-art DNNs for automated data processing of one-, two- and three-color data. Many researchers are exploiting the advantages of smFRET but often avoid three-color assays due to the intrinsic complexity of the experiments and time-consuming analysis. Thereby, they also miss out on additional information that can be extracted by adding another fluorophore to the system. In addition, the only other software we are aware of for manually analyzing multicolor smFRET traces is SMACKS (18,39). Unique features of Deep-LASI are the automated analysis of state pathways in kinetic data and its ability to automati-

cally analyze three-color smFRET data and extract absolute distances.

Application of Deep-LASI software on experimental data

Having introduced the Deep-LASI software, we now showcase its application to real three-color experimental data. For this, we show the results from experiments on a previously published L-shaped DNA origami structure (21). The L-shaped origami structure contains a flexible tether that can bind to three distinct states (Fig. 6 a). The tether is labeled with Cy3B at the 3' end and can bind to protruding strands placed at positions referred to as 6, 9 and 12 o'clock. The binding sites consist of complementary ssDNA strands of 7 nt length at 6 and 12 o'clock and 7.5-nt length (i.e. a strand with 8 nt containing a 1 bp mismatch) at 9 o'clock. Binding of the tether occurs by spontaneous base-pairing with the single-stranded protruding strands. To monitor the movement, we introduced two additional fluorophores, Atto488 and Atto647N, on the structure close to the complementary strands at 6 and 12 o'clock positions, respectively.

Data were collected using a three-color ALEX scheme of blue, green and red excitation, and collected on three separate EMCCD cameras. Approximately 8000 traces containing all three fluorophores were extracted from 100 movies. Using the automated neural network analysis, a trace classification was performed. The distribution of classes is shown in Fig. 6 b. Of the 7990 traces extracted from the data, a dynamic classification was most probable for 740 (or 9%) of the traces, which were then utilized for further analyses. Fig. 6 c shows a representative single-molecule dynamic trace alongside the apparent FRET efficiencies and kinetic analysis. From the selected traces, the software provides an overview of various parameters. Here, we show the total number of photons, the signal and background intensities, the brightness, the signal-to-noise ratio for the BB, GG and RR channels, as well as the time until the photobleaching of the corresponding fluorophore (Fig. 6 d). When the measurement is of sufficient quality, one can proceed with the analysis. Next, we analyzed here the regions of the traces selected for determination of the FRET correction factors for spectral crosstalk and direct excitation as well as differences in detection sensitivity (5,6,8) (exemplified in Fig. 4 e for the GR dye pair). These are necessary for determining

Photostatistics of the three individual detection channels. The histograms are automatically fit to a Gaussian or exponential functions and the fit parameter reported. Histograms represent the total number of detected photons before photobleaching, the total signal and background (normalized to the number of pixels in the particle mask), the background-corrected molecular brightness, the signal-to-noise ratio and the time until photobleaching of the respective fluorophore. (e) Distributions of the FRET correction factors extracted from the data for direct excitation, spectral crosstalk and differences in detection sensitivity (gamma) for the GR FRET pair. The mean, median and mode values for each histogram are indicated in red, blue and green, respectively. (f) DNN confidence output for the number of state classifier (*left*) and the state assignment classifier (*right*). (g) Apparent frame-wise FRET efficiency histograms are shown as shadowed plots for each dye pair with the three efficiency populations for the BG, BR, and GR pairs. The frame-wise weighted, state-wise corrected FRET histograms are outlined in black. (h) TDPs are shown illustrating the transitions detected for all three dye pairs using apparent state-wise FRET efficiencies. (a) Is reproduced under the terms of the Creative Common CC-BY 4.0 License (21). Copyright 2023, The Authors. Published by Springer Nature. To see this figure in color, go online.

accurate FRET efficiencies. As neural networks also classify the active fluorophores in each frame, it is possible to extract the maximum number of frames in the data that can be utilized for the individual factors. In addition, for three-color samples, it is also possible to utilize information from the dual-color-labeled complexes. Hence, the automated analysis is often more accurate than the corresponding manually analyzed traces. Traces that exhibit dynamics can then be further analyzed using either HMM or DNNs. **Fig. 6 c(iv–v)** depicts the Viterbi path of FRET efficiencies and state-probabilities returned from the DNN for the representative smFRET trace. As discussed above, the neural network also generates confidence level histograms (**Fig. 6 e**). The left histogram indicates a large number of traces containing dynamics between two of the states and a significant fraction of traces displaying all three expected states. The right histogram depicts the average confidence of the neural network in assigning the corresponding frames to the two and three states within each trace.

After completing the analysis, the results can be summarized using various tools. For the smFRET data shown here, we compare the frame-wise apparent FRET efficiency histograms (*colored*) and state-wise corrected FRET efficiency histograms (*black lines*) for each dye pair (**Fig. 6 f**). For the BG dye pair, three FRET states are well resolved with apparent efficiency values of 0.18, 0.73, and 0.48, corresponding to states 1, 2, and 3, respectively. As expected from the design and associated Förster radius values of 53 and 65 Å (40–42), the 9 and 12 o'clock positions are more difficult to discern for the BR and GR FRET pairs. Nevertheless, the GR shows three populations with 0.83, 0.22, and 0.30 in apparent FRET efficiency for the same states 1–3. Although applying FRET correction factors in 3cFRET experiments usually results in broadened FRET histograms, the neural network correctly extracts the designed state-wise averaged accurate FRET histograms. Having the additional dimensions available in 3cFRET, it is possible to distinguish the three states. Interestingly, as expected, the three FRET populations in BR converge into a single FRET state at a FRET efficiency of 0.44 upon correction.

Finally, the Deep-LASI software enables the creation of TDPs in separate windows for each dye pair (**Fig. 6 g**). The user can then manually select populations in the plot and fit the resulting dwell-time distributions to determine the underlying transition rates between the corresponding states. By fitting the dwell-time curves (**Fig. 6 h**), we identified state residency times of 0.65, 0.69 and 1.40 s for states 1, 2 and 3, respectively. As the same state trajectories exist for each dye-pair, the state transitions can be selected from the TDPs that provide the best contrast between the two corresponding states.

CONCLUSIONS

In conclusion, Deep-LASI is a software suite that allows for a straightforward and rapid analysis of single-molecule time trajectories. The software supports multiple data formats

from a variety of methods such as wide-field and confocal measurements. It allows for the convenient analysis of single-molecule data starting with multiple-channel registration, trace sorting and categorization, determination of the photobleaching steps, calculation of FRET correction factors, and kinetic analyses based on HMM or DNNs. Each step can be performed manually or automatically. It offers advanced functionalities for handling and interpreting single-molecule data in one, two and three colors such as the quantitative analysis of three-color smFRET data. By introducing Deep-LASI, we encourage researchers to exploit the capacities of single-molecule techniques without being concerned about the software environment or complicated, time-consuming analysis steps. As the field develops, the analysis requirements will change. Hence, the software is open source, inviting programming experts to extend the capabilities of Deep-LASI to address the expanding analysis needs of a rapidly growing research field.

DATA AND CODE AVAILABILITY

The software is publicly available as source code, requiring MATLAB or as a precompiled, standalone distribution for Windows or Mac OS, and accessible from a GitLab repository at: <https://gitlab.com/simon71/deeplasi>. A detailed manual can be found at: <https://deep-lasi-tutorial.readthedocs.io/en/latest/documentation.html>. The experimental data are provided in the Zenodo database (Zenodo: <https://zenodo.org/record/7561162>).

AUTHOR CONTRIBUTIONS

S.W. developed and implemented the deep learning algorithm, Deep-LASI, in the GUI-based image processing, data extraction and manual analysis software written by C.-B.S. P.A. collected the single-molecule TIRF data. S.W. and P.A. performed the deep-learning-assisted analyses. P.A. wrote the first draft of the manuscript and created the figures. All authors contributed to revising the manuscript. E.P. and D.C.L. supervised the project.

ACKNOWLEDGMENTS

We thank Philipp Messer for valuable input and support regarding implementation of wavelets in the software for particle detection. We thankfully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—project ID 201269156—SFB 1032 Project B03 (to D.C.L.) and individual grants PL696/4-1 (to E.P.). D.C.L. gratefully acknowledges funding from the Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the Federal Government and the Länder through the ONE MUNICH Project Munich Multiscale Biofabrication. D.C.L. and E.P. gratefully acknowledge the financial support of the Ludwig-Maximilians-Universität München via the Department of Chemistry, the Center for NanoScience (CeNS), and the LMUinnovativ program BioImaging Network (BIN).

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Ha, T., T. Enderle, ..., S. Weiss. 1996. Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl. Acad. Sci. USA.* 93:6264–6268.
- Petrosyan, R., A. Narayan, and M. T. Woodside. 2021. Single-Molecule Force Spectroscopy of Protein Folding. *J. Mol. Biol.* 433, 167207.
- Colson, L., Y. Kwon, ..., Y. Cho. 2023. Trends in Single-Molecule Total Internal Reflection Fluorescence Imaging and Their Biological Applications with Lab-on-a-Chip Technology. *Sensors.* 23, 7691.
- Roy, R., S. Hohng, and T. Ha. 2008. A practical guide to single-molecule FRET. *Nat. Methods.* 5:507–516.
- Hellenkamp, B., S. Schmid, ..., T. Hugel. 2018. Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nat. Methods.* 15:669–676.
- Agam, G., C. Gebhardt, ..., T. Cordes. 2023. Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins. *Nat. Methods.* 20:523–535.
- Ha, T., A. Y. Ting, ..., S. Weiss. 1999. Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proc. Natl. Acad. Sci. USA.* 96:893–898.
- Lerner, E., A. Barth, ..., S. Weiss. 2021. FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *Elife.* 10, e60416.
- Kapanidis, A. N., N. K. Lee, ..., S. Weiss. 2004. Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci. USA.* 101:8936–8941.
- Hildebrandt, L. L., S. Preus, and V. Birkedal. 2015. Quantitative single molecule FRET efficiencies using TIRF microscopy. *Faraday Discuss.* 184:131–142.
- Salem, C.-B., E. Ploetz, and D. C. Lamb. 2019. Probing dynamics in single molecules. In *Spectroscopy and Dynamics of Single Molecules*, pp. 71–115.
- Kudryavtsev, V., M. Sikor, ..., D. C. Lamb. 2012. Combining MFD and PIE for accurate single-pair Förster resonance energy transfer measurements. *ChemPhysChem.* 13:1060–1078.
- Götz, M., A. Barth, ..., S. Schmid. 2022. A blind benchmark of analysis tools to infer kinetic rate constants from single-molecule FRET trajectories. *Nat. Commun.* 13:5402.
- de Lannoy, C. V., M. Filius, ..., D. de Ridder. 2021. FRETboard: Semi-supervised classification of FRET traces. *Biophys. J.* 120:3253–3260.
- Huisjes, N. M., T. M. Retzer, ..., K. E. Duderstadt. 2022. Mars, a molecule archive suite for reproducible analysis and reporting of single-molecule properties from bioimages. *Elife.* 11, e75899.
- Kinz-Thompson, C. D., and R. L. Gonzalez, Jr. 2018. Increasing the Time Resolution of Single-Molecule Experiments with Bayesian Inference. *Biophys. J.* 114:289–300.
- Preus, S., S. L. Noer, ..., V. Birkedal. 2015. iSMS: single-molecule FRET microscopy software. *Nat. Methods.* 12:593–594.
- Schmid, S., M. Götz, and T. Hugel. 2016. Single-Molecule Analysis beyond Dwell Times: Demonstration and Assessment in and out of Equilibrium. *Biophys. J.* 111:1375–1384.
- Verma, A. R., K. K. Ray, ..., R. L. Gonzalez. 2023. Increasing the accuracy of single-molecule data analysis using tMAVEN. Preprint at bioRxiv. <https://doi.org/10.1101/2023.08.15.553409>.
- Heiss, G., E. Ploetz, ..., D. C. Lamb. 2019. Conformational changes and catalytic inefficiency associated with Mot1-mediated TBP–DNA dissociation. *Nucleic Acids Res.* 47:2793–2806.
- Wanninger, S., P. Asadiatouei, ..., D. C. Lamb. 2023. Deep-LASI: deep-learning assisted, single-molecule imaging analysis of multi-color DNA origami structures. *Nat. Commun.* 14:6564.
- Thomsen, J., M. B. Sletfjording, ..., N. S. Hatzakis. 2020. DeepFRET, a software for rapid and automated single-molecule FRET data classification using deep learning. *Elife.* 9:e60404.
- Li, J., L. Zhang, ..., N. G. Walter. 2020. Automatic classification and segmentation of single-molecule fluorescence time traces with deep learning. *Nat. Commun.* 11:5833.
- Liao, J., J. Qu, ..., J. Li. 2022. Deep-learning-based methods for super-resolution fluorescence microscopy. *J. Innov. Opt. Health Sci.* 16 (03)
- Yao, R., M. Ochoa, ..., X. Intes. 2019. Net-FLICS: fast quantitative wide-field fluorescence lifetime imaging with compressed sensing – a deep learning approach. *Light Sci. Appl.* 8:26.
- Huang, B., J. Li, ..., J. Qu. 2023. Enhancing image resolution of confocal fluorescence microscopy with deep learning. *PhotonIX.* 4, 2.
- Ragone, M., R. Shahabazian-Yassar, ..., V. Yurkiv. 2023. Deep learning modeling in microscopy imaging: A review of materials science applications. *Prog. Mater. Sci.* 138, 101165.
- Fablab. 2023. Deep-LASI Documentation. <https://deep-lasi-tutorial.readthedocs.io/en/latest/index.html>.
- Deep-LASI. 2023. Deep-LASI Software. <https://gitlab.com/simon71/deeplasi>.
- Reddy, B. S., and B. N. Chatterji. 1996. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Trans. Image Process.* 5:1266–1271.
- Mallat, S. G. 1989. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 11:674–693.
- Kapanidis, A. N., T. A. Laurence, ..., S. Weiss. 2005. Alternating Laser Excitation of Single Molecules. *Acc. Chem. Res.* 38:523–533.
- Daubechies, I. 1992. Ten Lectures on Wavelets. SIAM.
- Mallat, S. G. 2009. A Wavelet Tour of Signal Processing : The Sparse Way, 3rd ed. Elsevier/Academic Press, p. 805.
- Starck, J. L., and F. Murtagh. 2007. Astronomical Image and Data Analysis. Springer Berlin Heidelberg.
- Schreiber, J. 2018. Pomegranate: fast and flexible probabilistic modeling in python. *J. Mach. Learn. Res.* 18:1–6.
- Murphy, K. 1998. Hidden markov model (HMM) toolbox for Matlab. <http://www.cs.ubc.ca/~murphyk/Software/HMM/hmm.html>.
- Barth, A., L. Voith von Voithenberg, and D. C. Lamb. 2019. Quantitative Single-Molecule Three-Color Förster Resonance Energy Transfer by Photon Distribution Analysis. *J. Phys. Chem. B.* 123:6901–6916.
- Götz, M., P. Wortmann, ..., T. Hugel. 2016. A Multicolor Single-Molecule FRET Approach to Study Protein Dynamics and Interactions Simultaneously. *Methods Enzymol.* 581:487–516.
- Ploetz, E., E. Lerner, ..., T. Cordes. 2016. Förster resonance energy transfer and protein-induced fluorescence enhancement as synergistic multi-scale molecular rulers. *Sci. Rep.* 6, 33257.
- Atto-Tec. 2023. Förster Radius of Atto dyes. https://www.atto-tec.com/images/ATTO/Katalog_Preisliste/Katalog_2019_2020.pdf.
- Ross, J., P. Buschkamp, ..., P. Tinnefeld. 2007. Multicolor Single-Molecule Spectroscopy with Alternating Laser Excitation for the Investigation of Interactions and Dynamics. *J. Phys. Chem. B.* 111:321–326.

Paper 3.

Distance and Kinetic Tunability of Dynamic DNA Origami Structures
Examined at the Single-Molecule Level⁹⁹

Asadiatouei, P., Bohlen, J., Cole, F., Morella, F., Tinnefeld, P., Ploetz, E., Lamb, D.C.

In preparation for submission

Distance and Kinetic Tunability of Dynamic DNA Origami Structures Examined at the Single-Molecule Level

Pooyeh Asadiatouei, Johann Bohlen, Fiona Cole, Fabio Morella, Philip Tinnefeld, Evelyn Ploetz, Don C. Lamb

Keywords

DNA Origami structure, nanotechnology, single-molecule spectroscopy, single-molecule FRET, multi-color FRET

Abstract

DNA nanotechnology has revolutionized molecular engineering, enabling the construction of dynamic, programmable architectures with near-atomic precision. Among these, DNA origami structures stand out as molecular rulers—offering unparalleled control over spatial organization at the nanoscale. By strategically positioning fluorophores within these scaffolds, Förster resonance energy transfer (FRET) can be used to probe molecular distances, conformational changes and dynamics of the origami and attached molecular components, as well as hybridization dynamics with single-molecule sensitivity, and can even be tuned to response to external triggers. However, despite the inherent programmability of DNA origami structures, kinetic and structural uncertainties arise due to sequence (dis-)placement, environmental factors, and dye-DNA interactions. Here, we employ single-molecule FRET (smFRET) in solution and on surface-immobilized molecules to benchmark the spatial and temporal precision of DNA origami structures as kinetic rulers. We systematically investigated the binding dynamics of single-stranded DNA (ssDNA) tethers to designed docking sites along an L-shaped DNA origami scaffold, tuned docking site positioning, fluorophore identity, sequence and hybridization length. While DNA origami provides remarkable tunability and reproducibility at the nanoscale, we find that local environmental effects—including fluorophore charge and docking site microstructures—can modulate hybridization kinetics by up to a factor of two. These insights refine the use of dynamic DNA origami as molecular benchmarks for fluorescence-based assays and highlights design parameters that must be considered when engineering DNA-based nanodevices for precision and sensing applications.

Introduction

DNA nanotechnology has transformed molecular engineering, enabling the precise design of self-assembled nanostructures in two- and three-dimensional configurations with programmable dynamics and motion^{1, 2}. While early applications of DNA origami structures focused on creating static, well-defined architectures³, recent advances have unlocked their potential as dynamic molecular systems—capable of controlled motion¹, reconfiguration⁴, and molecular sensing⁵. By leveraging sequence-specific hybridization⁶, strand displacement reactions^{7, 8} and external stimuli^{9, 10} (like temperature, pH shifts and molecular crowding), dynamic DNA origamis can function as responsive nanodevices for applications ranging from molecular computing¹¹ and drug delivery¹² to high-precision biophysical measurements¹³, including force¹⁴ and affinity⁸ probing at the single-molecule level. A particularly exciting applications of dynamic DNA origami structures lies in nanoscale metrology, where these structures act as molecular ruler with defined kinetic and spatial properties. Their ability to undergo programmable conformational changes makes them ideal for benchmarking distance measurements in super-resolution imaging (e.g., STORM, DNA-PAINT)¹⁵ and fluorescence-based assays (e.g., Counting, or FRET)¹⁶. Unlike rigid molecular scaffolds, dynamic DNA origami provide kinetic programmability¹⁷, allowing the study of reaction dynamics¹⁸, control of accessibility¹⁹, and spatiotemporal regulation of molecular interactions²⁰. Beyond metrology, DNA-based nanomachines—such as DNA walkers²¹, switches²², and molecular robots²³—demonstrate the potential of dynamic DNA origami in molecular computing and synthetic biology. With their combination of spatial precision, kinetic control, and responsiveness, dynamic DNA origami represents a next-generation toolkit for interrogating molecular interactions and designing functional nanodevices.

DNA is a versatile building block due to its predictable base-pairing that can be exploited to program and control nanoscale motion. Yet, a critical factor in dynamic DNA origami function is the hybridization kinetics of single-stranded DNA (ssDNA). These can be monitored by incorporating small fluorescent molecules on or near target strands. However, incorporating fluorescent labels into DNA strands or origami structures may introduce potential side effects that can impact the energy landscape and conformational dynamics of the DNA origami structures. Certain fluorophores are known for unwanted dye-DNA strand interactions that can stabilize DNA double strands, be responsible for dye-dependent conformational dynamics on single DNA strands or lead to duplex dissociation²⁴. Previous studies show that cyanine dyes, especially when positioned terminally, tend to alter their photophysical properties and inhibit dynamics more than the rhodamine-, coumarin-, or xanthene-based dyes²⁵. Usually, terminally-labeled fluorophores stabilize hybridization by both stacking and electrostatic interactions with positively-charged dyes interacting stronger with the negatively-charged DNA than the negatively-charged dyes²⁶.

When using DNA origami probes with responses that rely on transient binding and hybridization of ssDNA, there is a potential for secondary structure formation like a stem loop or a triple helix. This can occur either when the single strands diffuse in solution or when they are tethered to a DNA origami structure, even for short strands. Such secondary structures can alter hybridization kinetics, especially by reducing the association rate k_{on} , which can be problematic for techniques such as DNA PAINT²⁷. Considering all potential intrinsic interactions of DNA strands on origami structures along with extrinsic factors such as fluorophores, can limit the precision with which the kinetics and positioning, and hence the functionality are predicted.

To investigate the effect of above-mentioned external influences on DNA hybridization kinetics, we designed and measured a number of L-shaped DNA origami structures that offers a programmable arrangement of protruding strands and biotin anchors on the backside for immobilization²⁸. The short side of the L-shaped DNA origami structure contains the breadboard portion used for programming the kinetics and, therefore, by

design, is not sensitive to the coverslip surface. The structure contains a flexible, single-stranded DNA (ssDNA) tether with complementary ssDNA docking strands in the near vicinity. The kinetics are then modified by changing the length of the complementary sequence. To monitor the dynamics, we use single-molecule Förster Resonance Energy Transfer (smFRET) experiments. Given that DNA nanostructures are typically smaller than the diffraction limit of light, the smFRET approach allows us to probe the conformation and dynamics of biomolecules on the nanometer scale. Additionally, they provide important insights into the structural heterogeneities of the nanostructures, detect rare mechanistic events, and separate out defective or non-functional molecules.

In this work, we employ smFRET techniques both on immobilized structures and in solution to investigate the critical aspects of dynamic DNA origami design and labeling. We varied the type and number of attached fluorophores, the labeling position and the length of the complementary region of the docking strands to investigate their influences on the FRET efficiency and kinetic rates. We measured the hybridization and dissociation kinetics as well as the fluorescence lifetime and anisotropy information to monitor the rotational flexibility and potential quenching for each fluorophore. Our results reveal that positively charged dyes can interact and destabilize the DNA hybridization, but that the local environmental of the origami around the docking strands have an inherent effect on their kinetics. This limits the programmability of the dynamics to within roughly a factor of 2. Our data show that these dynamic DNA origami nanostructures are excellent reference structures for two-, and multicolor FRET and especially for FRET dynamics similar to how rigid structures are commonly used as standards for super-resolution microscopy ²⁹.

Results and Discussion

DNA origami structure design

L-shaped DNA origami nanostructures containing a fluorescently labeled, ssDNA pointer with a length of 19 nucleotides (nt) were used throughout this study. For most structures, the protruding single-stranded DNA pointer was labeled with Cy3B at the 3'-end and placed between up to three surrounding docking strands. The complementary sequence length to the pointer was varied in order to control the number of involved binding states as well as associated kinetic rates (Figure 1). Two other dyes (Atto488 and Atto647N) were bound to the origami surface in close proximity to the two docking strands at 6 and 12 o'clock positions to probe the position of the tether. The docking strands were placed within reach of the tether to allow transient binding to the different states. In addition, four biotin groups were attached to the bottom of origami structures for surface immobilization.

To determine whether the positions of the blue and red dyes impact the tether binding and the associated FRET states, we designed two versions of the origami structures where the position of the fluorophores are switched. In version 1 (V1), the red dye is placed close to the 12 o'clock binding position (up) and the blue dye close to the 6 o'clock binding position (down) (Figure 1a, the left zoomed-out panel). The reverse arrangement was prepared for version 2 (V2) of the samples (Figure 1a, the right zoomed-out panel). To make a meaningful comparison between the different structures, the expected FRET states for each origami design were numbered regarding the YR FRET efficiency values of the first examined version (V1) with the highest value being state 1 (S1). In other words, for all structures, the tether is bound to the top docking site (S1, 12 o'clock position) and state 2 to the low position (S2, 6 o'clock position).

The tether dynamics can be tuned by varying the position and length of complementary binding sequences. As shown before, the binding stability is also influenced by the presence of a nucleic acid mismatch ³³. The kinetics of the tether were studied using four different complementary sequence lengths of 7 nt with a 1-nucleotide mismatch (referred to as 6.5 nt as a mismatch offers more possible interactions than a 6 nt structure), 7 nt, 8 nt with a 1-nucleotide mismatch (referred to as 7.5 nt) and 8 nt for both the V1 and V2 designs. To complement the three-color experiments, a set of two-color labeled DNA origami structures with identical binding positions were prepared for each dye pair, BY, BR and YR, for both arrangements of V1 and V2. Figure 1b shows a schematic of the V1 samples only. All six two-color origami structures contained the 7.5 nt complementary binding sequence length.

Another L-shaped DNA origami structure was designed containing a third docking strand at either 9 o'clock or 11 o'clock (state 3, S3; Figure 1c). The third binding strand contains 7.5 nt complementary bases whereas the 6 and 12 o'clock positions have a complementary length of 7 nt. Two-color (YR pair) samples in the V1 orientation for the 6-9-12 and 6-11-12 clocks are included for comparison to the three-color experiments (Figure 1d).

Additional structures were also generated as controls (Figure 1e). To investigate whether the pointer has a favored position or directional bias in the unbound state, we made two- and three-color origami samples without any complementary docking strands for the tether. We refer to these samples as the no-binding structures. Both V1 and V2 versions of the sample were prepared. Figure 1e shows the schematic of this last design for the V1 structures. As a final design, we changed the positioning of all three docking strands to be closer to tether, equally distributed and containing the same complementary 7 nt binding sequences. The tether was labeled with Atto542, and Cy5 was used as the acceptor dye on the origami surface (Figure 1f).

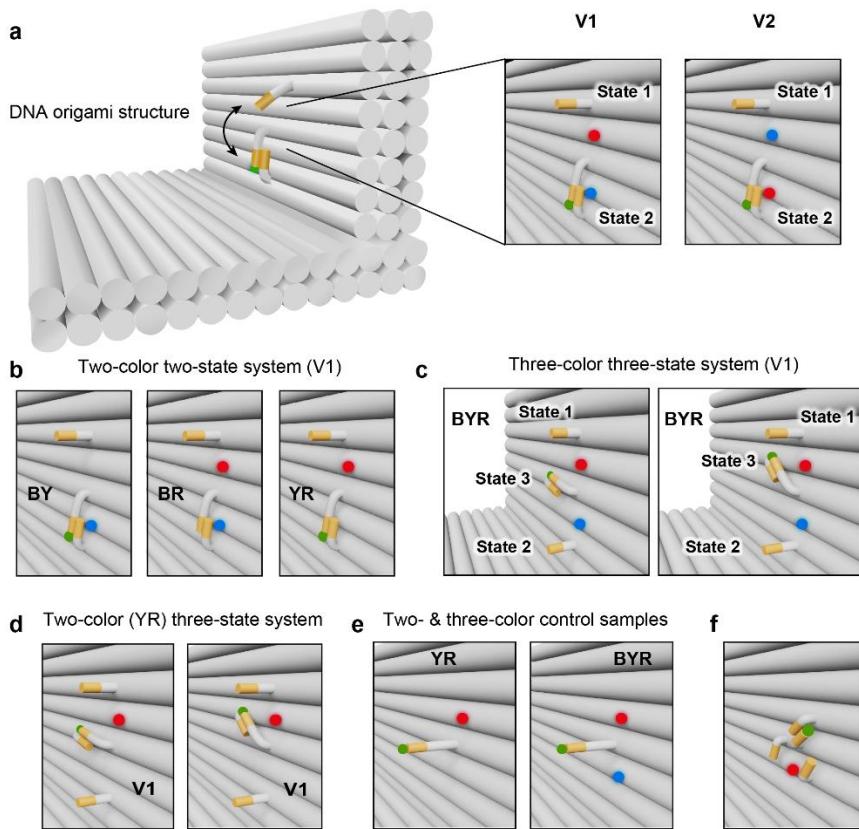


Figure 1. Overview of the L-shaped DNA origami nanostructures used in this study. (a) A schematic of the overall L-shaped DNA origami structure. A flexible tether labeled with Cy3B extrudes from the upper arm and can transiently bind to either of the docking strands at the 6 or 12 o'clock positions. Two versions of the structure were created where the Atto488 and Atto647N fluorophores were either attached with the red dye near the docking strand at 12 o'clock and the blue dye near 6 o'clock (V1) or vice versa (V2). (b) Schematics of the two-color variants of the origami structure in panel a with two transient binding states for the V1 samples showing the BY (left), BR (middle) and YR (right) samples. (c) Schematics of three-color, three-state DNA origami structures for the V1 labeling design with binding strands at the 6-9-12 (left) and 6-11-12 positions (right). (d)

Two-color, V1 variants of the three-state origami structures shown in panel c. (e) Additional control V1 DNA origami structures with two- (left) and three-color (right) labels but no docking strands. (f) A three-state system where the binding sites are symmetrically distributed about the tether.

Three- and two-color single-molecule FRET measurements on the surface using smTIRF

First, we measured the dynamics of the tether on the immobilized three-color, two-state DNA origami structures using prism-based TIRF microscopy. Figure 2a shows a schematic of a smTIRF assay on a three-color DNA origami structure using a biotin-streptavidin-biotin linkage for immobilization. The measurements were conducted on V1 of the three-color DNA origami structure designs with the blue and red fluorophores placed close to the two complementary binding strands at the 6 and 12 o'clock positions respectively. To determine the FRET efficiencies for each dye pair and investigate how the different number of complementary nucleotides affect the kinetics of the system, we prepared and measured four similar constructs containing a different number of complementary base pairs available on the docking strands (6.5, 7, 7.5 and 8 nt). A schematic of the docking interaction for the 7.5 nt complementary stand length is shown in Figure 2b. The obtained FRET efficiencies were corrected for spectral crosstalk of the donor emission into acceptor channel, direct excitation of the acceptor by donor excitation laser and by the detection efficiency correction factor (gamma factor). Dynamic switching between the different docking strands can be directly observed in the single-molecule FRET traces (Figure 2c). When the BY FRET efficiency is high the YR efficiency is low and vice versa.

A histogram of the FRET efficiencies observed in the dynamic traces are shown in Figure 2d for the various constructs. The corrected FRET values are averaged over each transition and weighted by the number of frames per transition, which we refer to as the frame-wise-weighted state-wise averaged FRET histograms and are plotted for each dye pair (BY, BR and YR). The results and correction factors are summarized in Supplementary Table S1. As expected for the two BY and YR dye pairs, two distinct FRET populations are

observed, which are related to the transient binding of the tether to either of the two docking strands. For the 6.5 nt sample, the kinetic rates approach that of the camera exposure time, leading to a rise in FRET values at intermediate FRET efficiencies³⁴. However, the FRET values for the low and high FRET states are similar for all constructs (S1 = 37-50 % for BY and 80-81 % for YR; S2 = 78-82 % and 6-11 % for BY and YR respectively). For the BR dye pair, we observe differences in the uncorrected FRET values for states 1 and 2 (Figure 2c, third panel), but they converge to a single peak at ~50 % FRET efficiency upon correction (Figure 2d, middle panel). The variation of ~±5 % in the corrected peak FRET efficiencies is within error given the amount of correction needed for the three-color FRET experiments and the high sensitivity of the FRET efficiency at values around 50 % (the variation in distances is < 2 Å).

Using our Deep-learning approach, we also analyzed the kinetics of the different DNA origami structures³⁵. Every transition is plotted on the transition density plot (TDP) which forms clusters around the values for states 1 and 2. Upon choosing a cluster, the dwell time distributions for the selected transitions are plotted and fit to a mono-exponential model (Figure 2e and 2f). As we showed earlier, the kinetic rates can be adjusted via the number of complementary bases³⁵ and adding a C-A mismatch also has an influence on the kinetic rates (Figure 2b, Supplementary Table S1). Interestingly, the data consistently shows longer dwell times for S2 compared to S1 (Figure 2f). The dwell-time ratio of S2 to S1 is 1.2, 1.5, 1.4 and 1.5 for the 6.5 nt, 7 nt, 7.5 nt and 8 nt samples respectively. This is consistent with the difference in populations of S1 and S2 observed in the frame-wise-weighted state-wise averaged smFRET histograms. A Gaussian fit to the two populations yields ratios of 1, 1.6, 1.3 and 1.5 for the 6.5 nt, 7 nt, 7.5 nt and 8 nt samples respectively. Hence, the tether spends more time bound to the lower position (6 o'clock, state 2) than the top binding strand (12 o'clock, state 1).

One possible explanation for the difference in binding rates of S1 and S2 could be due to interactions between dyes or between the dye and the DNA docking strand. To investigate this possibility, we prepared and measured two-color, two-state origami structures on the surface for the 7.5 nt constructs (Figure 2g). The results for the three dye-pair combinations are shown in Figure 2h. For the BY dye pair, two FRET populations are observed with FRET efficiencies of 28 % and 77 % (Figure 2h, left panel) and the YR FRET pair shows two FRET states with values of 81 % and 9 % (Figure 2h, right panel). The BR FRET pair exhibits a single FRET population with a FRET efficiency of 43 % (Figure 2h, middle panel). The FRET efficiencies are in excellent agreement with the three-color samples. Here again, the two-color samples exhibit longer dwell times when the tether is bound to the lower binding position. If we look at the kinetics, the dwell times are $\tau_1 = 1.40 \pm 0.02$ s and $\tau_2 = 2.80 \pm 0.003$ s, and $\tau_1 = 1.80 \pm 0.02$ s and $\tau_2 = 2.80 \pm 0.02$ s for the BY and YR samples respectively. For the YR sample, the kinetics are the same as the three-color sample. Interestingly, we observed an increase in the kinetic transition rate from S1 to S2 in the absence of the red dye, shifting from 1.80 ± 0.02 s to 1.40 ± 0.02 s. Hence, we see different rates depending on whether or not the red fluorophore is present on the DNA origami structure. This is not surprising as it has been shown that dyes can interact with DNAs²⁵. However, the asymmetry in dwell times between S1 and S2 gets even larger in the absence of the red dye.

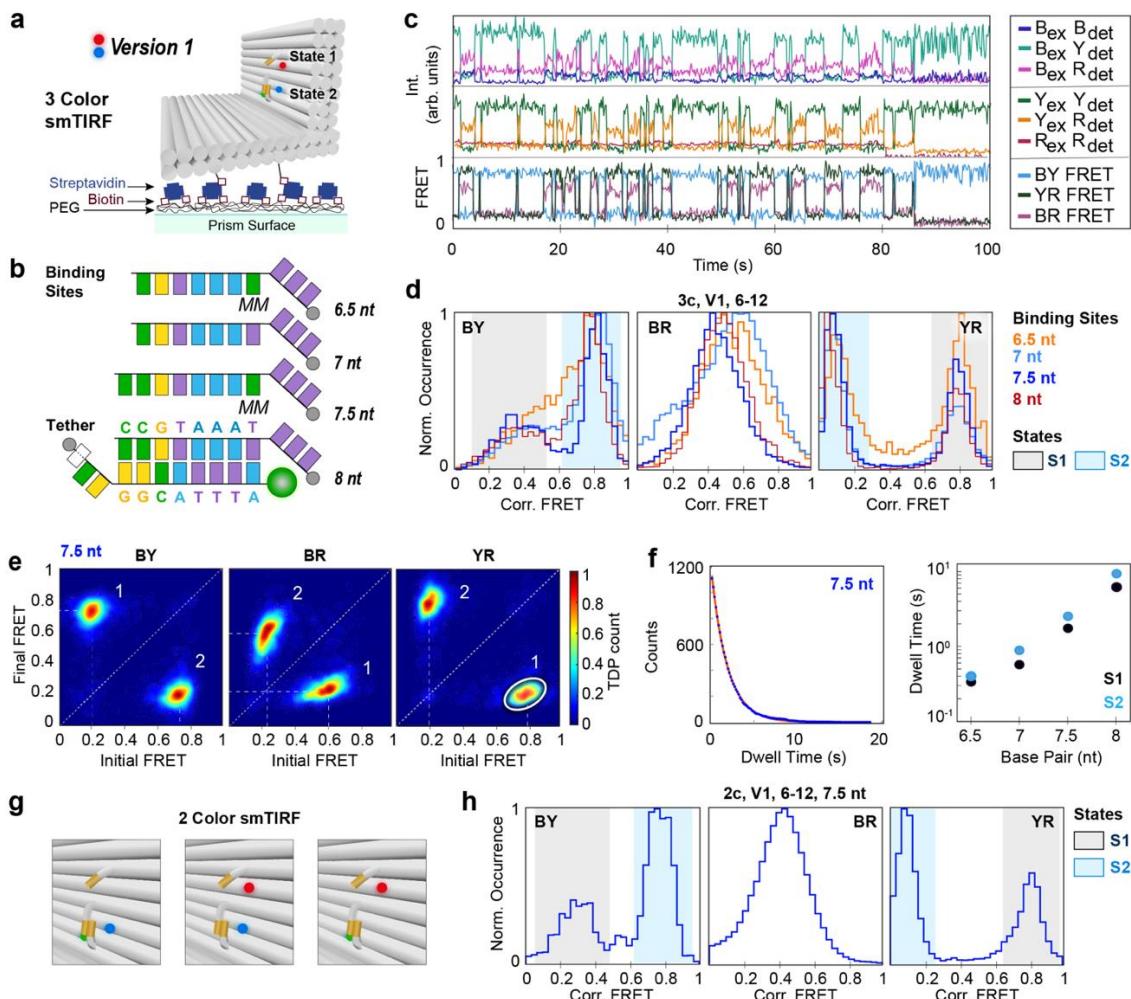


Figure 2. SmTIRF results for the three- and two-color, two-state V1 L-shaped DNA origami structures. (a) A schematic showing the smTIRF assay for measuring immobilized DNA origami structures on the passivated prism surface. (The dimensions are changed for clarity). The origami is labeled with Cy3B on the tether, Atto647N on the top and Atto488 at the bottom. (b) A schematic of the tether being bound to a series of complementary docking strands of 7 to 8 nucleotides. Base pairing constructs with or without a CA mismatching nucleotide were used to provide a range of kinetic rates (6.5, 7, 7.5 and 8 nt). (c) Representative single-molecule intensity and FRET trace of a 3c, V1, 6-12 DNA origami structure with 7.5 nt binding sites exhibits dynamic switching between the different docking strands. Top: Intensity traces of the blue, yellow and red dyes after blue excitation. Middle: Intensity traces of yellow and red dyes after yellow excitation together with red emission after red excitation. Lower: FRET efficiency traces for all three dye pairs. (d) Frame-wise-weighted state-wise averaged corrected FRET efficiency histograms for the BY (left panel), BR (middle panel) and YR (right panel) dye pairs for complementary nucleotides lengths of 6.5 (orange), 7 (blue), 7.5 (dark blue) and 8 (red). The dynamic BY pairs exhibit two populations with low FRET values ($\sim 35\%$) for S1 and a high FRET value around 80 % corresponding to S2. The static BR pair gives one broad histogram centered at around 50 %. The dynamic YR pair shows two populations with FRET efficiencies of around 80 % and 10 % for states 1 and 2 respectively. (e) Transition density plots showing the interconversion events between the two detected states for each dye pair. For each transition, the initial and final apparent FRET value is depicted to make the transition clusters. (f) Left: By selecting transitions within individual clusters (e.g. shown for the S1 \rightarrow S2 transition in panel e, highlighted in white), a distribution of dwell times for the individual transitions can be calculated. The data points are shown in blue, the mono-exponential fit in red. Right: Dwell-times for S1 and S2 as a function of complementary nucleotides on the docking strands. The dwell-times were determined from a mono-exponential fit to the dwell-time distribution. (g) Zoomed-in schematics of two-color, two-state origami structures containing one dye pair of BY (left), BR (middle) and YR (right). (h) Frame-wise-weighted, state-wise averaged corrected FRET efficiency histograms from smTIRF experiments with two-color ALEX excitation for each dye pair of BY (left), BR (middle) and YR (right) for the two-color structures containing 7.5 nt complementary base pairs. The dynamic BY pair gives two populations of 28 % and 77 % for S1 and S2 respectively. The static BR pair gives one population centered at 43 %. The dynamic YR pair shows two populations with FRET efficiencies of 81 % and 9 % for S1 and S2 respectively.

Three- and two-color single-molecule FRET measurements of DNA origami structures with switched fluorophores on the surface (smTIRF)

To further investigate the impact of the dyes on the kinetic rates, we performed smFRET experiments using TIRF excitation on immobilized three- and two-color structures using a second version of the origami structure (V2) where the labeling positions of Atto488 and Atto647N were swapped (Figure 3a). Again, the traces show kinetics between two states (Figure 3b) and the resulting FRET histograms show high and low FRET efficiency states for the BY and YR pairs but the FRET efficiency values for the two states are shifted closer together. The corresponding corrected FRET values for the three-color V2 samples after fitting the frame-wise-weighted state-wise averaged histograms are 85-90 % and 33-40 % for the BY FRET pair corresponding to S1 and S2 (Figure 3c, left) and 17-27 % and 58-61 % for the YR pair (Figure 3c, right) respectively. As the two dye pairs have different Förster Radii, a change in FRET efficiency values is expected upon exchanging of the blue and red dyes. However, the changes in FRET values cannot be explained purely by the differences in the assumed Förster Radii of the two dye pairs as discussed later. For the BR pair, a single FRET population with a FRET efficiency of 52-58 % was observed (Figure 3c, middle). The FRET value is very similar for V1 and V2, meaning that the distance between the blue and red dyes on the origami structures are not changing significantly. As the positions of the dyes are switched between V1 and V2, this suggests that the position of the blue and red dyes on the two versions of the L-shaped origami samples are similar.

We also analyzed the kinetics of the V2 samples using the TDPs for each dynamic pair (Figure 3d). The dwell times we measured for the V2 samples are very similar to the V1 samples for all four constructs (e.g., for the 7.5 nt sample: $\tau_1 = 1.70 \pm 0.02$ s and $\tau_2 = 2.50 \pm 0.02$ s, see Supplementary Table S1). Again, we observe longer dwell times for S2 and the ratio of dwell times between 1.3-1.7 (Figure 3e). Hence, we can reliably tune the dynamic rates via the number of complementary nucleotides, but there is still an intrinsic difference between the dwell times of S1 and S2.

Next, we measured two-color versions of L-shaped DNA origami structures with the V2 geometry with complementary docking strands of 7.5 nt. Figure 3f shows the zoomed-in design of the two-color, two-state origami structures for the BY (Figure 3f, left), BR (Figure 3f, middle) and YR (Figure 3f, right) constructs. For the BY and YR samples, dynamics between two FRET populations were observed with corrected FRET values of 86 % and 13 % (Figure 3g, left), and 28 % and 64 % (Figure 3g, right) for the BY and YR FRET pairs, respectively. Again, for the BR FRET pair, we obtained a single FRET population with 46 % FRET efficiency (Figure 3g, middle). A comparison of the dwell times between the three- and two-color V2 samples showed excellent agreement for the YR sample. However, as for the BY V1 sample, we see a decrease in the dwell time of the docking strand located near where the red dye would be. This again supports the idea that the presence of Atto647N near the docking strand influences the kinetics. However, the presence of Atto647N does not explain the consistent asymmetry in the dwell times of S1 and S2 for the same binding sequence.

So far, we have only discussed molecules that exhibit dynamic behavior. However, up to roughly half of the molecules can be static. Hence, we also investigated the conformation of the static DNA origami structures (Supplementary Note S6). Both S1 and S2 are present for all samples. In V1, S1 is the dominate state although we see an increase in the number of static molecules with increasing docking strand length (with the exception of the 6.5 nt sample, which may have other reasons). This bias disappears in two color sample, when Atto647N is not present. For V2, S2 now becomes the dominate state observed for static origami structures. Hence, we conclude that sticking of the Cy3B-labeled tether can occur to either state, leading to static structures. The probability increases as the length of the complementary docking strands increase. Furthermore, we see a

bias in the conformation of the static structures towards the docking strand that is in the vicinity of Atto647N, clearly indicating an interaction between Atto647N and Cy3B.

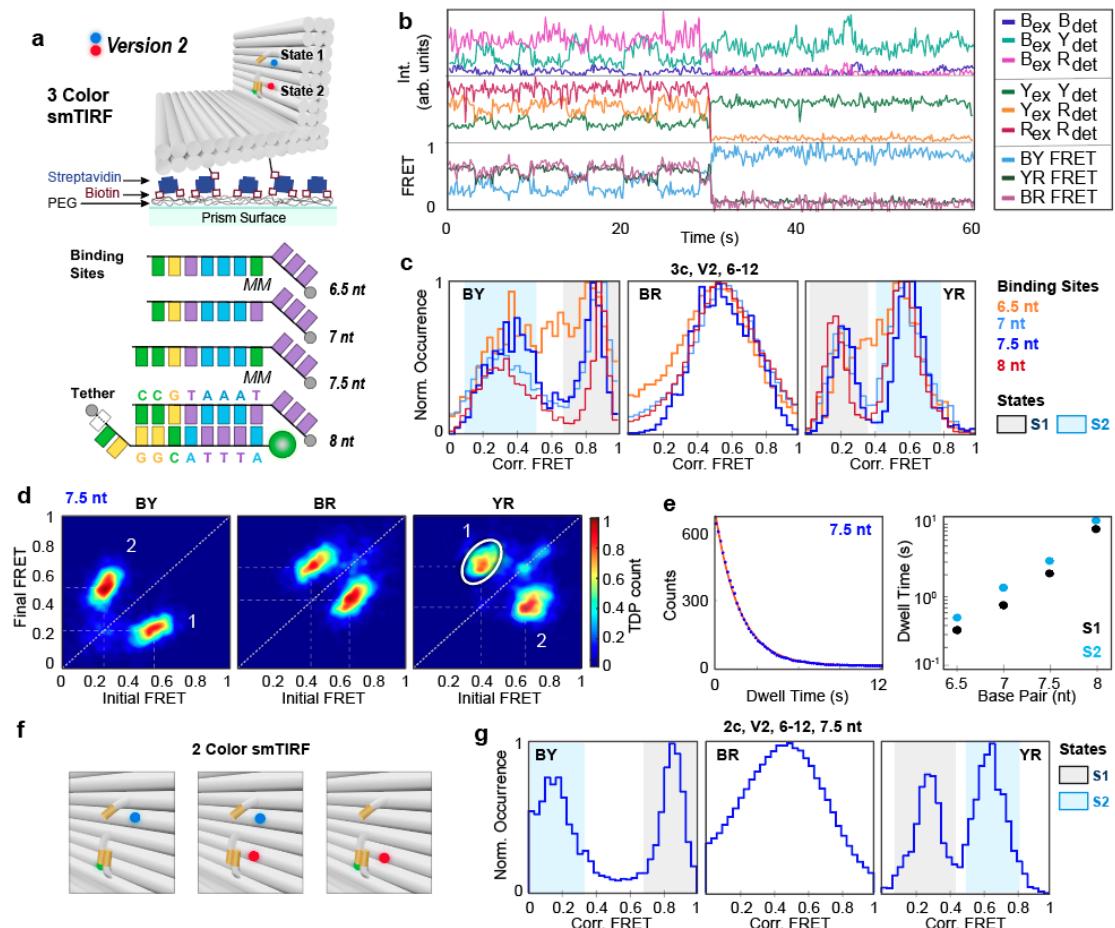


Figure 3. SmTIRF results of three- and two-color, two-state V2 L-shaped DNA origami structures. (a) Schematic of the smTIRF assay. DNA origami structures are immobilized on the surface of a passivated prism and labeled with Cy3B on the tether, Atto488 on the top and Atto647N at the bottom. Temporary binding of the tether to state 1 and state 2 is achieved via base-pairing between 7 and 8 nucleotides along the pointer and a series of complementary docking strands (6.5, 7, 7.5, and 8 nt). (b) Representative single-molecule intensity and FRET trace of a 3c, V2 6-12 DNA origami structure with 7.5 nt binding sites shows dynamic switching between the different docking strands. Top: Intensity traces of blue, yellow and red dyes after blue excitation. Middle: Intensity traces of yellow and red dyes after yellow excitation together with red emission after red excitation. Bottom: FRET traces for all three dye pairs. (c) Frame-wise-weighted, state-wise averaged corrected FRET efficiency histograms from smTIRF experiments with three-color ALEX excitation for the BY (left), BR (middle) and YR (right) dye pairs for structures with 6.5 nt (orange), 7 nt (blue), 7.5 nt (dark blue) and 8 nt (red) complementary docking strands. The dynamic BY pair shows two populations with 85-90 % FRET efficiency for S1 and 33-40 % corresponding to S2. The static BR pair gives one broad histogram centered at around 52-58 % for different constructs. The dynamic YR pair shows two populations with FRET efficiencies of around 17-27 % and 58-61 % for S1 and S2 respectively. (d) Transition density plots showing the transitions between the two detected states for each dye pair. For each transition, the initial FRET value of the transition is plotted versus the final apparent FRET value. (e) Left: The dwell-time distribution for the S1 → S2 transition for the YR FRET pair (highlighted in white in panel d). The data points are shown in blue, the mono-exponential fit in red. Right: Dwell-times for S1 and S2 as a function of complementary nucleotide length of the docking strands. The dwell-times were determined from a mono-exponential fit to the dwell-time distributions. (f) Zoomed-in schematics of two-color, two-state origami structures containing the BY (left), BR (middle) and YR (right) dye pairs. (g) Frame-wise-weighted, state-wise averaged, corrected FRET efficiency histograms from smTIRF experiments with two-color ALEX excitation for the BY (left), BR (middle) and YR (right) dye pairs for the two-color structures containing 7.5 nt complementary bases. The BY pair shows two populations of 86 % and 13 % for S1 and S2 respectively. The static BR pair exhibits one population centered at 46 %. The YR pair shows two populations with FRET efficiencies of 28 % and 64 % for S1 and S2, respectively.

Three- and two-color smFRET measurements of DNA origami structures in solution

We further investigated the L-shaped DNA origami structures with single-molecule burst experiments in solution. Multiparameter fluorescence detection with pulsed interleaved excitation (MFD-PIE)³⁶ additionally provides fluorescence lifetime and anisotropy information. This allows us to investigate fast dynamics, possible quenching as well as the rotational flexibility of the fluorophores. It is also a good control that immobilization of DNA origami structures to the surface does not influence the structure and dynamics of the structures³⁷. The resulting smFRET histograms for each dye pair of the V1 and V2 samples are summarized in Figure 4. The obtained apparent FRET efficiencies were corrected by applying the FRET-specific correction factors (see Material and Methods). As expected, the corrected FRET values from the measurements in solution are in excellent agreement with respect to the corresponding ones from the smTIRF experiments (Supplementary Table S2.1). This confirms that attachment of the DNA origami structures to the surface does not influence the geometry and functionality of the origami near the fluorophores.

With the lifetime information available from MFD-PIE experiments, we look for the presence of sub-millisecond dynamics by plotting the FRET Efficiency calculated using intensity versus the donor fluorescence lifetime (Supplementary Figure S7.1-S7.4)³⁸. For DNA structures with the 6.5 nt and 7 nt docking strands, we see clear evidence of dynamics, meaning we can observe transitions between the bound tethers, which were too quick to be observed in the TIRF experiments. The lifetime of Atto647N is also independent of FRET efficiency, indicating that there is no significant, state-dependent quenching of the fluorophore.

From the MFD-PIE data, we next calculate the steady-state and time-resolved anisotropy (Supplementary Figure S3.1). Atto488 shows a quick anisotropy decay with a steady-state anisotropy of ~ 0.11 - 0.17 across various structures. The decay of Cy3B is much slower with a higher residual anisotropy. The average steady-state anisotropy is approximately 0.2. Cy3B has no net charge but, when bound to the docking strands, is very close to the DNA origami surface. When looking at the burst-wise anisotropy, the high FRET state (in V1) shows a higher decay in anisotropy (Supplementary Figures S7.1-S7.4, panel i) suggesting more rotational freedom. There is very little difference in the burst-wise anisotropy observed for V2. A steady-state anisotropy value of 0.2 is relatively high, but still acceptable for most FRET experiments³⁹. Most problematic is the red fluorophore, Atto647N. Atto647N always shows the highest steady-state anisotropy values among the three dyes (usually around 0.30). This could be due to the hydrophobic nature of the fluorophore or its electrostatic interactions with the negatively charged surface of DNA origami's backbone while the dye has a net positive charge. The high anisotropy values suggest that Atto647N is often stuck to the DNA surface and not able to freely rotate along its linker axis. Such behavior has been observed previously which also depends largely on the surrounding nano-environment⁴⁰. Thus, distances calculated between fluorophore-pairs that include Atto647N have a high uncertainty.

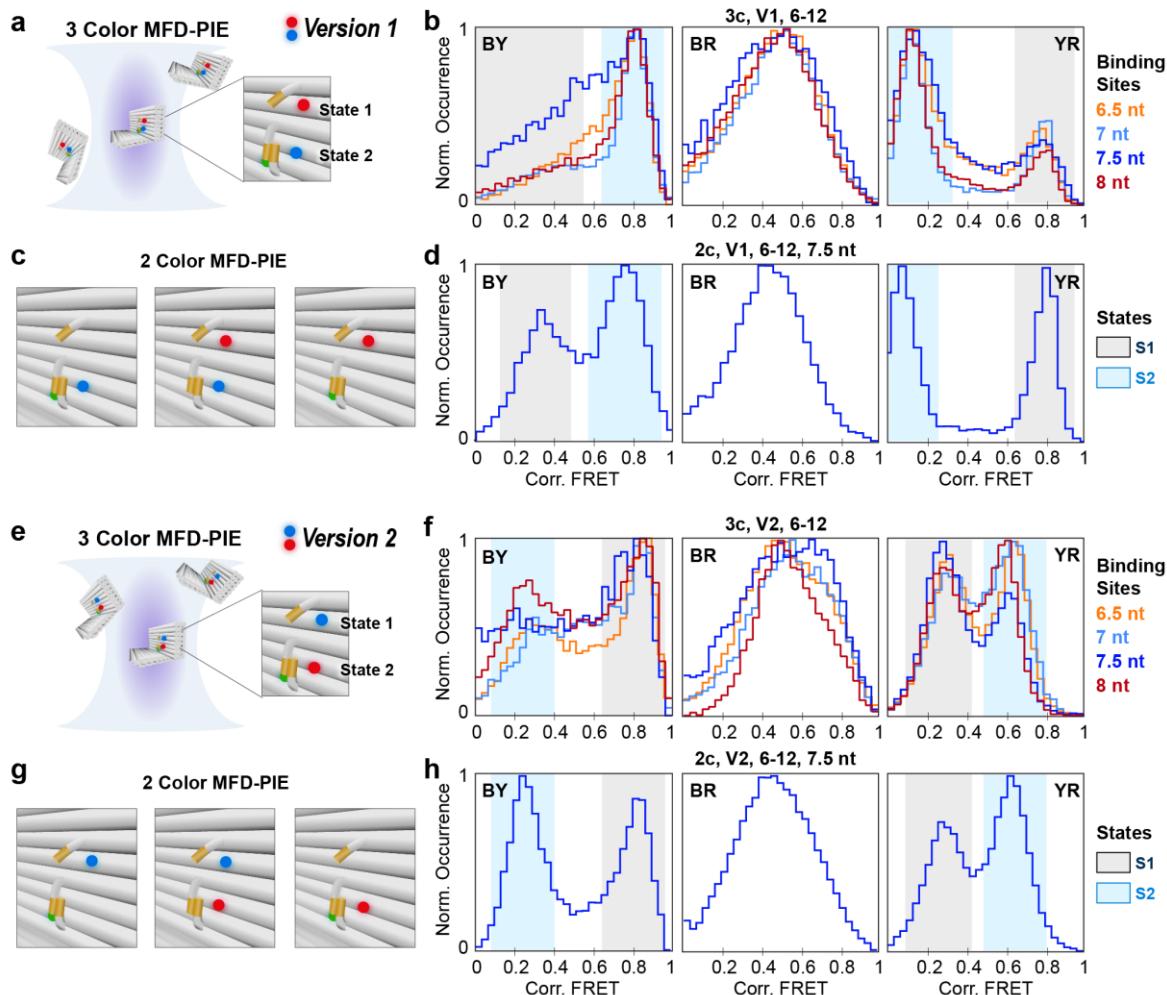


Figure 4. SmFRET results using MFD-PIE for the three- and two-color, two-state L-shaped DNA origami structures. (a) A schematic illustration of the MFD-PIE assay for measuring the FRET efficiency of freely diffusing DNA origami structures as they transverse the confocal volume (The dimensions are changed for clarity). The origami is labeled with two- or three-fluorophores having Cy3B on the tether, Atto647N on the top and/or Atto488 at the bottom (i.e. V1). (b) Corrected smFRET efficiency histograms for the BY (left), BR (middle) and YR (right) dye pairs with 6.5 nt (orange), 7 nt (blue), 7.5 nt (dark blue) and 8 nt (red) complementary docking nucleotides. The BY and YR dye pairs show two clear populations whereas the BR dye pair shows a single peak. (c) Zoomed-in schematics of the two-color, two-state origami structures containing either BY (left), BR (middle) and YR (right) dye pairs. (d) Corrected smFRET efficiency histograms from two-color MFD-PIE experiments for the BY (left), BR (middle) and YR (right) combinations of the two-color V1 structures containing 7.5 nt complementary base pairs. (e) A schematic illustration of the MFD-PIE assay for measuring the FRET efficiency of freely diffusing DNA origami V2 structures as they transverse the confocal volume (The dimensions are changed for clarity). (f) Corrected FRET efficiency histograms from three-color MFD-PIE experiments for the BY (left), BR (middle) and YR (right) dye pairs for V2 structures containing 6.5 nt (orange), 7 nt (blue), 7.5 nt (dark blue) and 8 nt (red) complementary nucleotides on the docking strands. As for V1, the BY and YR dye pairs show two clear populations whereas the BR dye pair shows a single peak. (g) Zoomed-in schematics of the two-color, two-state origami structures containing either the BY (left), BR (middle) and YR (right) dye pairs. (h) Corrected FRET efficiency histograms from two-color, MFD-PIE experiments for the BY (left), BR (middle) and YR (right) dye pairs of V2 structures containing 7.5 nt complementary base pairs.

We also measured the two-color variants with the 7.5 nt docking strands for V1 and V2 in solution (Figure 4d,h). The results are similar to those from the three-color experiments but the FRET histograms are better resolved as less corrections are necessary for determined the corrected FRET efficiencies. From the two-color

experiments, we see a slight quenching of the Cy3B fluorophore when bound to S2, but it is only a minor effect. In both the three-color and two-color experiments, there is still a preference for docking to S2.

Three- and two-color single-molecule FRET of DNA origami structures without any binding sites

One explanation for the obvious preference towards binding to the lower position of the origami structure is that the average position of the unbound tether is biased towards the lower docking strand. To test this hypothesis, we performed experiments on three- and two-color samples (V1 and V2) that did not contain complementary binding strands to see how the free tether behaves in such a situation. The corresponding results are shown in Figure 5. In all cases, we see one dominate FRET state. For the BY FRET pair, we measure a corrected FRET efficiency of 70 %. For the YR FRET pair, a FRET value of about 42 % is observed. As expected, the values for both FRET pairs are between the values detected for S1 and S2. As expected, the value for the BR FRET pair is unchanged. When switching the blue and red labels (V2), we measure very similar distributions as observed for V1. There is a slight shift in the FRET efficiency of the YR sample. This shift is also observable for the two-color YR samples (Figure 5c).

As the BY distances appear unchanged between V1 and V2 whereas the YR distances do show differences, we also measured the rotational freedom of the dyes using MFD-PIE (Supplementary Note 7.4 and 7.8). The overall results from the MFD-PIE experiments align well with the surface measurements. The population is slightly shifted from the static FRET line, indicating that the tether is flexible and the FRET efficiency is a dynamically averaged population (Supplementary Figures S7.11-7.12, 7.21-7.22). The time-resolved anisotropy measurements show a clear decrease in the anisotropy of Atto488 and Cy3B (Supplementary Figure S3.1), whereas the anisotropy of the Atto647N is still high. Thus, distances from the BY dye pair are more reliable than for the YR dye pair. As the corrected smFRET histograms of the BY FRET pair for both V1 and V2 are identical, we conclude that there is no bias in the inherent position of the tether towards binding sites S1 and S2. Also, the BR histograms are similar, suggesting that the distances between the fluorophores on the origami do not change when switched. Hence, we attribute the shift in the YR histograms between V1 and V2 to effects of the orientation of the red dye rather than a biased tether location (studied in more detail below). This suggests that the difference in binding rates comes from the environment surrounding the docking strands and/or the accessibility of the docking strands to the tether.

To check whether there is an intrinsic difference between docking strands due to the local environment, we also measured the binding of freely diffusing DNA strands labeled with Cy3B to the docking sites of S1 and S2. For this assay, we used the same two-state DNA origami structure (6-12 o'clock position, V1) without any tether (Supplementary Figure S4.1a). Binding events were measured using a Cy3B-labeled 8-nt imager strand with the same sequence as the normal tether. We had expected the dwell times to be significantly faster without the tether as the high effective local concentration of the binding strand (i.e. tether) was thought to lead to several dissociation and rebinding events that are too fast to be observed in the smTIRF experiments before successful dissociation from a docking site. Surprisingly, the dwell time distributions for binding to S1 and S2 were similar as for the DNA origami structures with the tether (Supplementary Note 4, Supplementary Figure S4). This suggests that, upon dissociation, the DNA tether switches to the opposite docking strand and does not rebind to the same strand.

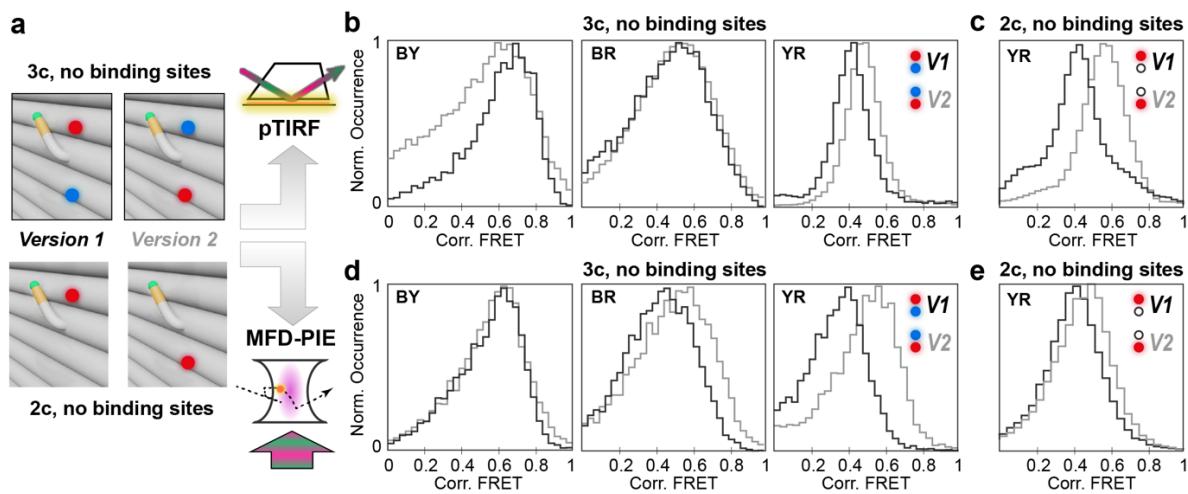


Figure 5. SmFRET experiments using TIRF and MFD-PIE for the three- and two-color L-shaped DNA origami structures with no complementary binding strands. (a) The schematic illustration showing DNA origami structures measured with smTIRF (b-c) and MFD-PIE (d-e). The origami structures are labeled with Cy3B on the tether, Atto647N on the top and Atto488 at the bottom, for V1, or with Atto488 and Atto647N switched for V2. In case of 2-color samples, Atto488 is missing. All designed constructs contain no complementary docking strands in the vicinity of the flexible tether. (b) Frame-wise-weighted state-wise averaged corrected FRET efficiency histograms from smTIRF experiments with three-color ALEX excitation for the BY (left), BR (middle) and YR (right) dye pairs of three-color nanostructures containing no docking strands. (c) Frame-wise-weighted, state-wise averaged corrected FRET efficiency histograms from smTIRF experiments with two-color ALEX excitation for the YR dye pair on constructs V1 and V2 without Atto488. (d) Burst-wise corrected FRET efficiency histograms from MFD-PIE experiments the BY (left), BR (middle) and YR (right) dye pairs for V1 and V2 constructs. (e) Burst-wise corrected FRET efficiency histograms from MFD-PIE experiments for the V1 and V2 constructs containing the YR dye pair only on a no-state nanostructures.

SmFRET measurements of DNA origami structures with three binding sites

Next, we tested how the kinetics change in the presence of a third docking strand. We designed both three- and two-color constructs with a third complementary docking strand around the protruding tether either at the 9 o'clock or 11 o'clock position while the positioning of the labels were kept the same. With this design, we tested how the same binding sequence behaves at different a position and whether the presence of a third docking site, referred to as state 3 (S3), influences the dwell times of S1 and S2. For these constructs, 7 nt complementary bases were used for the two original 6 o'clock and 12 o'clock positions whereas the third position (at 9 o'clock or 11 o'clock) was designed with a 7.5 nt complimentary sequence.

Results for both three-color, three-state samples (6-9-12 and 6-11-12 o'clock binding positions) are given in Figure 6. Although S1 and S3 have similar FRET efficiencies upon correction, they are clearly distinguishable in the apparent FRET histograms (Supplementary Note SN5). As expected, the corrected FRET efficiencies for S1 and S2 for all FRET pairs and the efficiency of the BR dye pair are all very similar to the two-state 6-12 o'clock (Figure 6, histogram outlines, Supplementary Tables S1.1 and S2.1). An additional state become observable for S3 with FRET values of 18 % and 13 % for the BY dye pair and both have 19 % for the YR dye pairs in the 6-9-12 and 6-11-12 origamis respectively (Figure 6, c and e, left and right panels).

The dwell time histograms for all states were extracted using the TDP of the apparent FRET efficiencies (Supplementary Figure S5.1). The dwell times can be reasonably approximated with a mono-exponential function yielding dwell times of 0.72 s, 0.98 s and 1.6 s for S1, S2 and S3 respectively. The dwell time for S3 is similar to the dwell times for S1 in the 7.5 nt constructs. Interestingly, the dwell times for S1 and S2 are almost unchanged in the presence of a third state with S1 again having a shorter residency time. We also verified

these results using two-labeled samples of Cy3B and Atto647N (Figure 6d and f). The FRET values and dwell-time distributions are both similar, although more difficult to separate (Supplementary Table S1.1). This supports the findings from the DNA-PAINT measurements and suggests that, upon escape from one of the docking sites, the tether moves to a different position and the probability of returning to the same binding site is lower.

We also investigated the dynamics and anisotropy of the three-state complexes using MFD-PIE (Supplementary Figure S5.2). From the E-tau plots (Supplementary Figures S7.9-S7.10, S7.19-S7.20), the occasional transition between the different states is visible. The anisotropy measurements for the 6-9-12 sample are shown in Supplementary Figure S3.1. All dyes behave similarly to the other constructs. The anisotropy of Cy3B for the 6-9-12 clock is still high, but slightly decreased from that of the 6-12 clock. This supports the conclusion that, in close proximity, Atto647N has an influence on Cy3B. With the additional binding site further away from Atto647N, the overall anisotropy is lowered. In summary, the results from MFD-PIE experiments corroborate those measured with immobilized molecules on the surface.

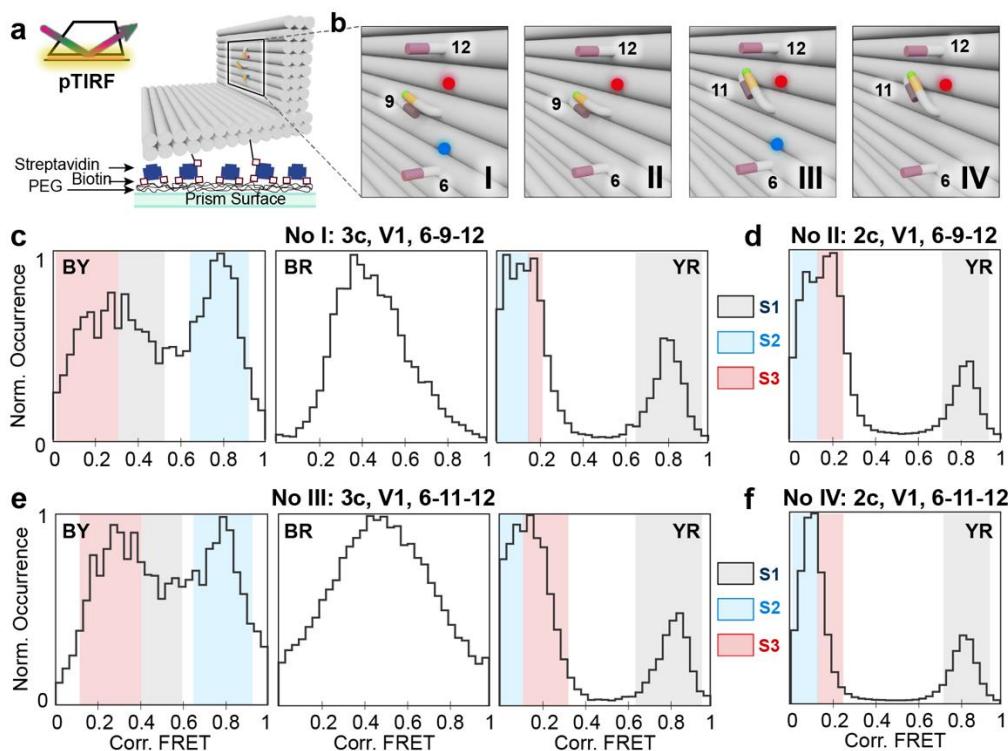


Figure 6. Single-molecule TIRF results for the three- and two-color, L-shaped DNA origami structures with three complementary docking strands. (a) The schematic showing the smTIRF assay for measuring immobilized DNA origami structures on the passivated prism surface. (b) Zoomed-in schematics of the four samples depicting the position and overhang of the docking strands and location of the fluorophores: Cy3B on the tether, Atto647N on the top and Atto488 at the bottom (in case of three-color samples). The three complementary single strands are located at 6, 9 and 12 o'clock positions (Structure I and II) and 6, 11 and 12 o'clock positions (Structure III and IV). Binding sites at position 6 and 12 share 7 matching nucleotides (highlighted in purple), while protruding strands at 9 and 11 o'clock share 7.5 nt matching nucleotides (highlighted in salmon red). (c) Frame-wise-weighted, state-wise averaged corrected FRET efficiency histograms from smTIRF experiments shown for the BY (left), BR (middle) and YR (right) FRET pairs on the three-color 6-9-12-state nanostructure (No. I). (d) Frame-wise weighted, state-wise averaged corrected FRET efficiency histogram for the YR dye pair on the two-color 6-9-12-state nanostructure (No. II). (e) Frame-wise-weighted, state-wise averaged corrected FRET efficiency histograms for the BY (left), BR (middle) and YR (right) FRET pairs on the three-color 6-11-12-state nanostructures (No. III). (f) Frame-wise weighted, state-wise averaged corrected FRET efficiency histogram for the YR dye pair on the two-color 6-11-12-state nanostructure (No. IV).

SmFRET experiments on DNA origami structures with three, symmetrically distributed docking strands

We have observed that the presence and absence of Atto647N influences the kinetics and that the local environment of the docking strands has an effect. Hence, we continued our investigation regarding the precision with which the binding kinetics can be controlled by designing a DNA origami structure with a symmetric distribution of 7 nt binding sites (1-5-9, Figure 7). Due to the influence of Atto674N on the binding kinetics, we replaced Atto647N with Cy5. We also checked the influence of the fluorophore on the tether by comparing the kinetics of Cy3B with that of Atto542. The results for the two structures were measured both on the surface and in solution.

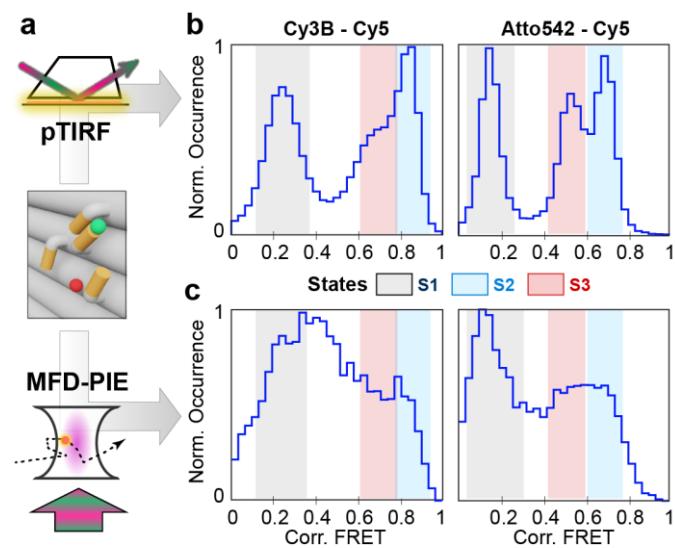


Figure 7. Single-molecule TIRF and MFD-PIE measurements of two-color-labeled L-shaped DNA origami structures carrying a small clock with three symmetrically oriented binding sites. (a) Schematics depicting the small clock design, which consists of a flexible ssDNA tether surrounded by three protruding ssDNA binding strands at the 1, 5 and 9 o'clock position roughly 6 nm apart from each other. The tether holds the donor dye, while the acceptor dye is positioned on the origami surface between the two lower complementary binding sites. The DNA origami structures were either immobilized on a prism surface for pTIRF measurements or investigated as freely diffusing molecules transversing the confocal volume of a MFD-PIE microscope. The colored arrows indicate the alternating laser excitation used in both assays. **(b)** Frame-wise weighted, state-wise averaged corrected FRET efficiency histograms for the two-color, three-state DNA origami samples labeled with Cy3B (left) or Atto542 (right) on the tether and Cy5 as a fixed acceptor. The histograms show three FRET populations. **(c)** Corrected FRET efficiency histograms for the symmetric two-color, three-state DNA origami samples labeled with Cy3B (left) or Atto542 (right) on the tether and Cy5 as a fixed acceptor. The histograms show at least three non-resolvable FRET populations.

The sample labeled with the Cy3B-Cy5 dye pair shows three corrected FRET efficiencies of 24 %, 84 % and 69 % for S1, S2 and S3 in the smTIRF measurements. The smFRET histograms from the MFD-PIE experiments are difficult to fit, but can be described with values 24 %, 80 % and 43 % which are similar at least for S1 and S2. Data from measurements with Atto542 are in better agreement with the corrected FRET efficiencies of 15 %, 70 % and 54 %, and 15 %, 67 % and 50 % for smTIRF and MFD-PIE experiments, respectively. By adjusting the Förster Radii, these measurements can yield very similar distances for the two different dye-pairs. However, the Förster Radii deviate significantly from the expected values. When considering the residual anisotropy of Atto542 (0.07), Cy3B (0.26) and Cy5 (0.40) (Supplementary Figure S3.1c), this could be due to orientational effects of the dyes. Considering the neutral electric charge of Cy3B and the net charge of -3 on Atto542 when they are bound to the flexible tether, the different values of residual anisotropies of the two dyes seem reasonable. Atto542 can rotate more freely compared to Cy3B due to the repulsive forces from the negative back bone of the protruding tether. Cy5 is again a very sticky red fluorophore with the highest residual anisotropy, which could be attributed to its net charge of +1 that attracts it to the negatively charged DNA.

The dwell-time distributions extracted from the TDPs after the smTIRF analysis yield dwell times of 0.60 s, 0.56 s and 0.42 s for the Cy3B - Cy5 construct. The dwell times of the first two states are very similar, but again there is an asymmetry of the third state with rates varying by a factor of 1.4. The value of 0.60 s for the 1 o'clock position corresponds well to the dwell time measured for the S1 state in the two-state clock (6-12) with 7 nt docking strands. When switching to Atto542 on the tether and 8 nt docking strands, all dwell times

increased to 0.86 s, 1.1 s and 0.97 s. The variation is less and the dwell times are faster than for the corresponding original 6-12 clock structure with the Cy3B tether. This suggests that the presence of Atto542 destabilizes the DNA hybridization due to the Coulomb repulsion between the dye and the DNA.

To look more closely into the effect of the fluorophore on the tether, we returned to the original two-state L-shaped DNA origami V1 structure and exchanged the Cy3B fluorophore on the tether for Atto542. A three-color structure (Atto542 on the tether, Atto488 at the 6 o'clock and Atto647N at 12 o'clock position), and two, two-color variants (with the Atto488-Atto542 and Atto542-Atto647N FRET pairs) were investigated. Both docking strands had a complementary base pairing length of 7.5 nt. The smTIRF results of these samples are summarized in Figure 8 and Supplementary Table S1.2.

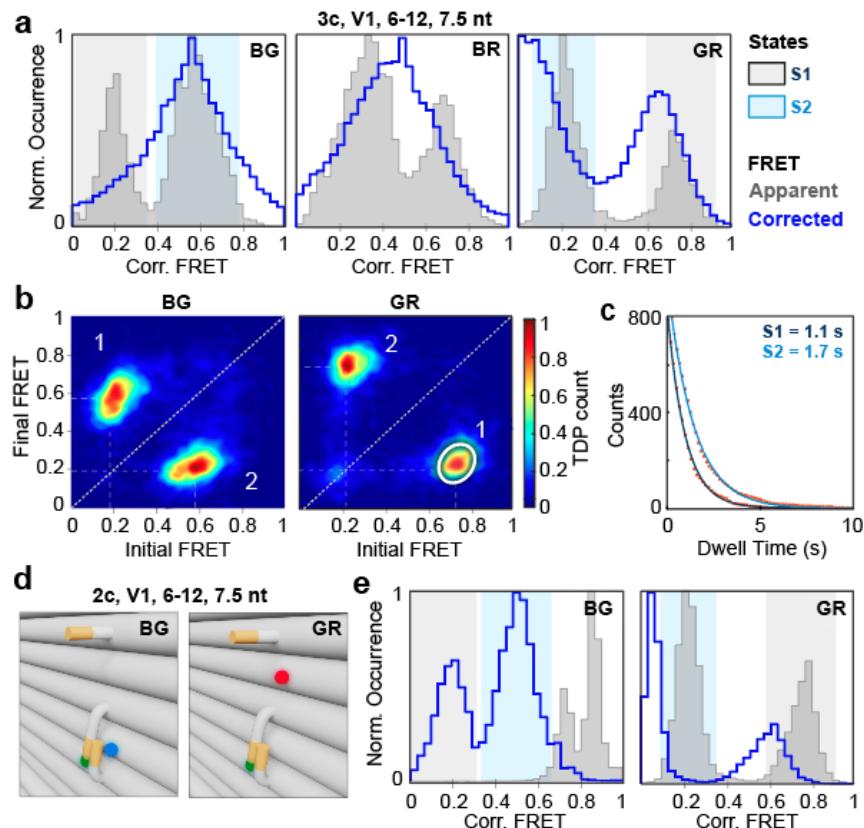


Figure 8. Single-molecule TIRF results for the three- and two-color, two-state V1 L-shaped DNA origami structures with Atto542 on the flexible tether. (a) Frame-wise weighted, state-wise averaged apparent (shadowed histograms) and corrected (outlined histograms) FRET efficiency histograms from smTIRF experiments for the BG (left), BR (middle) and GR (right) dye pairs for the measured structure with 7.5 complementary nucleotides. The dynamic BG and GR pairs exhibit two populations corresponding to the states S1 and S2. After correcting of the FRET efficiency for the BR pair, the histogram merges into a single broad peak. **(b)** Transition density plots showing the interconversion events between the two detected states for the BG (left) and GR (right) dye pairs. **(c)** The cumulative dwell-time distributions, calculated by selecting e.g. clusters from the GR TDP plot in panel b (as highlighted for the S1 \rightarrow S2 transition in white), reveal lifetimes of 1.1 s and 1.7 s for states 1 and 2, respectively, after fitting the data points with mono-exponential curves. **(d)** Zoomed-in schematics of the two-color, two-state origami structures containing either the BG (left) or GR (right) dye pair. **(e)** Frame-wise weighted, state-wise averaged apparent (shadowed histogram) and corrected (outlined histogram) FRET efficiency histograms from smTIRF experiments for the two-color structures containing the BG (left) or GR (right) FRET pairs.

The dynamic BG and GR FRET pairs each show two states with corrected FRET efficiencies of 12 % and 49 % and 64 % and 3 % respectively (Figure 8a, blue outline). The BG FRET populations are difficult to resolve, but the apparent FRET histogram clearly shows the two populations (Figure 8a, shadowed histograms). The static

BR pair shows one population with 44 % FRET efficiency similar to original structures. The TDPs based on the BG and GR apparent FRET efficiencies yield dwell times of 1.1 s and 1.7 s for S1 and S2, respectively. Hence, the asymmetry of the docking kinetics still remains, but the dwell times decrease. S1 becomes even faster than the dwell time for S1 for the two-color variant in the absence of a fluorophore in the vicinity (1.4 s). We also prepared and measured the two-color structures labeled with either the BG or GR FRET pairs Figure 8d,e. The corrected FRET efficiencies are very similar to those measured using the three-color constructs (Supplementary Table S1.2.). The dwell times of the two-color samples for S2 remain unchanged but the dwell time distribution for S1 in the absence of Atto488 becomes shorter (0.86 s). Hence, both the fluorophores on the structure and on the tether as well as the positioning of the docking strands has an influence on the designed kinetics.

Calculating the distances between the dyes for V1 and V2 two- and three-state samples

We also investigated the precision in which the locations of the components can be located. From the corrected FRET efficiencies, we can estimate the distances between the different fluorophores and compare them to the expected distances from the design⁴¹. The designed locations and measured distances are shown in Supplementary Figure S5.3. Equation 1 was used to calculate the distances between different dye pairs:

$$E = \frac{R_0^6}{R^6 + R_0^6} \quad \text{Eq. 1}$$

The parameters in equation 1 include the corrected FRET efficiency, E , the Förster radius, R_0 (the distance at which the FRET efficiency equals to 50 %) and the separation between the transition dipole moments of the donor and acceptor fluorophores, R . The calculated distances for all studied DNA origami nanostructures are summarized in Supplementary Tables S1.1 and S1.2. The measured separation between the blue and red dyes on all origami structures (~ 5.0 nm) are very consistent and match well with the expected separation of 6 nm. The distance between the tether binding to the 6 and 12 o'clock positions are very reproducible between all structures with separations of 6.92 ± 0.29 nm and 5.33 ± 0.08 nm to S2 for the blue and yellow dyes respectively for V1. Upon switching the blue and red dyes (V2), the distance between the tether and the blue dye is consistent whereas the separation between the tether (yellow) and the red dye varies a bit. However, the anisotropy values for both of these dyes are high with Atto647N showing almost no rotational averaging. As shown by Agam et al.³⁹, a high combined residual anisotropy for various dye pairs leads to higher absolute uncertainty in the distance between them. The fact that the red Atto dyes stick to origami surfaces has been demonstrated previously^{10, 40}. Hence, the small discrepancy is not surprising. Still, the determined distances are reproducible between the different structures and both V1 and V2 are within ca 1.0 - 1.5 nm in distance from each other.

For the three-state DNA origami structures 6-9-12 and 6-11-12, the distances to positions 6 and 12 are very consistent with each other and with the two-state structures. The distance to the 9 o'clock position also similar to what one expects from the design. In the case of the 6-11-12 structure, the 11 o'clock position is not what was expected for this position and is very similar to that of the 9 o'clock position. In summary, although DNA origami nanostructures are highly tunable and easy to engineer, there are uncertainties in the designed positioning and kinetic rates that can depend on the proximity of various fluorophores as well as the local environment.

Conclusions

DNA origami nanostructures are highly programmable structures with the ability to place molecules and tune kinetics with high precision. Such nano-devices are becoming advanced tools for biophysical studies, biosensing and diagnostics. As reference structures, DNA origamis are commonly used for super resolution microscopy but also offer advantages for FRET probes with high control over distances and kinetics ²⁹. To understand the degree of control we studied factors that influence the accuracy and precision of the programmable structures to allow researchers to construct finely-tuned nano-devices for their specific questions and tasks. As a test system, we choose the well-studied L-shaped DNA origami structure with a tether that fluctuates between two docking strands. The kinetics of these fluctuations can be tuned by adjusting the number of base pairings with the docking strands. The addition of mismatch allows additional fine-tuning of the kinetics by almost a factor of two.

Our study indicates that fluorophores, in the vicinity of the docking strand as well as on the tether, have a noticeable effect on the kinetic rate of switching. The presence of Atto647N near the docking strand stabilizes the hybridization. In the absence of Atto647N, we observed a decrease in the average dwell time for both the V1 and V2 structures. Also, the choice of fluorophore on the tether alters the kinetic rates. When using Atto542, which has a net charge of -3, the dwell times decrease as compared to the neutral Cy3B. This can be due to destabilization of the interaction via electrostatic repulsion of the fluorophore from the negatively charged origami structure. The hydrophobic and positively charged red dyes can also interact with the DNA origami surface. We observed strong sticking of Atto647N via the very high anisotropy values and slow decay times. Although interaction and influences of fluorophores on the kinetic rates is not unexpected, it is important to have an understanding of how strong these influences are.

However, even when completely removing fluorophores from the origami surface in the vicinity of the docking strands, it was not possible to obtain the same dwell times for the same sequences. When trying different locations, similar dwell times could be obtained in some positions but not all. Hence, there are still local factors such as the angle at which the docking strand leaves the structure and the sequence in the immediate vicinity that can influence the kinetic rates. In the case of the L-shaped origami, we found no bias in the average position of the tether in the absence of the docking strands. The kinetics of the system can be approximated with a mono-exponential. However, for measurements with high statics, a slightly non-exponential distribution of dwell times is observed suggesting different environments must be present.

What is very interesting is that the dwell time distributions did not change significantly when adding an additional state to the system. Initially, it was expected that the DNA duplex undergoes multiple dissociation and binding reactions that are too quickly to be observed in the smFRET experiments due to the high effective local concentration of the tether before switching to the other docking strand. Still, if another binding position is present, this should have an influence. When looking at the dwell times of the different docking strands using DNA-PAINT like experiments under identical buffer conditions in the absence of the tether, the dissociation rates did not greatly change. The asymmetry for the two states was still present indicating that the difference is due to differences in the docking strands themselves. The fact that the dissociation rates do not change, indicates that the tether switches location upon unbinding. Another way of saying it is that the probability of rebinding to the same state after dissociation is very low. One possible explanation is that, upon hybridization and formation of the double-strand DNA complex, the tether pulls the neighboring docking strands towards tether location. Hence, both the tether and docking strands are pulled out of their equilibrium position. Upon dissociation, they recoil such that the tether no longer returns to that previous state but binds

preferably to another state. The fact that the tether switches between states upon dissociation makes it possible to use this construct as a clock for various applications.

In summary, the programmability and tunability of DNA origami structures have led to their application in many fields from biophysical studies to diagnostics and drug delivery. However, when incorporating additional biomolecules into the origami structure, the influence of these molecules and their geometries on the functionality of the nano-device needs to be considered. Although many of these effects are small, a better understanding of these interactions allows for more precise designs. However, the necessary precision depends on the questions being addressed, the application and overall picture of the whole system⁴².

Materials and Methods

Chemicals

Chemicals were purchased from Sigma-Aldrich and used without further purification if not stated otherwise. Chemicals include acetic acid, agarose, ammonium persulfate, (3-aminopropyl-) triethoxysilane (APTES), biotin-poly(ethylene glycol)-silane (biotin-PEG, MW3000, PG2-BNSL-3k, Nanocs, NY; USA), bovine serum albumin (BSA; New England Biolabs, Ipswich, Ma, USA), Blue Juice gel loading buffer (ThermoFisher Scientific), ethylenediamine-tetraacetic acid sodium salt dihydrate (EDTA-Na₂ × 2H₂O), glycerol, magnesium chloride (MgCl₂ × 6H₂O), 2-[methoxy(polyethyleneoxy)propyl]trimethoxy-silane (mPEG, #AB111226, abcr; Germany), phosphate-buffered saline (PBS), protocatechuate 3,4-dioxygenase from *Pseudomonas* sp. (PCD), protocatechuic acid (PCA), streptavidin, sodium chloride, Tris base, Tris HCl, and 6-hydroxy-2,5,7,8-tetramethylchroman-2-carboxylic acid (Trolox) and beta-mercaptoethanol (βME).

All unmodified staple strands used for DNA origami structure folding are commercially available and were purchased from Integrated DNA Technologies®. Staple strands with modifications were obtained from Biomers (containing Biotin and Atto488) and Eurofins Genomics (with binding sites for Cy3b, Atto542 and Atto647N). The imager strands were purchased from Ella Biotech (with Cy3B at the 3' end). Sequences are provided in Supplementary Information S9.

Multi-color single-molecule TIRF setup for smFRET and PAINT measurements

Single-pair FRET experiments on surface-immobilized molecules were carried out on a home-built prism-type TIRF microscope as previously published⁴³. Four laser sources (Cobolt, Solna; Sweden) at 491 nm, 532 nm, 561 nm, and 640 nm are available and used for different combinations of two- and triple-color TIRF experiments. The resulting emission was collected by a 60× water immersion objective (60×/1.27 WI Plan Apo IR, Nikon) for prism-TIRF then cleaned up using notch filters (Stopline® Notch 488/647, AHF; Stopline® Notch 532, AHF). The red emission was separated from the blue/yellow emission by a dichroic mirror (630DCXR AHF; Germany) followed by separation of the blue and green/yellow emission (560DCXR AHF). The emission was spectrally filtered (AHF Analysentechnik, Tubingen, Germany) for the blue (ET525/50), yellow (HQ595/50) and red (ET685/40) collection channels and detected on three EMCCD cameras (Andor iXon (1×)/iXon Ultra (2×), Andor Technologies, Belfast; UK). For measurements with Atto542, a HQ585/65 emission filter was used. Data collection was done using the supplier's software Andor Solis (Version 4.29.30005.0; Oxford Instruments). All two- and three-color FRET experiments were carried out using msALEX⁴⁴, i.e. two- or three-excitation lasers were alternated framewise. Synchronization and alternation of the excitation laser sources as well as the frame-wise data acquisition on three separate cameras were achieved using a LabView-written program that controls a field programmable gate array (FPGA; NI cRIO-9073). While the program starts the measurement, the FPGA synchronizes the execution of the hardware via TTL pulses, i.e. it controls the excitation sources by direct modulation of the AOTF (491, 532/561, 640 nm) while simultaneously starting the data acquisition of the three cameras. For experiments on constructs with 6.5 nt, 7 nt, 7.5 nt and 8 nt complementary lengths of the docking strands, exposure times of 30 ms, 50 ms, 50 ms and 80 ms were used respectively and a 2.2 ms frame transfer time. Movies of 2000 (two-color) and 2400 (three-color) frames were collected. The laser powers were set to 28 mW (0.022 mm², 491 nm), 15 mW (0.011 mm², 532 nm), 16 mW (0.040 mm², 561 nm) and 10 mW (0.022 mm², 640 nm). The videos were analyzed afterward by a custom-written MATLAB program, Deep-LASI⁴⁵.

Single-molecule TIRF experiments of multi-color-labeled DNA Origami structures

Flow channels for single-molecule experiments were formed by sandwiching a pre-cut Nesco film (Nesco) channel between a coverslip and the surface-passivated custom-made quartz prism followed by melting the Nesco film to properly seal the chamber. The prism surface had been coated before with a biotin-PEG/mPEG layer to achieve surface passivation and prevent unspecific binding. The prism channel was first flushed with PBS and then incubated with a streptavidin solution (0.2 mg/mL) for 15 min. The channel was then washed 3× with PBS to remove unbound streptavidin. Afterwards, it was washed with the storage buffer suitable for the origamis stability (1× TAE, 12.5 mM MgCl₂ and 1 mM EDTA, pH=8.2). The DNA origami sample was added to the chamber after diluting to 40-50 pM with the storage buffer and immobilized to the prism surface via the biotin-streptavidin linkage. After 10-minute incubation, free origami structures were removed by rinsing the chamber 3× with the storage buffer.

To photo-stabilize the attached fluorophores on the DNA origami structure, an oxygen scavenging system based on PCA/PCD was used⁴⁶. The photo-stabilization buffer was prepared as follows: 1 µL of 100 mM Trolox/Ethanol solution was added to 97 µL storage buffer. The mixture was then aged by 254 and 366 nm UV radiation using a UV lamp (M&S Laborgeräte GmbH, UVAC-6U, 2 × 6 W) until an equal ratio of Trolox and Trolox-quinone was formed (typically around 6 minutes)⁴⁷. The ratio was determined using the two absorbance peaks of Trolox and Trolox-quinone measured on a nanodrop spectrophotometer (PeqLab Biotechnologie GmbH). Immediately before starting the TIRF experiments, 1 µL of 100 mM PCA in methanol and 1 µL of 100 mM PCD solution (50 % glycerol, 50 mM KCl, 100 mM Tris HCl, 1 mM EDTA-Na₂ × 2H₂O, pH=8) were added to the mixture of aged Trolox and storage buffer to obtain the photo-stabilization (imaging) buffer. When measuring DNA origami samples containing Atto488, 1 µL of 14.3 M βME was also added to the photo-stabilization buffer. In this case, Trolox was added to 96 µL of storage buffer to keep the same concentration. After flushing the sample chamber with the photo-stabilization buffer, it was sealed with a removable silicon sheet. A minimum of 5-minute waiting time preceded the data recordings. The photo-stabilization buffer was refreshed every 45 minutes until the end of the experiment.

Single-molecule TIRF combined with FRET-PAINT

For the DNA-PAINT experiments, the prism surface was passivated and the sample prepared as described above. The experiments were performed as described above for the smFRET experiments with the difference that, immediately before starting the TIRF experiments, 1 µL of 0.5 µM imager strand labeled with Cy3B was added to the photostabilization buffer. The FRET experiments were carried out using non-ALEX mode, i.e. only the 561 laser was used continuously for continuous illumination. An exposure time of 80 ms for a total of 1800 frames were used. The laser power was set to 15 mW.

Dual- and Triple-color MFD-PIE setup

Solution-based FRET experiments of freely diffusing DNA origami structures labeled with blue and/or yellow fluorophores were performed using a custom-built confocal setup as previously published⁴⁸. Briefly, samples were excited using pulsed-interleaved excitation at laser wavelengths of 485 nm⁴⁹, 560 nm and 640 nm (LDH-D-C-485, LDH-D-TA-560, LDH-D-C-640, PicoQuant, Berlin; Germany) with a 16.67 MHz alternation period for 2C (with one deactivated laser line) or 3C experiments. The average excitation intensities for dual- or triple-color measurements before the objective were 87 µW, 29 µW and 42 µW for the blue, yellow and red lasers respectively. The resulting fluorescence was collected by a 60x water immersion objective (Plan Apo IR 60x 1.27 WI, Nikon, Düsseldorf, Germany) and spectrally separated from the excitation beam by a polychroic mirror (zt405/488/561/633, AHF; Analysentechnik, Tübingen, Germany). The emission was first separated

using a polarizing beamsplitter (Thorlabs, Dachau, Germany) and then spectrally filtered (AHF Analysentechnik, Tübingen, Germany) for the blue (ET525/50), yellow (ET607/36) and red (ET670/30) detection range. Photons were then collected using photon counting avalanche photodiodes (2x COUNT-100B, LaserComponents, Olching, Germany; 4x SPCM- AQR-14, PerkinElmer, USA) and registered by independent but synchronized time-correlated single photon counting hardware (HydraHarp400, PicoQuant). The detector signal was recorded using a home-written program in C#.

Single-molecule two- and three-color FRET measurement in solution

All samples were measured in 8-well chamber slides (Nunc Lab-Tek, VWR) passivated with 1 mg/mL BSA in PBS for 10 min before each measurement. MFD-PIE experiments were carried out on freely diffusing DNA origami structures in the storage buffer (1x TAE, 12.5 mM MgCl₂ and 1 mM EDTA, pH=8.2) at concentrations of 10-30 pM at room temperature (23°C)³⁶. Each sample was recorded for 2 h or 3-4 h for dual- and triple-labeled DNA origami structures respectively. Data were analyzed using a home-written MatLAB based software, PIE Analysis with MatLAB⁵⁰.

Data analysis of smFRET TIRF data

For each detected molecule in the smTIRF assay, a time trace with the intensity and FRET signals are extracted. To perform accurate FRET efficiency calculations, correction factors for spectral crosstalk of the donor, direct excitation of the acceptor and difference in detection sensitivity need to be determined. For these, we used the procedure given in Hellenkamp et al.⁵¹. When possible, corrections extracted from individual traces were applied to that trace otherwise the median value from the measurement was applied. Values for the used correction factors from our smFRET TIRF measurements are given in **Supplementary Note S1**. The details of these steps are described in our previous publications^{35, 45}.

Data from the FRET-PAINT experiments were extracted using DeepLASI software after colocalizing the transient PAINT signal of the Cy3B-labeled imager strand (when attached to any of the docking strands) with the DNA origami labeled with Atto647N at the 12 o'clock position. Signals originating from the colocalizing molecules were extracted in the DD (donor) and DA (FRET) channels and sorted according to high and low FRET, respectively. Traces were analyzed manually. The underlying kinetics were analyzed by HMM and TDP.

In the solution experiments, bursts were extracted using an all photon burst search³⁴. Photons were considered by the algorithm to belong to a molecular burst when the local count rate within a sliding window exceeded a minimum total number of 50 photons having 30 consecutive photons within a time interval of 500 microseconds. Fluorescence photons arriving at the three detection channels (blue: B; yellow: Y; and red: R) were sorted according to the three excitation periods (blue: B; yellow: Y; and red: R) based on the photon arrival time. The apparent FRET efficiency between any of the two dyes is calculated from the measured intensities and corrected to provide accurate FRET values⁴⁸. The correction factors for the MFD-PIE experiments are given in **Supplementary Table S3**. The burst-wise fluorescence lifetimes were fit using one-, two- or three-exponentials depending on how many components were present in the data. Time resolved anisotropy was fit using a single exponential.

Author Contributions

P.A. collected the single-molecule TIRF and MFD data, performed the corresponding data analysis, wrote the first draft of the manuscript, and designed the figures. F.M. aided in collecting data. J.B. and F.C. prepared the DNA origami samples under the supervision of P.T. All authors contributed to revising the manuscript. E.P. and D.C.L. supervised the project.

Acknowledgments

We thankfully acknowledge the financial support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 201269156 – SFB 1032 Project B03 (to D.C.L.), Project A13 (to P.T.), individual grants to PL696/4-1 (to E.P.), TI 329/9-2, project number 267681426, TI 329/14-1, and TI 329/15-1 (to P.T.) and Germany's Excellence Strategy – EXC 089/1 – 390776260. D.C.L. and P.T. gratefully acknowledge funding from the Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the Federal Government and the Länder through the ONE MUNICH Project Munich Multiscale Biofabrication. P.T. acknowledges the support of BMBF (SIBOF, 03VP03891). D.C.L., P.T., and E.P. gratefully acknowledge the financial support of the Ludwig-Maximilians-Universität München via the Department of Chemistry, the Center for NanoScience (CeNS) and the LMUinnovativ program BioImaging Network (BIN).

References

- (1) Shi, X.; Pumm, A.-K.; Maffeo, C.; Kohler, F.; Feigl, E.; Zhao, W.; Verschueren, D.; Golestanian, R.; Aksimentiev, A.; Dietz, H.; et al. A DNA turbine powered by a transmembrane potential across a nanopore. *Nature Nanotechnology* **2024**, *19* (3), 338-344. DOI: 10.1038/s41565-023-01527-8. Harashima, T.; Otomo, A.; Iino, R. Rational engineering of DNA-nanoparticle motor with high speed and processivity comparable to motor proteins. *Nature Communications* **2025**, *16* (1), 729. DOI: 10.1038/s41467-025-56036-0.
- (2) Loretan, M.; Domljanovic, I.; Lakatos, M.; Ruegg, C.; Acuna, G. P. DNA Origami as Emerging Technology for the Engineering of Fluorescent and Plasmonic-Based Biosensors. *Materials (Basel)* **2020**, *13* (9). DOI: 10.3390/ma13092185. Kogikoski, S., Jr.; Ameixa, J.; Mostafa, A.; Bald, I. Lab-on-a-DNA origami: nanoengineered single-molecule platforms. *Chem Commun (Camb)* **2023**, *59* (32), 4726-4741. DOI: 10.1039/d3cc00718a.
- (3) Rothemund, P. W. Folding DNA to create nanoscale shapes and patterns. *Nature* **2006**, *440* (7082), 297-302. Douglas, S. M.; Dietz, H.; Liedl, T.; Högberg, B.; Graf, F.; Shih, W. M. Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **2009**, *459* (7245), 414-418. Han, D.; Pal, S.; Nangreave, J.; Deng, Z.; Liu, Y.; Yan, H. DNA origami with complex curvatures in three-dimensional space. *Science* **2011**, *332* (6027), 342-346. Ke, Y.; Ong, L. L.; Shih, W. M.; Yin, P. Three-dimensional structures self-assembled from DNA bricks. *science* **2012**, *338* (6111), 1177-1183.
- (4) Kim, M.; Lee, C.; Jeon, K.; Lee, J. Y.; Kim, Y.-J.; Lee, J. G.; Kim, H.; Cho, M.; Kim, D.-N. Harnessing a paper-folding mechanism for reconfigurable DNA origami. *Nature* **2023**, *619* (7968), 78-86. DOI: 10.1038/s41586-023-06181-7.
- (5) Centola, M.; Poppleton, E.; Ray, S.; Centola, M.; Welty, R.; Valero, J.; Walter, N. G.; Šulc, P.; Famulok, M. A rhythmically pulsing leaf-spring DNA-origami nanoengine that drives a passive follower. *Nature Nanotechnology* **2024**, *19* (2), 226-236. DOI: 10.1038/s41565-023-01516-x. Kogikoski, S., Jr.; Gonçalves Pontes, R.; Paschoalino, W. J.; Kubota, L. T. Opportunities for DNA Origami in Electrochemical Sensing Applications. *ACS Electrochemistry* **2025**. DOI: 10.1021/acselectrochem.4c00158. Büber, E.; Yaadav, R.; Schröder, T.; Franquelim, H. G.; Tinnefeld, P. DNA Origami Vesicle Sensors with Triggered Single-Molecule Cargo Transfer. *Angewandte Chemie International Edition* **2024**, *63* (49), e202408295. DOI: <https://doi.org/10.1002/anie.202408295> (accessed 2025/03/05).
- (6) Ijäs, H.; Nummelin, S.; Shen, B.; Kostainen, M. A.; Linko, V. Dynamic DNA Origami Devices: from Strand-Displacement Reactions to External-Stimuli Responsive Systems. In *International Journal of Molecular Sciences*, 2018; Vol. 19.
- (7) Yurke, B.; Turberfield, A. J.; Mills Jr, A. P.; Simmel, F. C.; Neumann, J. L. A DNA-fuelled molecular machine made of DNA. *Nature* **2000**, *406* (6796), 605-608. Seelig, G.; Soloveichik, D.; Zhang, D. Y.; Winfree, E. Enzyme-free nucleic acid logic circuits. *science* **2006**, *314* (5805), 1585-1588.
- (8) Grabenhorst, L.; Pfeiffer, M.; Schinkel, T.; Kümmerlin, M.; Brüggenthal, G. A.; Maglic, J. B.; Selbach, F.; Murr, A. T.; Tinnefeld, P.; Glembockyte, V. Engineering modular and tunable single-molecule sensors by decoupling sensing from signal output. *Nature Nanotechnology* **2025**, *20* (2), 303-310. DOI: 10.1038/s41565-024-01804-0.
- (9) Andersen, E. S.; Dong, M.; Nielsen, M. M.; Jahn, K.; Subramani, R.; Mamdouh, W.; Golas, M. M.; Sander, B.; Stark, H.; Oliveira, C. L. Self-assembly of a nanoscale DNA box with a controllable lid. *Nature* **2009**, *459* (7243), 73-76. Ijas, H.; Hakaste, I.; Shen, B.; Kostainen, M. A.; Linko, V. Reconfigurable DNA Origami Nanocapsule for pH-Controlled Encapsulation and Display of Cargo. *ACS Nano* **2019**, *13* (5), 5959-5967. DOI: 10.1021/acsnano.9b01857. Pumm, A. K.; Engelen, W.; Kopperger, E.; Isensee, J.; Vogt, M.; Kozina, V.; Kube, M.; Honemann, M. N.; Bertosin, E.; Langecker, M.; et al. A DNA origami rotary ratchet motor. *Nature* **2022**, *607* (7919), 492-498. DOI: 10.1038/s41586-022-04910-y. Buber, E.; Schröder, T.; Scheckenbach, M.; Dass, M.; Franquelim, H. G.; Tinnefeld, P. DNA Origami Curvature Sensors for Nanoparticle and Vesicle Size Determination with Single-Molecule FRET Readout. *ACS Nano* **2023**, *17* (3), 3088-3097. DOI: 10.1021/acsnano.2c11981.
- (10) Hubner, K.; Raab, M.; Bohlen, J.; Bauer, J.; Tinnefeld, P. Salt-induced conformational switching of a flat rectangular DNA origami structure. *Nanoscale* **2022**, *14* (21), 7898-7905. DOI: 10.1039/d1nr07793g.

(11) Severgnini, M.; Cremonesi, P.; Consolandi, C.; Caredda, G.; De Bellis, G.; Castiglioni, B. ORMA: a tool for identification of species-specific variations in 16S rRNA gene and oligonucleotides design. *Nucleic acids research* **2009**, *37* (16), e109-e109. Choi, H. M.; Beck, V. A.; Pierce, N. A. Next-generation in situ hybridization chain reaction: higher gain, lower cost, greater durability. *ACS nano* **2014**, *8* (5), 4284-4294.

(12) Jiang, Q.; Shang, Y.; Xie, Y.; Ding, B. DNA origami: from molecular folding art to drug delivery technology. *Advanced Materials* **2024**, *36* (22), 2301035. Kumar, M.; Jha, A.; Mishra, B. DNA-based nanostructured platforms as drug delivery systems. *Chem & Bio Engineering* **2024**, *1* (3), 179-198.

(13) Zhou, Y.; Dong, J.; Wang, Q. Fabricating higher-order functional DNA origami structures to reveal biological processes at multiple scales. *NPG Asia Materials* **2023**, *15* (1), 25. DOI: 10.1038/s41427-023-00470-3.

(14) Tayar, A. M.; Hagan, M. F.; Dogic, Z. Active liquid crystals powered by force-sensing DNA-motor clusters. *Proceedings of the National Academy of Sciences* **2021**, *118* (30), e2102873118. DOI: 10.1073/pnas.2102873118 (acccesed 2025/03/05).

(15) Raab, M.; Jusuk, I.; Molle, J.; Buhr, E.; Bodermann, B.; Bergmann, D.; Bosse, H.; Tinnefeld, P. Using DNA origami nanorulers as traceable distance measurement standards and nanoscopic benchmark structures. *Scientific Reports* **2018**, *8* (1), 1780. DOI: 10.1038/s41598-018-19905-x.

(16) Williams, N. D.; Landajuela, A.; Kasula, R. K.; Zhou, W.; Powell, J. T.; Xi, Z.; Isaacs, F. J.; Berro, J.; Toomre, D.; Karatekin, E.; et al. DNA-Origami-Based Fluorescence Brightness Standards for Convenient and Fast Protein Counting in Live Cells. *Nano Letters* **2020**, *20* (12), 8890-8896. DOI: 10.1021/acs.nanolett.0c03925. Chu, J.; Ejaz, A.; Lin, K. M.; Joseph, M. R.; Coraor, A. E.; Drummond, D. A.; Squires, A. H. Single-molecule fluorescence multiplexing by multi-parameter spectroscopic detection of nanostructured FRET labels. *Nature Nanotechnology* **2024**, *19* (8), 1150-1157. DOI: 10.1038/s41565-024-01672-8.

(17) Kamińska, I.; Bohlen, J.; Yaadav, R.; Schüler, P.; Raab, M.; Schröder, T.; Zähringer, J.; Zielonka, K.; Krause, S.; Tinnefeld, P. Graphene Energy Transfer for Single-Molecule Biophysics, Biosensing, and Super-Resolution Microscopy. *Advanced Materials* **2021**, *33* (24), 2101099. DOI: <https://doi.org/10.1002/adma.202101099> (acccesed 2025/03/05).

(18) Dong, R.; Aksel, T.; Chan, W.; Germain, R. N.; Vale, R. D.; Douglas, S. M. DNA origami patterning of synthetic T cell receptors reveals spatial control of the sensitivity and kinetics of signal activation. *Proceedings of the National Academy of Sciences* **2021**, *118* (40), e2109057118. DOI: 10.1073/pnas.2109057118 (acccesed 2025/03/05). Kern, N.; Dong, R.; Douglas, S. M.; Vale, R. D.; Morrissey, M. A. Tight nanoscale clustering of Fcγ receptors using DNA origami promotes phagocytosis. *eLife* **2021**, *10*, e68311. DOI: 10.7554/eLife.68311.

(19) Strauss, M. T.; Schueder, F.; Haas, D.; Nickels, P. C.; Jungmann, R. Quantifying absolute addressability in DNA origami with molecular resolution. *Nature Communications* **2018**, *9* (1), 1600. DOI: 10.1038/s41467-018-04031-z.

(20) Wassermann, L. M.; Scheckenbach, M.; Baptist, A. V.; Glembockyte, V.; Heuer-Jungemann, A. Full Site-Specific Addressability in DNA Origami-Templated Silica Nanostructures. *Advanced Materials* **2023**, *35* (23), 2212024. DOI: <https://doi.org/10.1002/adma.202212024> (acccesed 2025/03/05).

(21) Song, L.; Zhuge, Y.; Zuo, X.; Li, M.; Wang, F. DNA Walkers for Biosensing Development. *Advanced Science* **2022**, *9* (18), 2200327. DOI: <https://doi.org/10.1002/advs.202200327> (acccesed 2025/03/05).

(22) Daljit Singh, J. K.; Luu, M. T.; Abbas, A.; Wickham, S. F. J. Switchable DNA-origami nanostructures that respond to their environment and their applications. *Biophysical Reviews* **2018**, *10* (5), 1283-1293. DOI: 10.1007/s12551-018-0462-z.

(23) Thubagere, A. J.; Li, W.; Johnson, R. F.; Chen, Z.; Doroudi, S.; Lee, Y. L.; Izatt, G.; Wittman, S.; Srinivas, N.; Woods, D.; et al. A cargo-sorting DNA robot. *Science* **2017**, *357* (6356), eaan6558. DOI: 10.1126/science.aan6558 (acccesed 2025/03/05). Amir, Y.; Ben-Ishay, E.; Levner, D.; Ittah, S.; Abu-Horowitz, A.; Bachelet, I. Universal computing by DNA origami robots in a living animal. *Nature Nanotechnology* **2014**, *9* (5), 353-357. DOI: 10.1038/nnano.2014.58.

(24) Moreira, B. G.; You, Y.; Owczarzy, R. Cy3 and Cy5 dyes attached to oligonucleotide terminus stabilize DNA duplexes: predictive thermodynamic model. *Biophys Chem* **2015**, *198*, 36-44. DOI: 10.1016/j.bpc.2015.01.001.

(25) Jahnke, K.; Grubmuller, H.; Igaev, M.; Gopfrich, K. Choice of fluorophore affects dynamic DNA nanostructures. *Nucleic Acids Res* **2021**, *49* (7), 4186-4195. DOI: 10.1093/nar/gkab201.

(26) Hartmann, A.; Krainer, G.; Schlierf, M. Different fluorophore labeling strategies and designs affect millisecond kinetics of DNA hairpins. *Molecules* **2014**, *19* (9), 13735-13754. DOI: 10.3390/molecules190913735.

(27) Gao, Y.; Wolf, L. K.; Georgiadis, R. M. Secondary structure effects on DNA hybridization kinetics: a solution versus surface comparison. *Nucleic Acids Res* **2006**, *34* (11), 3370-3377. DOI: 10.1093/nar/gkl422. Jungmann, R.; Steinhauer, C.; Scheible, M.; Kuzyk, A.; Tinnefeld, P.; Simmel, F. C. Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami. *Nano Lett* **2010**, *10* (11), 4756-4761. DOI: 10.1021/nl103427w.

(28) Kaminska, I.; Bohlen, J.; Yaadav, R.; Schuler, P.; Raab, M.; Schroder, T.; Zahringer, J.; Zielonka, K.; Krause, S.; Tinnefeld, P. Graphene Energy Transfer for Single-Molecule Biophysics, Biosensing, and Super-Resolution Microscopy. *Adv Mater* **2021**, *33* (24), e2101099. DOI: 10.1002/adma.202101099. Krause, S.; Ploetz, E.; Bohlen, J.; Schuler, P.; Yaadav, R.; Selbach, F.; Steiner, F.; Kaminska, I.; Tinnefeld, P. Graphene-on-Glass Preparation and Cleaning Methods Characterized by Single-Molecule DNA Origami Fluorescent Probes and Raman Spectroscopy. *ACS Nano* **2021**, *15* (4), 6430-6438. DOI: 10.1021/acsnano.0c08383.

(29) Scheckenbach, M.; Bauer, J.; Zähringer, J.; Selbach, F.; Tinnefeld, P. DNA origami nanorulers and emerging reference structures. *APL Materials* **2020**, *8* (11). DOI: 10.1063/5.0022885.

(30) Hyeon, C.; Lee, J.; Yoon, J.; Hohng, S.; Thirumalai, D. Hidden complexity in the isomerization dynamics of Holliday junctions. *Nat. Chem.* **2012**. DOI: 10.1038/nchem.1463.

(31) Matsubara, H.; Fukunaga, H.; Saito, T.; Ikezaki, K.; Iwaki, M. A Programmable DNA Origami Nanospring That Reports Dynamics of Single Integrin Motion, Force Magnitude and Force Orientation in Living Cells. *ACS Nano* **2023**. DOI: 10.1021/acsnano.2c12545.

(32) Yuchao, C.; Xiangdan, M.; Huiting, L.; Haifeng, D. Engineering DNA walkers for bioanalysis. *Analytica Chimica Acta* **2022**, *1209*. DOI: 10.1016/j.aca.2021.339339.

(33) Bauer, J.; Reichl, A.; Tinnefeld, P. Kinetic Referencing Allows Identification of Epigenetic Cytosine Modifications by Single-Molecule Hybridization Kinetics and Superresolution DNA-PAINT Microscopy. *ACS Nano* **2024**, *18* (2), 1496-1503. DOI: 10.1021/acsnano.3c08451.

(34) Nir, E.; Michalet, X.; Hamadani, K. M.; Laurence, T. A.; Neuhauser, D.; Kovchegov, Y.; Weiss, S. Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. *J Phys Chem B* **2006**, *110* (44), 22103-22124. DOI: 10.1021/jp063483n.

(35) Wanninger, S.; Asadiatouei, P.; Bohlen, J.; Salem, C. B.; Tinnefeld, P.; Ploetz, E.; Lamb, D. C. Deep-LASI: deep-learning assisted, single-molecule imaging analysis of multi-color DNA origami structures. *Nat Commun* **2023**, *14* (1), 6564. DOI: 10.1038/s41467-023-42272-9.

(36) Kudryavtsev, V.; Sikor, M.; Kalinin, S.; Mokranjac, D.; Seidel, C. A.; Lamb, D. C. Combining MFD and PIE for accurate single-pair Forster resonance energy transfer measurements. *Chemphyschem* **2012**, *13* (4), 1060-1078. DOI: 10.1002/cphc.201100822.

(37) Gietl, A.; Holzmeister, P.; Grohmann, D.; Tinnefeld, P. DNA origami as biocompatible surface to match single-molecule and ensemble experiments. *Nucleic Acids Res.* **2012**, *40* (14). DOI: 10.1093/nar/gks326.

(38) Kalinin, S.; Valeri, A.; Antonik, M.; Felekyan, S.; Seidel, C. A. Detection of structural dynamics by FRET: a photon distribution and fluorescence lifetime analysis of systems with multiple states. *J Phys Chem B* **2010**, *114* (23), 7983-7995. DOI: 10.1021/jp102156t.

(39) Agam, G.; Gebhardt, C.; Popara, M.; Machtel, R.; Folz, J.; Ambrose, B.; Chamachi, N.; Chung, S. Y.; Craggs, T. D.; de Boer, M.; et al. Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins. *Nat Methods* **2023**, *20* (4), 523-535. DOI: 10.1038/s41592-023-01807-0.

(40) Hubner, K.; Joshi, H.; Aksimentiev, A.; Stefani, F. D.; Tinnefeld, P.; Acuna, G. P. Determining the In-Plane Orientation and Binding Mode of Single Fluorescent Dyes in DNA Origami Structures. *ACS Nano* **2021**, *15* (3), 5109-5117. DOI: 10.1021/acsnano.0c10259.

(41) Beckers, M.; Drechsler, F.; Eilert, T.; Nagy, J.; Michaelis, J. Quantitative structural information from single-molecule FRET. *Faraday Discuss* **2015**, *184*, 117-129. DOI: 10.1039/c5fd00110b. Adamczyk, A., K., Huijben, T., A., P., M., Kolataj, K., Zhu, F., Marie, R., Stefani, F., D., Acuna, G., P. Towards full control of molecular exciton energy transfer via FRET in DNA origami assemblies. *arXiv:2402.06292 [cond-mat.soft]* **2024**.

(42) Stein, I. H.; Schuller, V.; Bohm, P.; Tinnefeld, P.; Liedl, T. Single-molecule FRET ruler based on rigid DNA origami blocks. *Chemphyschem* **2011**, *12* (3), 689-695. DOI: 10.1002/cphc.201000781.

(43) Kopperger, E.; List, J.; Madhira, S.; Rothfischer, F.; Lamb, D. C.; Simmel, F. C. A self assembled nanoscale robotic arm controlled by electric fields. *Science* **2018**, *19* (6373), 296-301. DOI: 10.1126/science.aoa4284.

(44) Kapanidis, A. N.; Laurence, T. A.; Lee, N. K.; Margeat, E.; Kong, X.; Weiss, S. Alternating Laser Excitation of Single Molecules. *Acc Chem Res* **2005**, *38* (7), 523-533. DOI: 10.1021/ar0401348.

(45) Asadiatouei, P.; Salem, C. B.; Wanninger, S.; Ploetz, E.; Lamb, D. C. Deep-LASI, single-molecule data analysis software. *Biophys J* **2024**. DOI: 10.1016/j.bpj.2024.02.013.

(46) Aitken, C. E.; Marshall, R. A.; Puglisi, J. D. An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments. *Biophys J* **2008**, *94* (5), 1826-1835. DOI: 10.1529/biophysj.107.117689.

(47) Cordes, T. V., J.; Tinnefeld, P. On the Mechanism of Trolox as Antiblinking and Antibleaching Reagent. *Journal of the American Chemical Society* **2009**, *131* (14), 5018–5019.

(48) Barth, A.; Voith von Voithenberg, L.; Lamb, D. C. Quantitative Single-Molecule Three-Color Forster Resonance Energy Transfer by Photon Distribution Analysis. *J Phys Chem B* **2019**, *123* (32), 6901-6916. DOI: 10.1021/acs.jpcb.9b02967.

(49) Müller, B. K.; Zaychikov, E.; Bräuchle, C.; Lamb, D. C. Pulsed Interleaved Excitation. *Biophys. J.* **2005**, *89* (5), 3508-3522.

(50) Schrimpf, W.; Barth, A.; Hendrix, J.; Lamb, D. C. PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. *Biophys J* **2018**, *114* (7), 1518-1528. DOI: 10.1016/j.bpj.2018.02.035.

(51) Hellenkamp, B.; Schmid, S.; Doroshenko, O.; Opanasyuk, O.; Kuhnemuth, R.; Rezaei Adariani, S.; Ambrose, B.; Aznauryan, M.; Barth, A.; Birkedal, V.; et al. Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nat Methods* **2018**, *15* (9), 669-676. DOI: 10.1038/s41592-018-0085-0.

Supplementary Information

Distance and Kinetic Tunability of Dynamic DNA Origami Structures Examined at the Single-Molecule Level

Pooyeh Asadiatouei¹, Johann Bohlen¹, Fiona Cole¹, Fabio Morella¹, Philip Tinnefeld¹, Evelyn Ploetz¹,
Don C. Lamb^{1,*}

1. Department of Chemistry and Center for NanoScience (CeNS)
Ludwig-Maximilians-Universität München
Butenandtstr. 5-13, 81377 Munich, Germany

Corresponding author: d.lamb@lmu.de

Table of Content

| | |
|--|------------|
| SUPPLEMENTARY INFORMATION | 162 |
| SUPPLEMENTARY NOTE 1: SINGLE-MOLECULE TIRF RESULTS FOR ALL DNA ORIGAMI STRUCTURES | 164 |
| SUPPLEMENTARY NOTE 2: MFD-PIE RESULTS FOR ALL THREE- AND TWO-COLOR DNA ORIGAMI STRUCTURES..... | 169 |
| SUPPLEMENTARY NOTE 3: FLUORESCENCE LIFETIME AND ANISOTROPY | 171 |
| 3.1. STEADY-STATE ANISOTROPY | 171 |
| 3.2. ANISOTROPY DECAY PLOTS..... | 171 |
| SUPPLEMENTARY NOTE 4: KINETICS COMPARISON USING A DNA-PAINT ASSAY | 174 |
| 4.1. TWO-COLOR, SINGLE-MOLECULE FRET MEASUREMENTS OF IMMOBILIZED ORIGAMI STRUCTURES (SMTIRF) THAT DO NOT CONTAIN A TETHER..... | 174 |
| SUPPLEMENTARY NOTE 5: SINGLE-MOLECULE FRET MEASUREMENTS OF DNA ORIGAMI STRUCTURES WITH THREE BINDING SITES | 175 |
| 5.1. APPARENT FRET EFFICIENCIES AND KINETIC ANALYSES | 175 |
| SUPPLEMENTARY NOTE 6: STATIC ORIGAMI STRUCTURES | 177 |
| 6.1. ANALYSIS OF STATIC SMFRET TRACES FOR THREE- AND TWO-COLOR, TWO-STATE SMFRET MEASUREMENTS ON THE SURFACE (SMTIRF) | 177 |
| SUPPLEMENTARY NOTE 7: SUMMARY OF ALL SINGLE-MOLECULE MFD-PIE DATA..... | 179 |
| 7.1. THREE-COLOR, TWO-STATE CLOCKS (V1) WITH VARIABLE DOCKING STRAND LENGTH | 180 |
| 7.2. THREE-COLOR, TWO-STATE CLOCKS (V2) WITH VARIABLE DOCKING STRAND LENGTH | 184 |
| 7.3. THREE-COLOR, THREE-STATE CLOCKS (V1) WITH DIFFERENT BINDING POSITIONS | 188 |
| 7.4. TRIPLE-COLOR DNA ORIGAMIS FOR CONTROL | 190 |
| 7.5. DUAL-COLOR 2 STATE CLOCKS (V1) WITH 7.5. NT BINDING SITES AND DIFFERENT DYE PAIR COMBINATIONS | 192 |
| 7.6. DUAL-COLOR 2 STATE CLOCKS (V2) WITH 7.5. NT BINDING SITES AND DIFFERENT DYE PAIR COMBINATIONS | 195 |
| 7.7. DUAL-COLOR 3 STATE CLOCKS (V1) WITH DIFFERENT BINDING POSITIONS | 198 |
| 7.8. DUAL-COLOR DNA ORIGAMIS FOR CONTROL..... | 200 |
| 7.9. SMALL CLOCK WITH 3 STATES | 202 |
| SUPPLEMENTARY NOTE 8: TWO- AND THREE-COLOR CORRECTED FRET EFFICIENCIES..... | 204 |
| SUPPLEMENTARY NOTE 9: STAPLE STRANDS FOR MULTI-COLORED DNA ORIGAMI CLOCKS | 205 |
| 9.1. STAPLE STRANDS OF THE L-SHAPED ORIGAMI STRUCTURE..... | 205 |
| 9.2. OVERVIEW OF DNA ORIGAMI STRUCTURES WITH LARGE CLOCK DESIGN | 212 |
| 9.3. STAPLE STRANDS OF DNA ORIGAMI STRUCTURES WITH LARGE CLOCK DESIGN..... | 214 |
| 9.4. OVERVIEW OF DNA ORIGAMI STRUCTURES WITH SMALL CLOCK DESIGN | 216 |
| 9.5. STAPLE STRANDS OF DNA ORIGAMI STRUCTURES FOR THE SMALL CLOCK DESIGN..... | 216 |
| SUPPLEMENTARY REFERENCES..... | 218 |

Supplementary Note 1: Single-molecule TIRF results for all DNA origami structures

The following tables summarize the results for all two- and three-color DNA origami datasets discussed in this work that were measured using TIRF microscopy. From the single-molecule traces, corrected smFRET histograms were extracted and analyzed to provide information over the number of states and corresponding kinetics. For each data set, the table provides information regarding the correction factors used for calculating corrected FRET efficiencies, the corrected FRET efficiency values, the ratio of FRET populations obtained by Gaussian fitting models, the corresponding dwell times, the calculated distances in Ångstroms and the fraction of static traces. For details, see Supplementary Note 8.

Supplementary Table S1.1. Experimental smTIRF results for three- and two-color L-shaped DNA origami structures with the tether labeled with Cy3B for different numbers of states and complementary base pairing. The correction factors for determining corrected FRET efficiencies; spectral crosstalk (CT), direct excitation (DE) and detection efficiency (Gamma) are given along with the corrected FRET efficiencies for state 1, state 2 and state 3 when present followed by the ratio of the populations. The dwell times for each state were extracted from the generated transition density plot (TDP) of each dye pair and fit assuming mono-exponential kinetics. Distances were calculated using Eqn 1 with R_0 values of 67 Å, 50 Å and 64 Å for BY, BR and YR respectively. Statistics regarding the number of static and dynamic (all - # static) smFRET traces are given along with the fraction of static molecules in the respective states. The errors represent the uncertainty in the parameters estimated from the Jacobian matrix of the fit.

| smTIRF | FRET Pair | CT | DE | Gamma | Corrected FRET Efficiency (%) (S1 – S2) or (S1 - S2 - S3) | Fraction of the relative State populations (%) | Dwell Time t ₁ (s) | Dwell Time t ₂ (s) | Dwell Time t ₃ (s) | Distances (Å) | # of static to all sm-traces | Static Corrected FRET (%) (S1 – S2) |
|------------------------|-----------|------|------|-------|---|--|-------------------------------|-------------------------------|-------------------------------|---------------|------------------------------|-------------------------------------|
| 2 binding sites | | | | | | | | | | | | |
| 3 Color 2 States | V1 | | | | | | | | | | | |
| 6.5 nt | BY | 0.47 | 0.17 | 1.70 | 45 - 78 | 36 - 64 | 0.39 ± 0.01 | 0.46 ± 0.01 | NA | 66.0 - 53.4 | 548 / 1030 (53 %) | 44 - 77 |
| | BR | 0.04 | 0.01 | 1.30 | 54 | NA | | | | 48.7 | | NA |
| | YR | 0.13 | 0.10 | 0.77 | 81 - 11 | 51 - 49 | | | | 50.7 - 91.5 | | 81 - 30 |
| 7 nt | BY | 0.47 | 0.21 | 2.42 | 50 - 82 | 49 - 51 | 0.65 ± 0.01 | 0.99 ± 0.01 | NA | 67.3 - 52.3 | 160 / 2366 (7 %) | 40 - 83 |
| | BR | 0.06 | 0.07 | 1.32 | 55 | NA | | | | 48.4 | | NA |
| | YR | 0.14 | 0.10 | 1.11 | 80 - 9 | 39 - 61 | | | | 51.3 - 97.1 | | 81 - 7 |
| 7.5 nt | BY | 0.45 | 0.23 | 1.68 | 37 - 81 | 44 - 56 | 1.90 ± 0.02 | 2.70 ± 0.03 | NA | 73.5 - 52.9 | 209 / 710 (29 %) | 35 - 82 |
| | BR | 0.03 | 0.01 | 1.12 | 45 | NA | | | | 51.7 | | NA |
| | YR | 0.12 | 0.09 | 0.71 | 81 - 9 | 43 - 57 | | | | 50.7 - 95.0 | | 81 - 9 |
| 8 nt | BY | 0.48 | 0.24 | 1.43 | 40 - 78 | 50 - 50 | 6.40 ± 0.04 | 9.50 ± 0.10 | NA | 70.0 - 54.5 | 1427 / 2972 (48 %) | 33 - 77 |
| | BR | 0.04 | 0.01 | 0.91 | 49 | NA | | | | 50.3 | | NA |
| | YR | 0.14 | 0.10 | 0.89 | 80 - 6 | 40 - 60 | | | | 51.3 - 102.2 | | 80 - 5 |
| 2 Color 2 States | BY | 0.43 | 0.27 | 1.62 | 28 - 77 | 41 - 59 | 1.40 ± 0.02 | 2.80 ± 0.01 | NA | 74.6 - 55.0 | 154 / 260 (59 %) | 24 - 70 |
| | BR | 0.06 | 0.05 | 1.37 | 43 | NA | NA | NA | NA | 51.7 | 100 % | NA |
| 7.5 nt | YR | 0.13 | 0.11 | 0.78 | 81 - 9 | 41 - 59 | 1.80 ± 0.02 | 2.80 ± 0.02 | NA | 50.7 - 94.3 | 527 / 2467 (21 %) | 80 - 12 |
| 3 Color 2 States | V2 | | | | | | | | | | | |
| 6.5 nt | BY | 0.50 | 0.18 | 1.52 | 85 - 40 | 37 - 63 | 0.30 ± 0.01 | 0.44 ± 0.01 | NA | 50.4 - 72.0 | 537 / 886 (61 %) | 47 - 21 |
| | BR | 0.05 | 0.01 | 1.27 | 52 | NA | | | | 49.3 | | NA |
| | YR | 0.13 | 0.11 | 0.84 | 27 - 60 | 32 - 68 | | | | 71.6 - 60.4 | | 25 - 53 |
| 7 nt | BY | 0.50 | 0.20 | 1.08 | 89 - 33 | 35 - 65 | 0.64 ± 0.01 | 1.10 ± 0.01 | NA | 47.5 - 75.7 | 283 / 1096 (27 %) | 86 - 32 |
| | BR | 0.05 | 0.01 | 1.08 | 58 | NA | | | | 47.4 | | NA |
| | YR | 0.14 | 0.10 | 0.99 | 20 - 58 | 34 - 66 | | | | 81.4 - 61.2 | | 21 - 55 |
| 7.5 nt | BY | 0.75 | 0.22 | 2.7 | 86 - 38 | 35 - 65 | 1.70 ± 0.02 | 2.50 ± 0.02 | NA | 49.7 - 73.0 | 565 / 1061 (53 %) | 84 - 37 |
| | BR | 0.14 | 0.01 | 2.70 | 56 | NA | | | | 48.0 | | NA |
| | YR | 0.16 | 0.12 | 1.0 | 23 - 61 | 36 - 64 | | | | 79.0 - 60.0 | | 26 - 58 |
| 8 nt | BY | 0.77 | 0.26 | 2.0 | 90 - 33 | 33 - 67 | 6.60 ± 0.06 | 8.60 ± 0.10 | NA | 46.7 - 75.7 | 1754 / 2923 (60 %) | 89 - 31 |
| | BR | 0.14 | 0.01 | 2.25 | 57 | NA | | | | 47.7 | | NA |
| | YR | 0.16 | 0.09 | 1.07 | 17 - 58 | 35 - 65 | | | | 84.1 - 61.2 | | 21 - 55 |
| 2 Color 2 States | BY | 0.54 | 0.29 | 1.44 | 80 - 17 | 39 - 61 | 1.50 ± 0.01 | 2.60 ± 0.01 | NA | 49.7 - 92.4 | 1401 / 2306 (61 %) | 83 - 17 |
| | BR | 0.05 | 0.04 | 2.46 | 46 | NA | NA | NA | NA | 51.4 | 100 % | NA |
| 7.5 nt | YR | 0.15 | 0.04 | 0.94 | 28 - 64 | 40 - 60 | 1.80 ± 0.02 | 2.90 ± 0.05 | NA | 75.6 - 58.7 | 452 / 992 (46 %) | 30 - 62 |

| smTIRF | FRET Pair | CT | DE | Gamma | Corrected FRET Efficiency (%) (S1 – S2) or (S1 - S2 - S3) | Fraction of the relative State populations (%) | Dwell Time t ₁ (s) | Dwell Time t ₂ (s) | Dwell Time t ₃ (s) | Distances (Å) | # of static to all sm-traces | Static Corrected FRET (%) (S1 – S2) |
|-------------------------|--------------|------|------|-------|---|--|-------------------------------|-------------------------------|-------------------------------|---------------------|------------------------------|-------------------------------------|
| 3 binding sites | | | | | | | | | | | | |
| 3 Color 3 States | V1 | | | | | Based on app. FRET | | | | | | |
| 3 Color 6-9-12 | BY | 0.50 | 0.22 | 1.25 | 39 - 79 - 18 | 27 - 33 - 40 | 0.72 ± 0.02 | 0.99 ± 0.02 | 1.60 ± 0.01 | 72.5 - 54.0 - 86.7 | | |
| | BR | 0.04 | 0.01 | 1.55 | 44 | NA | | | | 52.1 | | |
| | YR | 0.15 | 0.13 | 0.84 | 82 - 9 - 19 | 27 - 19 - 54 | | | | 50.2 - 82.3 - 95.0 | | |
| 3 Color 6-11-12 | BY | 0.53 | 0.22 | 1.40 | 56 - 80 - 30 | 22 - 55 - 23 | 0.76 ± 0.03 | 0.91 ± 0.03 | 1.50 ± 0.02 | 64.6 - 53.4 - 77.5 | | |
| | BR | 0.05 | 0.01 | 1.17 | 49 | NA | | | | 50.3 | | |
| | YR | 0.15 | 0.11 | 0.84 | 83 - 9 - 20 | 23 - 77 (S2&3) | | | | 49.6 - 95.0 - 81.4 | | |
| 2 Color 3 States | YR / 6-9-12 | 0.14 | 0.09 | 0.80 | 83 - 6 - 19 | 19 - 81 (S2&3) | 0.60 ± 0.01 | 1.10 ± 0.02 | 1.50 ± 0.03 | 49.6 - 102.2 - 82.3 | | |
| | YR / 6-11-12 | 0.13 | 0.07 | 0.85 | 83 - 9 | 26 - 74 (S2&3) | 0.65 ± 0.01 | 0.86 ± 0.01 | 1.40 ± 0.03 | 49.6 - 95.0 | | |
| smTIRF | FRET Pair | CT | DE | Gamma | Corrected FRET Efficiency (%) (S1 – S2) or (S1 - S2 - S3) | Fraction of the relative State populations (%) | Dwell Time t ₁ (s) | Dwell Time t ₂ (s) | Dwell Time t ₃ (s) | Distances (Å) | # of static to all sm-traces | Static Corrected FRET (%) (S1 – S2) |
| No binding sites | | | | | | | | | | | | |
| | V1 | | | | | | | | | | | |
| 3 Color No State | BY | 0.49 | 0.25 | 1.56 | 41-70 | NA | NA | NA | NA | 71.5-58.4 | | |
| | BR | 0.05 | 0.01 | 1.25 | 49 | | | | | 50.3 | | |
| | YR | 0.15 | 0.13 | 0.75 | 42-10 | | | | | 68.2-93.2 | | |
| 2 Color No State | YR | 0.13 | 0.07 | 0.81 | 41-15 | NA | NA | NA | NA | 68.6-86.3 | | |
| | V2 | | | | | | | | | | | |
| 3 Color No State | BY | 0.53 | 0.25 | 1.29 | 67-31 | NA | NA | NA | NA | 59.8-76.9 | | |
| | BR | 0.06 | 0.01 | 1.40 | 51 | | NA | NA | NA | 49.7 | | |
| | YR | 0.15 | 0.13 | 1.08 | 49 | | NA | NA | NA | 65.0 | | |
| 2 Color No State | YR | 0.13 | 0.07 | 0.92 | 23-58 | NA | NA | NA | NA | 79.0-61.2 | | |

Table S1.2. Experimental smTIRF results for three- and two-color L-shaped DNA origami structures with 8 nt complementary base pairs and a Cy3B labeled imaging strand. The apparent FRET efficiencies (no corrected FRET values could be determined from these experiments) and dwell times for states 1 and 2 are provided. The errors represent the uncertainty in the parameters estimated from the Jacobian matrix of the fit.

| PAINT Assay | V1 | App. FRET (%) (S1 – S2) | Based on app. FRET | Dwell Time t ₁ (s) | Dwell Time t ₂ (s) |
|------------------|---------------------|----------------------------|-----------------------|----------------------------------|----------------------------------|
| 2 States 8 nt | YR Cy3B-Atto647N | 27 - 40/89 | 60 - 40 | 1.0 ± 0.04 | 1.8 ± 0.02 |

Table S1.3. Experimental smTIRF results for three- and two-color L-shaped DNA origami V1 structures with Atto542 on the tether and 7.5 nt complementary base pairs. The correction factors for determining corrected FRET efficiencies; spectral crosstalk (CT), direct excitation (DE) and detection efficiency (Gamma) are given along with the corrected FRET efficiencies for state 1, state 2 and state 3 when present followed by the ratio of the populations. The dwell times for each state were extracted from the generated transition density plot (TDP) of each dye pair and fit assuming mono-exponential kinetics. Distances were calculated using Eqn 1 with R₀ values of 64 Å, 50 Å and 63 Å for BG, BR and GR respectively. Statistics regarding the number of static and dynamic (all - # static) smFRET traces are given along with the fraction of static molecules in the respective states. The errors represent the uncertainty in the parameters estimated from the Jacobian matrix of the fit.

| smTIRF | FRET Pair | CT | DE | Gamma | Corrected FRET (%) (S1 – S2) | Ratio of Populations | Dwell Time t ₁ (s) | Dwell Time t ₂ (s) | Distance (Å) | # of static to all sm- traces | Static Corrected FRET (%) (S1 – S2) |
|-------------------------------|--------------|------|------|-------|---------------------------------|-------------------------|----------------------------------|----------------------------------|-----------------|----------------------------------|---|
| 3 Color 2 States | V1 | | | | #(State X) / #(all) (%) | | | | | | |
| 7.5 nt | BG | 0.72 | 0.30 | 1.45 | 12 - 49 | 12 - 88 | 1.10 ± 0.02 | 1.70 ± 0.01 | 89.2 - 64.4 | 51 / 548 (9 %) | 17 - 47 - 80* |
| | BR | 0.08 | 0.05 | 2.60 | 44 | NA | | | 52.1 | | NA |
| | GR | 0.15 | 0.05 | 1.18 | 64 - 3 | 35 - 65 | | | 57.2 - 112.4 | | 61 - 2 - 27* |
| 2 Color 2 States 7.5 nt | BG | 0.74 | 0.37 | 1.82 | 19 - 51 | 36 - 64 | 1.10 ± 0.01 | 1.60 ± 0.00 | 81.5 - 63.6 | 1080 / 1433 (75 %) | 43** |
| | GR | 0.12 | 0.02 | 1.84 | 60 - 5 | 41 - 59 | 0.86 ± 0.01 | 1.60 ± 0.01 | 58.9 - 102.9 | 141 / 577 (24 %) | 64 - 17 |

*In the static histogram, we see a third population appearing. / **Not resolvable.

Table S1.4. Experimental smTIRF results for two-color, three-state L-shaped DNA origami structures with symmetrically distributed docking strands (small clock) using Cy5 as an acceptor on the DNA origami structure. The correction factors for determining corrected FRET efficiencies; spectral crosstalk (CT), direct excitation (DE) and detection efficiency (Gamma) are given along with the corrected FRET efficiencies for state 1, state 2 and state 3 followed by the ratio of the populations. The dwell times for each state were extracted from the generated transition density plot (TDP) of each dye pair and fit assuming mono-exponential kinetics. Distances were calculated using Eqn 1 with R₀ values of 72 Å for YR (Cy3B, Cy5) and 71 Å for GR (Atto542, Cy5) respectively. Statistics regarding the number of static and dynamic (all - # static) sm FRET traces are given along with the fraction of static molecules in the respective states. The errors represent the uncertainty in the parameters estimated from the Jacobian matrix of the fit.

| smTIRF | FRET Pair | CT | DE | Gamma | Corrected FRET Efficiency (%) (S1 - S2 - S3) | Fraction of the relative State populations (%) | Dwell Time t ₁ (s) | Dwell Time t ₂ (s) | Dwell Time t ₃ (s) | Distances (Å) | # of static to all sm-traces | Static Corrected FRET (%) (S1 - S2) |
|--------|-----------|----|----|-------|--|--|----------------------------------|----------------------------------|----------------------------------|------------------|---------------------------------|--|
|--------|-----------|----|----|-------|--|--|----------------------------------|----------------------------------|----------------------------------|------------------|---------------------------------|--|

| | | | | | | | | | | | | |
|------------------------------|----------------------------------|------|------|------|--------------|--------------|-----------------|-----------------|-----------------|--------------------|-------------------|---------------------|
| 2 Color Symmetric | YR (7 nt) Cy3B-Cy5 | 0.14 | 0.16 | 0.55 | 24 - 84 - 69 | 37 - 20 - 43 | 0.60 ± 0.03 | 0.56 ± 0.01 | 0.42 ± 0.03 | 87.3 - 54.6 - 63.0 | 457 / 6152 (7 %) | 25 - 66 (S2 & 3) |
| 3 States (SC) | YR (8 nt) Atto542-Cy5 | 0.07 | 0.06 | 0.75 | 15 - 70 - 54 | 38 - 20 - 42 | 0.86 ± 0.01 | 1.10 ± 0.01 | 0.97 ± 0.01 | 94.8 - 61.7 - 69.1 | 191 / 1993 (10 %) | 17 - 70 - 47 |

Supplementary Note 2: MFD-PIE results for all three- and two-color DNA origami structures

The following table summarizes the MFD-PIE results for various two- and three-color DNA origami datasets discussed in this work. For each data set, burst-wise smFRET histograms were generated. The correction factors used for calculating the corrected FRET efficiencies are given along with the ratio of FRET populations obtained by fitting the histograms with Gaussians.

Supplementary Table S2.1. Experimental MFD-PIE results for various three- and two-color L-shaped DNA origami structures having different numbers of states and complementary base pairs. The correction factors for determining corrected FRET efficiencies; spectral crosstalk (CT), direct excitation (DE) and detection efficiency (Gamma) are given along with the corrected FRET efficiencies for state 1, state 2 and state 3 when present followed by the ratio of the populations.

| MFD-PIE | FRET Pair | CT | DE | Gamma | Corrected FRET [%] (S1 - S2) | Ratio of Populations |
|-------------------------------|-----------|-------|-------|-------|---------------------------------|------------------------|
| 3 Color 2 States | V1 | | | | | #(S1 - S2) / (all) (%) |
| 7 nt, 1 mm | BY | 0.07 | 0.17 | 0.40 | 47 - 78 | 45 - 55 |
| | BR | 0.004 | 0.02 | 0.29 | 54 | NA |
| | YR | 0.11 | 0.08 | 0.73 | 77 - 13 | 31 - 69* |
| 7 nt | BY | 0.07 | 0.16 | 0.50 | 50 - 81 | 42 - 58 |
| | BR | 0.005 | 0.008 | 0.38 | 50 | NA |
| | YR | 0.10 | 0.08 | 0.76 | 80 - 12 | 30 - 70* |
| 8 nt, 1 mm | BY | 0.08 | 0.16 | 0.30 | 50 - 81 | 40 - 60 |
| | BR | 0.0 | 0.0 | 0.21 | 48 | NA |
| | YR | 0.11 | 0.05 | 0.71 | 80 - 14 | 35 - 65* |
| 8 nt | BY | 0.09 | 0.17 | 0.45 | 57 - 81 | 38 - 62 |
| | BR | 0.005 | 0.04 | 0.34 | 49 | NA |
| | YR | 0.11 | 0.08 | 0.76 | 80 - 11 | 25 - 75* |
| 2 Color 2 States 7.5 nt | BY | 0.08 | 0.15 | 0.52 | 37 - 77 | 47 - 53* |
| | BR | 0.003 | 0.03 | 0.28 | 42 | NA |
| | YR | 0.11 | 0.12 | 0.75 | 80 - 7 | 44 - 56* |
| 3 Color 2 States | V2 | | | | | |
| 7 nt, 1 mm | BY | 0.07 | 0.15 | 0.53 | 85 - 38 | 31 - 69* |
| | BR | 0.003 | 0.004 | 0.41 | 53 | NA |
| | YR | 0.11 | 0.09 | 0.78 | 32 - 63 | 52 - 48 |
| 7 nt | BY | 0.07 | 0.17 | 0.48 | 86 - 35 | 28 - 72* |
| | BR | 0.004 | 0.0 | 0.39 | 54 | NA |
| | YR | 0.11 | 0.10 | 0.82 | 28 - 59 | 46 - 54 |
| 8 nt, 1 mm | BY | 0.06 | 0.12 | 0.44 | 79 - 26 | 40 - 60* |
| | BR | 0.003 | 0.03 | 0.35 | 53 | NA |
| | YR | 0.12 | 0.16 | 0.79 | 28 - 60 | 58 - 42 |
| 8 nt | BY | 0.07 | 0.14 | 0.44 | 86 - 37 | 41 - 59* |
| | BR | 0 | 0.0 | 0.37 | 54 | NA |
| | YR | 0.11 | 0.11 | 0.84 | 27 - 57 | 48 - 52 |
| 2 Color 2 States 7.5 nt | BY | 0.07 | 0.15 | 0.52 | 83 - 25 | 42 - 58* |
| | BR | 0.002 | 0.01 | 0.33 | 48 | NA |
| | YR | 0.10 | 0.11 | 0.80 | 30 - 64 | 46 - 54 |

| | | | | | | |
|--|-----------------------|-------|-------|-------|--------------|--------------|
| 3 Color 3 States | V1 | | | | | |
| 3 Color 6-9-12 | BY | 0.14 | 0.16 | 0.32 | 53 - 78 - 18 | 60 - 27 - 13 |
| | BR | 0.003 | 0 | 0.21 | 43 | NA |
| | YR | 0.11 | 0.0 | 0.65 | 50 - 20 | 24 - 76 |
| 3 Color 6-11-12 | BY | 0.14 | 0.16 | 0.28 | 48 - 83 - 25 | 62 - 20 - 18 |
| | BR | 0.0 | 0 | 0.22 | 44 | NA |
| | YR | 0.10 | 0 | 0.78 | 72 - 24 | 13 - 87 |
| 2 Color 3 States | YR 6-9-12 | 0.100 | 0.053 | 0.820 | 77 - 26 | 21 - 79 |
| | YR 6-11-12 | 0.100 | 0.070 | 0.780 | 75 - 22 | 18 - 82* |
| | V1 | | | | | |
| 3 Color No State | BY | 0.09 | 0.14 | 0.57 | 44 - 65 | 40 - 60 |
| | BR | 0.002 | 0.0 | 0.36 | 45 | NA |
| | YR | 0.09 | 0.12 | 0.63 | 41 - 16 | 39 - 61 |
| 2 Color No State | YR | 0.105 | 0.036 | 0.820 | 17 - 43 | 12 - 88 |
| 3 Color No States | V2 | | | | | |
| 3 Color No State | BY | 0.07 | 0.10 | 0.51 | 63 - 40 | 51 - 49 |
| | BR | 0.0 | 0.01 | 0.42 | 51 | NA |
| | YR | 0.10 | 0.12 | 0.83 | 11 - 50 | 40 - 60 |
| 2 Color No State | YR | 0.105 | 0.120 | 0.90 | 20 - 49 | 23 - 77 |
| 2 Color Symmetric 3 States (SC) | YR Cy3B-Cy5 | 0.11 | 0.05 | 0.32 | 24 - 80 - 43 | 20 - 18 - 62 |
| | YR Atto542- Cy5 | 0.03 | 0.06 | 0.35 | 15 - 67 - 50 | 43 - 28 - 29 |

* The two populations do not make 100 % of the whole population distribution because of the contribution of a third, intermediate population due to a transition occurring between the two docked states. The third, intermediate population is excluded from the ratio calculation between the two main populations.

Supplementary Note 3: Fluorescence lifetime and anisotropy

3.1. Steady-State Anisotropy

From the MFD-PIE data, we have information regarding the anisotropy of the fluorophores. From the individual bursts, we calculate a molecule-wise, steady-state anisotropy value (Supplementary Figures S7.1-S7.24, bottom panels; Supplementary Table S3.1). These data are then fit using the Perrin equation:

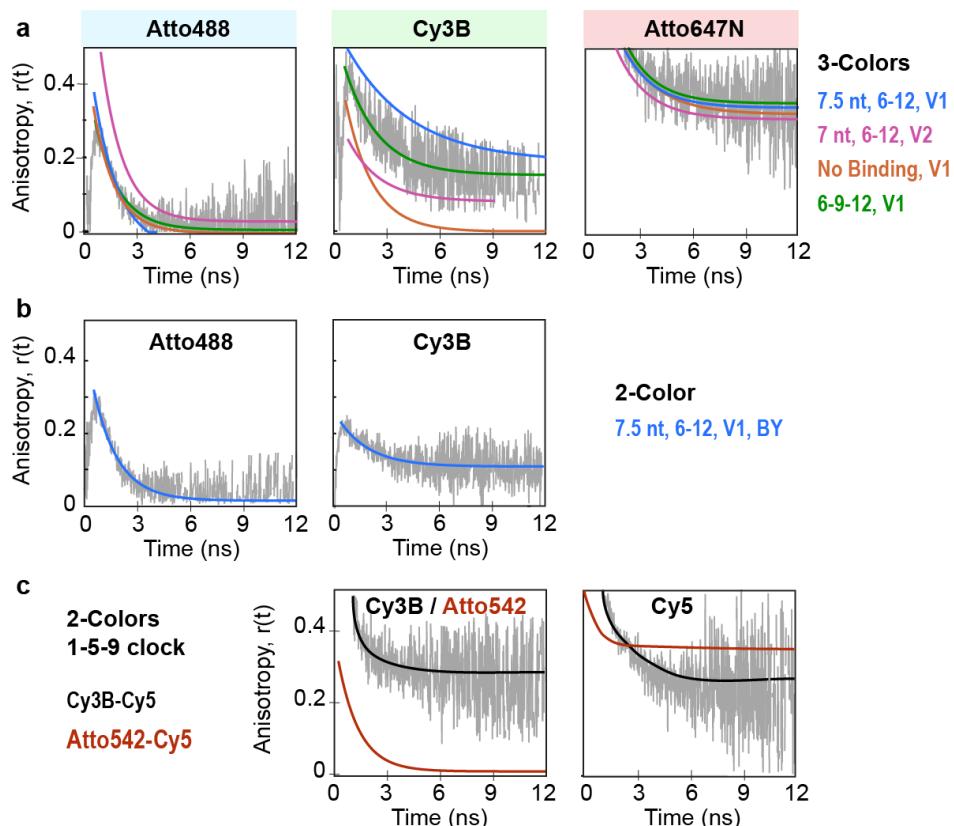
$$r(\tau) = \frac{r_0}{1 + \frac{\tau}{\rho}} \quad (\text{S1})$$

where the fundamental anisotropy, r_0 , is assumed to be 0.4 and ρ represents the rotational correlation time. The fluorescence lifetime for both the donor (undergoing FRET) and acceptor fluorophores (after direct excitation) are extracted from the experimental data. The fluorescence lifetime decay may include more than one population depending on how many FRET populations exist for each dye pair. In the event that more than one lifetime component is present, we utilized the average lifetime value. The rotational correlation time was then obtained by fitting the anisotropy-lifetime plot to Eqn S1.

3.2. Anisotropy decay plots

From the MFD-PIE data, we also have access to the time-resolved anisotropy decays of the dyes, which provides information regarding their rotational freedom. The results are summarized in this section (and in Supplementary Note 7). Supplementary Figure 3.1 (a and b) shows representative time-resolved anisotropy decays for selected three- and two-color DNA origami structures. Atto488 shows the fastest decay in all samples indicating that it has the highest rotational freedom among the labels. We attribute this to the mild repulsion between the dye (with a net charge of -1) and the negatively charged DNA origami. Cy3B on the tether shows different behaviors in different samples. As expected, Cy3B exhibits the fastest anisotropy decay on structures with no docking strands (Supplementary Figure S3.1a, orange curve). The anisotropy increases upon binding to docking sites with higher anisotropies for longer complementary docking strands. Interestingly, the time-resolved anisotropy decay for Cy3B in the two-color sample without Atto647N is faster. This suggests that there are additional interactions between Atto647N and Cy3B that increases its anisotropy. The anisotropy of Atto647N is extremely high, approaching the maximum value of 0.4, indicating that the dye with a net charge of +1 is sticking to the negatively charged DNA origami surface. This is consistent with previous observations that red Atto dyes stick to origami surfaces [1, 2].

On the two-color, three-state DNA origami structure (1-5-9), we also performed measurements where the fluorophore on the tether had been replaced with Atto542 and Atto647N was replaced with Cy5. Cy5, similar to Atto647N, has a positive net charge of +1 and an extremely high anisotropy (Supplementary Figure 3.1c, right panel). Atto542 has a faster decay of the time-resolved anisotropy relative to Cy3B indicating a higher rotational freedom (Supplementary Figure 3.1c, left panel). Due to the -3 net charge of Atto542, this is not unexpected.



Supplementary Figure S3.1. Fluorescence anisotropy decays for representative three- and two-color DNA origami structures. (a) The anisotropy decays of Atto488 (left panel), Cy3B (middle panel) and Atto647N (right panel) are shown for the following constructs: the three-color, 6-12, V1 structure with 7.5 nt docking strands (blue), the V2 structure with 7 nt docking strands (pink), the three-color, no-state V1 construct (orange) and the three-color, three-state V1 sample for the 6-9-12 o'clock (green). The solid lines represent the fit with a single exponential decay. For clarity, the anisotropy data is only shown for the V1 6-9-12 construct. Atto488 shows the fastest decay among the three fluorophores. Cy3B shows different decay times dependent on the details of the sample. Atto647N exhibits a very high anisotropy due to sticking of the dye to the origami surface. (b) Anisotropy decays of Atto488 (left panel) and Cy3B (right panel) for the two-color, 6-12, V1 structure with 7.5 nt docking strands. The blue dye shows the same rapidly decaying anisotropy curve whereas the anisotropy of the Cy3B-tether decreases in the absence of Atto647N. (c) Anisotropy decays of Cy3B, Atto542 and Cy5 on the two-color, three-state origami structures (1-5-9 clock). Left: Cy3B (black) shows a quick decay followed by a high residual anisotropy. In contrast, Atto542 (red) rapidly decays to zero, indicating a high amount of rotational flexibility. Right: Cy5, similar to Atto647N, has an extremely high anisotropy approaching 0.4.

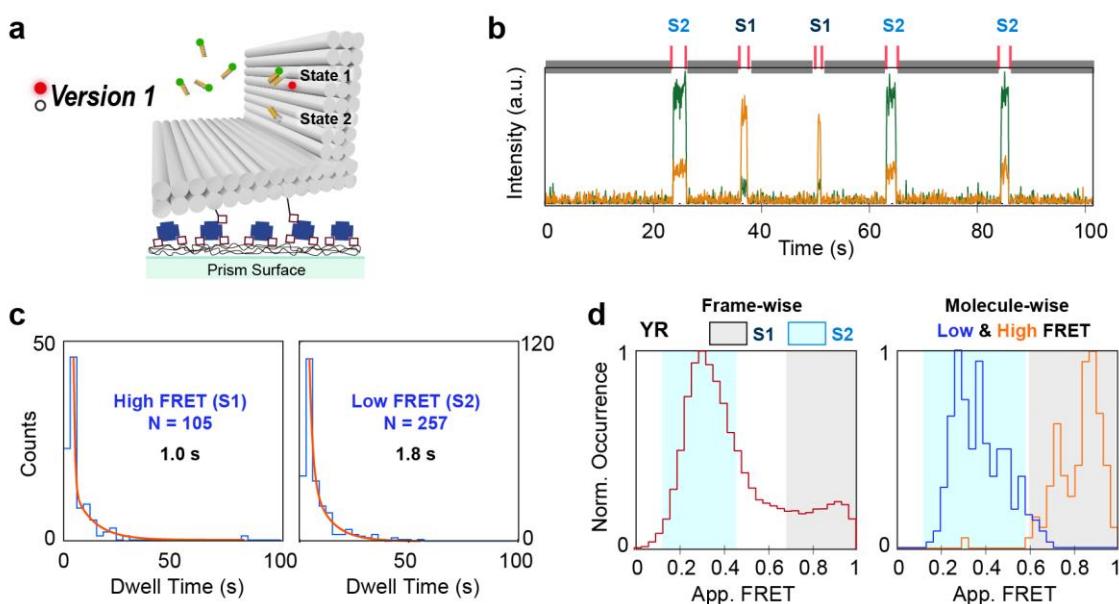
Supplementary Table S3.1. Fluorescence lifetime and steady-state anisotropy values for various constructs in this study. When two lifetime components were present, they are both listed.

| | Atto488 | | | Cy3B | | | Atto647N | |
|--------------------------|--------------------|-------------|------|---------------------------|-------------|------|-------------|------|
| | τ (ns) | ρ (ns) | r | τ (ns) | ρ (ns) | r | τ (ns) | r |
| 3 Color, 2 States | V1 | | | | | | | |
| 6.5 nt | 0.65 - 1.60 [1.13] | 0.98 | 0.16 | 1.20 - 2.35 [1.78] | 2.56 | 0.21 | 4.66 | 0.30 |
| 7 nt | 0.70 - 1.30 [1.0] | 1.50 | 0.25 | 1.03 - 2.40 [1.72] | 2.86 | 0.22 | 4.60 | 0.31 |
| 7.5 nt | 0.80 - 1.95 [1.38] | 0.79 | 0.12 | 1.0 - 1.99 [1.50] | 2.29 | 0.21 | 4.52 | 0.31 |
| 8 nt | 0.67 - 1.55 [1.11] | 1.89 | 0.15 | 1.12 - 2.38 [1.75] | 2.89 | 0.21 | 4.73 | 0.30 |
| 2 Color, 2 States | 0.94 - 2.36 [1.65] | 1.21 | 0.15 | 2.75 | 3.25 | 0.22 | | |
| 7.5 nt | 1.93 | 0.85 | 0.12 | | | | 4.30 | 0.32 |
| | | | | 1.0 - 2.50 [1.75] | 2.42 | 0.21 | 4.80 - 5.68 | 0.31 |
| | V2 | | | | | | | |
| 6.5 nt | 0.65 - 1.62 [1.14] | 0.95 | 0.16 | 0.88 - 1.40 [1.14] | 1.48 | 0.21 | 4.43 | 0.31 |
| 7 nt | 0.75 - 1.83 [1.29] | 1.43 | 0.17 | 1.48 - 1.94 [1.71] | 2.18 | 0.20 | 4.76 | 0.32 |
| 7.5 nt | 0.45 - 1.40 [0.93] | 0.88 | 0.16 | 0.87 - 1.31 [1.09] | 1.25 | 0.19 | 4.55 | 0.31 |
| 8 nt | 1.60 - 2.16 [1.88] | 1.28 | 0.15 | 1.45 - 1.96 [1.71] | 2.28 | 0.21 | 4.72 | 0.32 |
| 2 Color, 2 States | 0.86 - 2.41 [1.64] | 1.16 | 0.14 | 2.19 | 2.21 | 0.20 | | |
| 7.5 nt | 1.87 | 1.38 | 0.17 | | | | 4.05 | 0.32 |
| | | | | 1.14 - 1.50 [1.32] | 1.57 | 0.21 | 4.30 | 0.31 |
| 3 Color, 3 States | V1 | | | | | | | |
| 6-9-12 | 1.85 - 2.34 [2.10] | 0.87 | 0.12 | 2.14 - 2.55 [2.35] | 2.74 | 0.22 | 4.34 | 0.30 |
| 6-11-12 | 1.27 - 1.92 [1.60] | 1.11 | 0.14 | 2.08 - 2.45 [2.27] | 2.74 | 0.22 | 4.13 | 0.30 |
| 2 Color, 3 States | | | | | | | | |
| 6-9-12 | | | | 1.27 - 2.26 - 3.21 [2.25] | 2.71 | 0.22 | 4.33 | 0.30 |
| 6-11-12 | | | | 1.50 - 2.40 - 3.61 [2.50] | 2.69 | 0.21 | 4.37 | 0.31 |
| 3 Color, No State | 1.48 - 1.92 [1.70] | 1.0 | 0.14 | 1.85 - 2.19 [2.02] | 1.36 | 0.14 | 4.41 | 0.30 |
| YR, No State | | | | 1.82 - 2.09 [1.96] | 1.22 | 0.16 | 4.40 | 0.30 |
| | V2 | | | | | | | |
| 3 Color, No State | 1.37 - 1.73 [1.55] | 1.34 | 0.17 | 1.52 - 1.92 [1.72] | 1.30 | 0.18 | 4.50 | 0.32 |
| YR, No State | | | | 1.60 - 1.95 [1.78] | 1.87 | 0.21 | 4.40 | 0.32 |
| Symmetric Clock | | | | Cy3B | | | Cy5 | |
| | | | | 2.36 | 3.93 | 0.25 | 2.17 | 0.40 |
| | | | | Atto542 | | | Cy5 | |
| | | | | 2.1 - 3.0 [2.55] | 0.69 | 0.07 | 1.60 | 0.40 |

Supplementary Note 4: Kinetics comparison using a DNA-PAINT assay

4.1. Two-color, single-molecule FRET measurements of immobilized origami structures (smTIRF) that do not contain a tether

To check for fast transient dissociation and rebinding of the tether to the docking strand and to investigate the origin of the observed bias between the two binding strands, we measured binding to the docking sites of S1 and S2 using DNA PAINT imaging strands rather than a bound tether. We used a tether-free version of the same two-state DNA origami structure (6-12 o'clock position, V1) containing the 8 nt docking strands. To visualize FRET, we used an 8-nt imager strand with the same sequence as the normal tether labeled with Cy3B. The imager strands containing Cy3B were flushed into the chamber with the imaging buffer to have a total concentration of 5 nM of donor accessible to the structures. The experiment was carried out using continuous 561 nm excitation and the appearing FRET signals were investigated. In this experiment, we only investigated the state-specific binding and dissociation rates. Hence, ALEX excitation was not used in this experiment.

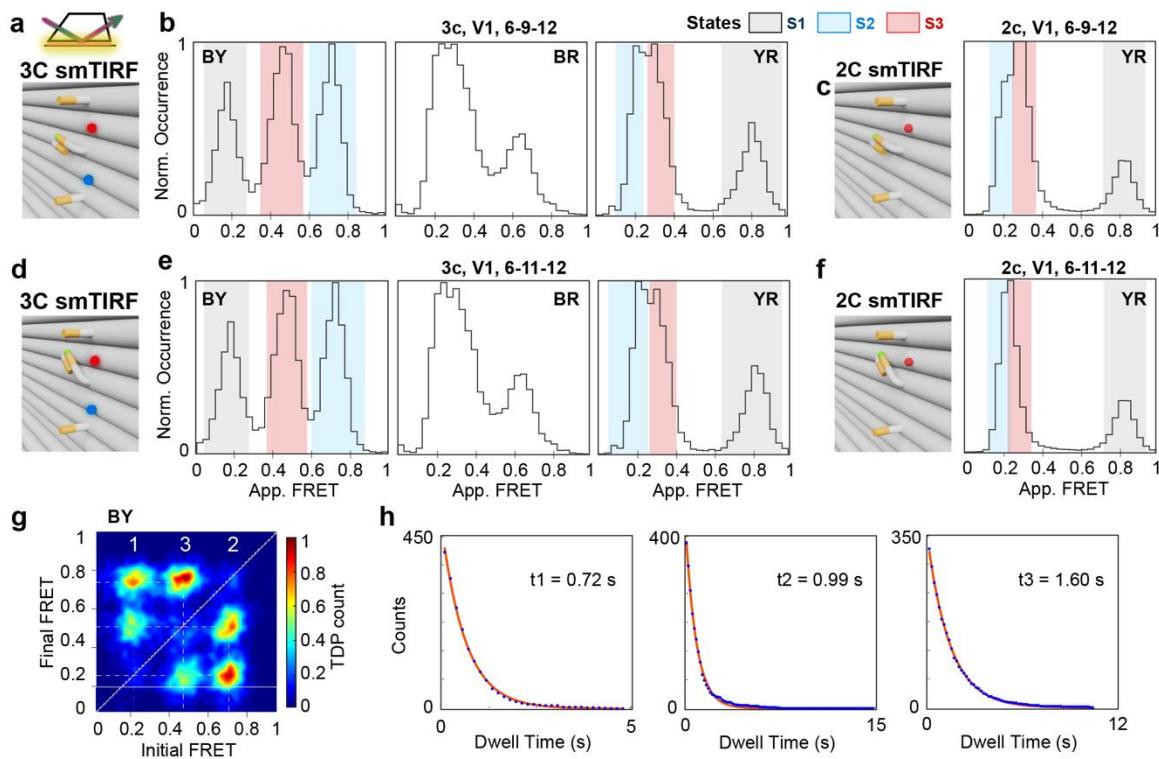


Supplementary Figure S4.1. Single-molecule TIRF results for the PAINT assay on a two-state, V1, L-shaped DNA origami structure. (a) A schematic showing the smTIRF / DNA-PAINT assay for measuring DNA binding to immobilized DNA origami structures on the passivated prism surface. (The dimensions are changed for clarity). The origami is labeled with Atto647N at the top (V1). Imager strands containing 8 complementary nucleotides to the docking strand were labeled with Cy3B and added at 5 nm concentration. (b) A single-molecule intensity trace showing the binding and unbinding events of imager strands to a DNA origami. The red lines show the start and end points of binding to the denoted docking state. The gray regions show the off time (time until binding). (c) The dwell-time distribution for the binding events corresponding to S1 and S2. The distribution was fit to a mono-exponential decay and dwell times of 1.0 s and 1.8 s were obtained, respectively. (d) Apparent frame-wise FRET efficiency histograms from smTIRF-PAINT experiments with only yellow excitation for the YR dye pair. The left panel shows the population distribution of states 1 and 2. The right panel shows the normalized separated histograms for each state. The large asymmetry in binding to the different states is an artifact of the extraction algorithm used for the analysis.

Supplementary Note 5: Single-molecule FRET measurements of DNA origami structures with three binding sites

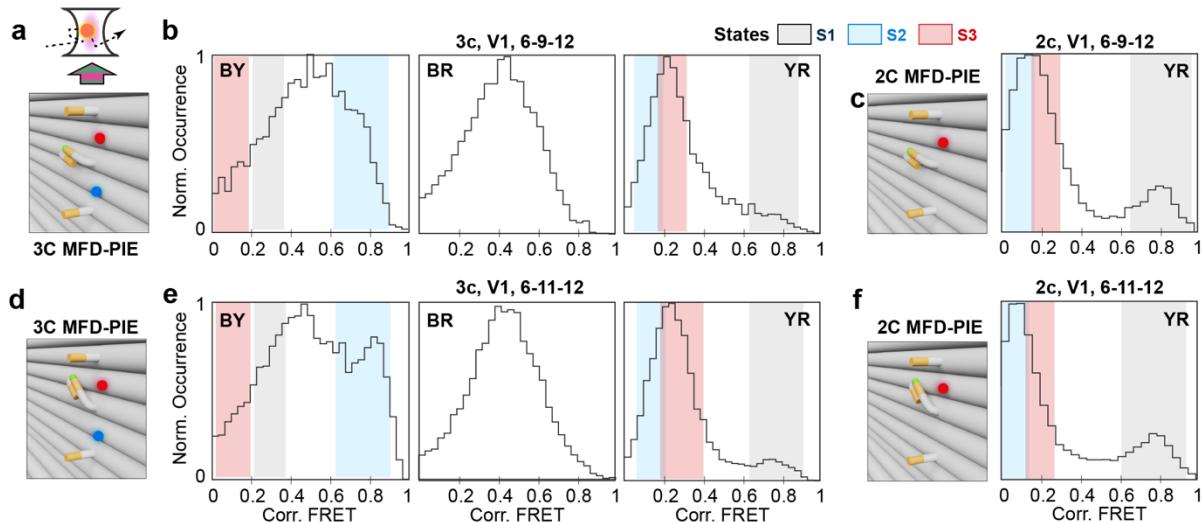
5.1. Apparent FRET efficiencies and kinetic analyses

In this section, we discuss the analysis of the apparent FRET histograms from smTIRF experiments on the three- and two-color DNA origami structures with three docking sites around the protruding tether with the third one located either at 9 or 11 o'clock position. While the corrected FRET efficiencies are needed for calculating distances between the different dye pairs, apparent FRET histograms may be better at resolving the different states and associated kinetics. In these constructs, docking strands at positions 6 and 12 are identical to those for the two-state clocks containing 7 nt docking strands. The third position at either 9 or 11 o'clock, was designed to have 7.5 nt complementary bases. The three states are clearly observed in the apparent FRET histograms especially for the BY dye pair (Supplementary Figure S5.1). As mentioned in the main manuscript, the kinetic analysis was performed on clusters from the TDP determined from the apparent FRET histograms and the dwell time distributions fit to a mono-exponential function (Supplementary Figure S5.1, g, h).



Supplementary Figure S5.1. Single-molecule TIRF results for the three- and two-color L-shaped DNA origami structures with three complementary binding strands. (a) A zoomed-in schematic of the investigated V1 DNA origami structures labeled with Cy3B on the tether, Atto647N on the top and Atto488 at the bottom. The three complementary docking strands are located at the 6, 9 and 12 o'clock positions (with 7 complementary nucleotides at position 6 and 12, and 7.5 nt at position 9). (b) The corresponding frame-wise weighted, state-wise averaged apparent FRET efficiency histograms from smTIRF experiments for the BY (left), BR (middle) and YR (right) FRET pairs. (c) Left: A schematic showing a zoomed-in DNA origami structure labeled with Cy3B on the

tether and Atto647N on the top. *Right*: Frame-wise weighted, state-wise averaged apparent FRET efficiency histogram from smTIRF experiments for the YR dye pair on the two-color 6-9-12-state nanostructures. **(d)** A zoomed-in schematic of the DNA origami labeled with three fluorophores. The three complementary docking strands are located at positions 6, 11 and 12 o'clock. **(e)** The corresponding frame-wise weighted, state-wise averaged apparent FRET efficiency histograms from smTIRF experiments for the BY (left), BR (middle) and YR (right panel). **(f)** *Left*: A schematic showing the zoomed-in DNA origami structure labeled with Cy3B on the tether and Atto647N on the top. *Right*: The corresponding frame-wise weighted, state-wise averaged apparent FRET efficiency histogram from smTIRF experiments for the YR dye pair. **(g)** The transition density plot (TDP) showing transitions between the two detected states for the BY dye pair. **(h)** By selecting the individual clusters in panel g, a distribution of dwell times for the individual transitions can be calculated. The dwell-time distributions and corresponding mono-exponential fits are shown for S1 (left), S2 (middle) and S3 (right).



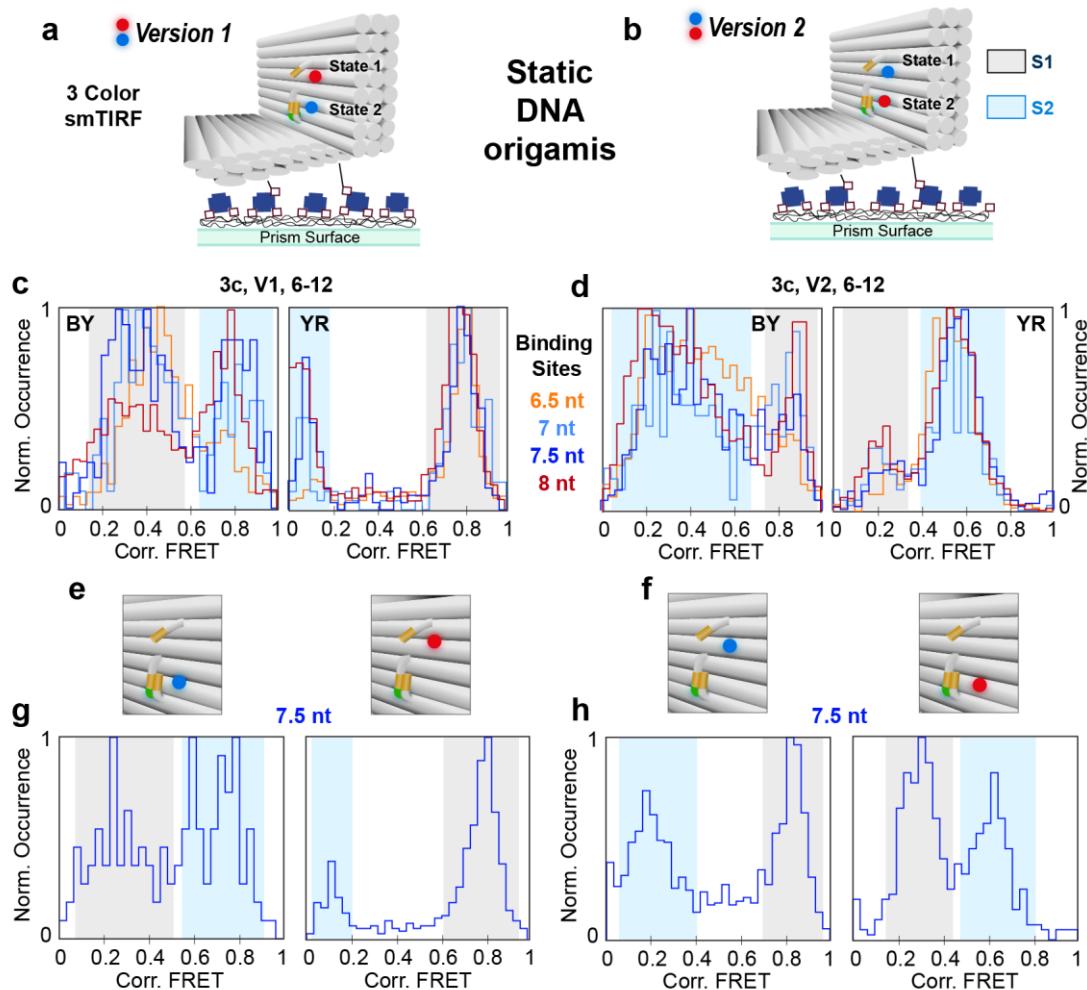
Supplementary Figure S5.2. Single-molecule MFD-PIE results for the three- and two-color, L-shaped DNA origami structures with three complementary docking strands. **(a)** A zoomed-in schematic of a diffusing DNA origami structure investigated by 3C MFD-PIE. The positions of the docking strands and location of the fluorophores are highlighted: Cy3B on the tether, Atto647N on the top and Atto488 at the bottom. The three complementary single strands are located at 6-9-12 and 6-11-12 o'clock positions (with 7 matching nucleotides at the 6 and 12 o'clock positions, and 7.5 nt at the 9 and 11 o'clock position). **(b)** Corrected FRET efficiency histograms from MFD-PIE experiments for the BY (left), BR (middle) and YR (right) FRET pairs on the three-color 6-9-12-state nanostructures. **(c)** *Left*: A zoomed-in schematic of the two-color DNA origami structure with docking strands as in panel a labeled with Cy3B on the tether and Atto647N at the upper position. *Right*: The corresponding corrected FRET efficiency histogram from MFD-PIE experiments for the YR dye pair. **(d)** A zoomed-in schematic of the DNA origami structure with three complementary single strands are located at the 6, 11 and 12 o'clock positions (with 7 matching nucleotides for the 6 o'clock and 12 o'clock positions, and 7.5 nt at 11 o'clock). **(e)** Corrected FRET efficiency histograms from MFD-PIE experiments for the BY (left), BR (middle) and YR (right) dye pairs on the three-color 6-11-12-state nanostructures. **(f)** *Left*: A zoomed-in schematic of the two-color DNA origami structure labeled with Cy3B on the tether and Atto647N on the top. The position and length of the docking strands are the same as in the three-color construct. *Right*: The corresponding corrected FRET efficiency histogram from MFD-PIE experiments for the YR dye pair.

Supplementary Note 6: Static origami structures

6.1. Analysis of static smFRET traces for three- and two-color, two-state smFRET measurements on the surface (smTIRF)

The DNA origami structures we are investigating are designed to be dynamic. So far, we focused on the dynamic traces that show transitions between the two or three incorporated FRET states. However, many of the detected FRET traces are of high quality but do not exhibit dynamics over the duration of the trace. To see if the static molecules have a preferred conformation, we also analyzed the smFRET histograms for static traces (Supplementary Figure 6.1). From the smFRET histograms, two populations are observable with corrected FRET values being very close to those measured for the dynamic molecules. Hence, it is apparent that static molecules can be observed in both conformations, S1 and S2. For the V1 samples, S1 is more populated than S2, although the ratio decreases with increasing docking strand length. The histograms of the V2 samples show a similar trend but with more molecules in state S2. This suggest that the origami structures are more likely to be static when binding to docking strands in the vicinity of Atto647N.

The number of static to dynamics traces (i.e. with respect to all high-quality, single-molecule traces increases with the length of the docking strand (with the exception of the 6.5 nt structures) from 7 % to \sim 50%. We attribute the high number of static traces for the 6.5 nt states to the increased difficulty in detecting high quality dynamic traces rather than an absolute increase in static molecules. As seen from the anisotropy data (Supplementary Note S3), Cy3B can interact with double-stranded DNA, stabilizing the binding. The slower the kinetics, the more time the dye has to interact and stabilize the hybridization.



Supplementary Figure S6.1. Single-molecule TIRF results for static traces from the three- and two-color, two-state V1 and V2 DNA origami structures. (a-b) Schematics depicting the smTIRF assay for measuring immobilized DNA origami structures on passivated prism surfaces for the (a) V1 and (b) V2 samples. (The dimensions are changed for clarity). The DNA origami is labeled with Cy3B on the tether, Atto647N (V1, on the top; V2 on the bottom) and Atto488 (V1 on the bottom, V2 on the top). The color legend in gray and blue indicates states S1 and S2, respectively, which will be used for the remainder of the figure. (c-d) Corrected smFRET efficiency histograms from smTIRF experiments for the BY (left) and YR (right) dye pairs for the docking strands containing complementary nucleotides of 6.5 (orange), 7 (blue), 7.5 (dark blue) and 8 (red) from the (c) V1 and (d) V2 samples. (e-f) Zoomed-in schematics of the two-color, two-state DNA origami structures with 7.5 nt complementary docking strands discussed in panels g and h. (g-h) Corrected smFRET efficiency histograms from two-color smTIRF experiments for the BY (left) and YR (right) dye pairs for (g) V1 and (h) V2 samples.

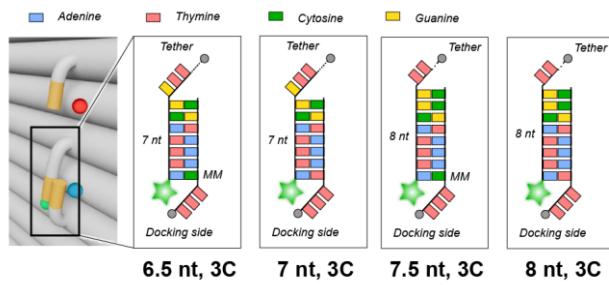
Supplementary Note 7: Summary of all single-molecule MFD-PIE data

In this section, we show the important plots resulting from three- and two-color MFD-PIE experiments for all the measured DNA origami structures.

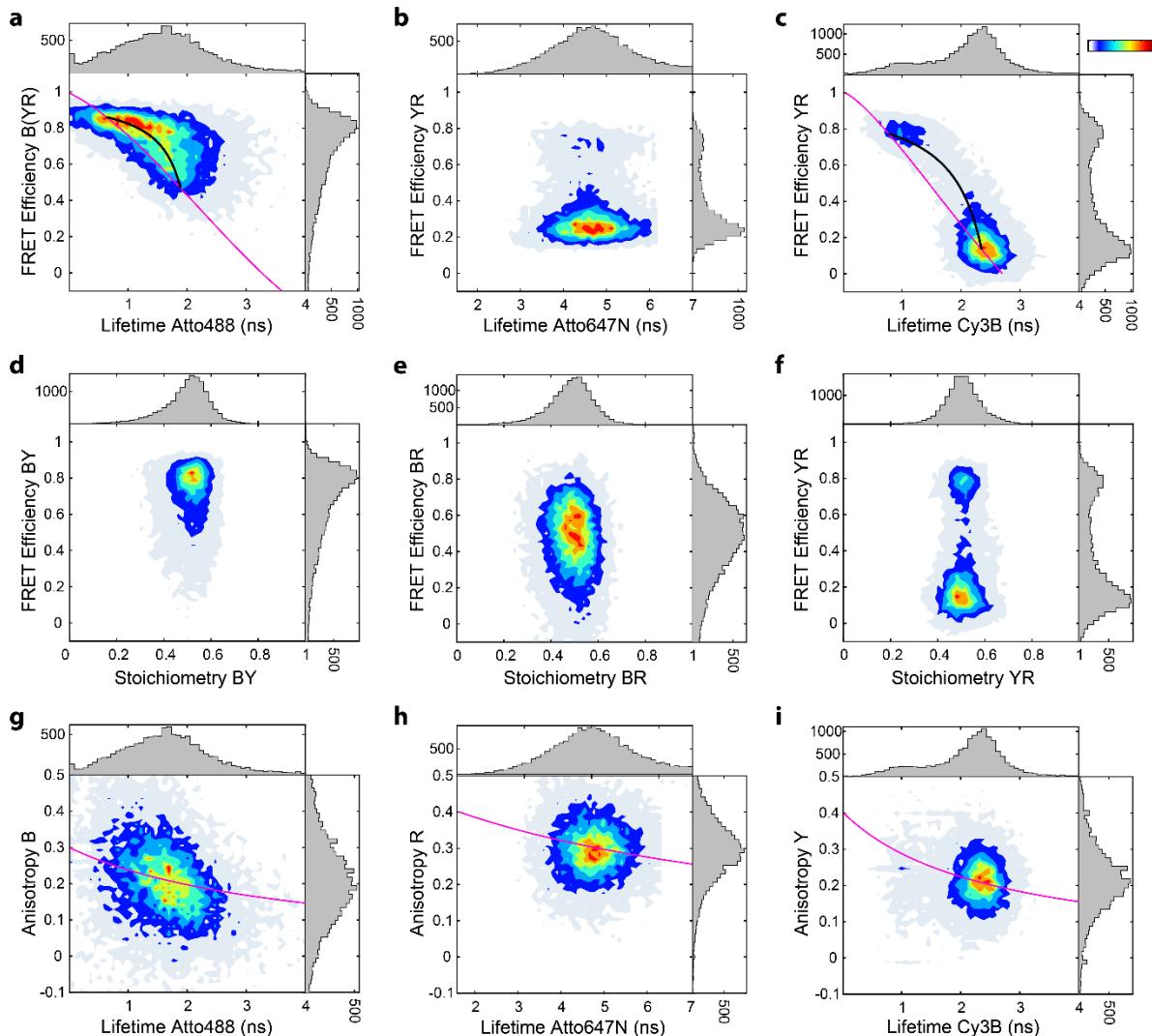
3-color experiments. The data from each three-color sample are summarized in a single figure with nine panels.

2-color experiments. For two-color samples, the important information extracted in the MFD-PIE experiments are summarized in five panels.

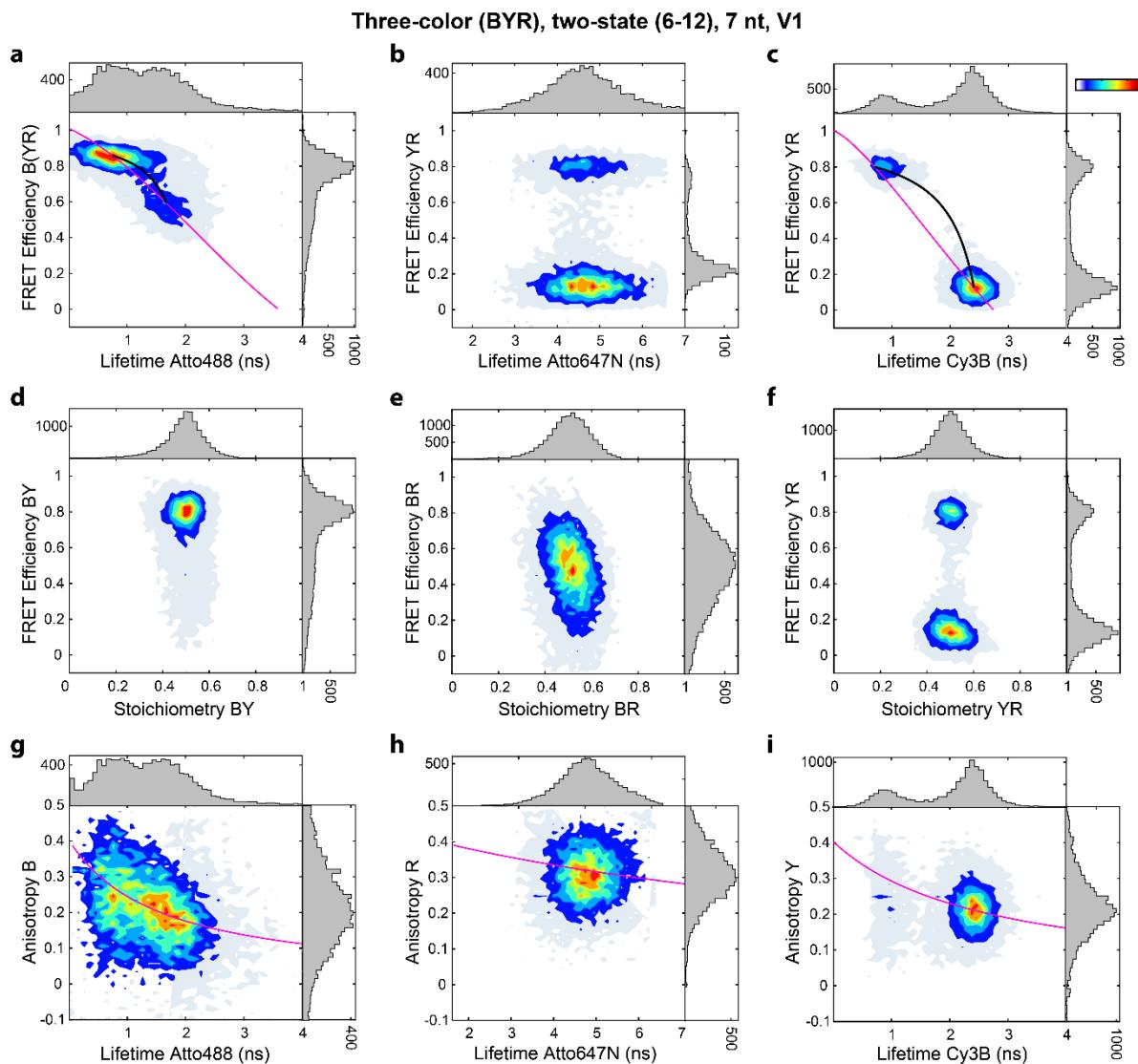
7.1. Three-color, two-state clocks (V1) with variable docking strand length



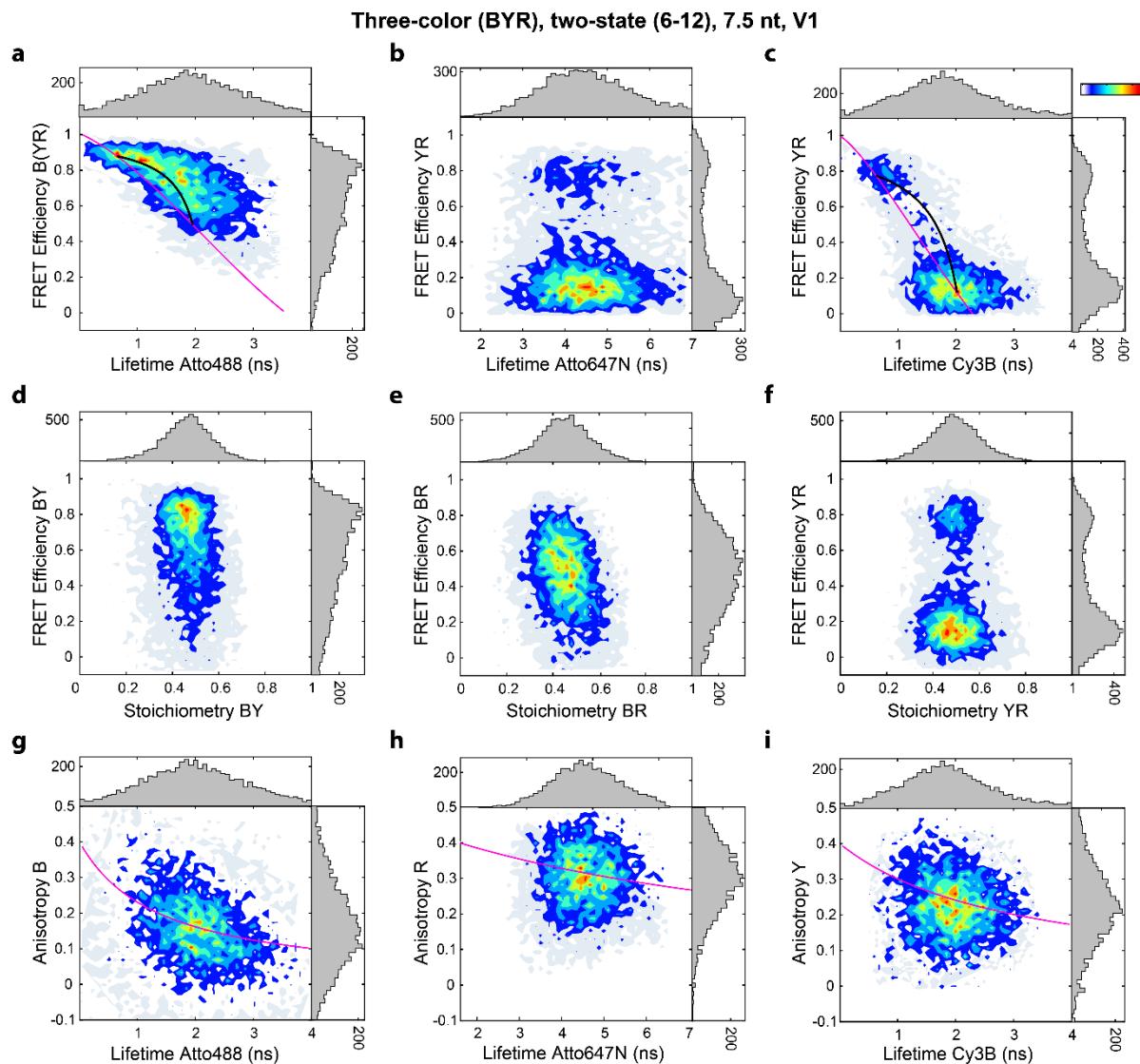
Three-color (BYR), two-state (6-12), 6.5 nt, V1



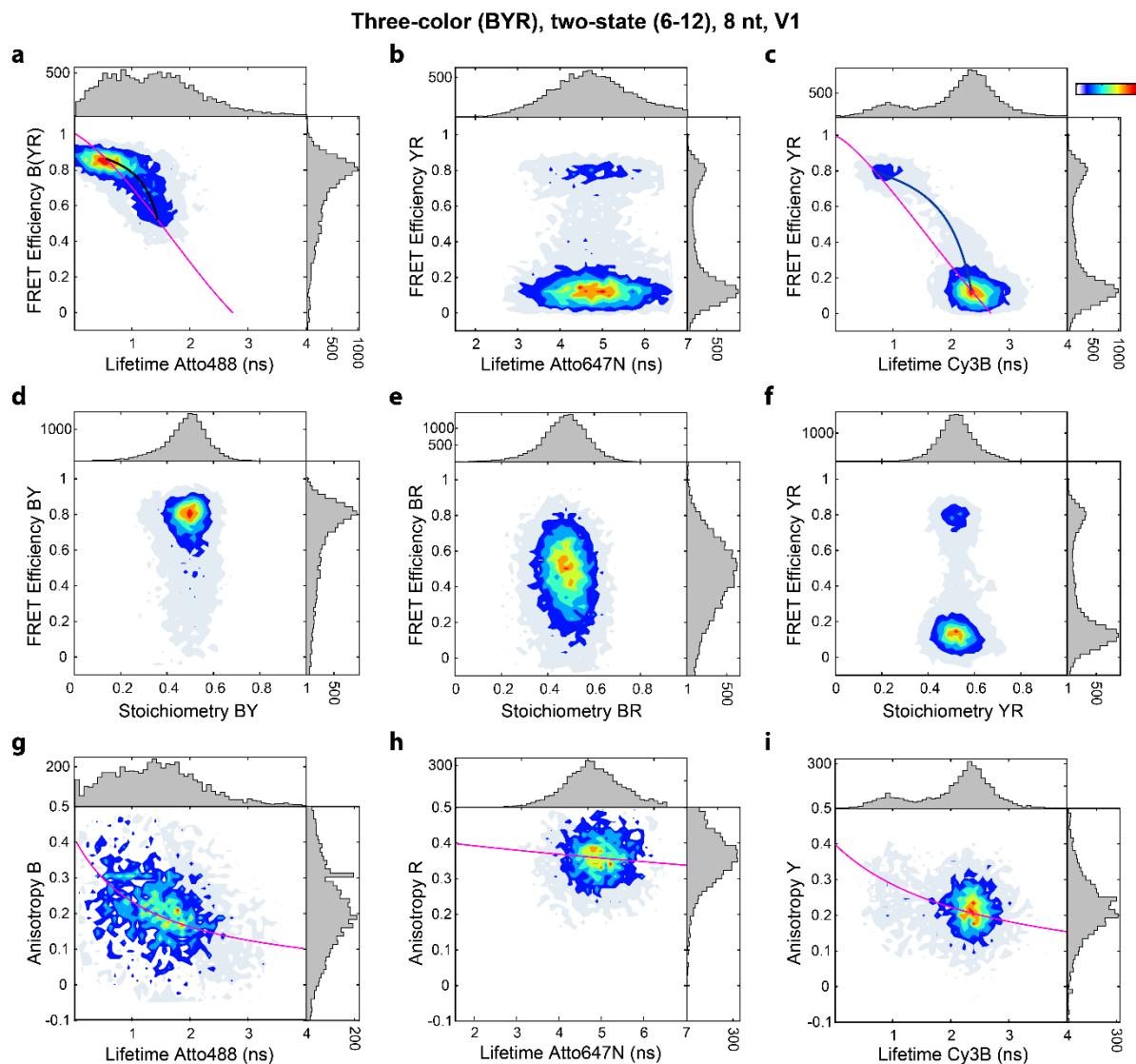
Supplementary Figure S7.1. Overview of 3c MFD-PIE results obtained for the triple-labeled 6-12 clock (Version 1), with 6.5 nt binding sites. (a-c) FRET efficiency histograms are plotted (a) for B \rightarrow YR as a function of the lifetime of Atto488 with blue excitation, and the YR FRET efficiency plotted as a function of (b) the lifetime of Atto647N measured with red excitation and (c) the lifetime of Cy3B with yellow excitation. The static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. (d-f) Stoichiometry versus FRET efficiency for the three dye pairs: (d) BY, (e) BR and (f) YR. (g-i) Anisotropy versus fluorescence lifetime histograms for (g) Atto488 with blue excitation, (h) Atto647N with red excitation and (i) Cy3B with yellow excitation. The magenta line is the corresponding fit to the Perrin Equation (Eq S1).



Supplementary Figure S7.2. Overview of 3c MFD-PIE results obtained for the triple-labeled 6-12 clock (Version 1), with 7 nt binding sites. (a-c) FRET efficiency histograms are plotted (a) for B \rightarrow YR as a function of the lifetime of Atto488 with blue excitation, and the YR FRET efficiency plotted as a function of (b) the lifetime of Atto647N measured with red excitation and (c) the lifetime of Cy3B with yellow excitation. The static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. (d-f) Stoichiometry versus FRET efficiency for the three dye pairs: (d) BY, (e) BR and (f) YR. (g-i) Anisotropy versus fluorescence lifetime histograms for (g) Atto488 with blue excitation, (h) Atto647N with red excitation and (i) Cy3B with yellow excitation. The magenta line is the corresponding fit to the Perrin Equation (Eq S1).

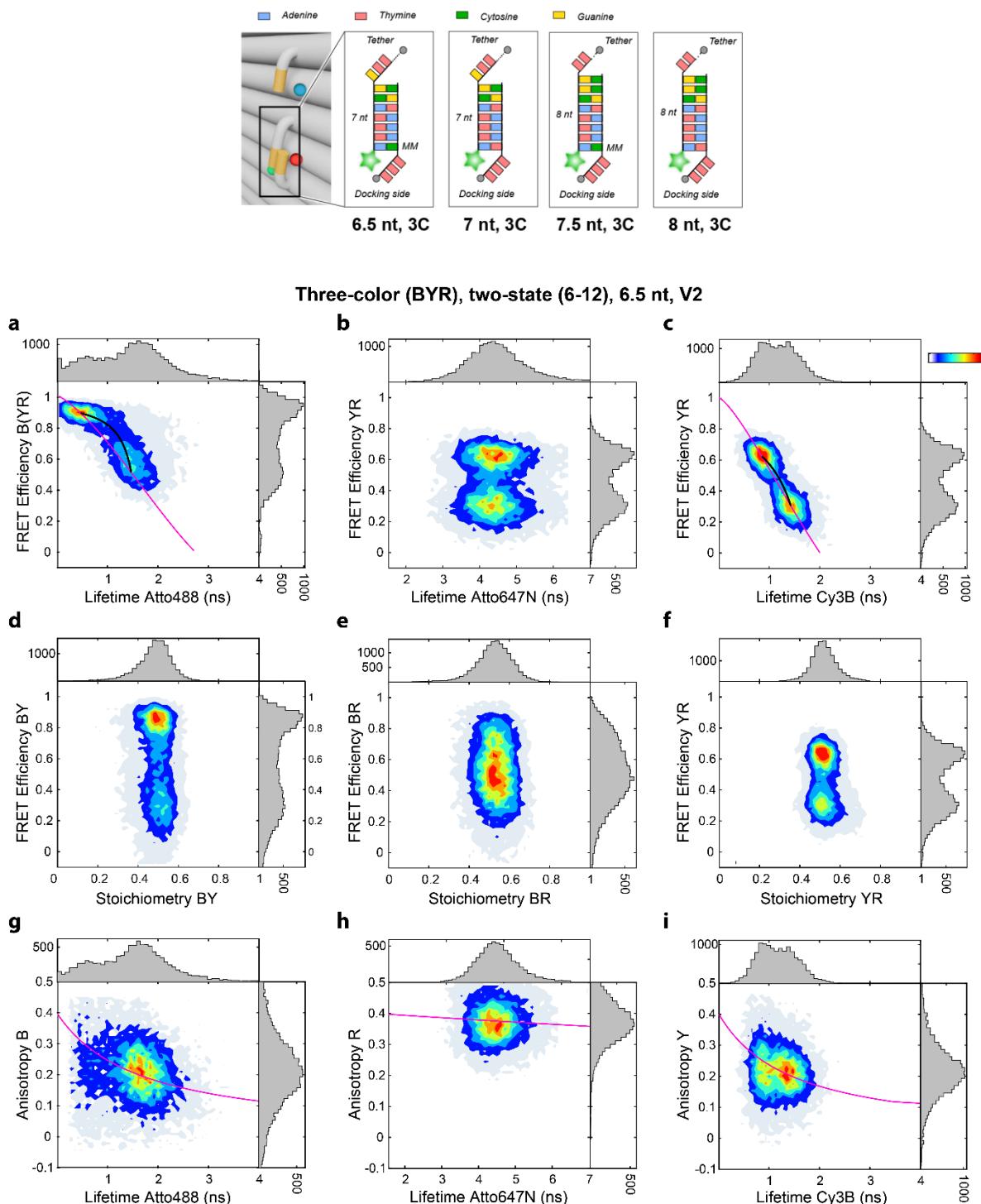


Supplementary Figure S7.3. Overview of 3c MFD-PIE results obtained for the triple-labeled 6-12 clock (Version 1), with 7.5 nt binding sites. (a-c) FRET efficiency histograms are plotted (a) for B \rightarrow YR as a function of the lifetime of Atto488 with blue excitation, and the YR FRET efficiency plotted as a function of (b) the lifetime of Atto647N measured with red excitation and (c) the lifetime of Cy3B with yellow excitation. The static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. (d-f) Stoichiometry versus FRET efficiency for the three dye pairs: (d) BY, (e) BR and (f) YR. (g-i) Anisotropy versus fluorescence lifetime histograms for (g) Atto488 with blue excitation, (h) Atto647N with red excitation and (i) Cy3B with yellow excitation. The magenta line is the corresponding fit to the Perrin Equation (Eq S1).

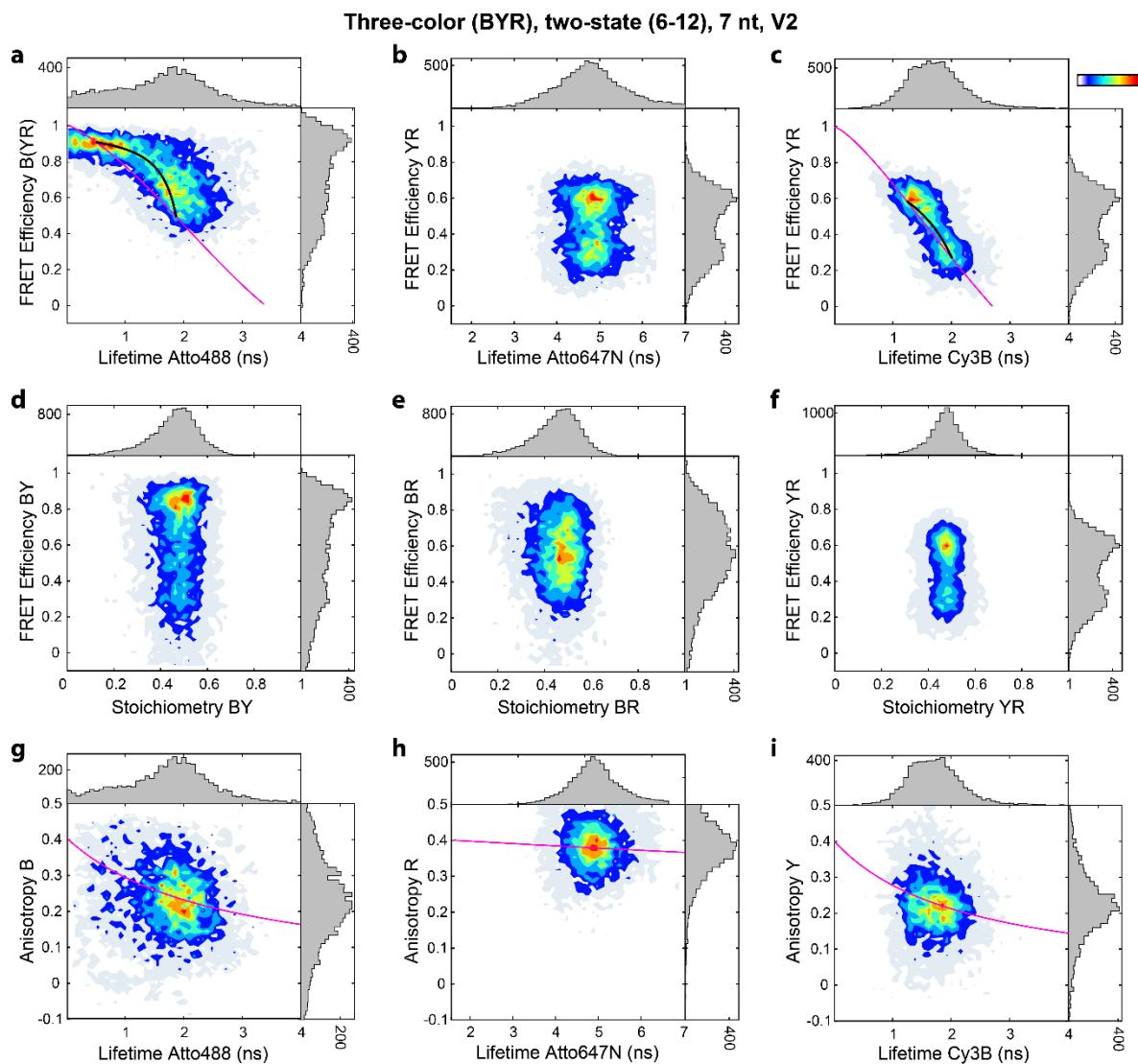


Supplementary Figure S7.4. Overview of 3c MFD-PIE results obtained for the triple-labeled 6-12 clock (Version 1), with 8 nt binding sites. (a-c) FRET efficiency histograms are plotted (a) for B \rightarrow YR as a function of the lifetime of Atto488 with blue excitation, and the YR FRET efficiency plotted as a function of (b) the lifetime of Atto647N measured with red excitation and (c) the lifetime of Cy3B with yellow excitation. The static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. (d-f) Stoichiometry versus FRET efficiency for the three dye pairs: (d) BY, (e) BR and (f) YR. (g-i) Anisotropy versus fluorescence lifetime histograms for (g) Atto488 with blue excitation, (h) Atto647N with red excitation and (i) Cy3B with yellow excitation. The magenta line is the corresponding fit to the Perrin Equation (Eq S1).

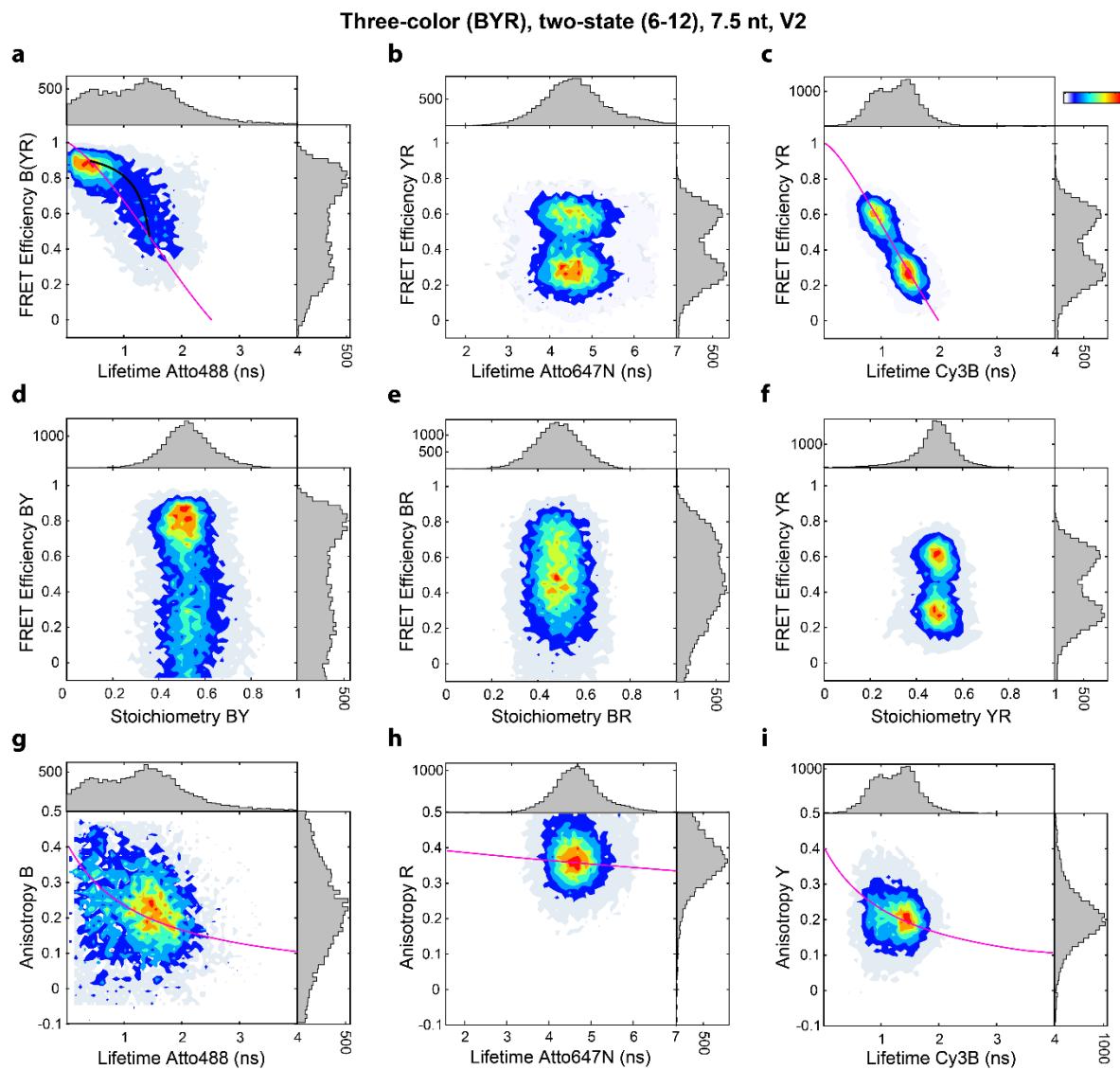
7.2. Three-color, two-state clocks (V2) with variable docking strand length



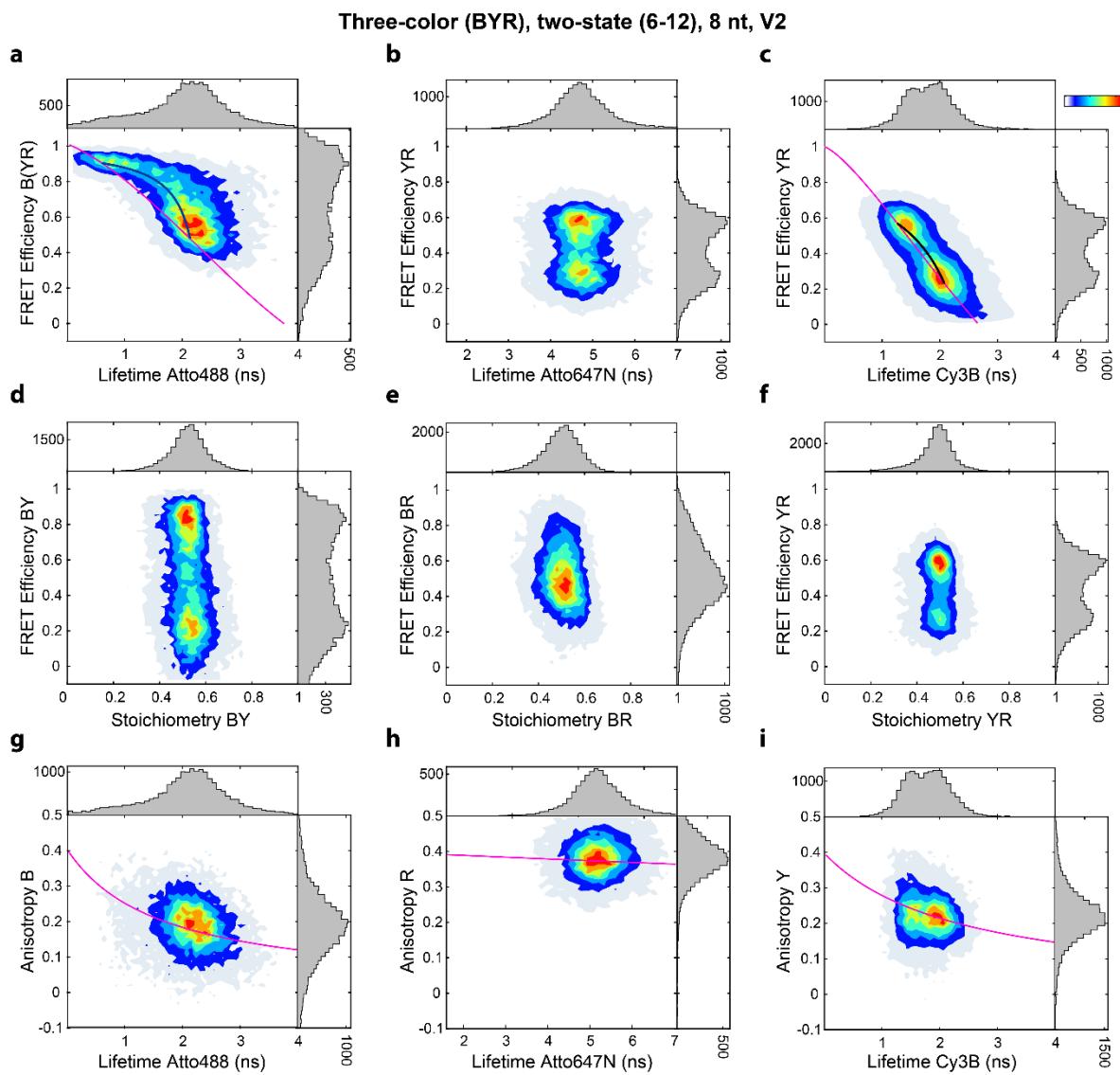
Supplementary Figure S7.5. Overview of 3c MFD-PIE results obtained for the triple-labeled 6-12 clock (Version 2), with 6.5 nt binding sites. (a-c) FRET efficiency histograms are plotted (a) for B → YR as a function of the lifetime of Atto488 with blue excitation, and the YR FRET efficiency plotted as a function of (b) the lifetime of Atto647N measured with red excitation and (c) the lifetime of Cy3B with yellow excitation. The static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. (d-f) Stoichiometry versus FRET efficiency for the three dye pairs: (d) BY, (e) BR and (f) YR. (g-i) Anisotropy versus fluorescence lifetime histograms for (g) Atto488 with blue excitation, (h) Atto647N with red excitation and (i) Cy3B with yellow excitation. The magenta line is the corresponding fit to the Perrin Equation (Eq S1).



Supplementary Figure S7.6. Overview of 3c MFD-PIE results obtained for the triple-labeled 6-12 clock (Version 2), with 7 nt binding sites. (a-c) FRET efficiency histograms are plotted (a) for B \rightarrow YR as a function of the lifetime of Atto488 with blue excitation, and the YR FRET efficiency plotted as a function of (b) the lifetime of Atto647N measured with red excitation and (c) the lifetime of Cy3B with yellow excitation. The static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. (d-f) Stoichiometry versus FRET efficiency for the three dye pairs: (d) BY, (e) BR and (f) YR. (g-i) Anisotropy versus fluorescence lifetime histograms for (g) Atto488 with blue excitation, (h) Atto647N with red excitation and (i) Cy3B with yellow excitation. The magenta line is the corresponding fit to the Perrin Equation (Eq S1).

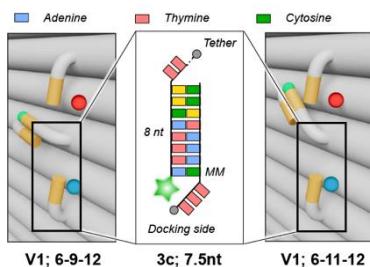


Supplementary Figure S7.7. Overview of 3c MFD-PIE results obtained for the triple-labeled 6-12 clock (Version 2), with 7.5 nt binding sites. (a-c) FRET efficiency histograms are plotted (a) for B \rightarrow YR as a function of the lifetime of Atto488 with blue excitation, and the YR FRET efficiency plotted as a function of (b) the lifetime of Atto647N measured with red excitation and (c) the lifetime of Cy3B with yellow excitation. The static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. (d-f) Stoichiometry versus FRET efficiency for the three dye pairs: (d) BY, (e) BR and (f) YR. (g-i) Anisotropy versus fluorescence lifetime histograms for (g) Atto488 with blue excitation, (h) Atto647N with red excitation and (i) Cy3B with yellow excitation. The magenta line is the corresponding fit to the Perrin Equation (Eq S1).

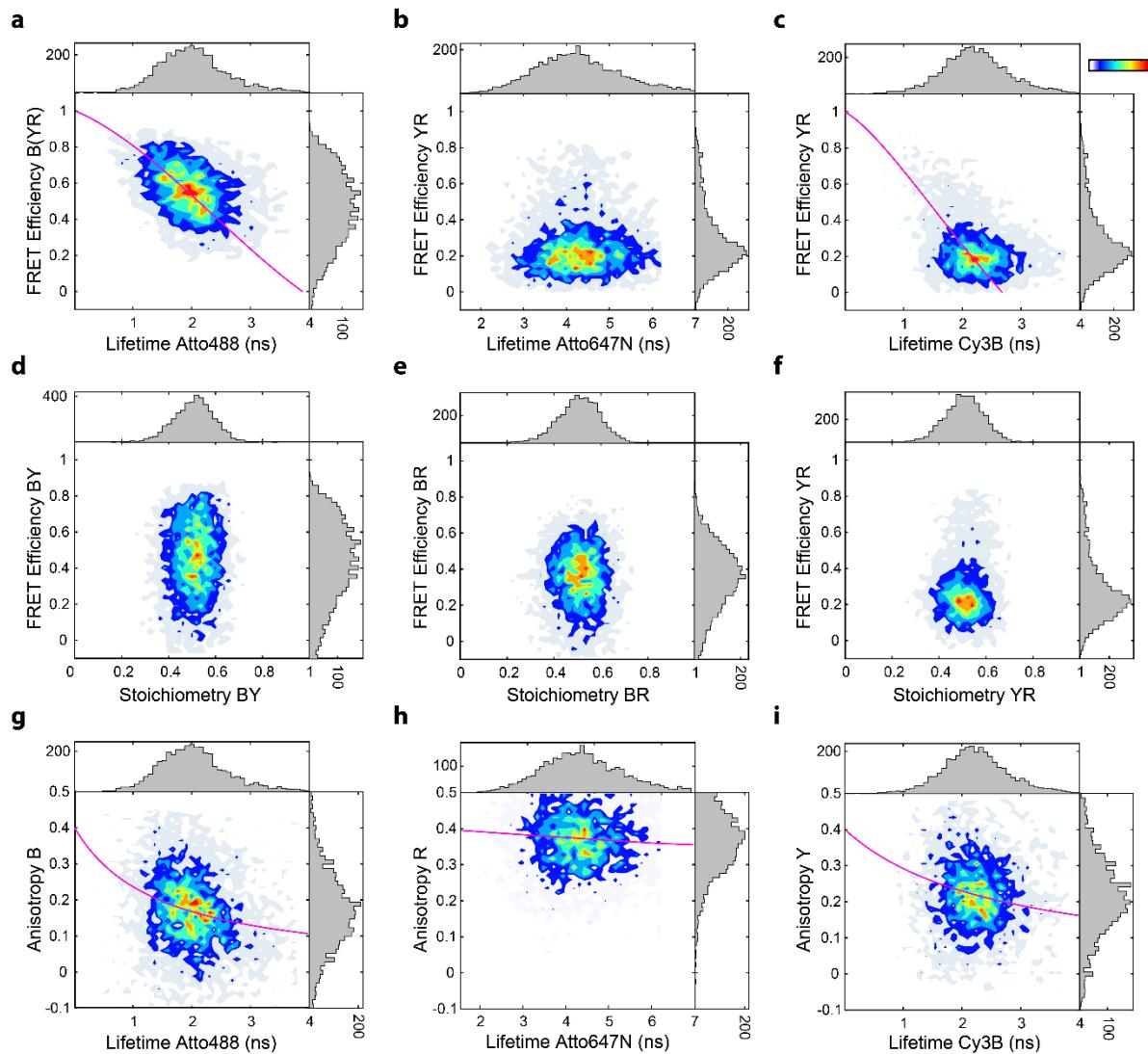


Supplementary Figure S7.8. Overview of 3c MFD-PIE results obtained for the triple-labeled 6-12 clock (Version 2), with 8 nt binding sites. (a-c) FRET efficiency histograms are plotted (a) for B \rightarrow YR as a function of the lifetime of Atto488 with blue excitation, and the YR FRET efficiency plotted as a function of (b) the lifetime of Atto647N measured with red excitation and (c) the lifetime of Cy3B with yellow excitation. The static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. (d-f) Stoichiometry versus FRET efficiency for the three dye pairs: (d) BY, (e) BR and (f) YR. (g-i) Anisotropy versus fluorescence lifetime histograms for (g) Atto488 with blue excitation, (h) Atto647N with red excitation and (i) Cy3B with yellow excitation. The magenta line is the corresponding fit to the Perrin Equation (Eq S1).

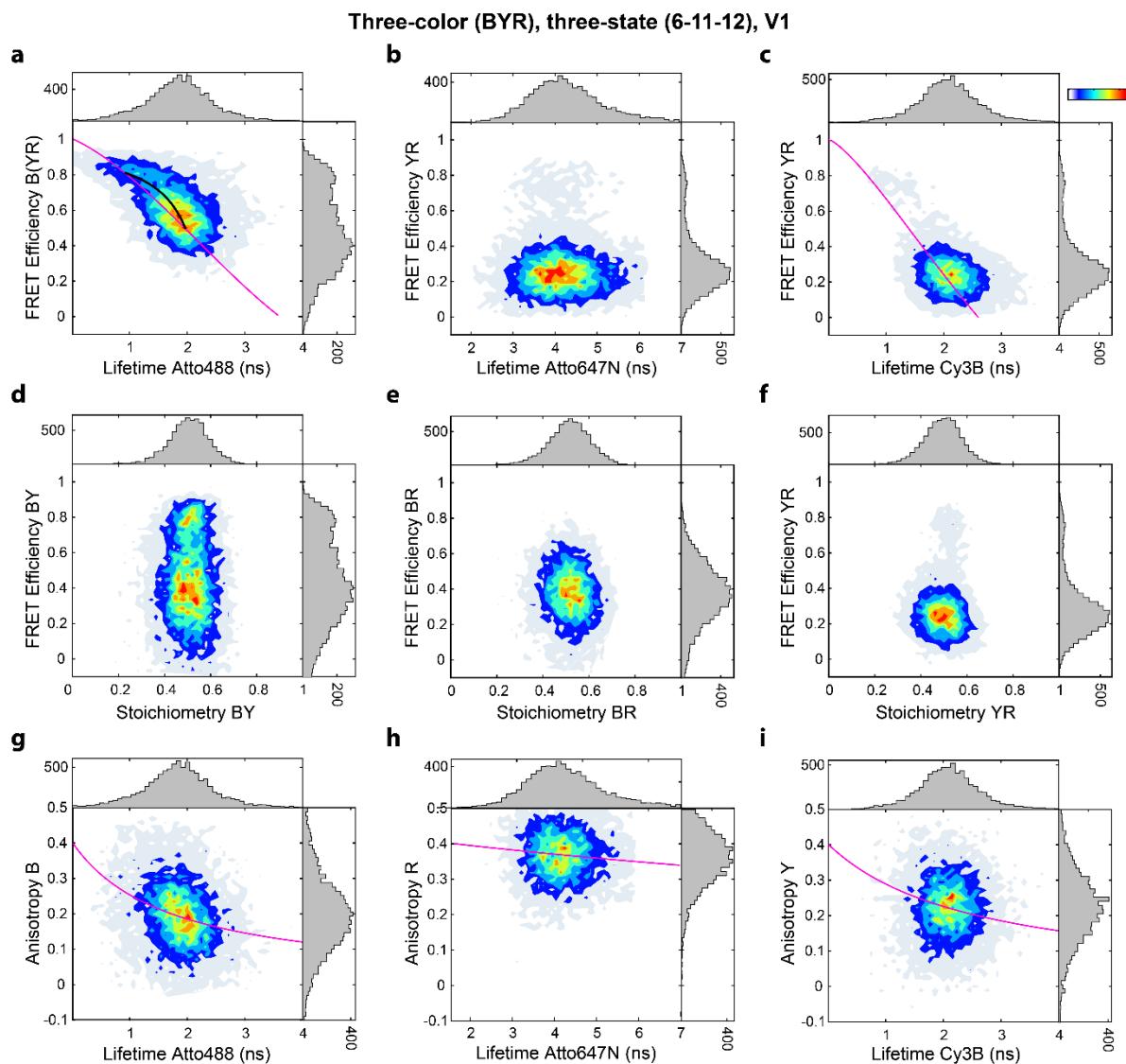
7.3. Three-color, three-state clocks (V1) with different binding positions



Three-color (BYR), three-state (6-9-12), V1

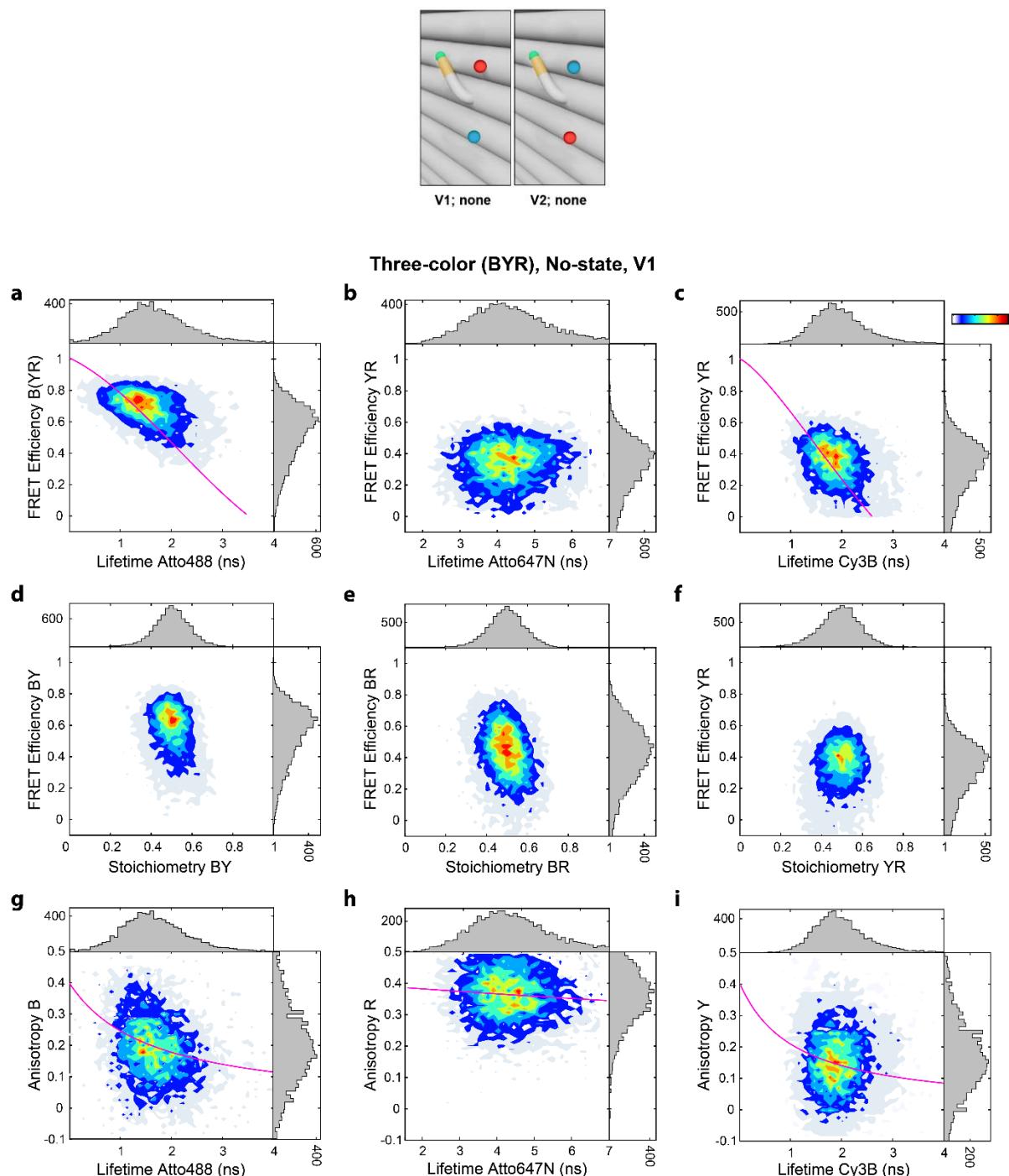


Supplementary Figure S7.9. Overview of 3c MFD-PIE results obtained for the triple-labeled 6-9-12 clock (Version 1). (a-c) FRET efficiency histograms are plotted (a) for B \rightarrow YR as a function of the lifetime of Atto488 with blue excitation, and the YR FRET efficiency plotted as a function of (b) the lifetime of Atto647N measured with red excitation and (c) the lifetime of Cy3B with yellow excitation. The static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. (d-f) Stoichiometry versus FRET efficiency for the three dye pairs: (d) BY, (e) BR and (f) YR. (g-i) Anisotropy versus fluorescence lifetime histograms for (g) Atto488 with blue excitation, (h) Atto647N with red excitation and (i) Cy3B with yellow excitation. The magenta line is the corresponding fit to the Perrin Equation (Eq S1).

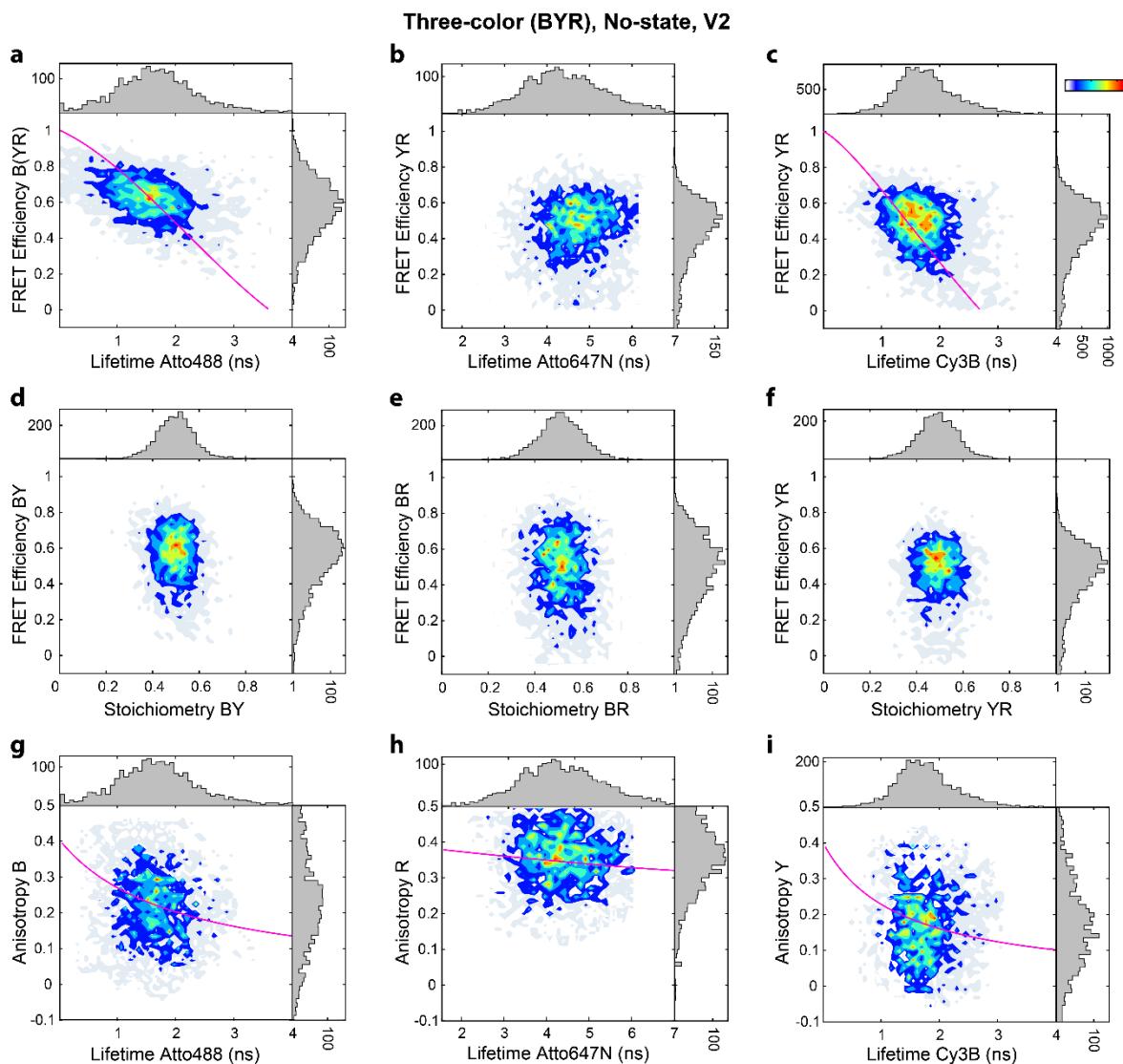


Supplementary Figure S7.10. Overview of 3c MFD-PIE results obtained for the triple-labeled 6-11-12 clock (Version 1). (a-c) FRET efficiency histograms are plotted (a) for B \rightarrow YR as a function of the lifetime of Atto488 with blue excitation, and the YR FRET efficiency plotted as a function of (b) the lifetime of Atto647N measured with red excitation and (c) the lifetime of Cy3B with yellow excitation. The static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. (d-f) Stoichiometry versus FRET efficiency for the three dye pairs: (d) BY, (e) BR and (f) YR. (g-i) Anisotropy versus fluorescence lifetime histograms for (g) Atto488 with blue excitation, (h) Atto647N with red excitation and (i) Cy3B with yellow excitation. The magenta line is the corresponding fit to the Perrin Equation (Eq S1).

7.4. Triple-color DNA origamis for control

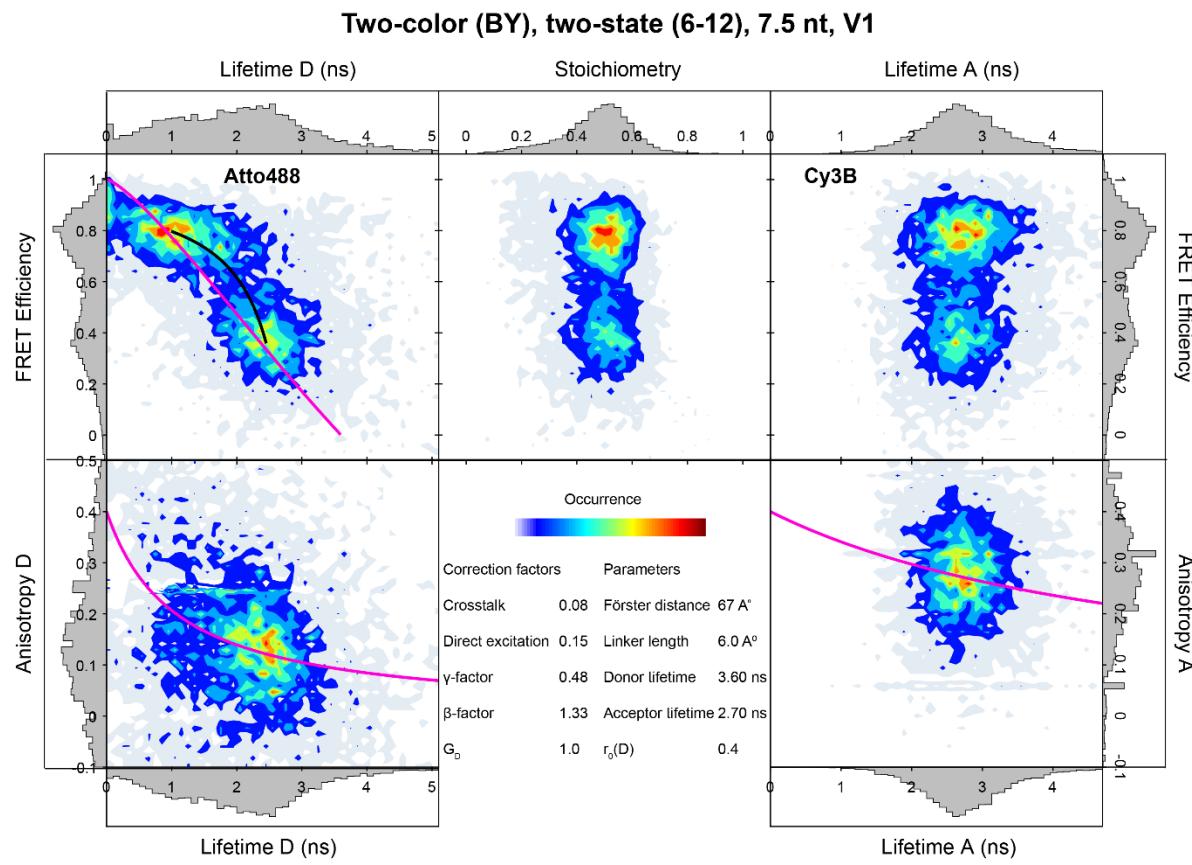


Supplementary Figure S7.11. Overview of 3c MFD-PIE results obtained for the triple-labeled DNA origami structure (Version 1) without any binding sites. (a-c) FRET efficiency histograms are plotted (a) for B \rightarrow YR as a function of the lifetime of Atto488 with blue excitation, and the YR FRET efficiency plotted as a function of (b) the lifetime of Atto647N measured with red excitation and (c) the lifetime of Cy3B with yellow excitation. The static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. (d-f) Stoichiometry versus FRET efficiency for the three dye pairs: (d) BY, (e) BR and (f) YR. (g-i) Anisotropy versus fluorescence lifetime histograms for (g) Atto488 with blue excitation, (h) Atto647N with red excitation and (i) Cy3B with yellow excitation. The magenta line is the corresponding fit to the Perrin Equation (Eq S1).

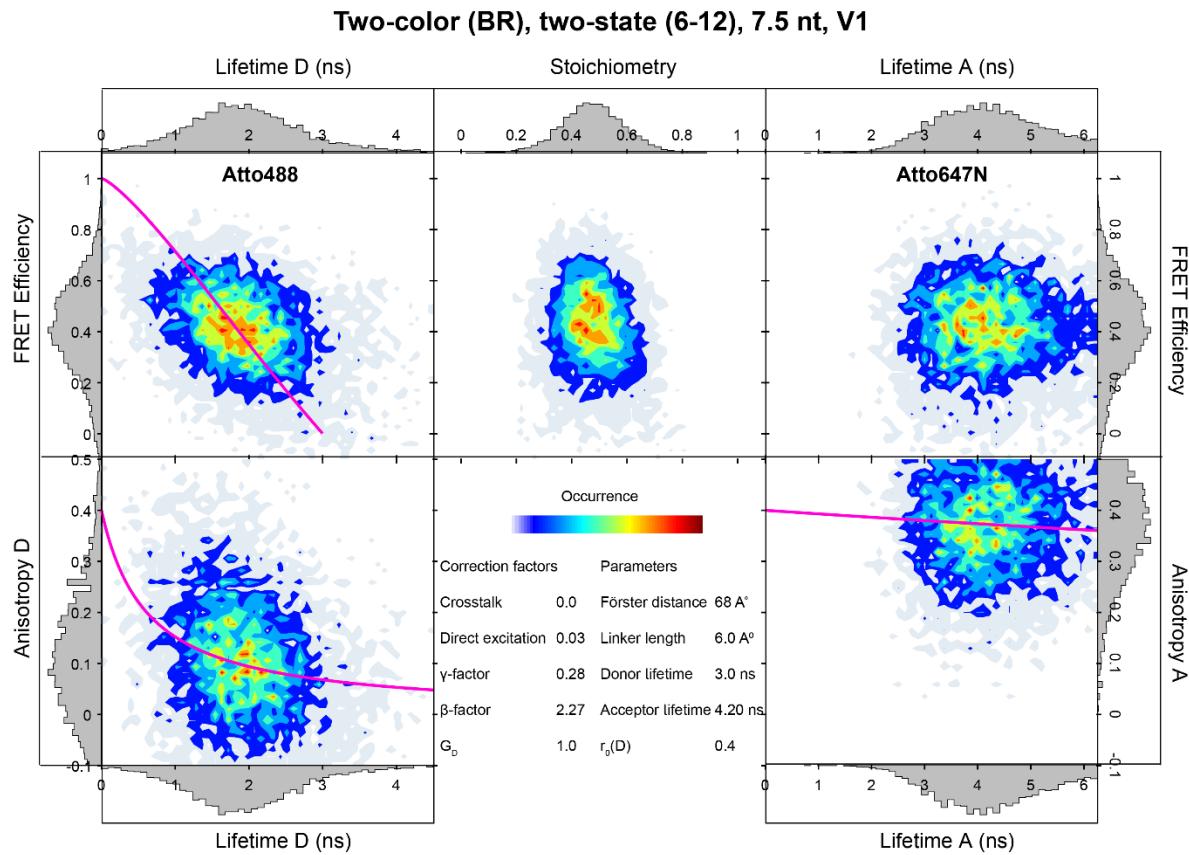


Supplementary Figure S7.12. Overview of 3c MFD-PIE results obtained for the triple-labeled DNA origami structure (Version 2) without any binding sites. (a-c) FRET efficiency histograms are plotted (a) for B \rightarrow YR as a function of the lifetime of Atto488 with blue excitation, and the YR FRET efficiency plotted as a function of (b) the lifetime of Atto647N measured with red excitation and (c) the lifetime of Cy3B with yellow excitation. The static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. (d-f) Stoichiometry versus FRET efficiency for the three dye pairs: (d) BY, (e) BR and (f) YR. (g-i) Anisotropy versus fluorescence lifetime histograms for (g) Atto488 with blue excitation, (h) Atto647N with red excitation and (i) Cy3B with yellow excitation. The magenta line is the corresponding fit to the Perrin Equation (Eq S1).

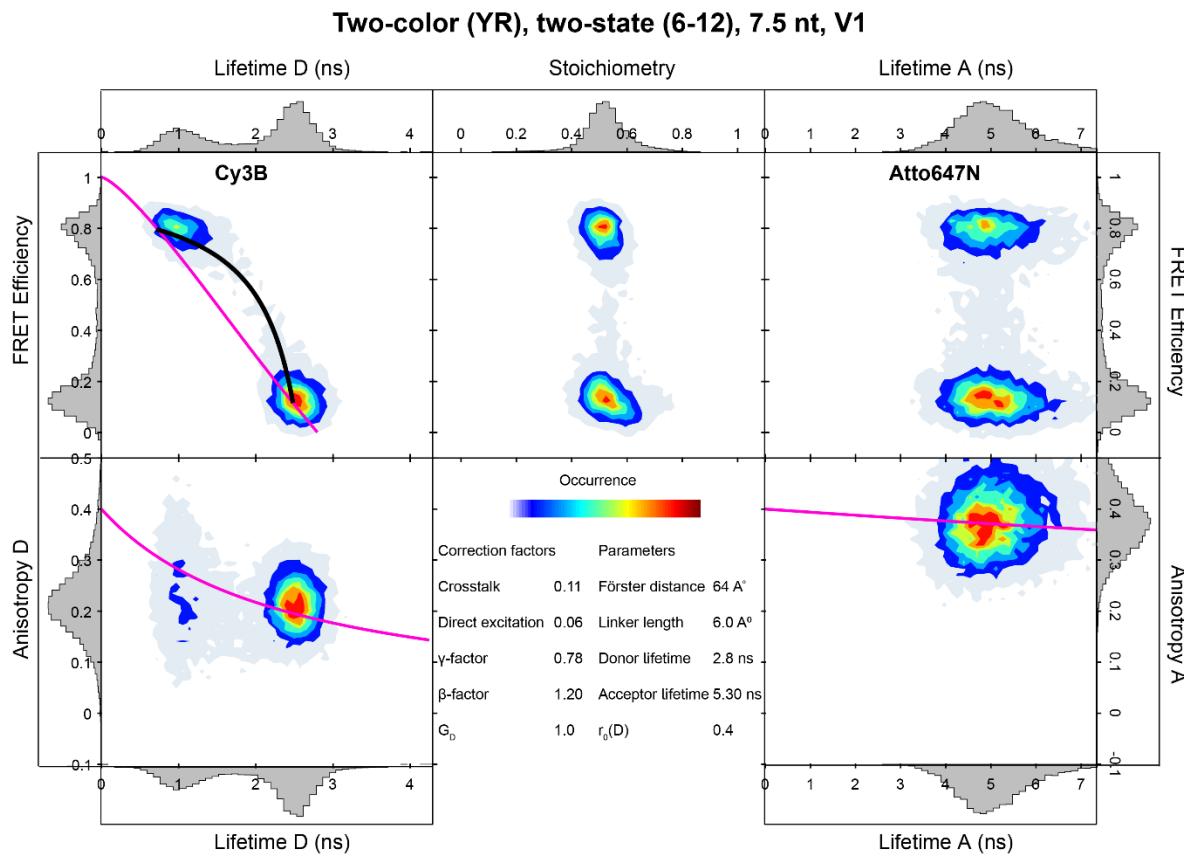
7.5. Dual-color 2 State clocks (V1) with 7.5. nt binding sites and different dye pair combinations



Supplementary Figure S7.13. Overview of 2c MFD-PIE results obtained for the 6-12 clock (Version 1) with 7.5 nt binding sites labeled with Atto488 and Cy3B. Top row: The FRET efficiency for the FRET pair is plotted as a function of (left) Donor lifetime, (middle) stoichiometry and (right) and acceptor lifetime. In the left panel, the static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. Bottom row: The anisotropy of the (left) donor and (right) acceptor molecules are plotted as a function of the respective lifetimes. A fit of the burst-wise steady-state anisotropy to the Perrin Equation (Eq S1) is shown in magenta. Middle: A summary of the parameters extracted from the data.

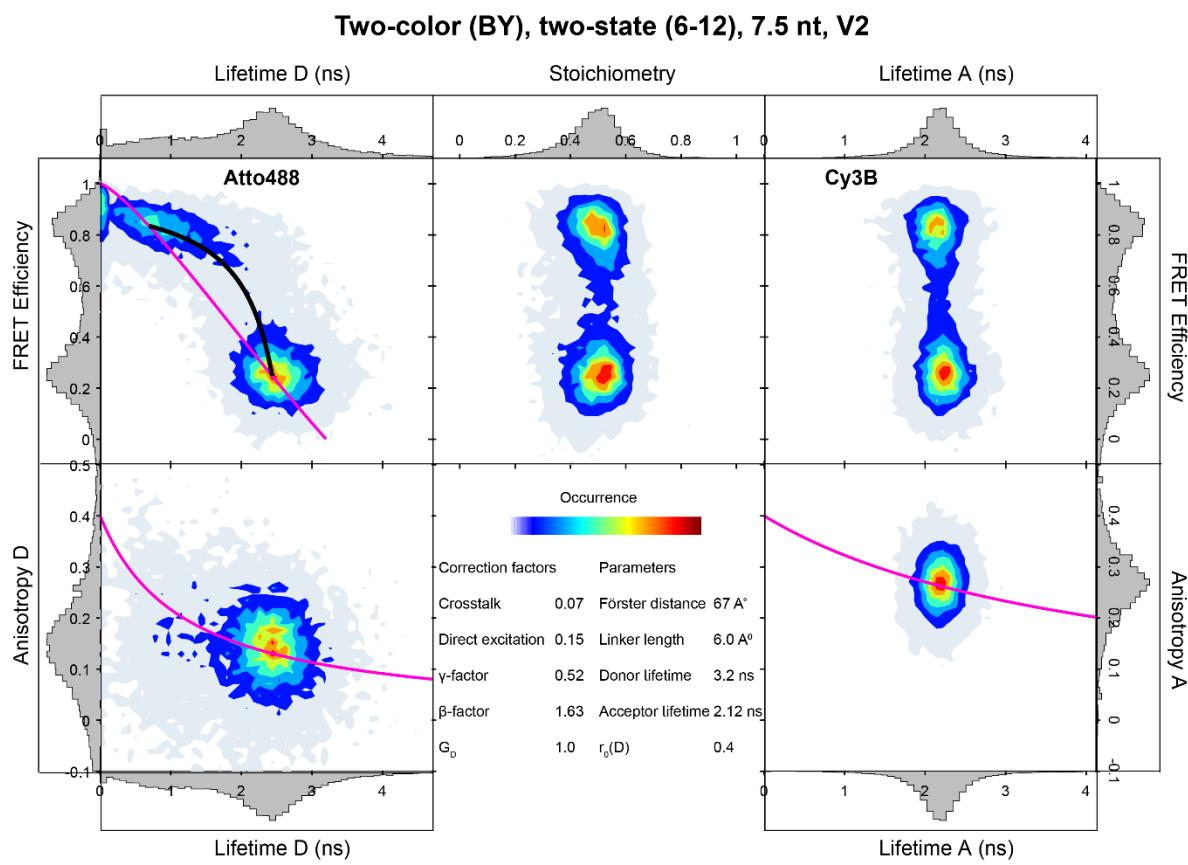


Supplementary Figure S7.14. Overview of 2c MFD-PIE results obtained for the 6-12 clock (Version 1) with 7.5 nt binding sites labeled with Atto488 and Atto647N. Top row: The FRET efficiency for the FRET pair is plotted as a function of (left) Donor lifetime, (middle) stoichiometry and (right) and acceptor lifetime. In the *left* panel, the static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. Bottom row: The anisotropy of the (left) donor and (right) acceptor molecules are plotted as a function of the respective lifetimes. A fit of the burst-wise steady-state anisotropy to the Perrin Equation (Eq S1) is shown in magenta. Middle: A summary of the parameters extracted from the data.

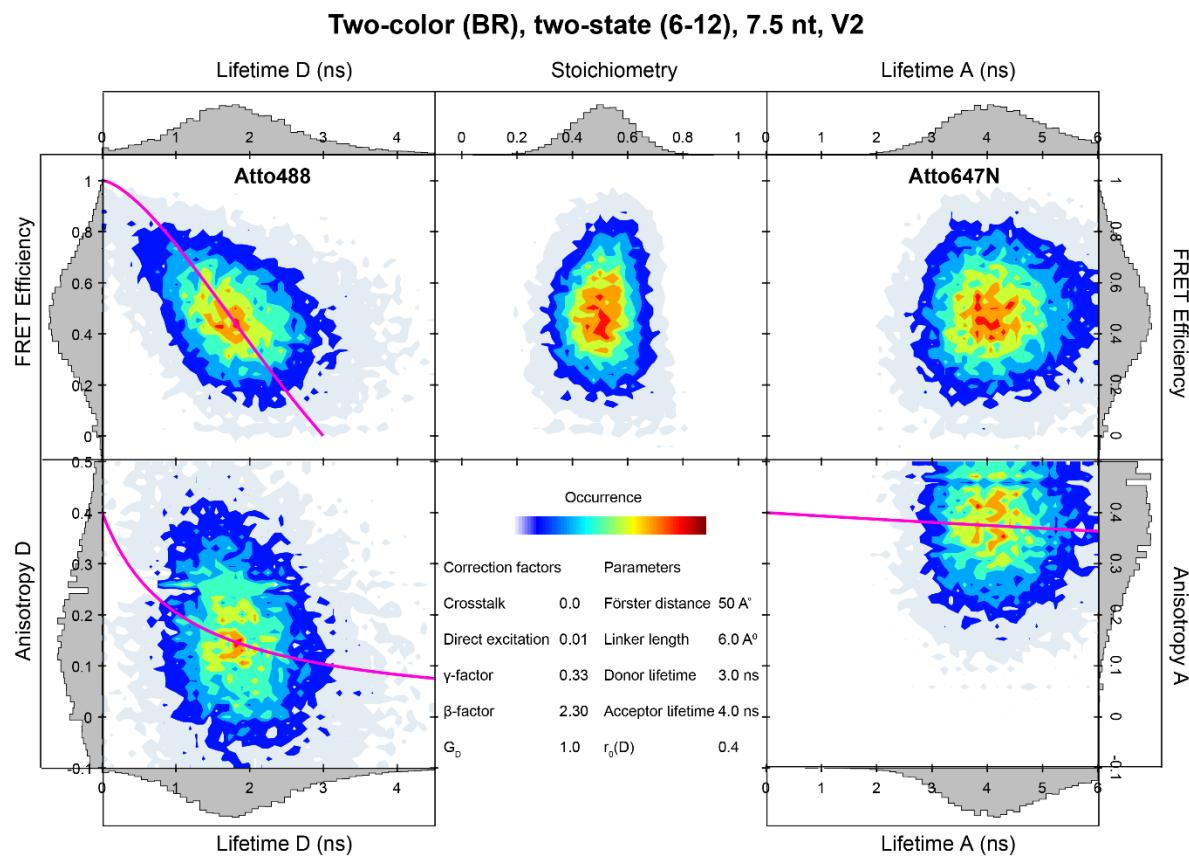


Supplementary Figure S7.15. Overview of 2c MFD-PIE results obtained for the 6-12 clock (Version 1) with 7.5 nt binding sites labeled with Cy3B and Atto647N. Top row: The FRET efficiency for the FRET pair is plotted as a function of (left) Donor lifetime, (middle) stoichiometry and (right) and acceptor lifetime. In the left panel, the static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. Bottom row: The anisotropy of the (left) donor and (right) acceptor molecules are plotted as a function of the respective lifetimes. A fit of the burst-wise steady-state anisotropy to the Perrin Equation (Eq S1) is shown in magenta. Middle: A summary of the parameters extracted from the data.

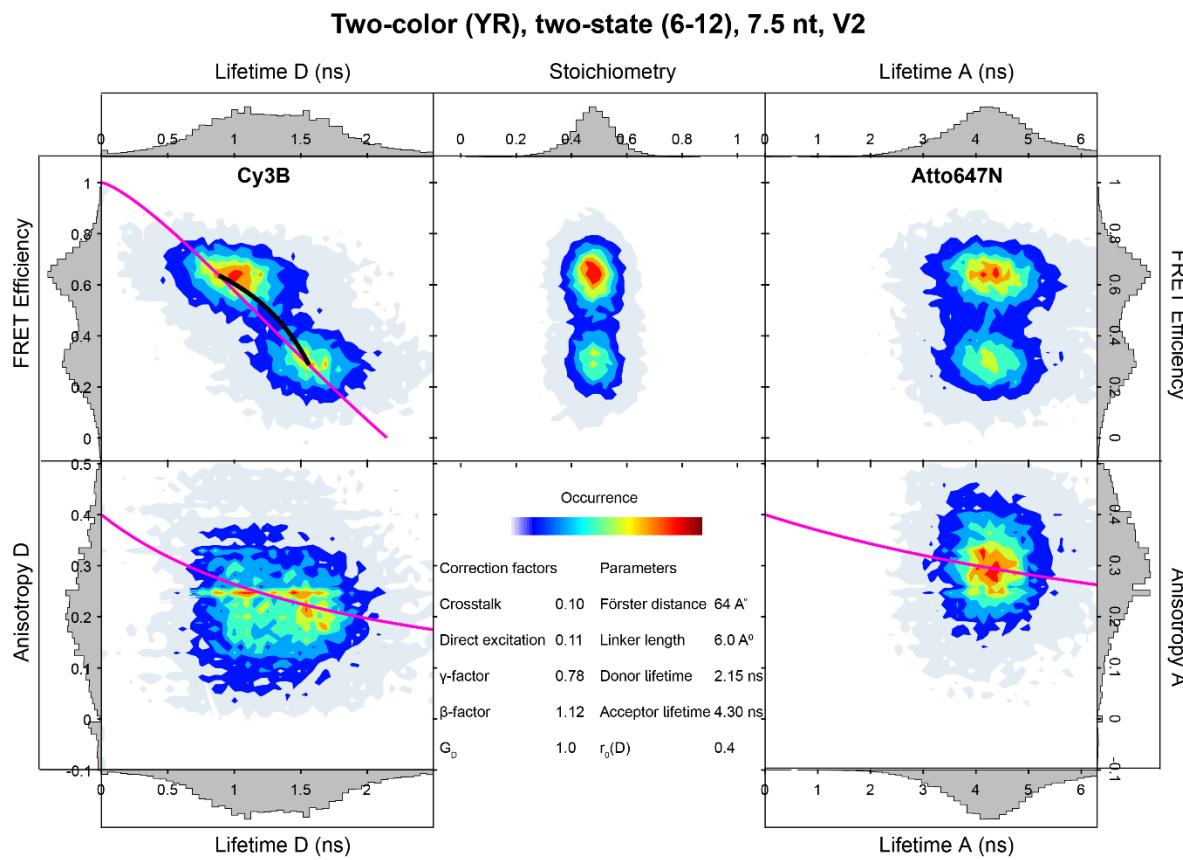
7.6. Dual-color 2 State clocks (V2) with 7.5. nt binding sites and different dye pair combinations



Supplementary Figure S7.16. Overview of 2c MFD-PIE results obtained for the 6-12 clock (Version 2) with 7.5 nt binding sites labeled with Atto488 and Cy3B. Top row: The FRET efficiency for the FRET pair is plotted as a function of (left) Donor lifetime, (middle) stoichiometry and (right) and acceptor lifetime. In the left panel, the static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. Bottom row: The anisotropy of the (left) donor and (right) acceptor molecules are plotted as a function of the respective lifetimes. A fit of the burst-wise steady-state anisotropy to the Perrin Equation (Eq S1) is shown in magenta. Middle: A summary of the parameters extracted from the data.

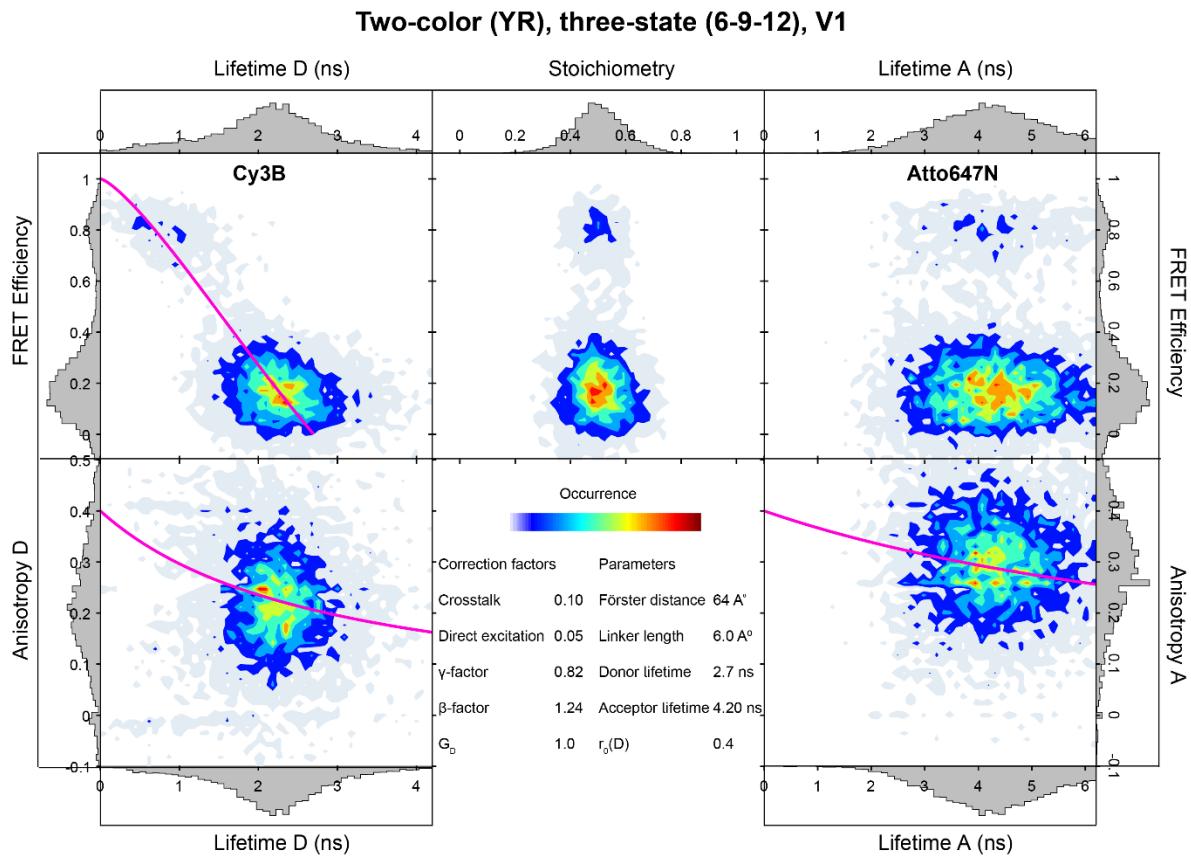


Supplementary Figure S7.17. Overview of 2c MFD-PIE results obtained for the 6-12 clock (Version 2) with 7.5 nt binding sites labeled with Atto488 and Atto647N. Top row: The FRET efficiency for the FRET pair is plotted as a function of (left) Donor lifetime, (middle) stoichiometry and (right) and acceptor lifetime. In the *left* panel, the static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. Bottom row: The anisotropy of the (left) donor and (right) acceptor molecules are plotted as a function of the respective lifetimes. A fit of the burst-wise steady-state anisotropy to the Perrin Equation (Eq S1) is shown in magenta. Middle: A summary of the parameters extracted from the data.

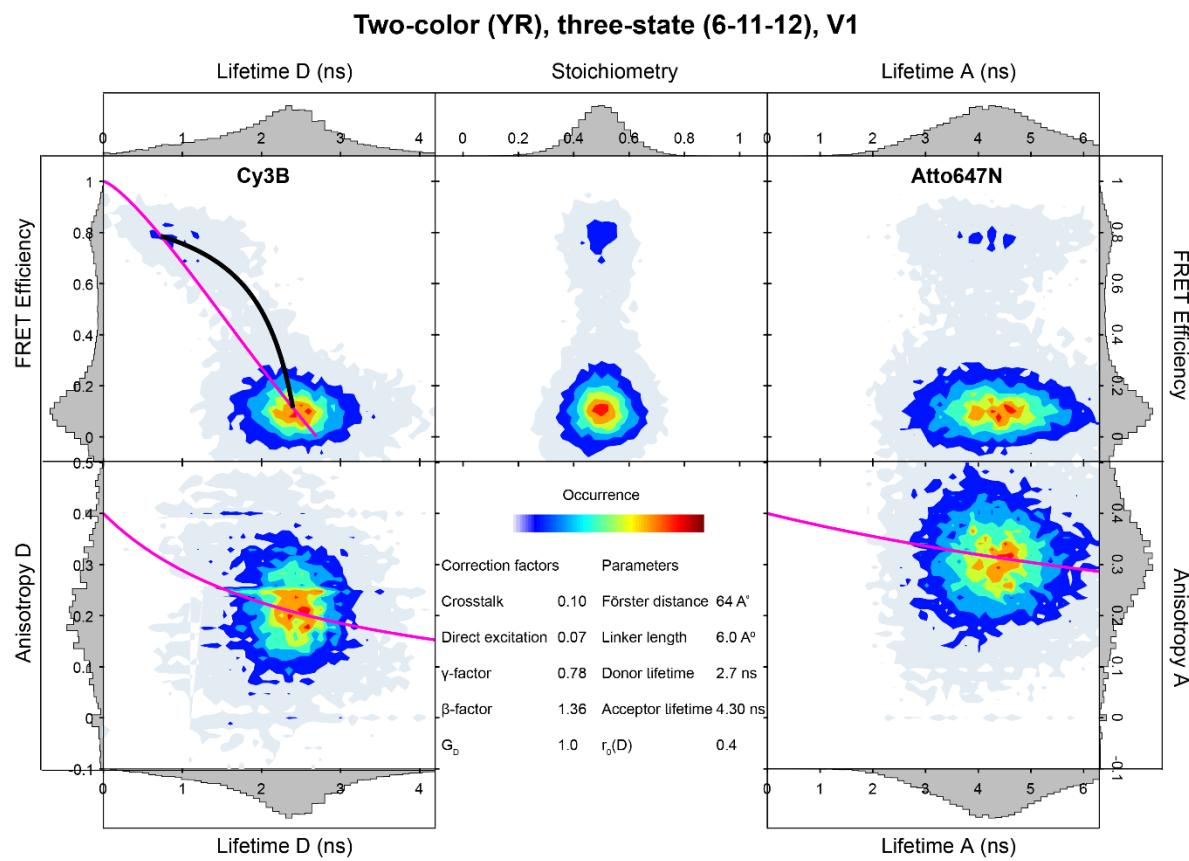


Supplementary Figure S7.18. Overview of 2c MFD-PIE results obtained for the 6-12 clock (Version 2) with 7.5 nt binding sites labeled with Cy3B and Atto647N. Top row: The FRET efficiency for the FRET pair is plotted as a function of (left) Donor lifetime, (middle) stoichiometry and (right) and acceptor lifetime. In the left panel, the static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. Bottom row: The anisotropy of the (left) donor and (right) acceptor molecules are plotted as a function of the respective lifetimes. A fit of the burst-wise steady-state anisotropy to the Perrin Equation (Eq S1) is shown in magenta. Middle: A summary of the parameters extracted from the data.

7.7. Dual-color 3 State clocks (V1) with different binding positions

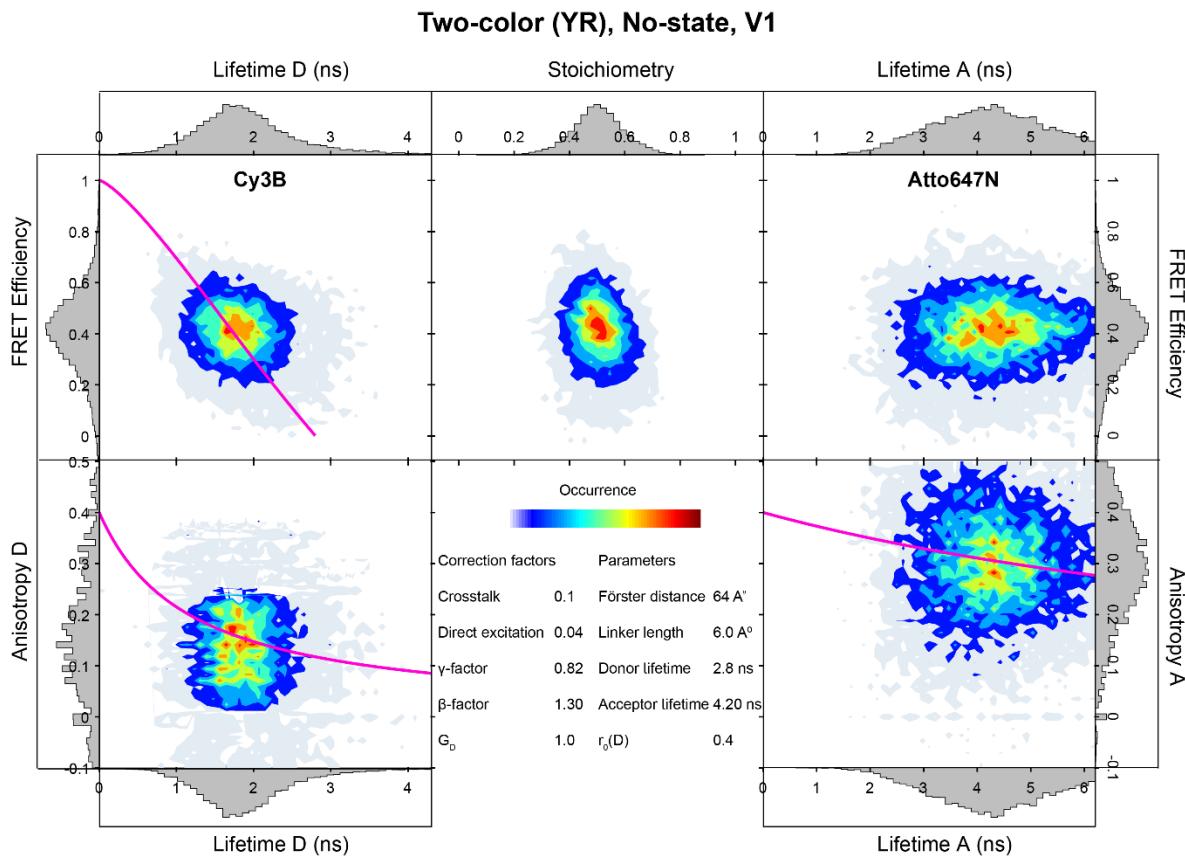


Supplementary Figure S7.19. Overview of 2c MFD-PIE results obtained for the dual-labeled 6-9-12 clock (Version 1). *Top row:* The FRET efficiency for the FRET pair is plotted as a function of (*left*) Donor lifetime, (*middle*) stoichiometry and (*right*) acceptor lifetime. In the *left* panel, the static FRET line is plotted in magenta. For details, see [3]. *Bottom row:* The anisotropy of the (*left*) donor and (*right*) acceptor molecules are plotted as a function of the respective lifetimes. A fit of the burst-wise steady-state anisotropy to the Perrin Equation (Eq S1) is shown in magenta. Middle: A summary of the parameters extracted from the data.

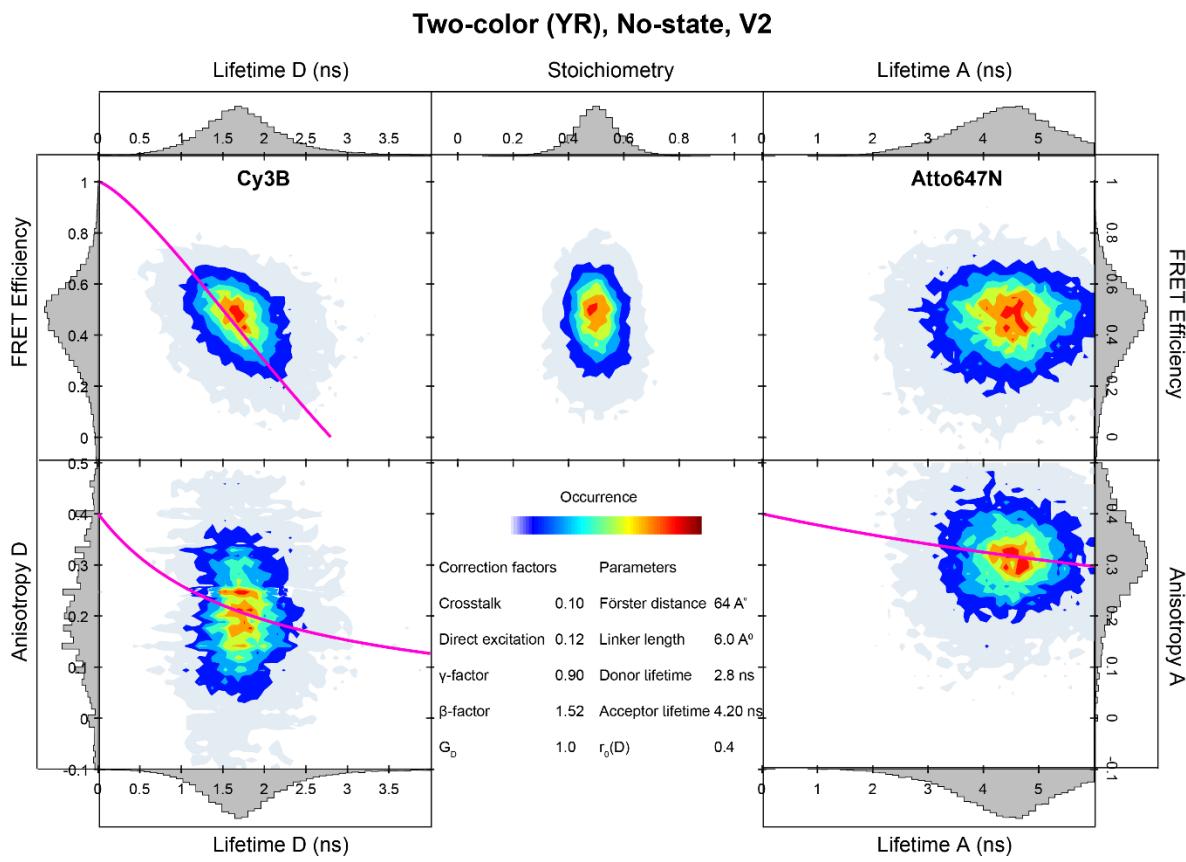


Supplementary Figure S7.20. Overview of 2c MFD-PIE results obtained for the dual-labeled 6-11-12 clock (Version 1). Top row: The FRET efficiency for the FRET pair is plotted as a function of (left) Donor lifetime, (middle) stoichiometry and (right) and acceptor lifetime. In the left panel, the static FRET line is plotted in magenta and a dynamic FRET line is shown in black. For details, see [3]. Bottom row: The anisotropy of the (left) donor and (right) acceptor molecules are plotted as a function of the respective lifetimes. A fit of the burst-wise steady-state anisotropy to the Perrin Equation (Eq S1) is shown in magenta. Middle: A summary of the parameters extracted from the data.

7.8. Dual-color DNA origamis for control

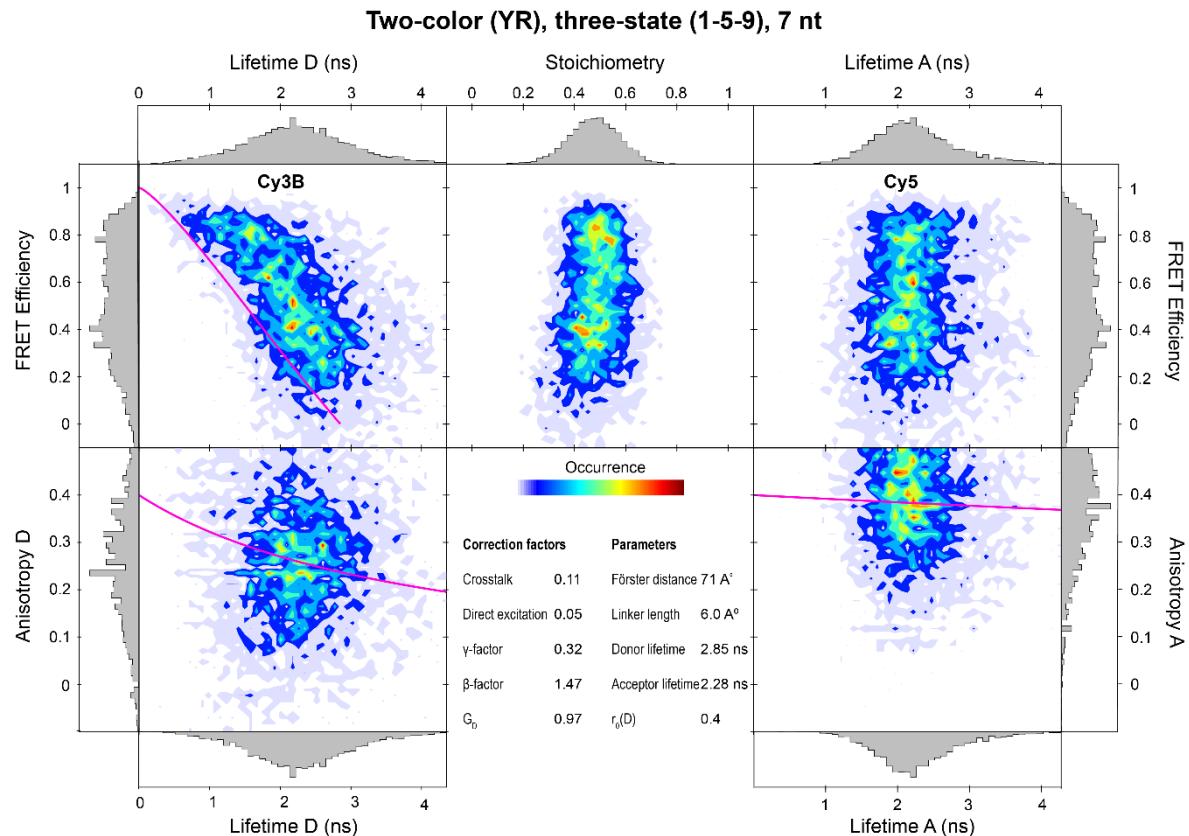


Supplementary Figure S7.21. Overview of 2c MFD-PIE results obtained for the dual-labeled DNA origami structure (Version 1) without any binding sites. Top row: The FRET efficiency for the FRET pair is plotted as a function of (left) Donor lifetime, (middle) stoichiometry and (right) and acceptor lifetime. In the left panel, the static FRET line is plotted in magenta. For details, see [3]. Bottom row: The anisotropy of the (left) donor and (right) acceptor molecules are plotted as a function of the respective lifetimes. A fit of the burst-wise steady-state anisotropy to the Perrin Equation (Eq S1) is shown in magenta. Middle: A summary of the parameters extracted from the data.

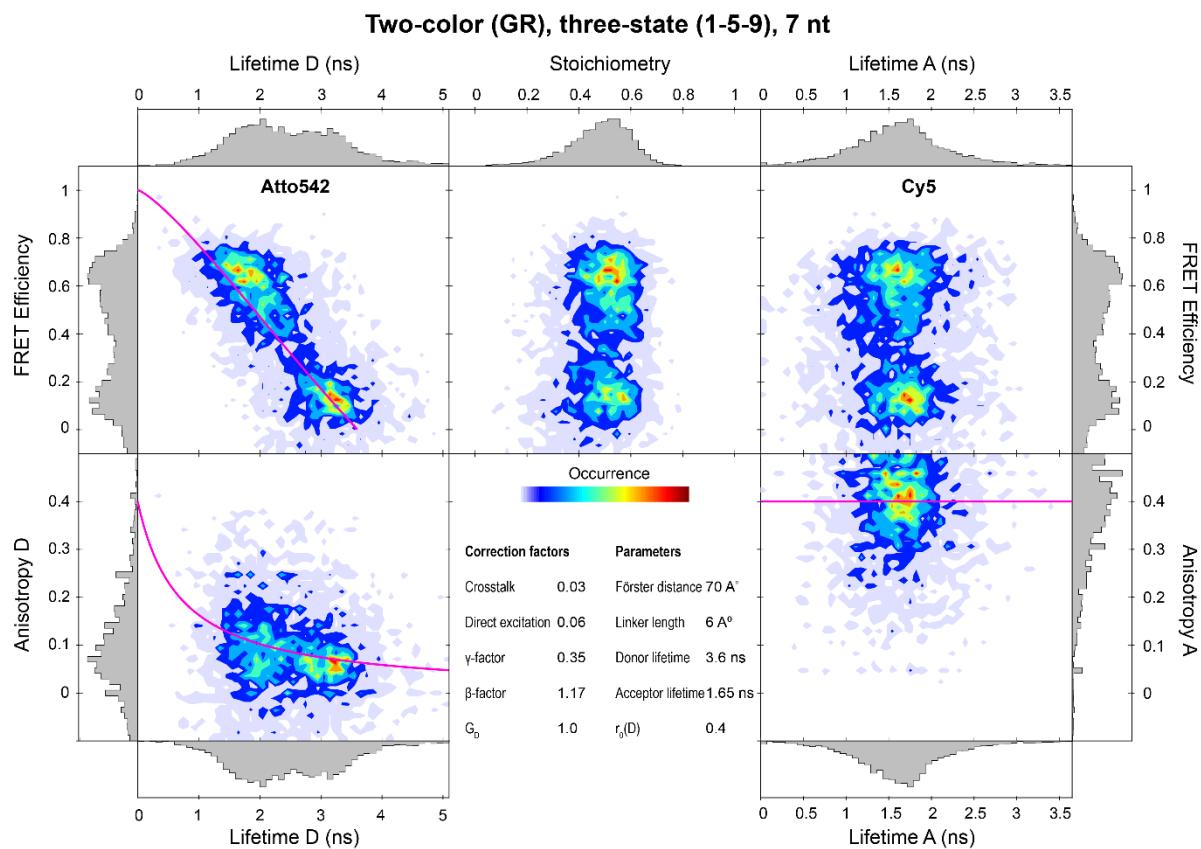


Supplementary Figure S7.22. Overview of 2c MFD-PIE results obtained for the dual-labeled DNA origami structure (Version 2) without any binding sites. *Top row:* The FRET efficiency for the FRET pair is plotted as a function of (left) Donor lifetime, (middle) stoichiometry and (right) and acceptor lifetime. In the left panel, the static FRET line is plotted in magenta. For details, see [3]. *Bottom row:* The anisotropy of the (left) donor and (right) acceptor molecules are plotted as a function of the respective lifetimes. A fit of the burst-wise steady-state anisotropy to the Perrin Equation (Eq S1) is shown in magenta. Middle: A summary of the parameters extracted from the data.

7.9. Small clock with 3 states



Supplementary Figure S7.23. Overview of 2c MFD-PIE results obtained for the dual-labeled small clock labeled with Cy3B and Cy5. Top row: The FRET efficiency for the FRET pair is plotted as a function of (left) Donor lifetime, (middle) stoichiometry and (right) and acceptor lifetime. In the left panel, the static FRET line is plotted in magenta. For details, see [3]. Bottom row: The anisotropy of the (left) donor and (right) acceptor molecules are plotted as a function of the respective lifetimes. A fit of the burst-wise steady-state anisotropy to the Perrin Equation (Eq S1) is shown in magenta. Middle: A summary of the parameters extracted from the data.



Supplementary Figure S7.24. Overview of 2c MFD-PIE results obtained for the dual-labeled small clock labeled with Atto542 and Cy5. *Top row:* The FRET efficiency for the FRET pair is plotted as a function of (*left*) Donor lifetime, (*middle*) stoichiometry and (*right*) and acceptor lifetime. In the *left* panel, the static FRET line is plotted in magenta. For details, see [3]. *Bottom row:* The anisotropy of the (*left*) donor and (*right*) acceptor molecules are plotted as a function of the respective lifetimes. A fit of the burst-wise steady-state anisotropy to the Perrin Equation (Eq S1) is shown in magenta. Middle: A summary of the parameters extracted from the data.

Supplementary Note 8: Two- and three-color corrected FRET efficiencies

For two- and three-color DNA origami structures, bursts were extracted having a minimum of $L = 50$ (two color) or 100 (three color) photons per burst and at least $M = 10$ photons per time window of $T = 500 \mu\text{s}$ using an all-photon burst search algorithm. After extraction, an additional ALEX-2CDE filtering of 0-16 was applied [4]. The signal in each detection channel I_{raw} was then corrected for background using the average intensity of the buffer reference sample, $\langle \text{bg} \rangle$, and the burst duration: t_D :

$$I = I_{\text{raw}} - \langle \text{bg} \rangle t_D \quad (\text{S2})$$

The FRET efficiency between the yellow and red fluorophores can be determined directly during the yellow excitation cycle after applying the correction factors for crosstalk ct_{YR} from the donor into the acceptor channel, for direct excitation de_{YR} of the acceptor fluorophore by the donor-exciting laser and the γ_{YR} detection correction factor. The latter counts for the differences in detection efficiency η and quantum yield ϕ of the different fluorophores and is calculated as

$$\gamma_{YR} = \frac{\eta_R \phi_R}{\eta_Y \phi_Y} \quad (\text{S3})$$

Using the determined correction factors, the corrected YR FRET efficiency can be calculated using:

$$E_{YR} = \frac{I_{YR} - ct_{YR} I_{YY} - de_{YR} I_{RR}}{\gamma_{YR} I_{YY} + I_{YR} - ct_{YR} I_{YY} - de_{YR} I_{RR}} \quad (\text{S4})$$

For three-color experiments, the correction factors for crosstalk ct_{XY} , direct excitation de_{XY} and detection efficiency γ_{XY} need to be determined for the other dye combinations: BY and BR. While photons in the yellow channel only need to be corrected for crosstalk of the blue fluorophore and direct excitation of the yellow fluorophore, the photons in the red channel require correction for the crosstalk from both blue and yellow emission into the red detection channel, direct excitation of the red fluorophore and the energy transfer to the red dye from direct excitation of the yellow fluorophore. The corresponding FRET intensities after blue excitation is given as:

$$I_{BY,corr} = I_{BY} - ct_{BY} I_{BB} - de_{BY} I_{YY} \quad (\text{S5})$$

$$I_{BR,corr} = I_{BR} - de_{BR} I_{RR} - ct_{BR} I_{BB} - ct_{YR}(I_{BY} - ct_{BY} I_{BB}) - de_{BY} \frac{E_{BY}}{1 - E_{YR}} I_{YY} \quad (\text{S6})$$

Using these expressions, the burst-wise FRET efficiencies after blue excitation are given by:

$$E_{BY} = \frac{I_{BY,corr}}{\gamma_{BY} I_{BB}(1 - E_{YR}) + I_{BY,corr}} \quad (\text{S7})$$

$$E_{BR} = \frac{I_{BR,corr} - E_{YR}(\gamma_{YR} I_{BY,corr} + I_{BR,corr})}{\gamma_{BR} I_{BB} + I_{BR,corr} - E_{YR}(\gamma_{BR} I_{BB,corr} + \gamma_{YR} I_{BY,corr} + I_{BR,corr})} \quad (\text{S8})$$

Supplementary Note 9: Staple strands for multi-colored DNA origami clocks

Details of the design of the L-shaped DNA origamis are given in this Supplementary Note. The structures used in this study were based on an L-shaped design previously published by Tinnefeld et al. [5-8]. A list of the unlabeled staple strands is given in Supplementary Note SN9.1. We investigated two distinct designs for the clocks, one with docking strands at the 6 and 12 o'clock (and sometimes at a third position) and a smaller clock with symmetrically distributed docking strands (at the 1, 5 and 9 o'clock positions). That means, the distance between centered, protruding pointer and the docking positions is further or shorter for the different clocks. An overview of all designed DNA origamis for the larger clock are given in Supplementary Note SN9.2. The strand IDs for the modified staple strands used to introduce fluorescent dyes for smFRET, surface immobilization, and protruding strands of the tether and docking sites are listed in Supplementary Note SN9.3. The same information for the symmetric, small clock can be found in Supplementary Notes SN9.4 and SN9.5.

9.1. Staple strands of the L-shaped origami structure

The L-shaped DNA origami structures are made of 252 single DNA staple strands annealed to a circular complementary single DNA scaffold strand of 8064 nucleotides originating from M13mp18 bacteriophages.

Supplementary Table S9.1. Unmodified staple strands of the L-shaped DNA origami structure denoted from the 5' to 3' end. All oligonucleotides were purchased by Integrated DNA Technologies.

| Staple ID | Sequence (5' to 3') |
|-----------|---|
| L1 | ATCCAGAACATATTAGCCATCAGGAACGGT |
| L2 | CGTGCCTGTTCTCGCATCCAGCGCCGGGTTA |
| L3 | ATAATCAGAAAAGCCAACATCCACTGTAATA |
| L4 | CATAGGTCTGAGAGACAAATCGTCGAATTACC |
| L5 | AGAAACAGCTTAGAAGGAAGAAAAATCTACGATTTAAGCATATAAC |
| L6 | GCACCCCTCGTCAGGTACGTTAGTAATGAATAGTTAGCGTCAATCAT |
| L7 | ACGATAAACCTAAACAAAGAATACTAAACATTACCAACAAAGC |
| L8 | TGCTCATTCTATGCGTTAATAAAACGAACATATTGATTGGCTTTG |
| L9 | AAGGGAACCGGATATTCACTCATTTGACCCGTAATGCCATCGGAAC |
| L10 | CGGAATCTCAGGTCTGTTAAATATGCATGCGAACGAATCATTG |
| L11 | TGAATTACCACTGAAATGGAATTACGAGGCATATAGCGAGAGAACCCCC |
| L12 | CATTATACGGTTACCCATAACCCCTGAAATACAATGTTAAACAGGG |
| L13 | ATTCAATACAGTGATTGGCATCAGGACGTTGTAACATAAACCAAGACG |
| L14 | TAATAAGAAGAGCCACCTTATTAGCGTTGCCATTCAACAATAGAAA |
| L15 | GGCACCAAAACCAAAAGTAAGAGCAACACTATAGCAACGTAAATGCC |
| L16 | ATAAAAAATACGCGTTCTCCTTGTATAAGAGCTATAT |
| L17 | TACCAAGTAACGCTAACAGTTGCTATTTGCACCCATCCT |

| | |
|-----|--|
| L18 | GAGGGTAGTTGCAGGGTGTAAACAACTTACGCCTGGAAAGAG |
| L19 | AGAGCCGCAAACAAATGAGACTCCTCAAGAGATTAGCGGGCAGTAGCA |
| L20 | AGTTGATTAGCTGAAAAGAGTACCTTAATTGTTAACCGGACCTAA |
| L21 | TCGATAGCAGCACCGTAAATCACGTTTGCT |
| L22 | AAAGACAAATTAGCAAGTCACCAATGAAACCA |
| L23 | ATATTACCGCCAGCATTGACAGGCAAAATCA |
| L24 | TTTCCCTTACACTGGTTGC |
| L25 | CTCCAATCGTCTGAAATTT |
| L26 | TTTTGCCTGAGTAGAAGAA |
| L27 | TTTTCCCGACTTACAAAATAACAGTTTT |
| L28 | ATACGAAAGAAAATTATTCAATTAAAGGTGAATTT |
| L29 | TTTCTTACAAACAATTG |
| L30 | TTTTAAGTTACCAGGGTAATTGAGCTTT |
| L31 | TTTTAAACGATGCTGATGG |
| L32 | TTTCAGGGTGGTTTTCTT |
| L33 | ACAAAGTATGAGGAAGCTTGAGGACTAAAGATTT |
| L34 | CCGAATCTAAAGCATCTTT |
| L35 | AGATGAAGGGTAAAGTTTT |
| L36 | TTTCGCAAATGGTCAATAAACCATTAGATGC |
| L37 | TCGAAGATGATGAAACTTTT |
| L38 | AGAGCAAATCCTGTCCAGATACCGACAAAAGGTAATTT |
| L39 | TTCCGGAATCATAATTTTT |
| L40 | TTTTGGATTATTCAGAGAA |
| L41 | TGCGGCCAGAATGCGGTTTT |
| L42 | TCAGCAGCAACCGCAATTT |
| L43 | TTTAGAGCGGGAGCTAGAT |
| L44 | TTTGCTAATATCAGAGAGATAACCCGCCACCGCG |
| L45 | ACCTCGTCATAAACATTTT |
| L46 | TTTTTGAGGGGACGACGAC |
| L47 | TTTAACAGTACCTTTACA |
| L48 | TTTTGGCCTCCTGTATAA |
| L49 | TTTGGCGCATAGGCTGGCTAACGGTGTAAATTGT |
| L50 | TCACCGGAAGCATAAATTT |
| L51 | TTTATCATCGCCTGAACAGACCAATT |
| L52 | TTCATAGGGTTGAGTGT |
| L53 | TTTTAGCCCGAATAGCCTATTTCTTT |
| L54 | TTTCCCTCAGAGCCACCCCTCAGAAAGCGCTTA |
| L55 | TAGTAATAACATCACTTTT |
| L56 | TTTTTGTTCCAGTTGGAACAAGA |
| L57 | TTTCGGGCCGTTTCACGG |
| L58 | TTTATTGCTGAATATAACATTTTTT |
| L59 | TTTAGTAATTCAATCGCAAGACAATTT |
| L60 | TTTGAAATGCCAACGGCAGC |
| L61 | CAGATGAATATACAGTTTT |

| | |
|------|--|
| L62 | TTTCCATATTATTCACCAATCCAAGTCAGAGA |
| L63 | TTTTTATCACCGTCACAGCGTCAGTTT |
| L64 | TTTCTTTTACAACGGAGATTGTTT |
| L65 | TTTGTGTAGGTAAAGATTTC |
| L66 | TTTTTTTTTTAAACTAG |
| L67 | TTTGATTAAGACGCTGAGA |
| L68 | TTTGCATGGCGCTTT |
| L69 | ATTATAGCGTGTAAAGTAAATGTTTT |
| L70 | TAGTCAGAAAGCGGATT |
| L71 | TTTTAGACTGGCATCAGTTGAGATT |
| L72 | CATAATAATCGCGT |
| L73 | ATATATATAAAGCGACGACATCGGCTGCTTCCTTATCATT |
| L74 | AAAACGGTAATCGTTT |
| L75 | ACAAAATTATCATCATATT |
| L76 | TTTTCTGATTATCACGT |
| L77 | TTTCATATAAAAGAAAGCCAACATT |
| L78 | TTTGTGTAAAGCCTGGCG |
| L79 | TTTAAACATCAAGAAAAA |
| L80 | AATGCAATAGATTAAGGGCTTAGAGCTT |
| L81 | ACATAGCGATAGCTT |
| L82 | TTTTGCATAAAAGCCTGAGTAATT |
| L83 | CTGATAGCCCTAAACTT |
| L84 | GAAAGGAGCGGGCGCTAGGTT |
| L85 | TTTGCCTCAGAGCATAAAAGAAAATTAAGCAATAATT |
| L86 | TTTACTGTAGCCTCAGAACCGCCATT |
| L87 | TTTACGTTCCAGTAAGCGTCATACATGGCTTCAGTTAAT |
| L88 | AGTGTGCTGCAAGGC |
| L89 | TTAATTAAACCATACATACATAAAGGTGGCAATT |
| L90 | CCGTGCATCTGCCAGTT |
| L91 | TTTTAGGAATACCAACAGTAGTAATT |
| L92 | TTTATTGGGCTTGAGATGGCCAGAACGATT |
| L93 | TTTGAAACAACAAAGGAACACTGATT |
| L94 | TTTACTAGAAAAGCCTGTT |
| L95 | TTTCGACTTGATCGAGAGGGTTGATATAAGTATT |
| L96 | ATTTAGAAGTATTAGATT |
| L97 | TTTACCTTGCTGAACCAGG |
| L98 | TTTTCCAAGAACGGGTGCGAACCTT |
| L99 | TTTACGCATAATGAGAATAGAAAGTT |
| L100 | TTTAAACCGCAGAAAATT |
| L101 | TTTGTTCGTCACCAAGTACTGTACCGTAAT |
| L102 | CATGTTTACCAAGTCC |
| L103 | TTTGGAAATTGTGAGAGAT |
| L104 | TTTATTAAAGTTGGTAGC |
| L105 | TTTGGAACCTAAGTCTGAATT |

| | |
|------|---|
| L106 | TTTTATGCCATTAAAAATA |
| L107 | GAGCCGATATAACAACAACCATGCCCTTTTT |
| L108 | CGGCCTCGTTAGAATCTTT |
| L109 | TTTGCCTGGCAAGTAG |
| L110 | TAGTTGCCAGTTGCCGGAGGTTTGAAGATCAATAA |
| L111 | ATGGCTACAATCAACTGAGAGCCAGCAGCAAATGAAAAACGAACCTAATGCCTGGCAGA |
| L112 | TCATCAACAAGGCAAATATGTACCCGGTT |
| L113 | TTCAAATTTAGAAAAAACAGGAGCAAACAAGAGAATCGATGAAGGGTGAGATATTTA |
| L114 | CAACTAATGCAGACAGAGGGGCAACTG |
| L115 | GTACTATGGTTGCTTTTAGACACGCAAATT |
| L116 | TGTAGCTCACATTACCGTCGAAAGAC |
| L117 | ATCAAAAAGTCATAAAACCGAACACATTATCAACTTAGTAGAT |
| L118 | AACGTCAATAGACGGGAATACCCAAAAGAACAGACTCCGTTTTAT |
| L119 | GGAGGGAAAGAGCCAGCAATCAGTAGCGACAGACCAGAACCGCCTC |
| L120 | AGCGAACCGAAGCCTGGAGAATCACAAAGGCTATCAGGT |
| L121 | GCCCCCTGGTGTATCACCGTACTC |
| L122 | TACAGGCATTAATTAACCAATAGAACGCCATCAAAGTCATCAGAATTAGCCTAAATCG |
| L123 | TATTTTGAGAGATCTGCCATATTCCTCTACTCAATTGA |
| L124 | CATTGCCTGAGAGTCTTATGACCATAATCATTCAATT |
| L125 | CCAGCCAGCTTCCGGGTAATGGGTAACAC |
| L126 | ATCGGCAAAATCCCTACGTGGACTCCAACGT |
| L127 | CCTGCAGCCATAACGGGGTGTCCAGCATCAGC |
| L128 | GGGCCTTCGCTATTACGTTGACCTCACCG |
| L129 | ACCCCATGCCCTATTTCTGTATGGGTTAGTTAACAGCAGCTGA |
| L130 | GTTGTACCAACCTCATAAAGGCCGGAGACAG |
| L131 | TCTTCTGGCTGAATAATGCTCATTAGTAACAT |
| L132 | CTCTCACGGAAAAGAACGGATAAAACGACG |
| L133 | ACGCCAGATGACGGGGCGCGTAGCCCCAGC |
| L134 | TTAATTCATGTTCTATAACTATATGTAATGCTGATGTCATAAGAACCTTGACAAAATT |
| L135 | TTTCATCGAATAATATCCAGCTACAATACTCCAGCAATTCTTACAG |
| L136 | AATAAGTTAGCAAAACGCAATAATAACGAGAATTAAAGCCAA |
| L137 | GACCGTGTGATAAATACAAATTCT |
| L138 | ACAAGAACCGAACTGATGTTACTAGCCGGAAAGACAGCACTACGAA |
| L139 | ATCAAACCTAAATTCTGGAAGGGCCATATCA |
| L140 | CGCTGGCACACGGGAGACGCAGAACAGCGG |
| L141 | GAGAACATTAAATTACAGGTAGAAAG |
| L142 | CTGCGCGGCTAACTCACAATTCCACACAACATACGAGTACCGGGCTCTGTGGGTGTTAG |
| L143 | CCGAGTAAGCCAACAGGGTACCGCATTGCAA |
| L144 | AAACGGCGCAAGCTTGAAGGGCGATCGGTGC |
| L145 | CAAAAGAATAAAATACCCAGCGATTACCAAGCGCGAA |
| L146 | CTTAATTGAGACCGGAAACAGGTCAGGATTAGAGGTGGCA |
| L147 | GCCAGTGCATTGACCCACCGCTCTGGTGCC |
| L148 | CCAGAACGGAGCCCAATCAAGTTG |
| L149 | CCCCCTGCGCCGCTTAGCTGTTCTGTGT |

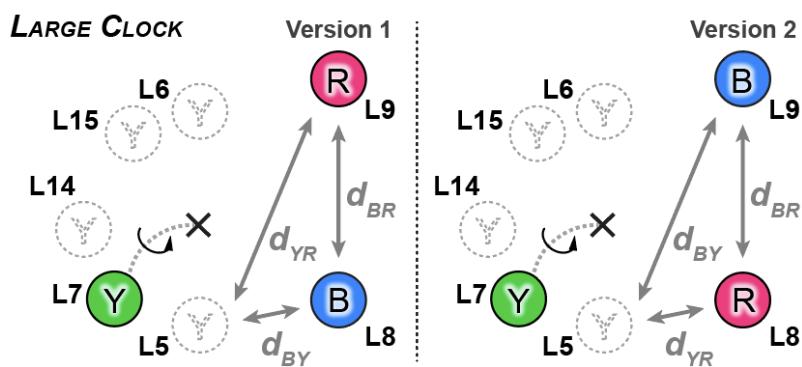
| | |
|------|---|
| L150 | GGAAACCAGGCAAAGCGTACATAAGTGAGTGA |
| L151 | AAATCAACACGTGGCATCAGTATTCTCAATCC |
| L152 | AGGAGGTGGCGGATAAGTATTAGAGGCTAAATCCTCTACAGGAG |
| L153 | GACAGATGGACCTTCATCAAGAGCCCTGAC |
| L154 | CTGAGGCCAACGGCTACAGAGGTTCCATT |
| L155 | ATAACCTTATCAACAAAAATTGTATAACCTCC |
| L156 | AAATCAGCTCATTTTGAGCGAATAGGTCA |
| L157 | CACAGACATTCAAGGGATCTCAAAAAAAAGGTTCTAAAGCCGCTT |
| L158 | TAATAGTATTCTCGTGATTAAATTGTT |
| L159 | CGTTGGTAGTCACGACGCCAGCTGGCAAAGGGGATATCGGCCTGCGCATCGGCCAGCTT |
| L160 | CTTCTGACCTAAATTGCAGAGGCCAGAACGCAATTACG |
| L161 | GCTGCGCAACTGTTGGCAGACCTATTAGAAGG |
| L162 | AGAACGTTAACGGCGTAATGGTAAAGGTTCTTGCCTGGTGTGGCTTGCGCTT |
| L163 | TTAGTTGCCTGTTAGGTCAATTGCGGATAGGAAGCCGACTATT |
| L164 | AATTACATAGATTTCATAACGGATTGCC |
| L165 | AAACGGGGTTTGCTACATAACGCCAAAAAAAGGCTTGAATCTT |
| L166 | TGCGAATAATAATCGACAATGTTGGTCG |
| L167 | TTATACTTAGCACTAAAAGTTGTGCCGCCA |
| L168 | GCCGTACAATATAAAAGAAACCACAGAAGGAGCGGACTCGTATTACATTGTCAAATAT |
| L169 | ATTGCGTTAACACATTCAATTACCTGAGCAAAGGGAGAACAGGTTAACATGATGG |
| L170 | GGAGCCTTCACCCCTCAGAGCCACC |
| L171 | CCAGCTTACGGCTGAAACGTGCCCGTCTCGT |
| L172 | TTCGTAATCATGGTCATCCATCAGTTAAAGT |
| L173 | AACAGAGGTGAGGCAGACAAATTAAAAGGG |
| L174 | TTGAGTAAGCCACCCCTCAGAACCG |
| L175 | CAGTATGTTATTTGCGAAGCCCTTTAACATTGAGTTCTGAACA |
| L176 | GCCTGTTGCTCTGTTACCTTAACTTAA |
| L177 | ATAAACAAATCCCTAGTGAATTATCAAAT |
| L178 | CAAAGGGCTGTCGTGGCCCTGAGAGAGTT |
| L179 | CTCAAATGTTAGAAATGGAAGTTCACGCGCATTACTCAACTGGCT |
| L180 | CCCGCCGCGCTTAATGAAAGCCGGGAACGTG |
| L181 | TTCACCAGGTAGCAATGCCCTGCTGGTAAT |
| L182 | GTCGAAATCCCGACCTGCTCCACCAACTTTAGCATTC |
| L183 | TGATTGCTTGAATACAACAGAACATTGGGA |
| L184 | TTCTGAAACATGAAAGTGCAGGCCATTG |
| L185 | AACCGTTCACACGGAAATACCTACATTGACGCTAAACTATCACTTAAACAGGAG |
| L186 | CGTTGAAAATAGCAAGCCCAATA |
| L187 | CTTTGCGTTATTCAATGATATTCAACCGTT |
| L188 | AAATCCCGTAAAAAAACGTTTGGACTTGT |
| L189 | TATCATTGCGGAACATCCTGATATAAGAA |
| L190 | AAATTATTGGAAACAGCCATTGAAATCGC |
| L191 | GCAGCAAGCGGTCCACAAGTGTGAGGCCA |
| L192 | CCAACATGACGCTCAATGCCGGAGGAATACC |
| L193 | TATTTGTTAAAATTGGGTATATCAAAAC |

| | |
|------|---|
| L194 | TGTTGCCCTGCGGCTGATCAGATGCAGTGTCA |
| L195 | TGCGGGATAGCAGCGACGAGGCGCAGAGAAACGGCCGCGTAACGATC |
| L196 | TACCGATAGTTGCGCTTTCA |
| L197 | TCAAATCACCATCAATACGCAAGG |
| L198 | GTAAGAATAGTTGAAACTTCGCAAAACACCGC |
| L199 | ATTGCCCTCACCGCCCCAGCTGCTTGCCTG |
| L200 | AAGCGCATAATGAAACAGATATAGAAGGCTTAGCAAGCCTTATTACG |
| L201 | GGAATTAGGTAATTTCGGTATAGCCCCACCGGAACCACCAAC |
| L202 | GTTTCCCGTAGATGGCAGGAAGATCGCACT |
| L203 | GCGAGAAAAGGGATGACGAGCACGATAACGTGCTTTCACGCTGAAGAAAGC |
| L204 | GGGGCGCGCCCAATTCACTAAAGTACGGTGTACGAGAATAGCTCAA |
| L205 | GAAATTGTTATCCGCTCACATTAAATTATGA |
| L206 | TTTTTAATGCACGTACAAGTTACCCATTCA |
| L207 | CAATTCAATAGATAATAAATCCTTGC |
| L208 | CCTCAGAGCACAAGAAGAAAAAGTAAGCAG |
| L209 | CGCTCACTATCAGACGGTCCGTGAGCCTC |
| L210 | GCAGAGGCGAATTATTTCTATTGCTATTAA |
| L211 | TTAGAGCTATCCTGAGGCTGGTTAGGGCGC |
| L212 | GCCAGTACGTTATAAGGCGTTAATAAGAATAAACACAAAT |
| L213 | AACGTTATTAAATTACAACATACTAGTTGGC |
| L214 | GCCGGGCGCGTTGCGCCGCTGACCCCTTG |
| L215 | CTGCAACAGTGCCACGTATCTGGTAGATTAGA |
| L216 | TAAAGTTAGAACCGCTAATTGATCGGGGTTAAGTTGGCCTG |
| L217 | GAAACAACCGGGCGCCACAGGGGGCTTACTGACTTCTCCACGTACAGACGCCAGG |
| L218 | GTCCACTAAACCGCGGGACGGCAACAGCTG |
| L219 | GGAACCCAAAACATAAACAGTTAGCG |
| L220 | ATCGGCCTAAAGAATAAATCAAAGAATAGCCGAGACCAGTGAGGGAGAGGGTGCCTA |
| L221 | ACAGTTGAGGATCCCCAGATAGAACTGAAAGC |
| L222 | CCGAACCGCAAGAAAGCAATAGCTATCTACTACAATCCGATTGAG |
| L223 | GCAGTTGGCGGTTGTCAGTTATGGAAGGAG |
| L224 | GCGATTAAAGGAAGGGCGCGTAACCACCA |
| L225 | TGTACTGGTAATAAGTTAGTG |
| L226 | CAAATCGTCAGCGTGGTGCCATCCCACGCAA |
| L227 | TCTTACCATAAAGCCATAATTAGAATGGTTAGGGTAGC |
| L228 | AGGCGAAAATCCTGTTGTCATCACCCCCGAT |
| L229 | GCCTAATTATCATATGATAAGAGATTAGTTAATTCTAT |
| L230 | TTTCATGGCATATTGACGGCACCG |
| L231 | CTAGCTGATAAAATTACAGTAGGG |
| L232 | CCCTGAACAAATAAGAAACGCGAGGGCGT |
| L233 | CACATCCTCAGCGGTGGTATGAGCCGGTCAC |
| L234 | CAGGAAAACGCTCATACCAGTAAATTGTA |
| L235 | CCACCCCTGTTAGGAAGGATGTCCTTCCAGCAGACGATTACAGCT |
| L236 | CAAACCCCTTAGTCTTACCAAGCAGAAGATAA |
| L237 | GGCTTAGGTTGGGTTAAGCTAATGATTTCGA |

| | |
|------|---|
| L238 | CCGTCGGAGTAGCATTCAAAAACAGGAAGATT |
| L239 | ATGAGTGACCTGTGCAGTTCTGCCAGCACG |
| L240 | CCGGCAAATCGGCGAAGTGGTGAAGGGATAG |
| L241 | ACAAGAAATAGGAATCCAATAGCAAGCAAATATAGCAGCATCCTGAA |
| L242 | CCATTACCAAGGGCGACATCTTCATAGGCAGAAAGAATAGGTTGAG |
| L243 | TGGAGCCGGCCTCGGGTACATCGACATAAAA |
| L244 | CACTCATGAAACCACCTTAAATCAAGATTGAGCGTCTTTGTTT |
| L245 | GTATAAGCAAATATTTAGATAAGTAACAACG |
| L246 | AGGAAACCGAGGACGTAGAAAAAGTACCG |
| L247 | CGGGAAACGAAAAACCTGATGGTGGTCCGAA |
| L248 | AGCATGTACGAGAACAACTCCGGTATTCTAAGAACGATTTCCAGA |
| L249 | ACATTCTGAAGAGTCTCCGCCAGCAGCTCGAA |
| L250 | GGGGTCATTGCAGGCGGGATTGACTAAAATA |
| L251 | TGCTTCGAGGTGAATCTCCAAAA |
| L252 | CAGTACCATTAGTACCCAGTCCCCGTATAAATTGATGAATTAAG |

9.2. Overview of DNA Origami structures with large clock design

DNA origami clocks with large radius (~6 nm) were designed by placing a central ssDNA tether strand at position L7 (Supplementary Figure S9.1), surrounded by ssDNA binding sites with varying lengths. An overview of all designed structures is given in Supplementary Table S9.2 in which we specify their nomenclature and all introduced modifications. For the two-state DNA origami structures, we introduced docking sites at the 6 o'clock (staple strand L5) and 12 o'clock (staple strand L6) positions, respectively. To create three-state systems, an additional docking site was introduced either at the 9 o'clock (staple strands L14) or 11 o'clock (staple strand L15) position. For implementing the 9 and 11 o'clock binding sites, the staple strands L12 and L13 were additionally replaced by three smaller staples, referred to as L12-13-I, L12-13-II and L12-13-III (Supplementary Table S9.3).



Supplementary Figure S9.1: Schematic of designed clock on the L-shaped DNA origami structures. Clocks with two states have binding sites at the 6 (L6) and 12 (L5) o'clock position. The three-state origami clocks have an additional binding site at either the 9 o'clock (L14) or 11 o'clock (L15) position. To monitor the binding position of the tether and the expected dynamics, smFRET experiments were performed by placing dyes at the L8 and L9 positions near the 6 and 12 o'clock docking strands respectively and with Cy3B placed at the 3' end of the pointer attached to position L7. In V1, DNA origami structures are labeled at L9 by Atto647N and at position L8 with Atto488. In V2, the position of the blue and red fluorophores are exchanged.

Binding rates were tuned by introducing tethers at the L7 position with varying base-pairing regions containing complementary docking lengths of 6.5 nt (i.e. 7 nt with 1 bp mismatch), 7 nt, 7.5 nt (i.e. 8 nt with 1 bp mismatch) and 8 nt. To monitor the dynamics and docking position of the ssDNA tether, up to three additional fluorophores – Atto488, Cy3B and Atto647N – were incorporated into the structures by replacing the unlabeled ssDNA L7, L8 and L9 (Supplementary Figure S9.1) with strands containing the appropriate labels (Supplementary Table S9.4). The tether was always labeled at the 3' end with Cy3B.

We refer to DNA origami structures as **Version 1** (V1) where the red fluorophore, Atto647N, is attached at position (12 o'clock) and the blue donor dye, Atto488, is at position L8 (6 o'clock). **Version 2** (V2, with Atto488 and Atto647N at inverse positions) were designed to explore the influence of local composition and asymmetries introduced by the DNA origami structure itself. All samples include biotinylated attachment sites at positions L1-L4 (Supplementary Table S9.5).

Supplementary Table S9.2. Overview of all designed DNA Origami structures for the large clock studied in this work, including nomenclature and strand IDs of incorporated modifications given in Supplementary Tables S9.3-9.5. Binding sites were introduced at L5, L6, L14 and L15 and fluorescent labels attached to L7, L8 and L9 positions. For bindings sites at L14 and L15, staple strands L12-L13 were replaced by oligonucleotides L12-L13-I until L12-L13-III.

| # | Name | Blue | Yellow | Red | Binding sites | | | Replaced |
|-----------|--------------------------|--------|--------|--------|---------------|----------|-----------|--------------|
| 1 | V1-BYR-(6/12)-6.5nt | 488-L8 | 561-L7 | 640-L9 | 6.5nt-L5 | 6.5nt-L6 | | |
| 2 | V1-BYR-(6/12)-7nt | 488-L8 | 561-L7 | 640-L9 | 7nt-L5 | 7nt-L6 | | |
| 3 | V1-BYR-(6/12)-7.5nt | 488-L8 | 561-L7 | 640-L9 | 7.5nt-L5 | 7.5nt-L6 | | |
| 4 | V1-BYR-(6/12)-8nt | 488-L8 | 561-L7 | 640-L9 | 8nt-L5 | 8nt-L6 | | |
| 5 | V1-BY-(6/12)-7.5nt | 488-L8 | 561-L7 | | 7.5nt-L5 | 7.5nt-L6 | | |
| 6 | V1-BR-(6/12)-7.5nt | 488-L8 | | 640-L9 | 7.5nt-L5 | 7.5nt-L6 | | |
| 7 | V1-YR-(6/12)-7.5nt | | 561-L7 | 640-L9 | 7.5nt-L5 | 7.5nt-L6 | | |
| 8 | V1-BYR-(6/9/12)-7/7.5nt | 488-L8 | 561-L7 | 640-L9 | 7nt-L5 | 7nt-L6 | 7.5nt-L14 | L12-13-I-III |
| 9 | V1-BYR-(6/11/12)-7/7.5nt | 488-L8 | 561-L7 | 640-L9 | 7nt-L5 | 7nt-L6 | 7.5nt-L15 | L12-13-I-III |
| 10 | V1-YR-(6/9/12)-7/7.5nt | | 561-L7 | 640-L9 | 7nt-L5 | 7nt-L6 | 7.5nt-L14 | L12-13-I-III |
| 11 | V1-YR-(6/11/12)-7/7.5nt | | 561-L7 | 640-L9 | 7nt-L5 | 7nt-L6 | 7.5nt-L15 | L12-13-I-III |
| 12 | V1-BYR | 488-L8 | 561-L7 | 640-L9 | | | | |
| 13 | V1-YR | | 561-L7 | 640-L9 | | | | |
| 14 | Paint-V1-(6/12)-no | | | 640-L9 | 8nt-L5 | 8nt-L6 | | |
| 15 | V2-BYR-(6/12)-6.5nt | 488-L9 | 561-L7 | 640-L8 | 6.5nt-L5 | 6.5nt-L6 | | |
| 16 | V2-BYR-(6/12)-7nt | 488-L9 | 561-L7 | 640-L8 | 7nt-L5 | 7nt-L6 | | |
| 17 | V2-BYR-(6/12)-7.5nt | 488-L9 | 561-L7 | 640-L8 | 7.5nt-L5 | 7.5nt-L6 | | |
| 18 | V2-BYR-(6/12)-8nt | 488-L9 | 561-L7 | 640-L8 | 8nt-L5 | 8nt-L6 | | |
| 19 | V2-BY-(6/12)-7.5nt | 488-L9 | 561-L7 | | 7.5nt-L5 | 7.5nt-L6 | | |
| 20 | V2-BR-(6/12)-7.5nt | 488-L9 | | 640-L8 | 7.5nt-L5 | 7.5nt-L6 | | |
| 21 | V2-YR-(6/12)-7.5nt | | 561-L7 | 640-L8 | 7.5nt-L5 | 7.5nt-L6 | | |
| 22 | V2-BYR | 488-L9 | 561-L7 | 640-L8 | | | | |
| 23 | V2-YR | | 561-L7 | 640-L8 | | | | |
| 24 | V1-BGR-(6/12)-7.5nt | 488-L8 | 542-L7 | 640-L9 | 7.5nt-L5 | 7.5nt-L6 | | |
| 25 | V1-BG-(6/12)-7.5nt | 488-L8 | 542-L7 | | 7.5nt-L5 | 7.5nt-L6 | | |
| 26 | V1-GR-(6/12)-7.5nt | | 542-L7 | 640-L9 | 7.5nt-L5 | 7.5nt-L6 | | |

9.3. Staple strands of DNA Origami structures with large clock design

Protruding strands and binding sites. The binding kinetics between the ssDNA tether and the protruding docking strands was controlled by varying the length of base-pairing between their complementary sequence, i.e. we designed four different DNA sequences at position L6 (12 o'clock) and L5 (6 o'clock; Supplementary Table S9.3).

Supplementary Table S9.3. Modification to the staple strands needed to incorporate single-stranded docking sites into the L-shaped DNA origami structures. The complementary docking sequences are highlighted in orange. The docking strands have a three-base extension (TTT) from the DNA origami structure. The pointer contains nine single-stranded thymine bases plus GC and the docking sequence, highlighted in grey. All sequences are given from the 5' to 3' end.

| ID | Sequence (5' to 3') | Supplier | Function | Replace |
|------------|--|----------------------|--|-----------|
| 6.5nt-L6 | GCACCCCTCCGTCAGGTACGTTAGTAAATGAATAGTTAGCGT CAATCATTTCAAATGC | Eurofins Genomics | Pos 12 catching site 7 nt <i>mismatch</i> | L6 |
| 7nt-L6 | GCACCCCTCCGTCAGGTACGTTAGTAAATGAATAGTTAGCGT CAATCATTTTAAATGC | Eurofins Genomics | Pos 12 catching site 7 nt | L6 |
| 7.5nt-L6 | GCACCCCTCCGTCAGGTACGTTAGTAAATGAATAGTTAGCGT CAATCATTTTCAAATGCC | Eurofins Genomics | Pos 12 catching site 8 nt <i>mismatch</i> | L6 |
| 8nt-L6 | GCACCCCTCCGTCAGGTACGTTAGTAAATGAATAGTTAGCGT CAATCATTTTAAATGCC | Eurofins Genomics | Pos 12 catching site 8 nt | L6 |
| 6.5nt-L5 | AGAACACAGCTTAGAAGGAAGAAAAATCTACGATTTAAG CATATAACTTTCAAATGC | Eurofins Genomics | Pos 6 catching site 7 nt mismatch | L5 |
| 7nt-L5 | AGAACACAGCTTAGAAGGAAGAAAAATCTACGATTTAAG CATATAACTTTTAAATGC | Eurofins Genomics | Pos 6 catching site 7 nt | L5 |
| 7.5nt-L5 | AGAACACAGCTTAGAAGGAAGAAAAATCTACGATTTAAG CATATAACTTTCAAATGCC | Eurofins Genomics | Pos 6 catching site 8 nt mismatch | L5 |
| 8nt-L5 | AGAACACAGCTTAGAAGGAAGAAAAATCTACGATTTAAG CATATAACTTTTAAATGCC | Eurofins Genomics | Pos 6 catching site 8 nt | L5 |
| L12-13-I | CATTATACCACTGATTGGCATCAGGACGTTAACATAA ACCAAGACG | Eurofins Genomics | Replacement for Pos 9 | L12 & L13 |
| L12-13-II | TAATAAGAAGAGCCACCCCTATTAGCGTT | Eurofins Genomics | Replacement for Pos 9 | L12 & L13 |
| L12-13-III | TTACCCATAACCCTCGAAATACAATGTTAACAGGG | Eurofins Genomics | Replacement for Pos 9 | L12 & L13 |
| 7.5nt-L14 | TGCCATTCAACAATAGAAAATTCATATGGTTTTTTCAAAT GCC | Eurofins Genomics | Pos 9 catching site 8 nt mismatch | L14 |
| 7.5nt-L15 | GGCACCAAAACCAAAAGTAAGAGCAACACTATAGCAACGT AAATCGCCTTTCAAATGCC | Eurofins Genomics | Pos 11 catching site 8 nt mismatch | L15 |

Fluorescent modification. To monitor the movement of the protruding single-stranded DNA, the pointer was labeled at the 3' end via Cy3B or Atto542. Its position and binding times at the different docking sites were monitored via smFRET by labeling the DNA origami platform with Atto488 and/or Atto647N.

Supplementary Table S9.4. Modified staple strands necessary for introducing fluorescent labels at the pointer and DNA origami surface. All sequences are given from the 5' to 3' end.

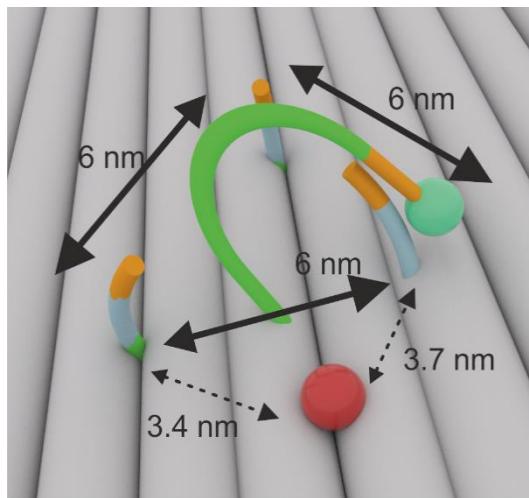
| ID | Sequence (5' to 3') | Supplier | Function | Replace |
|--------|---|----------------------|---------------------|---------|
| 488-L8 | TGC TCA TTC TXA TGC GTT AAT AAA ACG AAC TAT ATT CAT TGG CTT TTG; X = dT- Atto488 | biomers | Lower Label – V1 | L8 |
| 640-L9 | AAG GGA ACC GYA TAT TCA CTC ATC TTT GAC CCG TAA TGC CAT CGG AAC; Y = dT- Atto647N | Eurofins Genomics | Upper Label – V1 | L9 |
| 561-L7 | GGCACCAAAACCAAAAGTAAGAGCAACACTATAGCAACGT AAATGCCCTTTTTTCGG CY3B GCATTAA - Cy3B | Eurofins Genomics | Pointer - dye at 3' | L7 |
| 542-L7 | GGCACCAAAACCAAAAGTAAGAGCAACACTATAGCAACGT AAATGCCCTTTTTTCGG Atto542 GCATTAA - Atto542 | Eurofins Genomics | Pointer - dye at 3' | L7 |
| 488-L9 | AAG GGA ACC GXA TAT TCA CTC ATC TTT GAC CCG TAA TGC CAT CGG AAC; X = dT- Atto488 | biomers | Lower Label – V2 | L9 |
| 640-L8 | TGC TCA TTC TYA TGC GTT AAT AAA ACG AAC TAT ATT CAT TGG CTT TTG; Y = dT- Atto647N | Eurofins Genomics | Upper Label – V2 | L8 |

Surface immobilization. L-shaped DNA origami structures were designed for surface immobilization via anchor strands labeled with biotin attached at the 5' ends. These oligomers replace staple strands L1-L4.

Supplementary Table S9.5. Modified staple strands necessary to immobilize the L-shaped DNA origami structures to the surface of the prism via a biotin-avidin-biotin interaction. All oligonucleotides were purchased from biomers. All sequences are given from the 5' to 3' end.

| ID | Sequence (5' to 3') | Function | Replace |
|--------|---|-------------------------|---------|
| Bio-L1 | Biotin -ATCCAGAACAAATTAGTCATCAGGAACGGT | Biotin attachment at 5' | L1 |
| Bio-L2 | Biotin -CGTGCCTGTTCTCGCATCCAGCGCCGGGTTA | Biotin attachment at 5' | L2 |
| Bio-L3 | Biotin -ATAATCAGAAAAGCCCAACATCCACTGTAATA | Biotin attachment at 5' | L3 |
| Bio-L4 | Biotin -CATAGGTCTGAGAGACAAATCGTGAATTACC | Biotin attachment at 5' | L4 |

9.4. Overview of DNA Origami structures with small clock design



DNA origami structures were designed with a smaller radius with approximately 6 nm distances between the protruding docking strands. The single-stranded pointer, labeled with either Cy3B or Atto542, exits the origami in the middle between the three docking strands. To investigate the residence time of the tether at the three different states, an additional acceptor fluorophore, Cy5, was attached to the DNA origami surface between the 5 and 9 o'clock positions. The complementary base-pairing sequence at the 3' end of

the linker strands is 7 nt for the Cy3B/Cy5 labeled structure and 8 nt for the Atto542/Cy5 labeled structure.

Supplementary Table S9.6. Nomenclature of L-shaped DNA origami structures with a symmetric distribution of docking strands and a small clock diameter and strand IDs of incorporated modifications.

| Name | Green | Yellow | Red | Catchings Sites | | | Replacements |
|-------------------------|--------|--------|------|-----------------|-----------|-----------|--------------|
| YR-Pos1/5/9s-7nt | --- | 561-L7 | 640s | 7nt-Pos1s | 7nt-Pos5s | 7nt-Pos9s | Repl-1-10s |
| GR-Pos1/5/9s-8nt | 542-L7 | --- | 640s | 8nt-Pos1s | 8nt-Pos5s | 8nt-Pos9s | Repl-1-10s |

9.5. Staple strands of DNA Origami structures for the small clock design

DNA Origami structures with the symmetric distribution of docking strands and smaller diameter were derived from the L-shaped origami structure folded from strands given in **Supplementary Note SN9.1**. Clock-specific modifications are summarized in **Supplementary Table S9.7**. For the folding, the staples labeled with L5, L6, L8-L11 as well as L15-L23 were left out and replaced by strands for labeling, pointer, binding sites and unlabeled, connecting oligos. Docking sites with a 7 nt and 8 nt overhang, respectively, were introduced at position L1, L5 and L9. The pointer was introduced at Position L7 labeled with Cy3B or Atto542. For reading out the position of the pointer via smFRET, the designed DNA origami structures were labeled additionally via Cy5 (640S). For surface immobilization, biotin-labeled oligonucleotides Bio-L5-L8 were added (**Supplementary Table S9.8**).

Supplementary Table S9.7. Modified staple strands needed for producing the L-shaped DNA origami structure with a smaller diameter and a symmetric distribution of binding sites. To incorporate the ssDNA strands encoding the pointer, binding sites, and smFRET acceptor, the original staple strands L5, L6, L8-L11 as well as L15-L23 in the unmodified L-origami design were replaced with the corresponding sequences given below. All replacing sequences have been purchased by Eurofins Genomics. All sequences are given from the 5' to 3' end.

| ID | Sequence (5' to 3') | Function |
|-----------------------|--|----------------------------|
| Labeling sites | | |
| 561-L7 | GGCACCAAAACCAAAAGTAAGAGCAACACTATAGCA ACGTAATCGCCTTTTTTTT CGGGCATT - Cy3B | Pointer - dye at 3' |
| 542-L7 | GGCACCAAAACCAAAAGTAAGAGCAACACTATAGCA ACGTAATCGCCTTTTTTTT CGGGCATT - Atto542 | Pointer - dye at 3' |
| 64Os | AGAACACAGCTTAGAAGGAAGAAAAATCTACGATTTA - Cy5 | Cy5 at 3' |
| Replacement | | |
| Repl-1s | GAAAAGAGTACCTTAATTGTTAATTGGACCATAA | Replacement |
| Repl-2s | ATGCGTTAATAAACGAACTATATTCAATTGGCTTTG | Replacement |
| Repl-3s | AGCATATAACAGTTGATTAGCT | Replacement |
| Repl-4s | TGCTCATTCACTGAAATTACGAGGCATATAGCGAGAGAATCCCC | Replacement |
| Repl-5s | GTAATCGCAAAGACAAATTA | Replacement |
| Repl-6s | CGGAATCTCAGGCTGTTAAATATGCATGCG | Replacement |
| Repl-7s | ATATTCACTCATCTTGACCCGTAAATGCCATCGGAAC | Replacement |
| Repl-8s | GCAAGTCACCAATGAAACCATTGACAGGCAAATCA | Replacement |
| Repl-9s | GCACCCCTCGTCAGGTACGTTAGTAAATGAATAGTT | Replacement |
| Repl-10s | ATATTCACCGCCAGCATCGATAGCAGCACCGTAAAATACGTTTGCT | Replacement |
| Binding sites | | |
| 7nt-Pos5s | AACGAATCATTGTAATTACCTTTTT CAAATGC | Pos 5s catching site 7 nt |
| 7nt-Pos9s | GGCACCAAAACCAAAAGTAAGAGCAACACTATAGCAACTTTTT CAAATGC | Pos 5s catching site 7 nt |
| 7nt-Pos1s | AGCGTCAATCATAAGGGAACCGGTTTTT CAAATGC | Pos 5s catching site 7 nt |
| 8nt-Pos5s | AACGAATCATTGTAATTACCTTTTTTT CAAATGC | Pos 11s catching site 8 nt |
| 8nt-Pos9s | GGCACCAAAACCAAAAGTAAGAGCAACACTATAGCAACTTTTT CAAATGC | Pos 9s catching site 8 nt |
| 8nt-Pos1s | AGCGTCAATCATAAGGGAACCGGTTTTT CAAATGC | Pos 11s catching site 8 nt |

Supplementary Table S9.8. Modified staple strands needed for immobilization of the L-shaped DNA origami structures to the prism surface. For immobilization, a biotin-avidin-biotin linkage was used. All oligonucleotides were purchased from biomers. All sequences are given from the 5' to 3' end.

| ID | Sequence (5' to 3') | Function | Replace |
|----------------|---|-------------------------|---------|
| Bio-L16 | Biotin -ATAAAAATATCGCGTTCTCCTTTGATAAGAGCTATAT | Attachment Biotin at 5' | L16 |
| Bio-L17 | Biotin -TACCACTAACGCTAACAGTTGCTATTTGCACCCCATCCT | Attachment Biotin at 5' | L17 |
| Bio-L18 | Biotin -GAGGGTAGTTGCAGGGTGCTAACAACTTTCACGCCTGGAAAG AG | Attachment Biotin at 5' | L18 |
| Bio-L19 | Biotin -AGAGCCGCAAACAAATGAGACTCCTCAAGAGATTAGCGGGCAG TAGCA | Attachment Biotin at 5' | L19 |

Supplementary References

1. Hubner, K., et al., *Determining the In-Plane Orientation and Binding Mode of Single Fluorescent Dyes in DNA Origami Structures*. ACS Nano, 2021. **15**(3): p. 5109-5117.
2. Hubner, K., et al., *Salt-induced conformational switching of a flat rectangular DNA origami structure*. Nanoscale, 2022. **14**(21): p. 7898-7905.
3. Kalinin, S., et al., *Detection of structural dynamics by FRET: a photon distribution and fluorescence lifetime analysis of systems with multiple states*. J Phys Chem B, 2010. **114**(23): p. 7983-95.
4. Tomov, T.E., et al., *Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis*. Biophys J, 2012. **102**(5): p. 1163-73.
5. Krause, S., et al., *Graphene-on-Glass Preparation and Cleaning Methods Characterized by Single-Molecule DNA Origami Fluorescent Probes and Raman Spectroscopy*. ACS Nano, 2021. **15**(4): p. 6430-6438.
6. Kaminska, I., et al., *Graphene Energy Transfer for Single-Molecule Biophysics, Biosensing, and Super-Resolution Microscopy*. Adv Mater, 2021. **33**(24): p. e2101099.
7. Masullo, L.A., et al., *Pulsed Interleaved MINFLUX*. Nano Lett, 2021. **21**(1): p. 840-846.
8. Wanninger, S., et al., *Deep-LASI: deep-learning assisted, single-molecule imaging analysis of multi-color DNA origami structures*. Nat Commun, 2023. **14**(1): p. 6564.

Acknowledgments

Moving to Germany from Iran to pursue my PhD has profoundly shaped both my professional and personal growth. Over the past five years, I have not only achieved significant academic and scientific milestones, but I have also gained invaluable life experiences that have broadened my perspective on the world. This transformative journey would not have been possible without the support and collaboration of a remarkable group of intelligent and inspiring individuals. I feel deeply fortunate to be a part of such a dynamic and distinguished environment, and I am truly grateful for the opportunity to learn and grow here.

First and foremost, I want to sincerely thank you Don first for your patience and support throughout the long visa process. I am deeply grateful for the opportunity you provided me to join your group for my PhD despite my lack of prior experience in optics. Your patience in teaching me the fundamentals of optics and guiding me through the use of microscopy equipment—from lasers to detectors—gave me the confidence to manage and even further develop the advanced TIRF microscope. I truly appreciate the valuable insights you have shared, and for showing me how to approach problems with critical thinking and a mindset of continual learning.

I am deeply grateful to Philip Tinnefeld for the excellent collaboration between our two groups. Your insights and openness were invaluable to the success of our big DNA origami project. Thank you for providing decent origami samples so efficiently and incorporating our new labeling and assembly ideas in no time. I would also like to extend my sincere thanks to all my PhD defense committee for their time, expertise and thoughtful feedback, which have illuminated new perspectives on my research journey.

I am deeply grateful to you Evelyn, for all that I have learned under your guidance. You taught me crucial skills, from surface passivation and working with single-molecule sensitivity to mastering measurements on MFD setups and handling the complexities of three-color FRET data. Your initial ideas on the DNA origami project were pivotal in shaping the direction of my entire PhD. I truly appreciate how you consistently pushed the projects I was involved in, ultimately leading to the successful publications we were able to achieve together. Not to forget, thanks a lot for pushing Starnberg people for me!

Thank you, Simon, for your exceptional programming skills in developing the extensive Deep-LASI software and providing user-friendly interfaces that allowed me to efficiently navigate and analyze the continuous data streams. Your calm and positive spirit as a colleague greatly contributed to our success in tackling the complex three-color FRET data analysis!

I want to express my gratitude to Dr. Moritz Ehrl for his patience and support throughout the challenging process of securing my German residency permit at various stages during my PhD. Renewing my expiring permits was certainly a complex task that required significant time and effort from both of us. I truly appreciate the extensive work you put into preparing all the necessary documents.

I also want to thank you Ecenaz for being such a delightful and energetic officemate. I thoroughly enjoyed sharing the office with you, especially those times when it was just the two of us working with your favorite radio station playing cool music. Also traveling through the US together for the Biophysical Society meeting was a wonderful experience and an opportunity I greatly cherished.

I would also like to extend my gratitude to the entire wonderful AK Lamb team from both generations I've worked with. Thank you for creating such a warm and pleasant working environment and for making our lunch discussions both enjoyable and informative. Every gathering and celebration with you has been truly memorable and special.

Last but certainly not least, I would like to express my deep gratitude to my beloved family. Your unwavering support and love have been significant in helping me navigate the long and challenging journey of my PhD. Thank you for always believing in me and wishing the best for me. Thank you dad (baba Abas) for being my first teacher in life, I have learned a lot from you. I feel incredibly blessed to also have had my wonderful mom, Soraya, my brother, Peyman, and my sister, Paria by my side during the final stages of my PhD thesis and defense. And Paria, not everyone is as fortunate as I am to have a wonderful sibling close by to share experiences and feelings with. Your presence made the distance from home feel much smaller.

Bibliography

1. Maurice, J., A.M. Lett, C. Skinner, A. Lim, M. Richardson, A.P. Thomas, P.A. Summers, K. Vyas, A.W. Tadbier, R. Vilar, M.K. Kuimova, S. Miodragovic, N. Vergis, P. Kelly, M.F. Cordeiro, J. Hoare, A. Darzi, R. Goldin, M. Thursz, and A.J. Thompson, *Transcutaneous fluorescence spectroscopy as a tool for non-invasive monitoring of gut function: first clinical experiences*. *Sci Rep*, 2020. **10**(1): p. 16169.
2. Truong, L. and A.R. Ferre-D'Amare, *From fluorescent proteins to fluorogenic RNAs: Tools for imaging cellular macromolecules*. *Protein Sci*, 2019. **28**(8): p. 1374-1386.
3. Lerner, E., A. Barth, J. Hendrix, B. Ambrose, V. Birkedal, S.C. Blanchard, R. Borner, H. Sung Chung, T. Cordes, T.D. Craggs, A.A. Deniz, J. Diao, J. Fei, R.L. Gonzalez, I.V. Gopich, T. Ha, C.A. Hanke, G. Haran, N.S. Hatzakis, S. Hohng, S.C. Hong, T. Hugel, A. Ingargiola, C. Joo, A.N. Kapanidis, H.D. Kim, T. Laurence, N.K. Lee, T.H. Lee, E.A. Lemke, E. Margeat, J. Michaelis, X. Michalet, S. Myong, D. Nettels, T.O. Peulen, E. Ploetz, Y. Razvag, N.C. Robb, B. Schuler, H. Soleimaninejad, C. Tang, R. Vafabakhsh, D.C. Lamb, C.A. Seidel, and S. Weiss, *FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices*. *Elife*, 2021. **10**(60416).
4. Barth, A., L. Voith von Voithenberg, and D.C. Lamb, *Quantitative Single-Molecule Three-Color Forster Resonance Energy Transfer by Photon Distribution Analysis*. *J Phys Chem B*, 2019. **123**(32): p. 6901-6916.
5. Ploetz, E., E. Lerner, F. Husada, M. Roelfs, S. Chung, J. Hohlbein, S. Weiss, and T. Cordes, *Forster resonance energy transfer and protein-induced fluorescence enhancement as synergetic multi-scale molecular rulers*. *Sci Rep*, 2016. **6**: p. 33257.
6. Hellenkamp, B., S. Schmid, O. Doroshenko, O. Opanasyuk, R. Kuhnemuth, S. Rezaei Adariani, B. Ambrose, M. Aznauryan, A. Barth, V. Birkedal, M.E. Bowen, H. Chen, T. Cordes, T. Eilert, C. Fijen, C. Gebhardt, M. Gotz, G. Gouridis, E. Gratton, T. Ha, P. Hao, C.A. Hanke, A. Hartmann, J. Hendrix, L.L. Hildebrandt, V. Hirschfeld, J. Hohlbein, B. Hua, C.G. Hubner, E. Kallis, A.N. Kapanidis, J.Y. Kim, G. Krainer, D.C. Lamb, N.K. Lee, E.A. Lemke, B. Levesque, M. Levitus, J.J. McCann, N. Naredi-Rainer, D. Nettels, T. Ngo, R. Qiu, N.C. Robb, C. Rocker, H. Sanabria, M. Schlierf, T. Schroder, B. Schuler, H. Seidel, L. Streit, J. Thurn, P. Tinnefeld, S. Tyagi, N. Vandenberk, A.M. Vera, K.R. Weninger, B. Wunsch, I.S. Yanez-Orozco, J. Michaelis, C.A.M. Seidel, T.D. Craggs, and T. Hugel, *Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study*. *Nat Methods*, 2018. **15**(9): p. 669-676.
7. Weiss, S., *Fluorescence spectroscopy of single biomolecules*. *Science*, 1999. **283**(5408): p. 1676-83.
8. Kilic, S., S. Felekyan, O. Doroshenko, I. Boichenko, M. Dimura, H. Vardanyan, L.C. Bryan, G. Arya, C.A.M. Seidel, and B. Fierz, *Single-molecule FRET reveals multiscale chromatin dynamics modulated by HP1alpha*. *Nat Commun*, 2018. **9**(1): p. 235.
9. Kim, J.Y., C. Kim, and N.K. Lee, *Real-time submillisecond single-molecule FRET dynamics of freely diffusing molecules with liposome tethering*. *Nat Commun*, 2015. **6**: p. 6992.
10. Agam, G., C. Gebhardt, M. Popara, R. Machtel, J. Folz, B. Ambrose, N. Chamachi, S.Y. Chung, T.D. Craggs, M. de Boer, D. Grohmann, T. Ha, A. Hartmann, J. Hendrix, V. Hirschfeld, C.G. Hubner, T. Hugel, D. Kammerer, H.S. Kang, A.N. Kapanidis, G. Krainer, K. Kramm, E.A. Lemke, E. Lerner, E. Margeat, K. Martens, J. Michaelis, J. Mitra, G.G. Moya Munoz, R.B. Quast, N.C. Robb, M. Sattler, M. Schlierf, J. Schneider, T. Schroder, A. Sefer, P.S. Tan, J. Thurn, P. Tinnefeld, J. van Noort, S. Weiss, N. Wendler, N. Zijlstra, A. Barth, C.A.M. Seidel, D.C. Lamb, and T. Cordes, *Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins*. *Nat Methods*, 2023. **20**(4): p. 523-535.

11. Gotz, M., A. Barth, S.S. Bohr, R. Borner, J. Chen, T. Cordes, D.A. Erie, C. Gebhardt, M. Hadzic, G.L. Hamilton, N.S. Hatzakis, T. Hugel, L. Kisley, D.C. Lamb, C. de Lannoy, C. Mahn, D. Dunukara, D. de Ridder, H. Sanabria, J. Schimpf, C.A.M. Seidel, R.K.O. Sigel, M.B. Sletfjording, J. Thomsen, L. Vollmar, S. Wanninger, K.R. Weninger, P. Xu, and S. Schmid, *A blind benchmark of analysis tools to infer kinetic rate constants from single-molecule FRET trajectories*. *Nat Commun*, 2022. **13**(1): p. 5402.
12. Kudryavtsev, V., M. Sikor, S. Kalinin, D. Mokranjac, C.A. Seidel, and D.C. Lamb, *Combining MFD and PIE for accurate single-pair Forster resonance energy transfer measurements*. *Chemphyschem*, 2012. **13**(4): p. 1060-78.
13. Kapanidis, A.N.L., N. K.; Laurence, T. A.; Doose, S.; Margeat, E.; Weiss, S., *Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules*. *PNAS*, 2004. **101**(24): p. 8936-8941.
14. Schmid, S., M. Gotz, and T. Hugel, *Single-Molecule Analysis beyond Dwell Times: Demonstration and Assessment in and out of Equilibrium*. *Biophys J*, 2016. **111**(7): p. 1375-1384.
15. Kapanidis, A.N., Laurence, T. A., Lee, N. K., Margeat, E., Kong, X., Weiss, S., *Alternating Laser Excitation of Single Molecules*. *Acc Chem Res*, 2005. **38**(7): p. 523-33.
16. Lee, N.K., Kapanidis, A. N., Koh, H. R., Korlann, Y., Ho, S. O., Kim, Y., Gassman, N., Kim, S. K., Weiss, S., *Three-Color Alternating-Laser Excitation of Single Molecules: Monitoring Multiple Interactions and Distances*. *Biophys J*, 2007. **92**(1): p. 303-312.
17. Douglas, S.M., H. Dietz, T. Liedl, B. Hogberg, F. Graf, and W.M. Shih, *Self-assembly of DNA into nanoscale three-dimensional shapes*. *Nature*, 2009. **459**(7245): p. 414-8.
18. Buber, E., T. Schroder, M. Scheckenbach, M. Dass, H.G. Franquelim, and P. Tinnefeld, *DNA Origami Curvature Sensors for Nanoparticle and Vesicle Size Determination with Single-Molecule FRET Readout*. *ACS Nano*, 2023. **17**(3): p. 3088-3097.
19. R. Kosinski, J.M.P., E. C. Schöneweiß, Y. B. Ruiz-Blanco, I. Ponzo, K. Bravo-Rodriguez, M. Erkelenz, S. Schlücker, G. Uhlenbrock, E. Sanchez-Garcia, B. Saccà, *The role of DNA nanostructures in the catalytic properties of an allosterically regulated protease*. *Science Advances*, 2022. **8**.
20. Sacca, B., R. Meyer, M. Erkelenz, K. Kiko, A. Arndt, H. Schroeder, K.S. Rabe, and C.M. Niemeyer, *Orthogonal protein decoration of DNA origami*. *Angew Chem Int Ed Engl*, 2010. **49**(49): p. 9378-83.
21. Ijas, H., I. Hakaste, B. Shen, M.A. Kostiainen, and V. Linko, *Reconfigurable DNA Origami Nanocapsule for pH-Controlled Encapsulation and Display of Cargo*. *ACS Nano*, 2019. **13**(5): p. 5959-5967.
22. Lakowicz, J.R., *Principles of Fluorescence Spectroscopy*. 3 ed. 2010: Springer.
23. Duan, Y.-C., L.-L. Wen, Y. Gao, Y. Wu, L. Zhao, Y. Geng, G.-G. Shan, M. Zhang, and Z.-M. Su, *Fluorescence, Phosphorescence, or Delayed Fluorescence?—A Theoretical Exploration on the Reason Why a Series of Similar Organic Molecules Exhibit Different Luminescence Types*. *The Journal of Physical Chemistry C*, 2018. **122**(40): p. 23091-23101.
24. Baleizao, C., Berberan-Santos, M. N., *Thermally activated delayed fluorescence in fullerenes*. *Ann N Y Acad Sci*, 2008. **1130**: p. 224-34.
25. PARSON, W.W., *Modern Optical Spectroscopy*. 2009: Springer. 529.
26. Penfold, T.J., E. Gindensperger, C. Daniel, and C.M. Marian, *Spin-Vibronic Mechanism for Intersystem Crossing*. *Chem Rev*, 2018. **118**(15): p. 6975-7025.
27. Salem, C.-B., E. Ploetz, and D.C. Lamb, *Probing dynamics in single molecules*, in *Spectroscopy and Dynamics of Single Molecules*. 2019. p. 71-115.
28. Condon, E., U., *The Franck Condon Principle and Related Topics*. *American Journal of Physics*, 1947. **15**.
29. Stokes, G., G., *On the Change of Refrangibility of Light*. *Philosophical Transactions of the Royal Society*, 1852. **142**.

30. Pashley-Johnson, F., R. Munaweera, S.I. Hossain, S.C. Gauci, L. Delafresnaye, H. Frisch, M.L. O'Mara, F.E. Du Prez, and C. Barner-Kowollik, *How molecular architecture defines quantum yields*. Nat Commun, 2024. **15**(1): p. 6033.
31. Michalet, X., S. Weiss, and M. Jager, *Single-molecule fluorescence studies of protein folding and conformational dynamics*. Chem Rev, 2006. **106**(5): p. 1785-813.
32. Loretan, M., I. Domljanovic, M. Lakatos, C. Ruegg, and G.P. Acuna, *DNA Origami as Emerging Technology for the Engineering of Fluorescent and Plasmonic-Based Biosensors*. Materials (Basel), 2020. **13**(9).
33. Mutze, J., Iyer, V., Macklin, J. J., Colonell, J., Karsh, B., Petrasek, Z., Schwille, P., Looger, L. L., Lavis, L. D., Harris, T. D., *Excitation spectra and brightness optimization of two-photon excited probes*. Biophys J, 2012. **102**(4): p. 934-44.
34. Tian, Y., Halle, J., Wojdyl, M., Sahoo, D., Scheblykin, I. G., *Quantitative measurement of fluorescence brightness of single molecules*. Methods Appl Fluoresc, 2014. **2**(3): p. 035003.
35. Muller, B.K., Zaychikov, E., Brauchle, C., Lamb, D. C., *Pulsed interleaved excitation*. Biophys J, 2005. **89**(5): p. 3508-22.
36. Schrimpf, W., A. Barth, J. Hendrix, and D.C. Lamb, *PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data*. Biophys J, 2018. **114**(7): p. 1518-1528.
37. Pal, N., *Single-Molecule FRET: A Tool to Characterize DNA Nanostructures*. Front Mol Biosci, 2022. **9**: p. 835617.
38. Deniz, A.A.D., M.; Grunwell, J. R.; Ha, T.; faulhaber, A. E.; Chemla, D. S.; Weiss, S.; Schultz, P. G., *Single-pair fluorescence resonance energy transfer on freely diffusing molecules: Observation of Förster distance dependence and subpopulations*. Proc. Natl. Acad. Sci. USA, 1999. **96**: p. 3670–3675.
39. Bandyopadhyay, D. and P.P. Mishra, *Decoding the Structural Dynamics and Conformational Alternations of DNA Secondary Structures by Single-Molecule FRET Microspectroscopy*. Front Mol Biosci, 2021. **8**: p. 725541.
40. Bartels, K., T. Lasitza-Male, H. Hofmann, and C. Low, *Single-Molecule FRET of Membrane Transport Proteins*. Chembiochem, 2021. **22**(17): p. 2657-2671.
41. Hartmann, A., K. Sreenivasa, M. Schenkel, N. Chamachi, P. Schake, G. Krainer, and M. Schlierf, *An automated single-molecule FRET platform for high-content, multiwell plate screening of biomolecular conformations and dynamics*. Nat Commun, 2023. **14**(1): p. 6511.
42. Kong, M. and E.C. Greene, *Mechanistic Insights From Single-Molecule Studies of Repair of Double Strand Breaks*. Front Cell Dev Biol, 2021. **9**: p. 745311.
43. Lichtman, J.W. and J.A. Conchello, *Fluorescence microscopy*. Nat Methods, 2005. **2**(12): p. 910-9.
44. Dunst, S. and P. Tomancak, *Imaging Flies by Fluorescence Microscopy: Principles, Technologies, and Applications*. Genetics, 2019. **211**(1): p. 15-34.
45. Jensen, E.C., *Types of imaging, Part 2: an overview of fluorescence microscopy*. Anat Rec (Hoboken), 2012. **295**(10): p. 1621-7.
46. St Croix, C.M., Shand, S. H., Watkins, S. C., *Confocal microscopy: comparisons, applications, and problems*. Biotechniques, 2005. **39**(6 Suppl): p. S2-5.
47. Nakano, A., *Spinning-disk Confocal Microscopy — A Cutting-Edge Tool for Imaging of Membrane Traffic*. Cell Struct Funct., 2002.
48. Xu, C., M. Nedergaard, D.J. Fowell, P. Friedl, and N. Ji, *Multiphoton fluorescence microscopy for in vivo imaging*. Cell, 2024. **187**(17): p. 4458-4487.
49. Wu, J., N. Ji, and K.K. Tsia, *Speed scaling in multiphoton fluorescence microscopy*. Nature Photonics, 2021. **15**(11): p. 800-812.
50. Sanderson, J., *Multi-Photon Microscopy*. Curr Protoc, 2023. **3**(1): p. e634.

51. Ward, A.E., V. Kiessling, O. Pornillos, J.M. White, B.K. Ganser-Pornillos, and L.K. Tamm, *HIV-cell membrane fusion intermediates are restricted by Serincs as revealed by cryo-electron and TIRF microscopy*. *J Biol Chem*, 2020. **295**(45): p. 15183-15195.
52. Joyce, P., S. Joemetsa, S. Isaksson, S. Hossain, P. Larsson, C. Bergstrom, and F. Hook, *TIRF Microscopy-Based Monitoring of Drug Permeation Across a Lipid Membrane Supported on Mesoporous Silica*. *Angew Chem Int Ed Engl*, 2021. **60**(4): p. 2069-2073.
53. Mattheyses, A.L., S.M. Simon, and J.Z. Rappoport, *Imaging with total internal reflection fluorescence microscopy for the cell biologist*. *J Cell Sci*, 2010. **123**(Pt 21): p. 3621-8.
54. Eigen M., R., R., *Sorting single molecules: Application to diagnostics and evolutionary biotechnology*. *Proc. Natl. Acad. Sci.*, 1994. **13**.
55. Kohl, T., Schwille, P., *Fluorescence correlation spectroscopy with autofluorescent proteins*. *Adv Biochem Eng Biotechnol*, 2005. **95**: p. 107-42.
56. Day, C.A. and M. Kang, *The Utility of Fluorescence Recovery after Photobleaching (FRAP) to Study the Plasma Membrane*. *Membranes (Basel)*, 2023. **13**(5).
57. Sobakinskaya, E., M. Schmidt Am Busch, and T. Renger, *Theory of FRET "Spectroscopic Ruler" for Short Distances: Application to Polyproline*. *J Phys Chem B*, 2018. **122**(1): p. 54-67.
58. Schneckenburger, H., P. Weber, M. Wagner, S. Enderle, B. Kalthof, L. Schneider, C. Herzog, J. Weghuber, and P. Lanzerstorfer, *Combining TIR and FRET in Molecular Test Systems*. *Int J Mol Sci*, 2019. **20**(3).
59. Schirripa Spagnolo, C. and S. Luin, *Choosing the Probe for Single-Molecule Fluorescence Microscopy*. *Int J Mol Sci*, 2022. **23**(23).
60. Marras, A.E., L. Zhou, H.J. Su, and C.E. Castro, *Programmable motion of DNA origami mechanisms*. *Proc Natl Acad Sci U S A*, 2015. **112**(3): p. 713-8.
61. Rothemund, P.W., *Folding DNA to create nanoscale shapes and patterns*. *Nature*, 2006. **440**(7082): p. 297-302.
62. Woo, S. and P.W.K. Rothemund, *Self-assembly of two-dimensional DNA origami lattices using cation-controlled surface diffusion*. *Nature Communications*, 2014. **5**(1).
63. Weiden, J. and M.M.C. Bastings, *DNA origami nanostructures for controlled therapeutic drug delivery*. *Current Opinion in Colloid & Interface Science*, 2021. **52**.
64. Xavier, P.L. and A.R. Chandrasekaran, *DNA-based construction at the nanoscale: emerging trends and applications*. *Nanotechnology*, 2018. **29**(6): p. 062001.
65. Kumar, M., A. Jha, and B. Mishra, *DNA-Based Nanostructured Platforms as Drug Delivery Systems*. *Chem & Bio Engineering*, 2024. **1**(3): p. 179-198.
66. Selnihhin, D., S.M. Sparvath, S. Preus, V. Birkedal, and E.S. Andersen, *Multifluorophore DNA Origami Beacon as a Biosensing Platform*. *ACS Nano*, 2018. **12**(6): p. 5699-5708.
67. Ochmann, S.E., H. Joshi, E. Buber, H.G. Franquelim, P. Stegemann, B. Sacca, U.F. Keyser, A. Aksimentiev, and P. Tinnefeld, *DNA Origami Voltage Sensors for Transmembrane Potentials with Single-Molecule Sensitivity*. *Nano Lett*, 2021. **21**(20): p. 8634-8641.
68. Wang, Y., I. Baars, I. Berzina, I. Rocamonde-Lago, B. Shen, Y. Yang, M. Lolaico, J. Waldvogel, I. Smyrlaki, K. Zhu, R.A. Harris, and B. Höglberg, *A DNA robotic switch with regulated autonomous display of cytotoxic ligand nanopatterns*. *Nature Nanotechnology*, 2024. **19**(9): p. 1366-1374.
69. Ramezani, H. and H. Dietz, *Building machines with DNA molecules*. *Nat Rev Genet*, 2020. **21**(1): p. 5-26.
70. Liu, X., F. Zhang, X. Jing, M. Pan, P. Liu, W. Li, B. Zhu, J. Li, H. Chen, L. Wang, J. Lin, Y. Liu, D. Zhao, H. Yan, and C. Fan, *Complex silica composite nanomaterials templated with DNA origami*. *Nature*, 2018. **559**(7715): p. 593-598.
71. Shen, B., Linko, V., Tappio, K., Pikker, S., Lemma, T., Gopinath, A., Gothelf, K., V., Kostiainen, M., A., Toppari, J., J., *Plasmonic nanostructures through DNA-assisted lithography*. *Sci. Adv.* , 2018. **4**.

72. Teschome, B., S. Facsko, T. Schonherr, J. Kerbusch, A. Keller, and A. Erbe, *Temperature-Dependent Charge Transport through Individually Contacted DNA Origami-Based Au Nanowires*. *Langmuir*, 2016. **32**(40): p. 10159-10165.

73. Bonde, S., R.A.M. Osmani, R. Trivedi, V. Patravale, M. Angolkar, A.G. Prasad, and A.A. Ravikumar, *Harnessing DNA origami's therapeutic potential for revolutionizing cardiovascular disease treatment: A comprehensive review*. *Int J Biol Macromol*, 2024. **270**(Pt 1): p. 132246.

74. Amir, Y., E. Ben-Ishay, D. Levner, S. Ittah, A. Abu-Horowitz, and I. Bachelet, *Universal computing by DNA origami robots in a living animal*. *Nature Nanotechnology*, 2014. **9**(5): p. 353-357.

75. Wang, D., Y. Fu, J. Yan, B. Zhao, B. Dai, J. Chao, H. Liu, D. He, Y. Zhang, C. Fan, and S. Song, *Molecular logic gates on DNA origami nanostructures for microRNA diagnostics*. *Anal Chem*, 2014. **86**(4): p. 1932-6.

76. Kopperger, E., J. List, S. Madhira, F. Rothfischer, D.C. Lamb, and F.C. Simmel, *A self-assembled nanoscale robotic arm controlled by electric fields*. *Science*, 2018. **19**(6373): p. 296-301.

77. Yan, W., S. Li, M. Deguchi, Z. Zheng, D. Rus, and A. Mehta, *Origami-based integration of robots that sense, decide, and respond*. *Nat Commun*, 2023. **14**(1): p. 1553.

78. Andersen, E.S., M. Dong, M.M. Nielsen, K. Jahn, R. Subramani, W. Mamdouh, M.M. Golas, B. Sander, H. Stark, C.L. Oliveira, J.S. Pedersen, V. Birkedal, F. Besenbacher, K.V. Gothelf, and J. Kjems, *Self-assembly of a nanoscale DNA box with a controllable lid*. *Nature*, 2009. **459**(7243): p. 73-6.

79. Bujold, K.E., J.C.C. Hsu, and H.F. Sleiman, *Optimized DNA "Nanosuitcases" for Encapsulation and Conditional Release of siRNA*. *J Am Chem Soc*, 2016. **138**(42): p. 14030-14038.

80. Jiang, Q., Y. Shang, Y. Xie, and B. Ding, *DNA Origami: From Molecular Folding Art to Drug Delivery Technology*. *Adv Mater*, 2023: p. e2301035.

81. Koirala, D., P. Shrestha, T. Emura, K. Hidaka, S. Mandal, M. Endo, H. Sugiyama, and H. Mao, *Single-molecule mechanochemical sensing using DNA origami nanostructures*. *Angew Chem Int Ed Engl*, 2014. **53**(31): p. 8137-41.

82. Wang, D., C. Vietz, T. Schroder, G. Acuna, B. Lalkens, and P. Tinnefeld, *A DNA Walker as a Fluorescence Signal Amplifier*. *Nano Lett*, 2017. **17**(9): p. 5368-5374.

83. Liu, X.R., I.Y. Loh, W. Siti, H.L. Too, T. Anderson, and Z. Wang, *A light-operated integrated DNA walker-origami system beyond bridge burning*. *Nanoscale Horiz*, 2023. **8**(6): p. 827-841.

84. Kim, M., C. Lee, K. Jeon, J.Y. Lee, Y.J. Kim, J.G. Lee, H. Kim, M. Cho, and D.N. Kim, *Harnessing a paper-folding mechanism for reconfigurable DNA origami*. *Nature*, 2023. **619**(7968): p. 78-86.

85. Kucinic, A., C.M. Huang, J. Wang, H.J. Su, and C.E. Castro, *DNA origami tubes with reconfigurable cross-sections*. *Nanoscale*, 2023. **15**(2): p. 562-572.

86. Tang, Z., Z.X. Yin, X. Sun, J.Z. Cui, J. Yang, and R.S. Wang, *Dynamically NAND gate system on DNA origami template*. *Comput Biol Med*, 2019. **109**: p. 112-120.

87. Wang, S., Z. Zhou, N. Ma, S. Yang, K. Li, C. Teng, Y. Ke, and Y. Tian, *DNA Origami-Enabled Biosensors*. *Sensors (Basel)*, 2020. **20**(23).

88. Kogikoski, S., Jr., J. Ameixa, A. Mostafa, and I. Bald, *Lab-on-a-DNA origami: nanoengineered single-molecule platforms*. *Chem Commun (Camb)*, 2023. **59**(32): p. 4726-4741.

89. Stein, I.H., C. Steinhauer, and P. Tinnefeld, *Single-molecule four-color FRET visualizes energy-transfer paths on DNA origami*. *J Am Chem Soc*, 2011. **133**(12): p. 4193-5.

90. Stein, I.H., V. Schuller, P. Bohm, P. Tinnefeld, and T. Liedl, *Single-molecule FRET ruler based on rigid DNA origami blocks*. *Chemphyschem*, 2011. **12**(3): p. 689-95.

91. Dutta, P.K., R. Varghese, J. Nangreave, S. Lin, H. Yan, and Y. Liu, *DNA-directed artificial light-harvesting antenna*. *J Am Chem Soc*, 2011. **133**(31): p. 11985-93.

92. Bartnik, K., A. Barth, M. Pilo-Pais, A.H. Crevenna, T. Liedl, and D.C. Lamb, *A DNA Origami Platform for Single-Pair Forster Resonance Energy Transfer Investigation of DNA-DNA Interactions and Ligation*. *J Am Chem Soc*, 2020. **142**(2): p. 815-825.

93. Cole, F., J. Zähringer, J. Bohlen, T. Schröder, F. Steiner, M. Pfeiffer, P. Schüler, F.D. Stefani, and P. Tinnefeld, *Super-resolved FRET and co-tracking in pMINFLUX*. *Nature Photonics*, 2024. **18**(5): p. 478-484.

94. Adamczyk, A., K., Huijben, T., A., P., M., Kolataj, K., Zhu, F., Marie, R., Stefani, F., D., Acuna, G., P., *Towards full control of molecular exciton energy transfer via FRET in DNA origami assemblies*. arXiv:2402.06292 [cond-mat.soft], 2024.

95. Wu, Y., R.D. Tilley, and J.J. Gooding, *Challenges and Solutions in Developing Ultrasensitive Biosensors*. *J Am Chem Soc*, 2019. **141**(3): p. 1162-1170.

96. Elliott, A.D., *Confocal Microscopy: Principles and Modern Practices*. *Current Protocols in Cytometry*, 2019. **92**(1).

97. Kühnemuth, R., Seidel, C. A. M., *Principles of Single Molecule Multiparameter Fluorescence Spectroscopy*. *Single Molecules*, 2001. **2**(4): p. 251-254.

98. Eyal Nir, X.M., Kambiz M. Hamadani, Ted A. Laurence, Daniel Neuhauser, Yevgeniy Kovchegov, and Shimon Weiss, *Shot-Noise Limited Single-Molecule FRET Histograms: Comparison between Theory and Experiments*. *The Journal of Physical Chemistry B*, 2006. **110**(44): p. 22103-22124.

99. Asadiatouei, P., J. Bohlen, F. Cole, F. Morella, P. Tinnefeld, E. Ploetz, and D.C. Lamb, *Distance and Kinetic Tunability of Dynamic DNA Origami Structures Examined at the Single-Molecule Level*. (In Preparation).

100. Veiksina, S., S. Kopanchuk, and A. Rinken, *Fluorescence anisotropy assay for pharmacological characterization of ligand binding dynamics to melanocortin 4 receptors*. *Anal Biochem*, 2010. **402**(1): p. 32-9.

101. Cheow, L.F., R. Viswanathan, C.S. Chin, N. Jennifer, R.C. Jones, E. Guccione, S.R. Quake, and W.F. Burkholder, *Multiplexed analysis of protein-ligand interactions by fluorescence anisotropy in a microfluidic platform*. *Anal Chem*, 2014. **86**(19): p. 9901-8.

102. Fish, K.N., *Total Internal Reflection Fluorescence (TIRF) Microscopy*. *Curr Protoc*, 2022. **2**(8): p. e517.

103. Fish, K.N., *Total internal reflection fluorescence (TIRF) microscopy*. *Curr Protoc Cytom*, 2009. **Chapter 12**: p. Unit12 18.

104. Li, Q., C.N. Hulleman, R.J. Moerland, E. Mailvaganam, S. Ganapathy, D. Brinks, S. Stallinga, and B. Rieger, *Waveguide-based total internal reflection fluorescence microscope enabling cellular imaging under cryogenic conditions*. *Opt Express*, 2021. **29**(21): p. 34097-34108.

105. Martin-Fernandez, M.L., C.J. Tynan, and S.E. Webb, *A 'pocket guide' to total internal reflection fluorescence*. *J Microsc*, 2013. **252**(1): p. 16-22.

106. Agnarsson, B., Ingthorsson, S., Gudjonsson, T., Leosson1, K., *Evanescence-wave fluorescence microscopy using symmetric planar waveguides*. *Opt Express*, 2009. **17**.

107. Priyadarshi, A., F.T. Dullo, D.L. Wolfson, A. Ahmad, N. Jayakumar, V. Dubey, J.-C. Tinguely, B.S. Ahluwalia, and G.S. Murugan, *A transparent waveguide chip for versatile total internal reflection fluorescence-based microscopy and nanoscopy*. *Communications Materials*, 2021. **2**(1).

108. Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, *Fiji: an open-source platform for biological-image analysis*. *Nat Methods*, 2012. **9**(7): p. 676-82.

109. Roy, R., S. Hohng, and T. Ha, *A practical guide to single-molecule FRET*. *Nat Methods*, 2008. **5**(6): p. 507-16.

110. Walter, N.G., C.Y. Huang, A.J. Manzo, and M.A. Sobhy, *Do-it-yourself guide: how to use the modern single-molecule toolkit*. *Nat Methods*, 2008. **5**(6): p. 475-89.

111. Lamichhane, R., Solem, A., Black, W., Rueda, D., *Single-molecule FRET of protein–nucleic acid and protein–protein complexes: Surface passivation and immobilization*. Methods, 2010. **52**(2): p. 192-200.
112. Asadiatouei, P., C.B. Salem, S. Wanninger, E. Ploetz, and D.C. Lamb, *Deep-LASI, single-molecule data analysis software*. Biophys J, 2024.
113. Wanninger, S., P. Asadiatouei, J. Bohlen, C.B. Salem, P. Tinnefeld, E. Ploetz, and D.C. Lamb, *Deep-LASI: deep-learning assisted, single-molecule imaging analysis of multi-color DNA origami structures*. Nat Commun, 2023. **14**(1): p. 6564.
114. Thomsen, J., M.B. Sletfjerding, S.B. Jensen, S. Stella, B. Paul, M.G. Malle, G. Montoya, T.C. Petersen, and N.S. Hatzakis, *DeepFRET, a software for rapid and automated single-molecule FRET data classification using deep learning*. Elife, 2020. **9**.
115. Sikor, M., K. Mapa, L.V. von Voithenberg, D. Mokranjac, and D.C. Lamb, *Real-time observation of the conformational dynamics of mitochondrial Hsp70 by spFRET*. EMBO J, 2013. **32**(11): p. 1639-49.
116. Ha, T.E., T.; Ogletree, DF.; Chemla, DS.; Selvin, PR.; Weiss, S., *Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor*. Proc. Natl. Acad. Sci. U S A, 1996: p. 6264-8.
117. Cole, F., M. Pfeiffer, D. Wang, T. Schroder, Y. Ke, and P. Tinnefeld, *Controlled mechanochemical coupling of anti-junctions in DNA origami arrays*. Nat Commun, 2024. **15**(1): p. 7894.
118. Aitken, C.E., Marshall, R. A., Puglisi, J. D., *An oxygen scavenging system for improvement of dye stability in single-molecule fluorescence experiments*. Biophys J, 2008. **94**(5): p. 1826-35.
119. Cordes, T.V., J.; Tinnefeld, P., *On the Mechanism of Trolox as Antiblinking and Antibleaching Reagent*. Journal of the American Chemical Society, 2009. **131**(14): p. 5018–5019.